Statistika (MMS-1001)

Dr. Danardono, MPH

danardono@ugm.ac.id

Program Studi Statistika

Jurusan Matematika FMIPA UGM

Materi dan Jadual

Tatap Muka	Pokok Bahasan	Sub Pokok Bahasan
1.	Statistika Deskriptif	1. Tentang perkuliahan MMS-1001 2. Peran Statistika 3. Terminologi 4. Representasi Grafik
2.	Statistika Deskriptif	1. Distribusi Frekuensi 2. Ukuran Tengah 3. Ukuran Dispersi
3.	Peluang dan Variabel Random	1. Kejadian dan Peluang 2. Variabel Random dan Distribusinya 3. Harga Harapan, Variansi dan Sifat-Sifatnya
4.	Distribusi Variabel Random Diskret dan Kontinu	1. Distribusi Variabel Random Diskret 2. Distribusi Variable Random Kontinu

Materi dan Jadual

5.	Distribusi Sampling Statistik	1. Distribusi Sampling Statistik untuk Rerata 2. Pendekatan Normal untuk Binomial
6.	Inferensi Statistik	1. Estimasi Parameter 2. Uji Hipotesis
7.	Inferensi Statistik Satu Populasi Sembarang	1. Interval Konfidensi Untuk Mean 2. Uji Hipotesa Mean 3. Interval Konfidensi Untuk Proporsi 4. Uji Hipotesis Proporsi
8.	*Review* dan Ringkasan	
9.	Inferensi Statistik Satu Populasi Normal	1. Interval konfidensi Untuk Mean 2. Uji Hipotesis Mean 3. Hubungan Interval Konfidensi dan uji Hipotesis

Materi dan Jadual

10.	Inferensi Statistik Dua Populasi Sembarang	1. Interval konfidensi Untuk Selisih Mean Dua Populasi 2. Uji Hipotesis Selisih Mean Dua populasi 3. Interval konfidensi Untuk Selisih Proporsi Dua Populasi 4. Uji Hipotesis Selisih Proporsi Dua populasi
11.	Inferensi Statistik Dua Populasi Normal	1. Interval Konfidensi Perbandingan Variansi Dua Populasi 2. Uji Hipotesis Perbandingan Variansi Dua Populasi 3. Interval konfidensi Selisih Mean Dua Populasi 4. Uji Hipotesis Selisih Mean Dua Populasi
12.	Analisis Variansi Satu Arah	1. Dasar-dasar ANAVA 2. Tabel ANAVA dan Uji F 3. Pembandingan Ganda

Materi dan Jadual

13.	Analisis Regresi Linear Sederhana	1. Dasar-dasar Model Regresi 2. Estimasi Model regresi 3. Analisis Korelasi 4. Inferensi dalam Regresi
14.	*Review* dan Ringkasan	

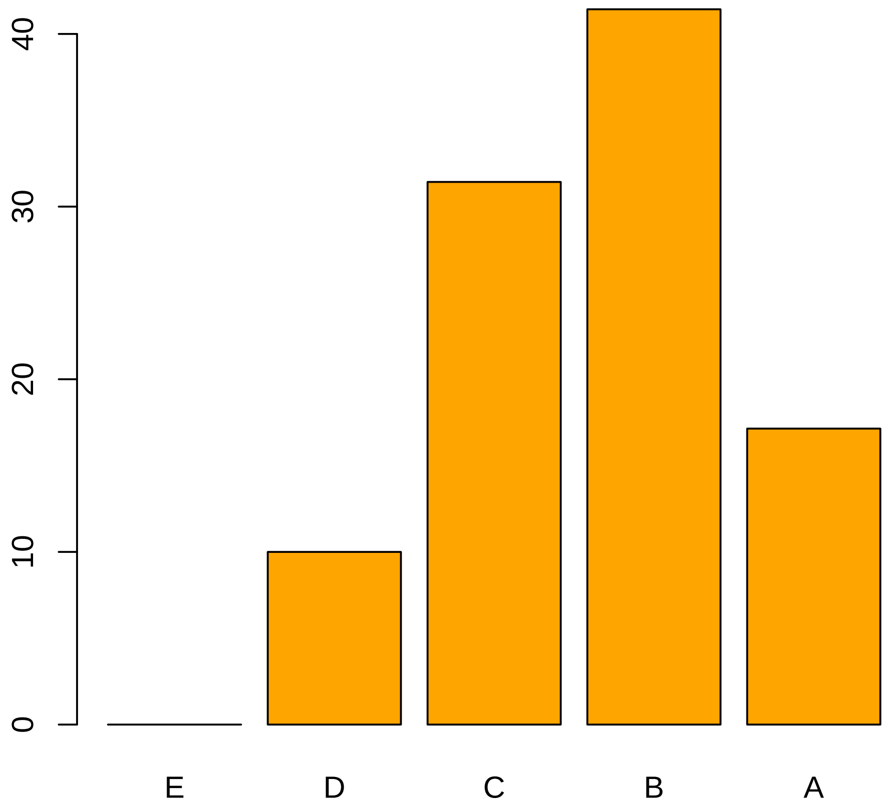
Buku teks:

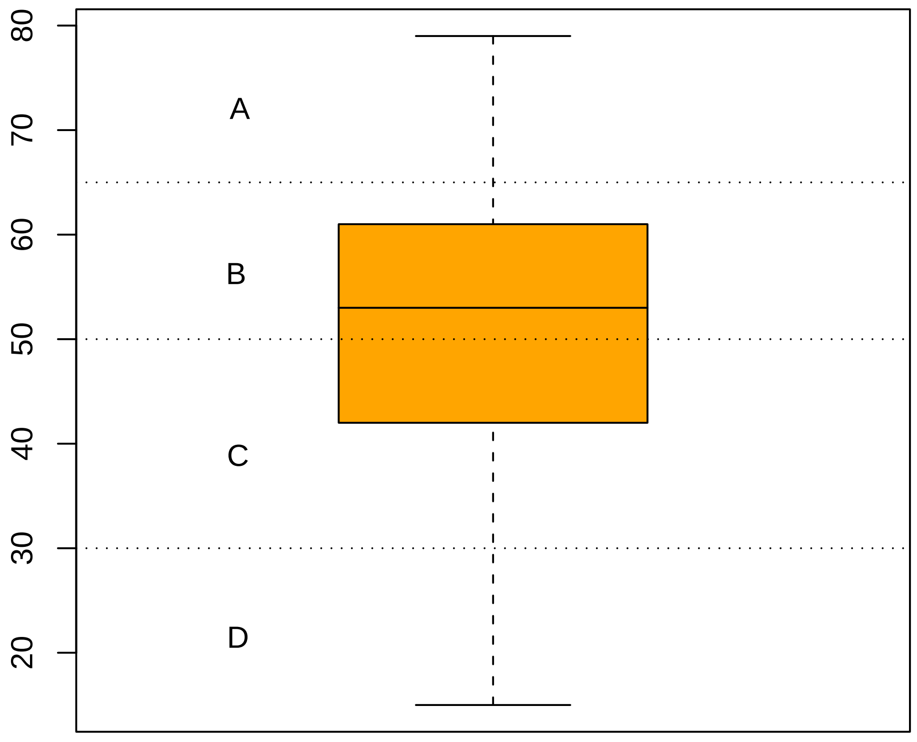
Soejoeti, Z. (1984). *Buku Materi Pokok Metode Statistik I*, Universitas Terbuka, Penerbit Karunika, Jakarta.

Gunardi, A. Rakhman (2004). *Metode Statistika*, FMIPA UGM, Yogyakarta.

Penilaian

Nilai MMS-1001 tahun 2005 (140 mahasiswa)





Penilaian

No	Unsur Penilaian	Prosentase
1.	Ujian Akhir	
2.	Sisipan	
3.	Tugas	
4.	Kuis	

Data

Penghasilan mingguan 40 buruh bangunan di suatu kota (dalam ribuan rupiah):

58	72	64	65	67	92	55	51	69	73
64	59	65	55	75	56	89	60	84	68
74	67	55	68	74	43	67	71	72	66
62	63	83	64	51	63	49	78	65	75

Data

Hasil pengukuran keasaman (PH) dari 35 kolam di suatu daerah:

6,4	6,6	6,2	7,2	6,2	8,1	7,0
7,0	5,9	5,7	7,0	7,4	6,5	6,8
7,0	7,0	6,0	6,3	5,6	6,3	5,8
5,9	7,2	7,3	7,7	6,8	5,2	5,2
6,4	6,3	6,2	7,5	6,7	6,4	7,8

Data

Tinggi (cm) dan berat badan (kg) 10 orang mahasiswa:

Mahasiswa	Tinggi	Berat
1	170	70
2	162	65
3	169	59
4	165	62
5	171	67
6	170	65
7	168	60
8	163	61
9	166	63
10	172	64

Data

Banyaknya penjualan telepon seluler di suatu toko:

Merek	Banyak penjualan
Sony-Ericsson	72
Motorola	60
Nokia	85
Samsung	54
LG	32
Siemens	64

Skala Pengukuran

Skala	Yang dapat ditentukan untuk dua pengamatan sembarang
Nominal	persamaan (klasifikasi)
Ordinal	persamaan dan urutan
Interval	persamaan, urutan dan jarak (satuan pengukuran)
Rasio	persamaan, urutan, jarak dan rasio (titik nol yang murni ada)

Skala Pengukuran

Contoh:

- Nominal: jenis pekerjaan, warna
- Ordinal: kepangkatan, tingkat pendidikan
- Interval: tahun kalender (Masehi, Hijriyah), temperatur (Celcius, Fahrenheit)
- Rasio: berat, panjang, isi

Statistika Deskriptif

Metode atau cara-cara yang digunakan untuk meringkas dan menyajikan data dalam bentuk tabel, grafik atau ringkasan numerik data.

Grafik *Stem-and-leaf*

- Untuk menunjukkan bentuk distribusi data
- Data berupa angka dengan minimal dua digit

Contoh (Data penghasilan buruh):

```
4 | 3 9
5 | 1 1 5 5 5 6 8 9
6 | 0 2 3 3 4 4 4 5 5 5 6 7 7 7 8 8 9
7 | 1 2 2 3 4 4 5 5 8
8 | 3 4 9
9 | 2
Stem= 10, Leaf = 1
```

Distribusi Frekuensi

Merupakan suatu tabel menunjukkan frekuensi kemunculan data atau frekuensi relatifnya yang berguna untuk meringkas data numerik maupun kategori.

- Untuk data diskret atau data kategori, banyaknya nilai yang dihitung kemunculannya biasanya sesuai dengan banyaknya nilai data yang berbeda dari data diskret atau kategori tersebut
- Untuk data kontinu, biasanya dibuat kelas interval 5-20 banyaknya.

Distribusi Frekuensi

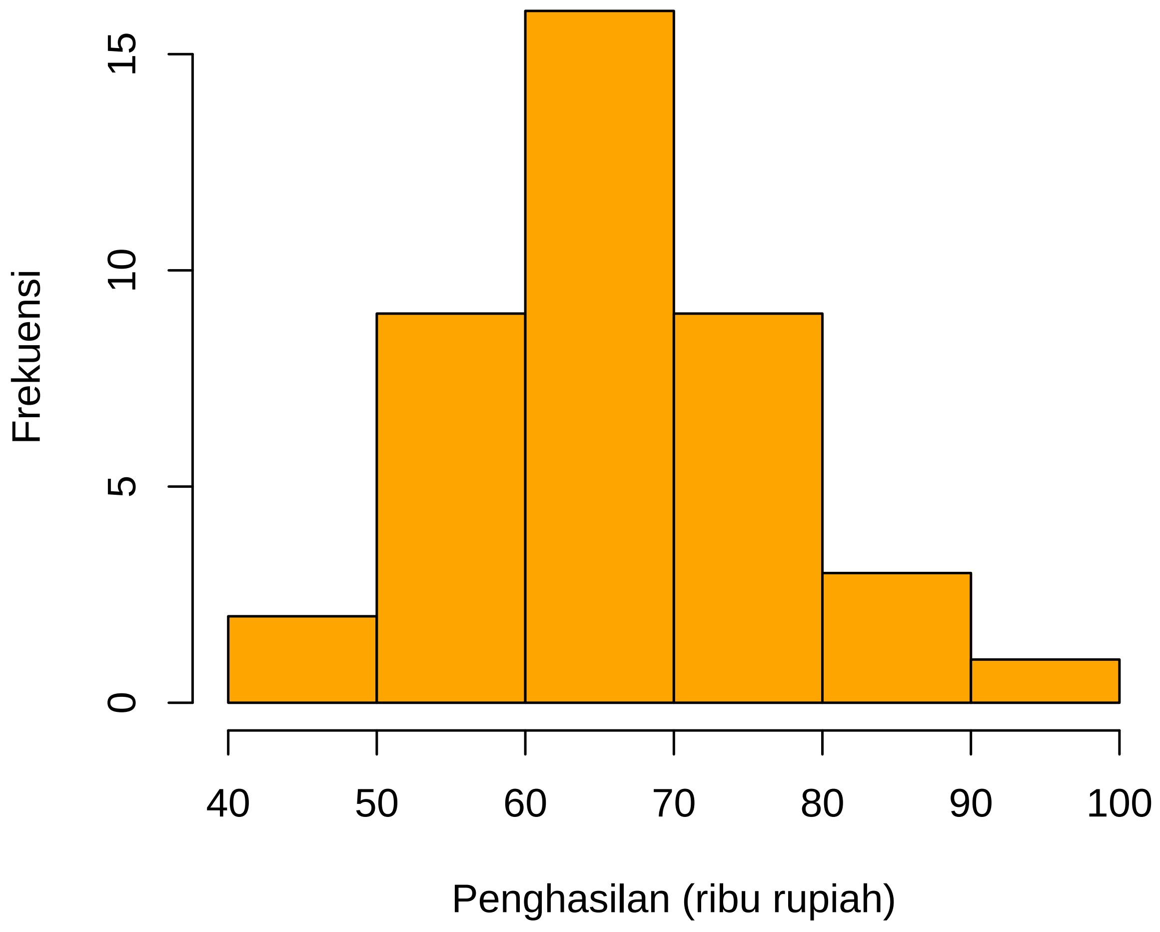
Contoh (Data penghasilan buruh):

Kelas	Frekuensi	Frekuensi Relatif	Frekuensi Relatif Kumulatif
$[40, 50)$	2	0,050	0,050
$[50, 60)$	8	0,200	0,250
$[60, 70)$	17	0,425	0,625
$[70, 80)$	9	0,225	0,900
$[80, 90)$	3	0,075	0,975
$[90, 100)$	1	0,025	1,000

Histogram

Representasi grafik dari distribusi frekuensi data kontinu.

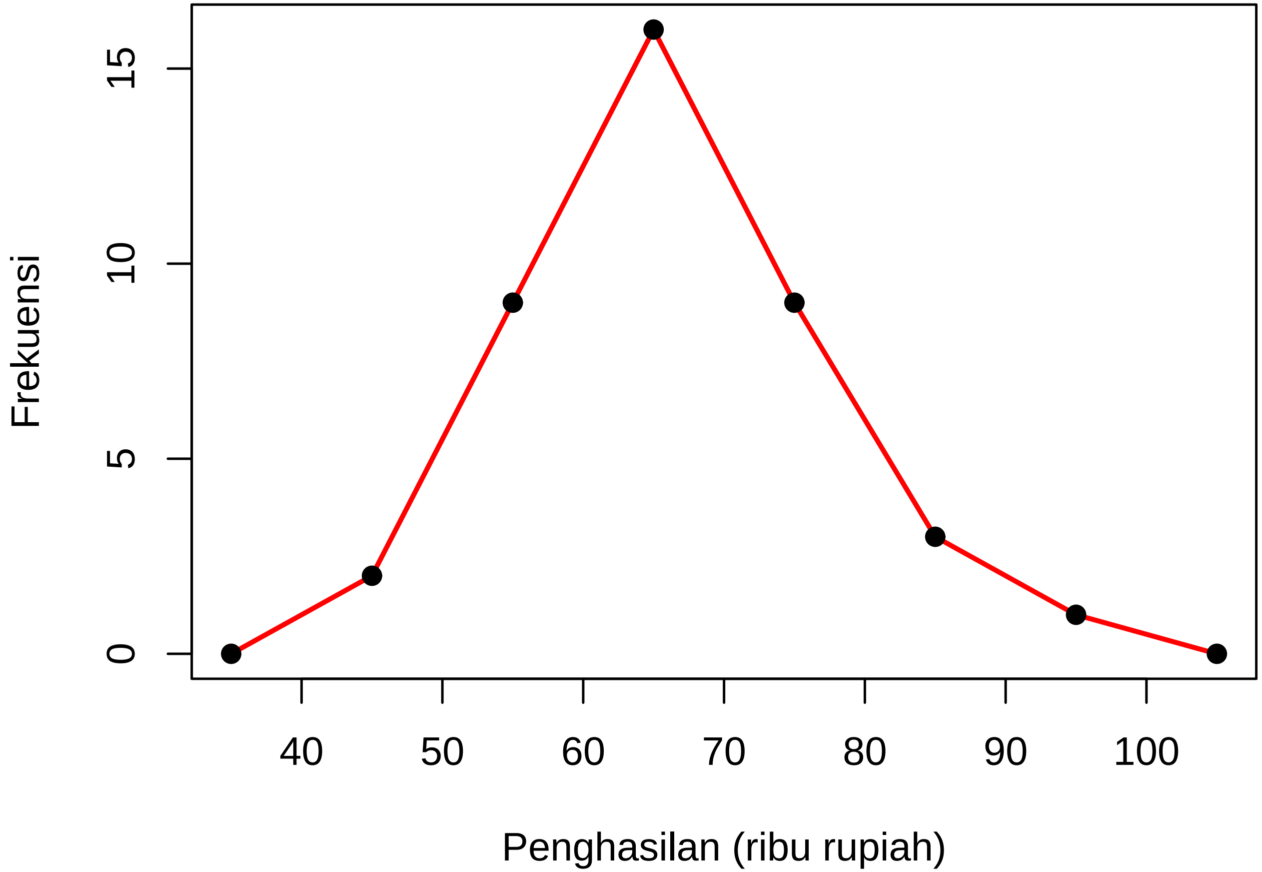
Contoh (Data penghasilan buruh):



Poligon Frekuensi

Representasi grafik dari distribusi frekuensi data kontinu dengan mengambil nilai tengah tiap kelas.

Contoh (Data penghasilan buruh):



Ogive Frekuensi Kumulatif

Plot frekuensi kumulatif dengan batas atas interval dari distribusi frekuensi.

Contoh (Data penghasilan buruh):

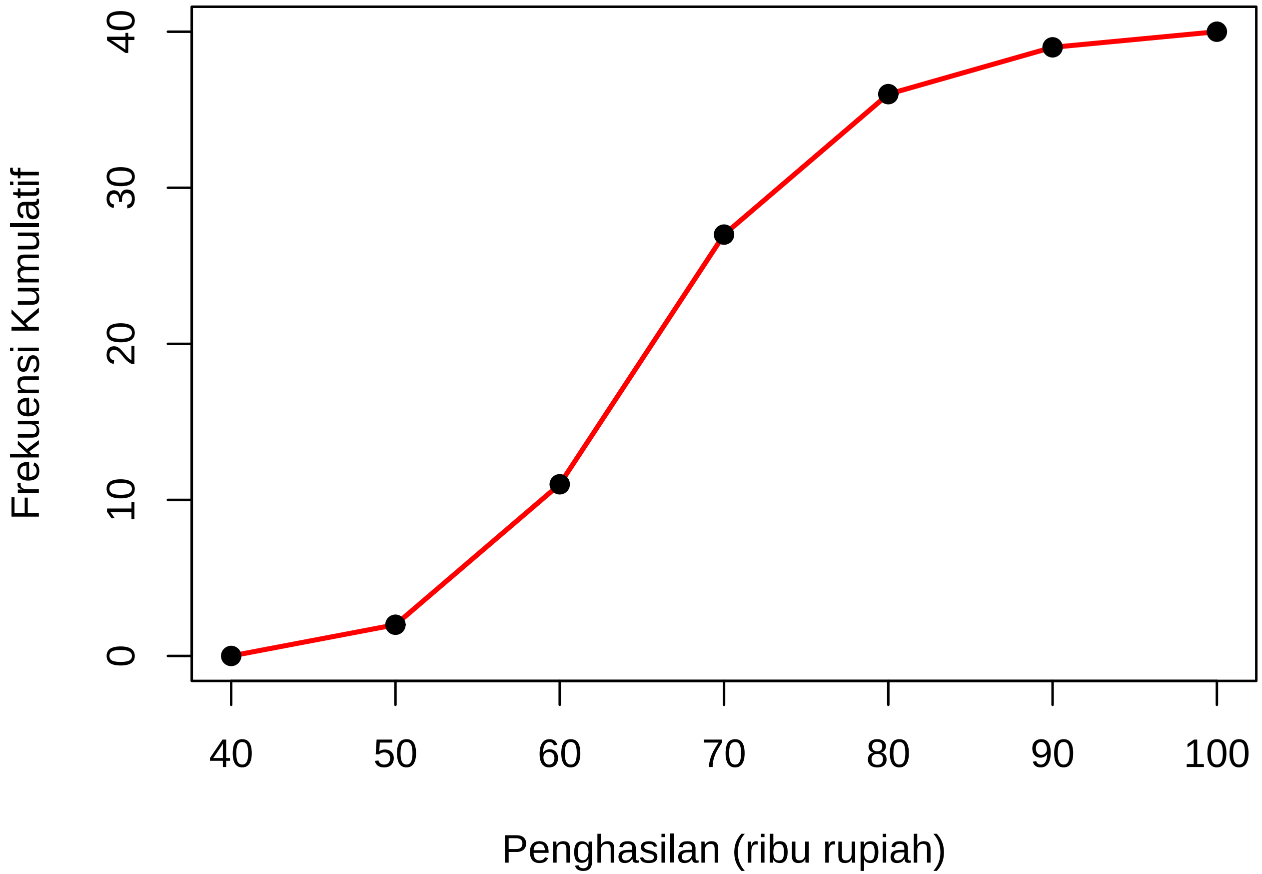


Diagram Batang

Representasi grafik dari distribusi frekuensi data diskret atau kategori.

Contoh (Data telepon seluler):

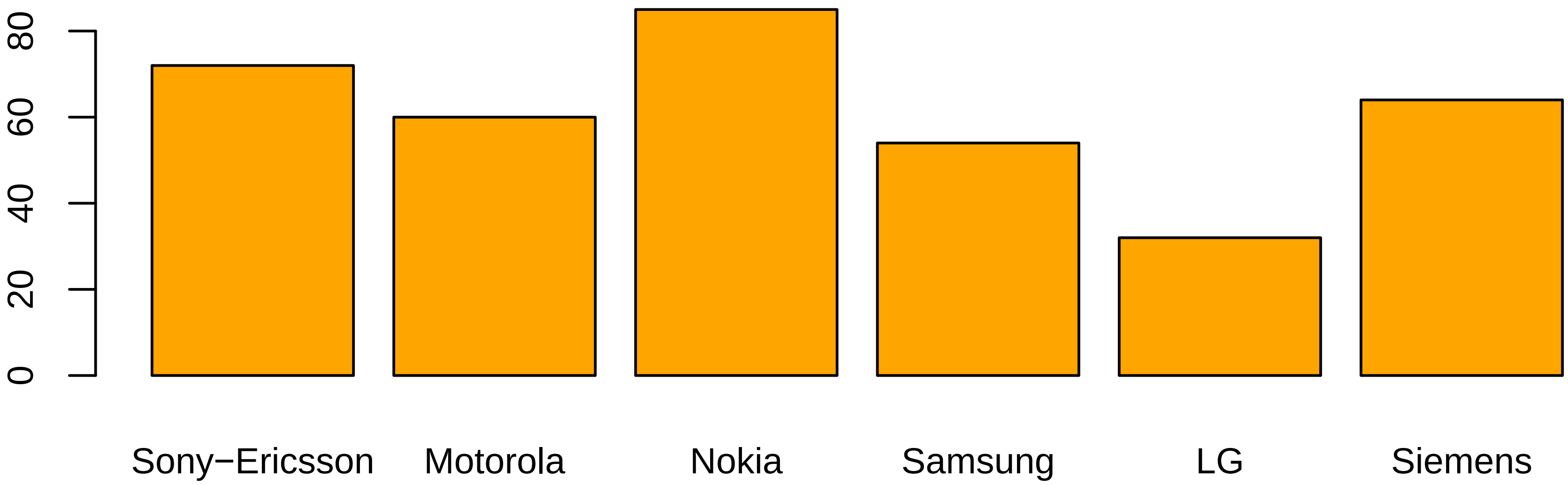
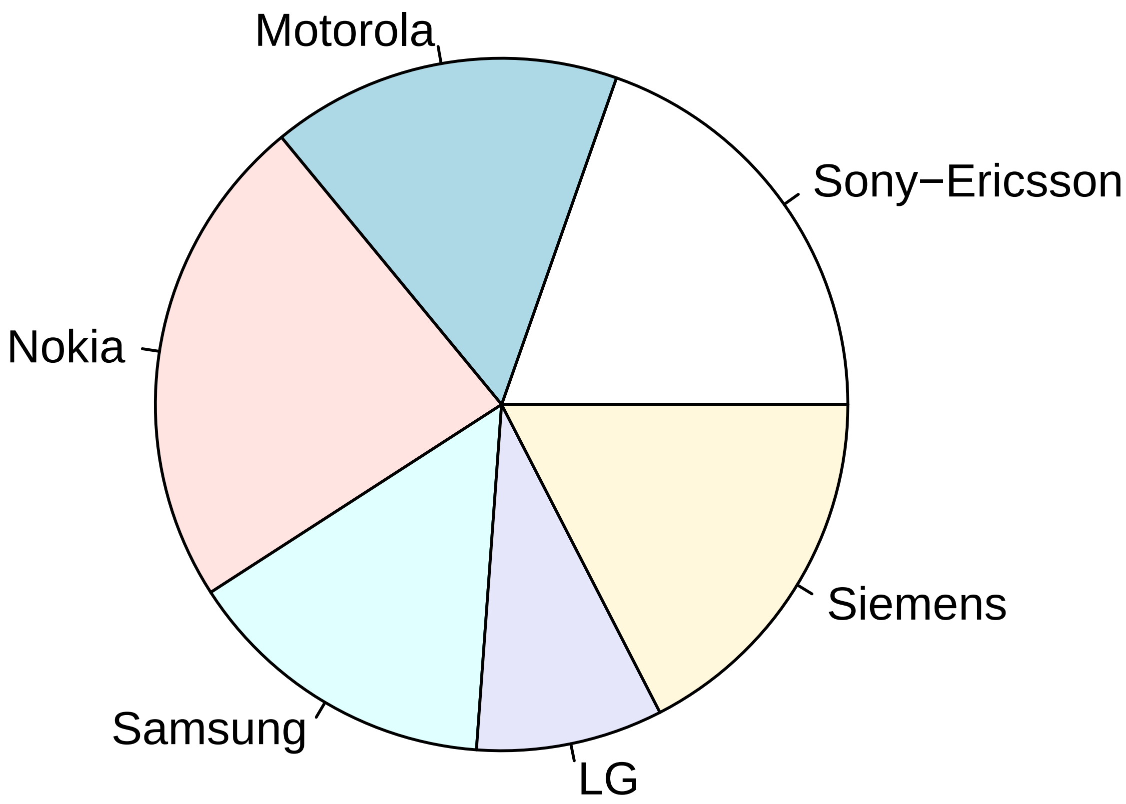


Diagram Lingkaran

Representasi grafik dari distribusi frekuensi data diskret atau kategori.

Contoh (Data telepon seluler):



Notasi Himpunan Data

Data statistik sering dilambangkan dengan huruf X, Y dilengkapi dengan indeks.

Notasi Himpunan Data

Data statistik sering dilambangkan dengan huruf X, Y dilengkapi dengan indeks.

Contoh (Data penghasilan buruh):
X: penghasilan mingguan buruh (dalam ribuan rupiah)
$X_1 = 58$; $X_2 = 72$; $X_{10} = 73$; $X_{40} = 75$;

Notasi Himpunan Data

Data statistik sering dilambangkan dengan huruf X, Y dilengkapi dengan indeks.

Contoh (Data tinggi dan berat mahasiswa):

X : tinggi mahasiswa (cm)
Y : berat mahasiswa (kg)

$X_1 = 170$; $Y_1 = 70$;
$X_7 = 1683$; $Y_7 = 60$;

Notasi Sigma

$$\sum_{i=1}^{n} X_i = X_1 + X_2 + \ldots + X_n$$

$$\sum_{i=1}^{n}\sum_{j=1}^{m} X_{ij} = X_{11} + X_{12} + \ldots + X_{nm}$$

Notasi Sigma

Beberapa aturan:

- Jika $X_i = k$, k suatu konstan, maka

$$\sum_{i=1}^{n} X_i = nk$$

Notasi Sigma

Beberapa aturan:

- Jika $X_i = k$, k suatu konstan, maka

$$\sum_{i=1}^{n} X_i = nk$$

- Jika k suatu konstan, maka

$$\sum_{i=1}^{n} kX_i = k\sum_{i=1}^{n} X_i$$

Notasi Sigma

Beberapa aturan:

- Jika $X_i = k$, k suatu konstan, maka

$$\sum_{i=1}^{n} X_i = nk$$

- Jika k suatu konstan, maka

$$\sum_{i=1}^{n} kX_i = k\sum_{i=1}^{n} X_i$$

-

$$\sum_{i=1}^{n}(X_i + Y_i) = \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} Y_i$$

Ringkasan Numerik

Ringkasan Numerik atau statistik:

- Data tunggal (tidak dikelompokkan), dengan n observasi dinotasikan sebagai

$$x_1, x_2, \ldots, x_n$$

- Data berkelompok (distribusi frekuensi), dengan k nilai tunggal dinotasikan sebagai

$$x_1, x_2, \ldots, x_k$$

  yang masing-masing mempunyai frekuensi

$$f_1, f_2, \ldots, f_k$$

  dengan $n = \sum_{i=1}^{k} f_i$ adalah total observasi

Mean Aritmetik

- Data tunggal:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- Data berkelompok:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} f_i x_i$$

Mean Terbobot

Misalkan $w_i \geq 0$ adalah bobot (*weight*) untuk data tunggal x_i

$$\bar{x}_w = \frac{1}{\sum_{i=1}^{n} w_i} \sum_{i=1}^{n} w_i x_i$$

Variansi

Data tunggal:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

atau

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i^2 - n\bar{x}^2)$$

Variansi

Data berkelompok:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n} f_i(x_i - \bar{x})^2$$

atau

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(f_i x_i^2 - n\bar{x}^2)$$

Peluang dan Variabel Random

Statistika Inferensial: Mengambil kesimpulan, inferensi atau generalisasi tentang suatu populasi berdasarkan informasi yang diperoleh dari sampel.

Peluang (probabilitas): Harga angka yang menunjukkan seberapa besar kemungkinan suatu peristiwa akan terjadi.

Peluang dan Variabel Random

Statistika Inferensial: Mengambil kesimpulan, inferensi atau generalisasi tentang suatu populasi berdasarkan informasi yang diperoleh dari sampel.

Peluang (probabilitas): Harga angka yang menunjukkan seberapa besar kemungkinan suatu peristiwa akan terjadi.

sangat tidak mungkin sangat mungkin

tidak mungkin mungkin ya mungkin tidak pasti

Peluang dan Variabel Random

Statistika Inferensial: Mengambil kesimpulan, inferensi atau generalisasi tentang suatu populasi berdasarkan informasi yang diperoleh dari sampel.

Peluang (probabilitas): Harga angka yang menunjukkan seberapa besar kemungkinan suatu peristiwa akan terjadi.

Peluang dan Variabel Random

Eksperimen (percobaan, *trial*): Prosedur yang dijalankan pada kondisi yang sama dan dapat diamati hasilnya (*outcome*).

Ruang sampel (semesta, *universe*: Himpunan semua hasil yang mungkin dari suatu eksperimen.

Peristiwa (kejadian, *event*): Himpunan bagian dari suatu ruang sampel.

Peluang dan Variabel Random

Contoh

Eksperimen	:	Pelemparan sebuah mata uang logam dua kali
Hasil	:	Sisi mata uang yang tampak
Ruang sampel	:	$S = \{\text{MM,MB,BM,BB}\}$ dengan M: sisi muka dan B: sisi belakang
Peristiwa	:	$A = \text{paling sedikit muncul satu belakang}$ $= \{\text{MB,BM,BB}\}$ $B = \text{muncul sisi yang sama}$ $= \{\text{MM,BB}\}$

Peluang dan Variabel Random

Contoh

Eksperimen	:	Sebuah biji kedelai ditanam
Hasil	:	Tumbuh atau tidak tumbuh
Ruang sampel	:	$S = \{\text{tidak tumbuh, tumbuh}\}$ atau $S = \{0, 1\}$
Peristiwa	:	$A = \text{biji kedelai tumbuh}$ $= \{1\}$

Peluang dan Variabel Random

Contoh

Eksperimen	:	Pemilihan seorang mahasiswa secara random dan dicatat IPnya
Hasil	:	Bilangan antara 0 sampai dengan 4
Ruang sampel	:	$S = \{0 \leq X \leq 4 \mid X \in \mathcal{R}\}$ Himpunan bilangan real antara 0 sampai dengan 4
Peristiwa	:	$A =$ IP di atas 2 $= \{2 \leq X \leq 4 \mid X \in \mathcal{R}\}$ $B =$ IP di bawah 1 $= \{0 \leq X \leq 1 \mid X \in \mathcal{R}\}$

Peluang dan Variabel Random

Contoh

Eksperimen	:	Sebuah dadu dilempar sekali
Hasil	:	
Ruang sampel	:	
Peristiwa	:	

Peluang dan Variabel Random

Contoh

Eksperimen	:	Sebuah dadu dilempar sekali
Hasil	:	mata dadu 1, 2, 3, 4, 5, 6
Ruang sampel	:	
Peristiwa	:	

Peluang dan Variabel Random

Contoh

Eksperimen	:	Sebuah dadu dilempar sekali
Hasil	:	mata dadu 1, 2, 3, 4, 5, 6
Ruang sampel	:	$S = \{1, 2, 3, 4, 5, 6\}$
Peristiwa	:	

Peluang dan Variabel Random

Contoh

Eksperimen	:	Sebuah dadu dilempar sekali
Hasil	:	mata dadu 1, 2, 3, 4, 5, 6
Ruang sampel	:	$S = \{1, 2, 3, 4, 5, 6\}$
Peristiwa	:	$A =$ muncul mata dadu genap

Peluang dan Variabel Random

Contoh

Eksperimen	:	Sebuah dadu dilempar sekali
Hasil	:	mata dadu 1, 2, 3, 4, 5, 6
Ruang sampel	:	$S = \{1, 2, 3, 4, 5, 6\}$
Peristiwa	:	$A = \text{muncul mata dadu genap}$
		$= \{2, 4, 6\}$

Peluang dan Variabel Random

Contoh

Eksperimen	:	Sebuah dadu dilempar sekali
Hasil	:	mata dadu 1, 2, 3, 4, 5, 6
Ruang sampel	:	$S = \{1, 2, 3, 4, 5, 6\}$
Peristiwa	:	$A =$ muncul mata dadu genap $= \{2, 4, 6\}$ $B =$ muncul mata dadu gasal $= \{1, 3, 5\}$

Peluang dan Variabel Random

Peluang Suatu Peristiwa
Definisi klasik, dengan menganggap tiap-tiap elemen ruang sampel S mempunyai peluang yang sama untuk terjadi. Peluang terjadinya peristiwa A,

$$P(A) = \frac{n(A)}{n(S)}$$

dengan $n(A) =$ banyaknya anggota dalam peristiwa A, dan $n(S) =$ banyaknya anggota ruang sampel

Peluang dan Variabel Random

Peluang Suatu Peristiwa

Beberapa ketentuan:

- $0 \leq P(A) \leq 1$
- $P(S) = 1$ (peluang dari ruang sampel)
- $P(\emptyset) = 0$ (peluang dari peristiwa yang tidak akan pernah terjadi)
- $P(A) = 1 - P(A^c)$ (aturan komplemen)
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (aturan penjumlahan)
  Bila A dan B adalah kejadian yang saling asing,
  $A \cap B = \emptyset$, maka $P(A \cup B) = P(A) + P(B)$
- $P(B) = P(A \cap B) + P(A^c \cap B)$
  $A \cap B$ dan $A^c \cap B$ saling asing

Peluang dan Variabel Random

Peluang Suatu Peristiwa

Contoh

Sebuah dadu dilempar sekali. $S = \{1, 2, 3, 4, 5, 6\}$ dan $n(S) = 6$. Misal didefinisikan A : muncul mata dadu 3 dan B : muncul mata dadu bilangan prima $A = \{3\}$ dan $n(A) = 1$; $B = \{2, 3, 5\}$ dan $n(B) = 3$ dan

$$P(A) = \frac{n(A)}{n(S)} = \frac{1}{6}$$

dan

$$P(B) = \frac{n(B)}{n(S)} = \frac{3}{6} = \frac{1}{2}$$

Peluang dan Variabel Random

Peluang Bersyarat dan Independensi
Diketahui A dan B dua peristiwa dari ruang sampel S, dan $P(B) > 0$, maka peluang bersyarat terjadinya A jika diketahui B telah terjadi, ditulis $P(A \mid B)$, didefinisikan sebagai

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

Dua kejadian A dan B disebut kejadian independen jika

$$P(A \cap B) = P(A).P(B)$$

Peluang dan Variabel Random

Peluang Bersyarat dan Independensi

Contoh (Peluang Bersyarat)

Sepasang dadu dilempar bersama jika diketahui jumlah kedua mata dadu yang keluar adalah 6, hitunglah peluang bahwa satu diantara dua dadu tersebut adalah mata dadu 2.

$B = \{\text{jumlahan mata dadu adalah } 6\}$
$\quad = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$ dan
$A = \{\text{mata dadu 2 muncul dari salah satu dadu}\}$
$\quad = \{(2,4),(4,2)\}$

$$P(A \mid B) = \frac{n(A \cap B)}{n(B)} = \frac{2}{5}$$

Peluang dan Variabel Random

Peluang Bersyarat dan Independensi

Contoh (Peluang Bersyarat)

Peluang suatu penerbangan yang telah terjadwal teratur berangkat tepat waktu adalah $P(A) = 0,83$; peluang sampai tepat waktu adalah $P(B) = 0,82$; peluang berangkat dan sampai tepat waktu adalah $P(A \cap B) = 0,78$.

Peluang bahwa suatu pesawat sampai tepat waktu jika diketahui berangkat tepat waktu adalah

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0,78}{0,83} = 0,94$$

Peluang bahwa suatu pesawat berangkat tepat waktu jika diketahui sampai tempat waktu adalah

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0,78}{0,82} = 0,95$$

Peluang dan Variabel Random

Peluang Bersyarat dan Independensi

Contoh (independensi)

Suatu kota kecil mempunyai satu unit mobil pemadam kebakaran dan satu ambulans yang bekerja saling independen untuk keadaan darurat. Peluang mobil kebakaran siap saat diperlukan adalah 0.98. Peluang ambulans siap waktu diperlukan adalah 0.92. Dalam suatu kejadian kebakaran gedung, hitung peluang keduanya siap.

Misalkan A dan B menyatakan kejadian mobil pemadam kebakaran dan ambulans siap. Karena A dan B independen, peluang mobil pemadam kebakaran dan ambulans siap :

$$P(A \cap B) = P(A).P(B) = 0,98 \times 0,92 =, 9016$$

Peluang dan Variabel Random

Teorema Bayes

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A).P(B \mid A)}{P(A).P(B \mid A) + P(A^c).P(B \mid A^c)}$$

Secara umum jika kejadian $A_1, A_2, \ldots, A_k$ saling asing dan gabungannya $A_1 \cup A_2 \cup \ldots, \cup A_k = S$ dan kejadian $B = S \cap B$, maka

$$P(A_i \mid B) = \frac{P(A_i).P(B \mid A_i)}{\sum_{i=1}^{k} P(A_i).P(B \mid A_i)}$$

Peluang dan Variabel Random

Teorema Bayes

Contoh

Sebuah pabrik mempunyai 3 mesin A, B dan C yang memproduksi berturut-turut 60%, 30%, dan 10% dari total banyak unit yang diproduksi pabrik. Persentase kerusakan produk yang dihasilkan dari masing-masing mesin tersebut berturut-turut adalah 2%, 3% dan 4%. Suatu unit dipilih secara random dan diketahui rusak. Hitung probabilitas bahwa unit tersebut berasal dari mesin C.

Misal kejadian R adalah unit yang rusak, akan dihitung $P(C \mid R)$, yaitu probabilitas bahwa suatu unit diproduksi oleh mesin C dengan diketahui unit tersebut rusak.

Teorema Bayes

Contoh (lanjutan)

Dengan teorema Bayes, kejadian $P(A)$, $P(B)$ dan $P(C)$ adalah peluang (persentase produksi) dari masing-masing mesin; $P(R \mid A)$, $P(R \mid B)$ dan $P(R \mid C)$ adalah peluang (persentase kerusakan) dari masing-masing mesin.

$$
\begin{aligned}
P(C \mid R) &= \frac{P(C).P(R \mid C)}{P(A).P(R \mid A) + P(B).P(R \mid B) + P(C).P(R \mid C)} \\
&= \frac{(0,1)(0,04)}{(0,6)(0,02) + (0,3)(0,03) + (0,1)(0,04)} = \frac{4}{25}
\end{aligned}
$$

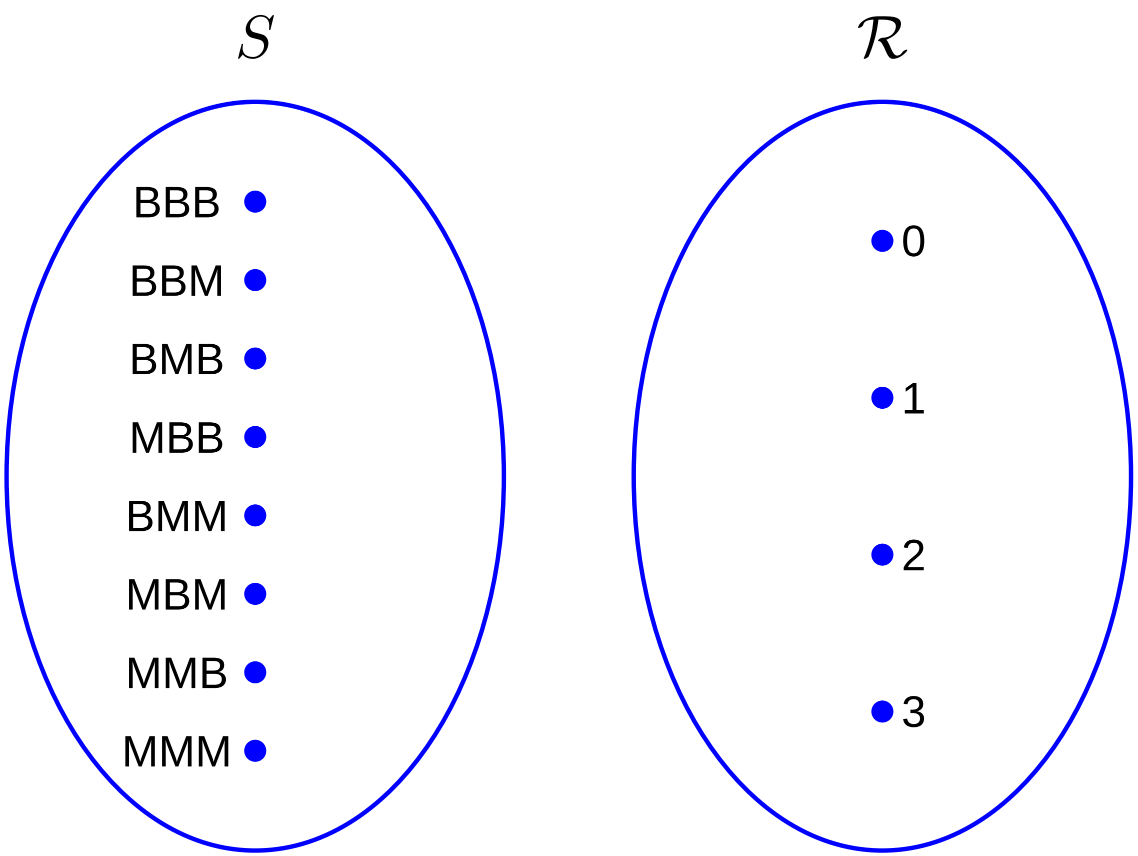
Peluang dan Variabel Random

Variabel Random
Variabel random adalah suatu cara memberi harga angka kepada setiap elemen ruang sampel, atau suatu fungsi bernilai real yang harganya ditentukan oleh setiap elemen dalam ruang sampel

Contoh
Eksperimen (proses random) melemparkan uang logam tiga kali, $S = \{\mathrm{BBB, BBM, BMB, MBB, BMM, MBM, MMB, MMM}\}$.
Didefinisikan variabel random X : banyak M (muka) muncul dalam pelemparan uang logam tiga kali.

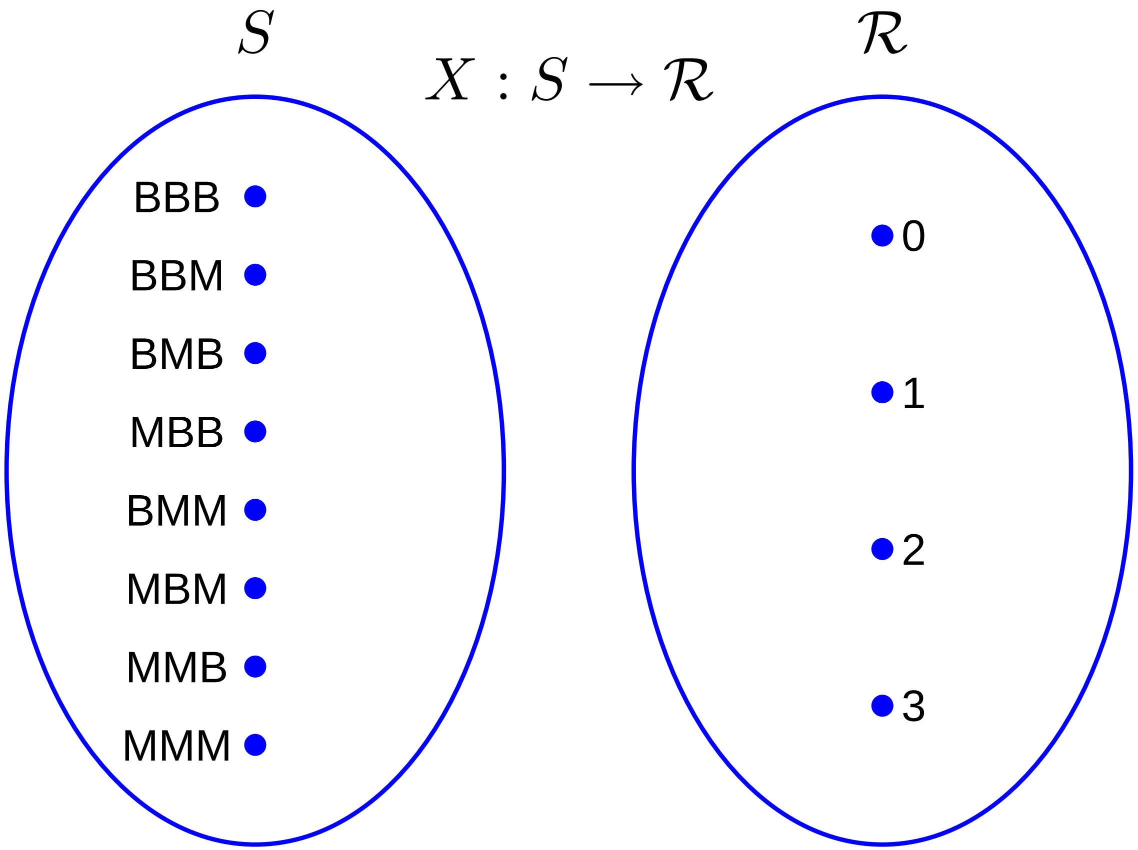
Peluang dan Variabel Random

Contoh (variabel random)



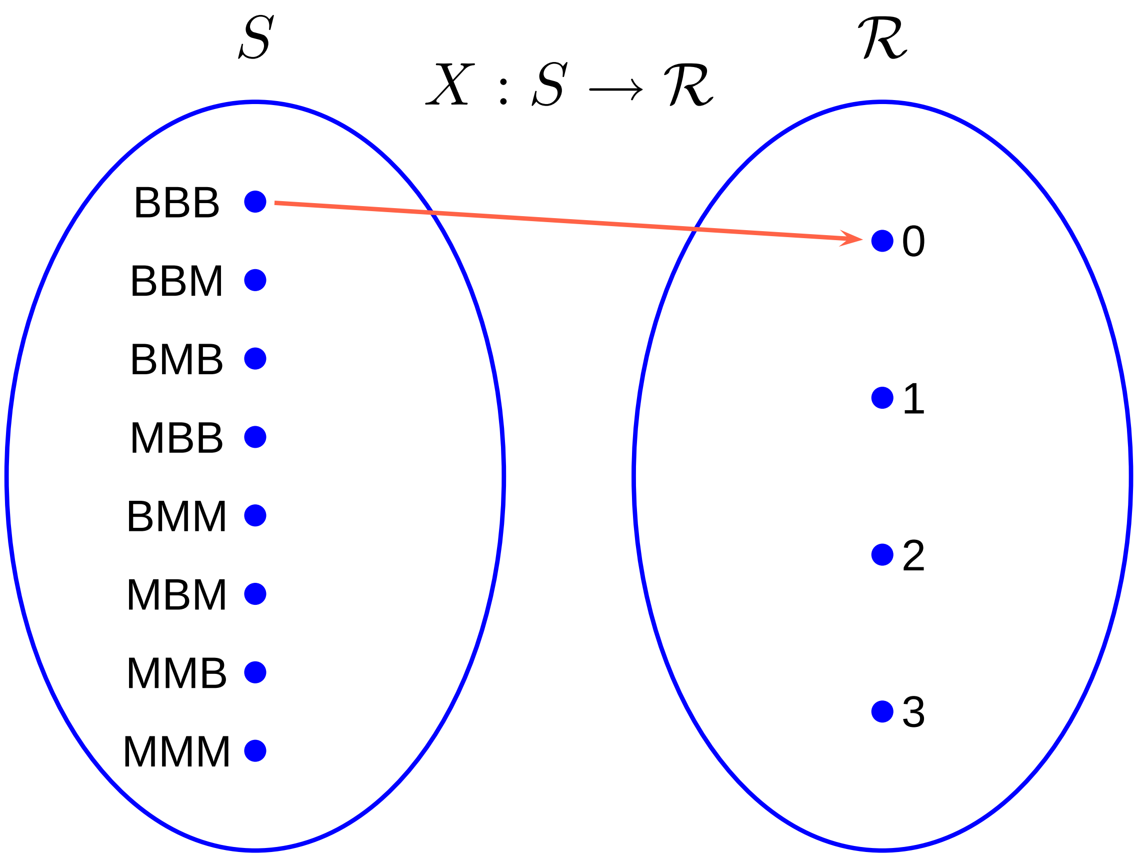
Peluang dan Variabel Random

Contoh (variabel random)



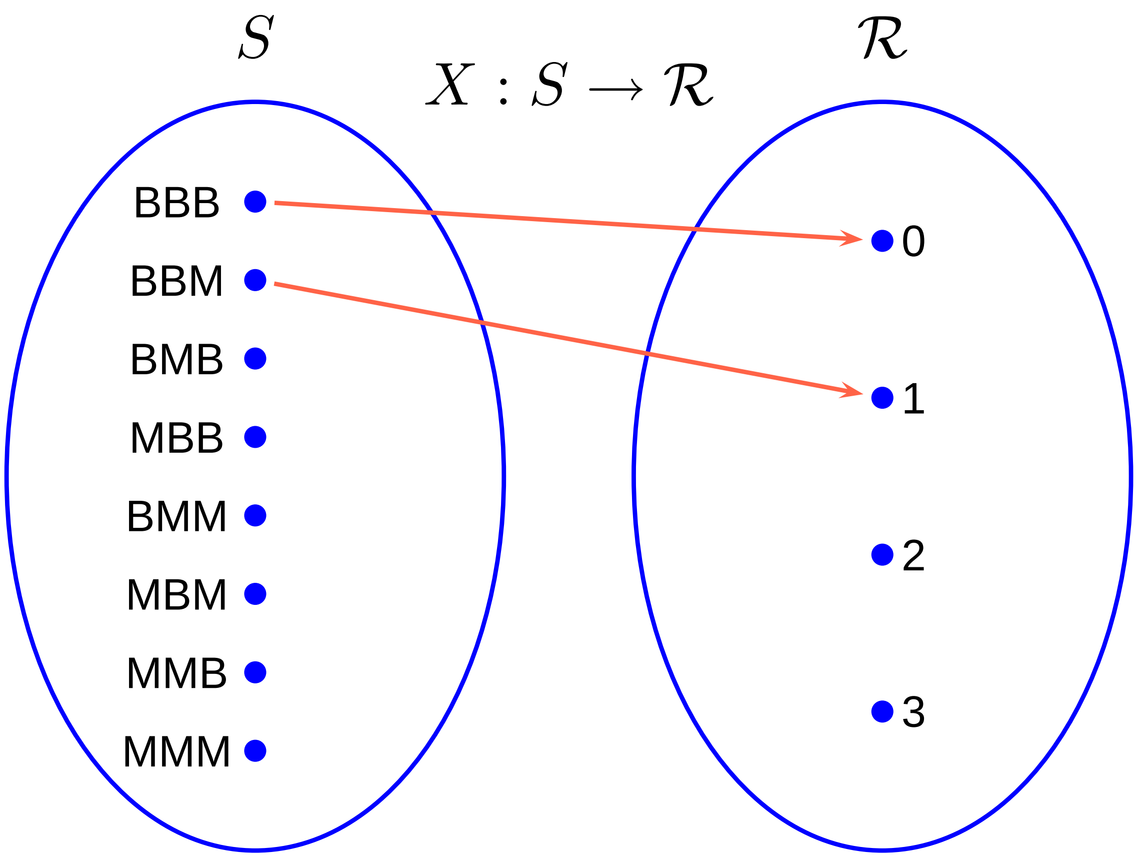
Peluang dan Variabel Random

Contoh (variabel random)



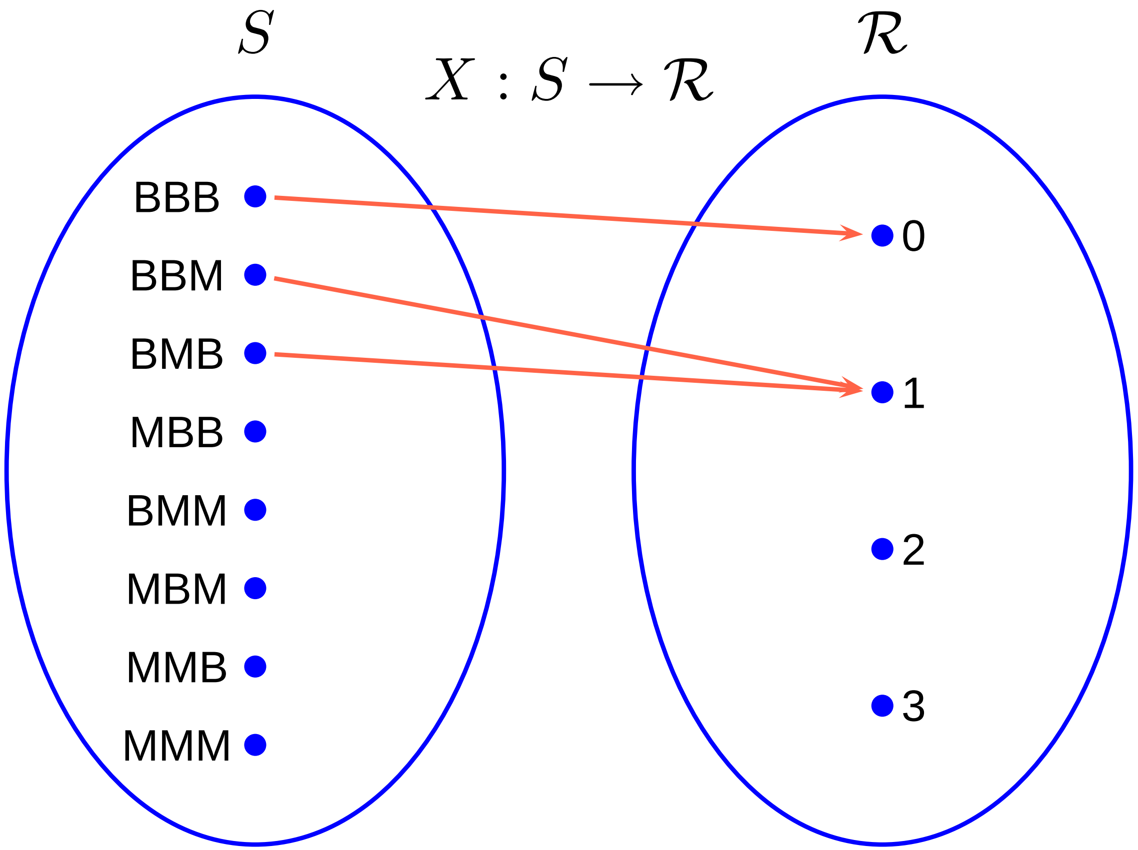
Peluang dan Variabel Random

Contoh (variabel random)



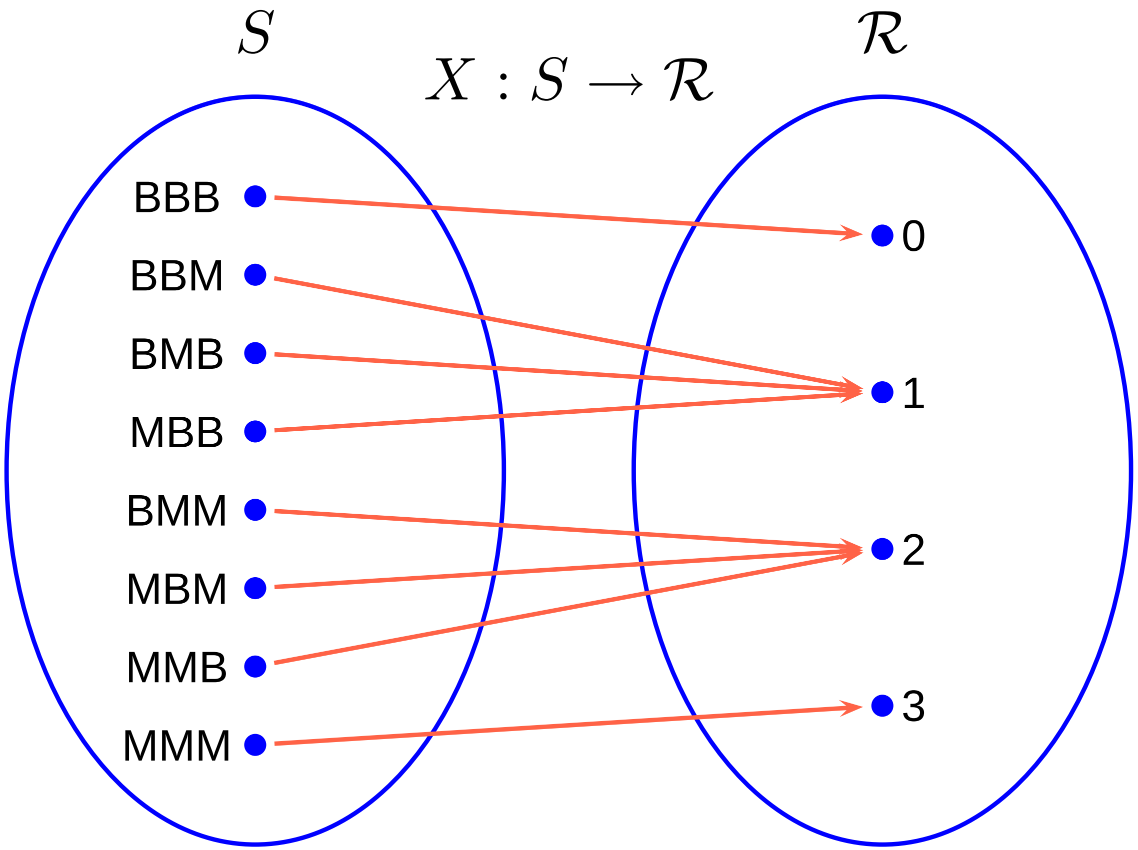
Peluang dan Variabel Random

Contoh (variabel random)



Peluang dan Variabel Random

Contoh (variabel random)



Peluang dan Variabel Random

Variabel random diskret: Suatu variabel random yang hanya dapat menjalani harga-harga yang berbeda yang berhingga banyaknya (sama banyaknya dengan bilangan bulat)

Variabel random kontinu: Suatu variabel random yang dapat menjalani setiap harga dalam suatu interval (tak berhingga banyaknya)

Distribusi Peluang: Model matematik yang menghubungkan semua nilai variabel random dengan peluang terjadinya nilai tersebut dalam ruang sampel. Distribusi peluang dapat direpresentasikan dalam bentuk fungsi, tabel, atau grafik. Distribusi peluang dapat dianggap sebagai frekuensi relatif jangka panjang.

Peluang dan Variabel Random

Distribusi Peluang Diskret

Fungsi $f(x)$ disebut sebagai fungsi peluang dari variabel random diskret X, jika untuk setiap harga x yang mungkin :

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$

Peluang untuk nilai x tertentu:

$$P(X = x) = f(x)$$

Distribusi kumulatif $F(x)$

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t)$$

Peluang dan Variabel Random

Distribusi Peluang Diskret

Distribusi peluang X dalam bentuk tabel:

Harga X	$P(X = x) = f(x)$
x_1	P_1
x_2	P_2
$\ldots$	$\ldots$
x_k	P_k

Peluang dan Variabel Random

Distribusi Peluang Diskret

Contoh

Distribusi banyaknya sisi muka yang muncul dalam pelemparan mata uang logam tiga kali.

Harga X	$P(X = x) = f(x)$
0	$1/8$
1	$3/8$
2	$3/8$
3	$1/8$
	$\sum P(x) = 1$

Peluang dan Variabel Random

Distribusi Peluang Kontinu (Fungsi Densitas)
Distribusi peluang untuk variabel random kontinu.
Fungsi $f(x)$ disebut sebagai fungsi densitas peluang dari variabel random kontinu X, jika untuk setiap harga x yang mungkin :

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$

Nilai peluang untuk interval tertentu

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Distribusi kumulatif F(x)

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du$$

Peluang dan Variabel Random

Distribusi Peluang Kontinu (Fungsi Densitas)

Contoh

Fungsi densitas suatu variabel random X

$$f(x) = \begin{cases} \frac{x}{2} & \text{untuk } 0 < x < 2 \\ 0 & \text{untuk } x \text{ yang lain} \end{cases}$$

Peluang dan Variabel Random

Harga harapan, Variansi dan sifat-sifatnya

Harga Harapan (Ekspektasi, *Expected Value*)

$$\mathrm{E}(X) = \begin{cases} \sum_x x f(x) & \text{bila } X \text{ diskret} \\ \\ \int_{-\infty}^{\infty} x f(x) dx & \text{bila } X \text{ kontinu} \end{cases}$$

$\mathrm{E}(X)$ sering ditulis sebagai μ_X atau μ

Variansi (*Variance*)

$$\begin{aligned} \mathrm{Var}(X) &= \mathrm{E}(X-\mu)^2 \\ &= \mathrm{E}(X^2) - \mu^2 \end{aligned}$$

Peluang dan Variabel Random

Harga harapan, Variansi dan sifat-sifatnya

Sifat-sifat Harga Harapan

- $\mathrm{E}(aX + b) = a\mathrm{E}(X) + b$, a, b konstan
- $\mathrm{E}\left[g(X) + h(X)\right] = \mathrm{E}\left[g(X)\right] + \mathrm{E}\left[h(X)\right]$

Sifat-sifat Variansi

$\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$, a, b konstan

Deviasi standar (akar dari variansi):

$\sigma_X = \sqrt{\mathrm{Var}(X)}$

Peluang dan Variabel Random

Dua Variabel Random

Ada dua variabel random yang diamati bersamaan dalam suatu eksperimen.

Contoh:
Sebuah mata uang logam dilemparkan tiga kali.
X: banyaknya M muncul dalam dua lemparan pertama
Y: banyaknya M muncul dalam lemparan ketiga

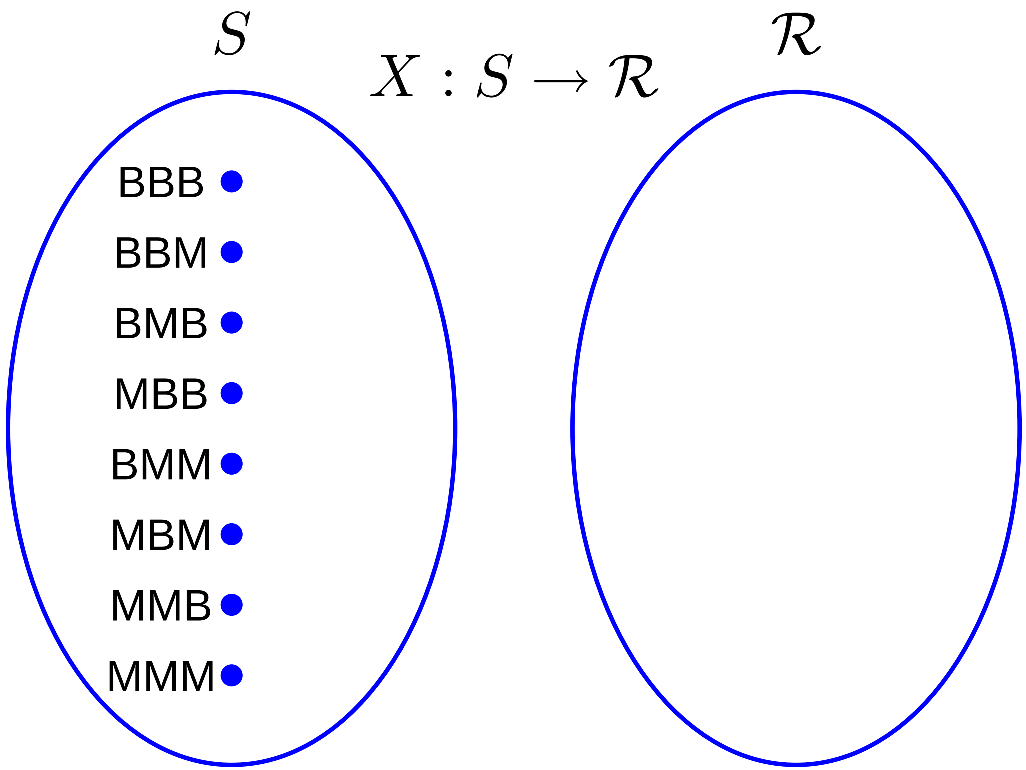
Distribusi peluang untuk dua variabel random disebut sebagai distribusi peluang bersama

Peluang dan Variabel Random

Contoh

Distribusi peluang variabel random X :

x	$P(X = x)$



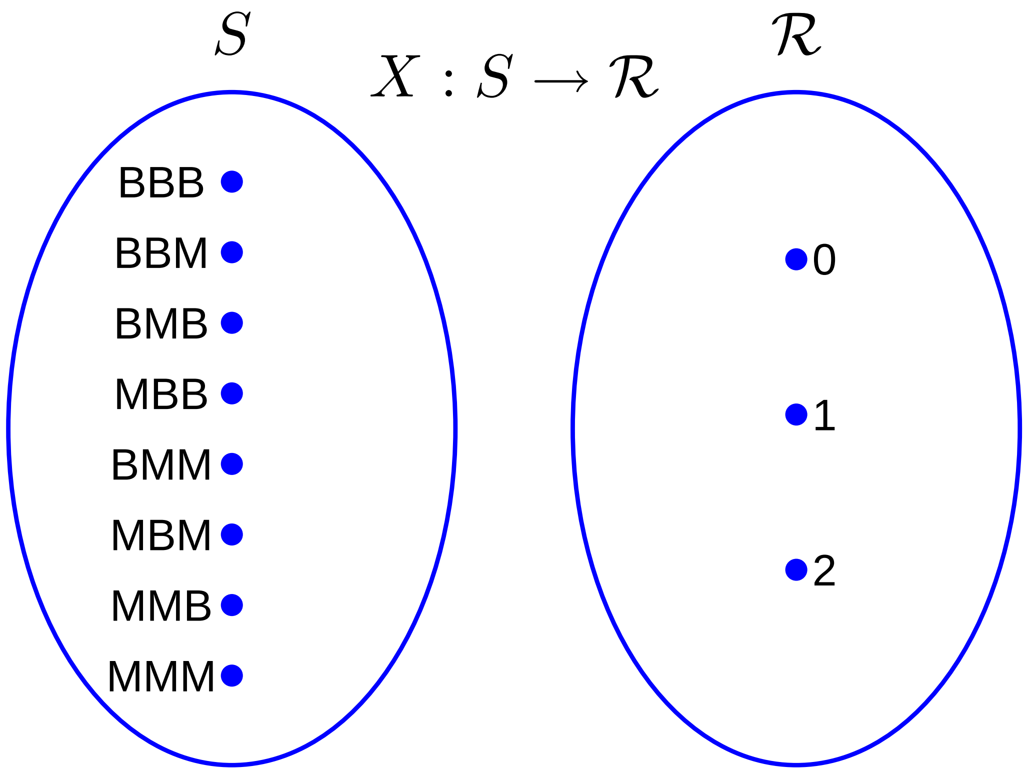
X : banyaknya M muncul dalam dua lemparan pertama

Peluang dan Variabel Random

Contoh

Distribusi peluang variabel random X :

x	$P(X = x)$
0	
1	
2	



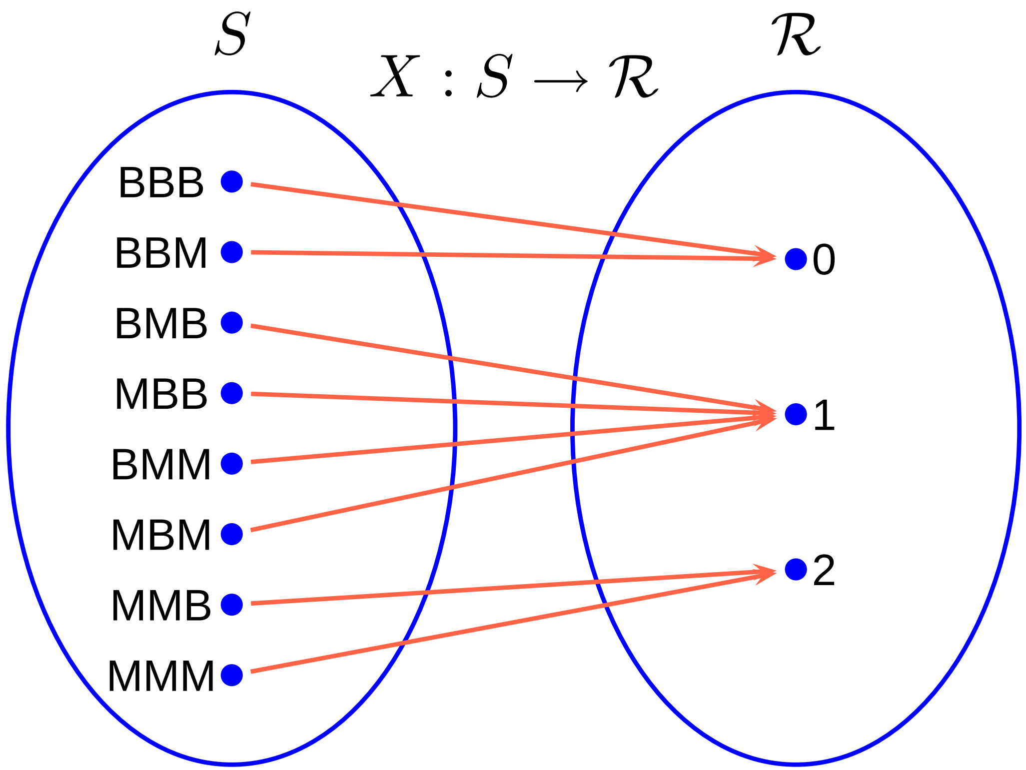
X : banyaknya M muncul dalam dua lemparan pertama

Peluang dan Variabel Random

Contoh

Distribusi peluang variabel random X :

x	$P(X = x)$
0	
1	
2	



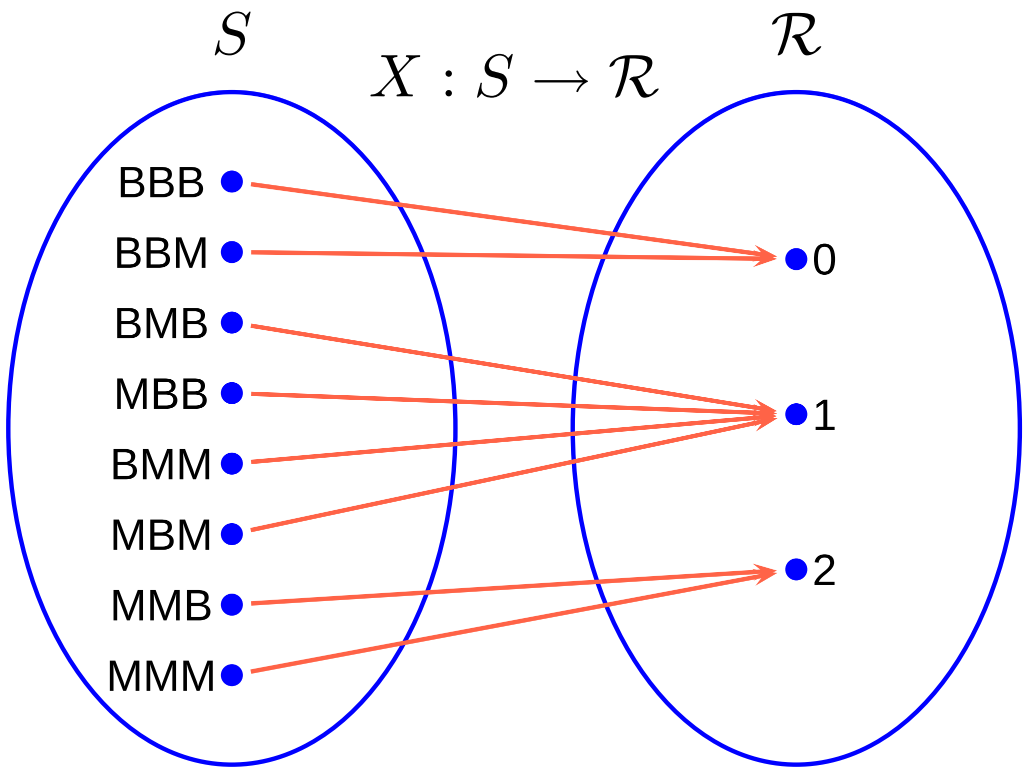
X : banyaknya M muncul dalam dua lemparan pertama

Peluang dan Variabel Random

Contoh

Distribusi peluang variabel random X :

x	$P(X = x)$
0	1/4
1	1/2
2	1/4



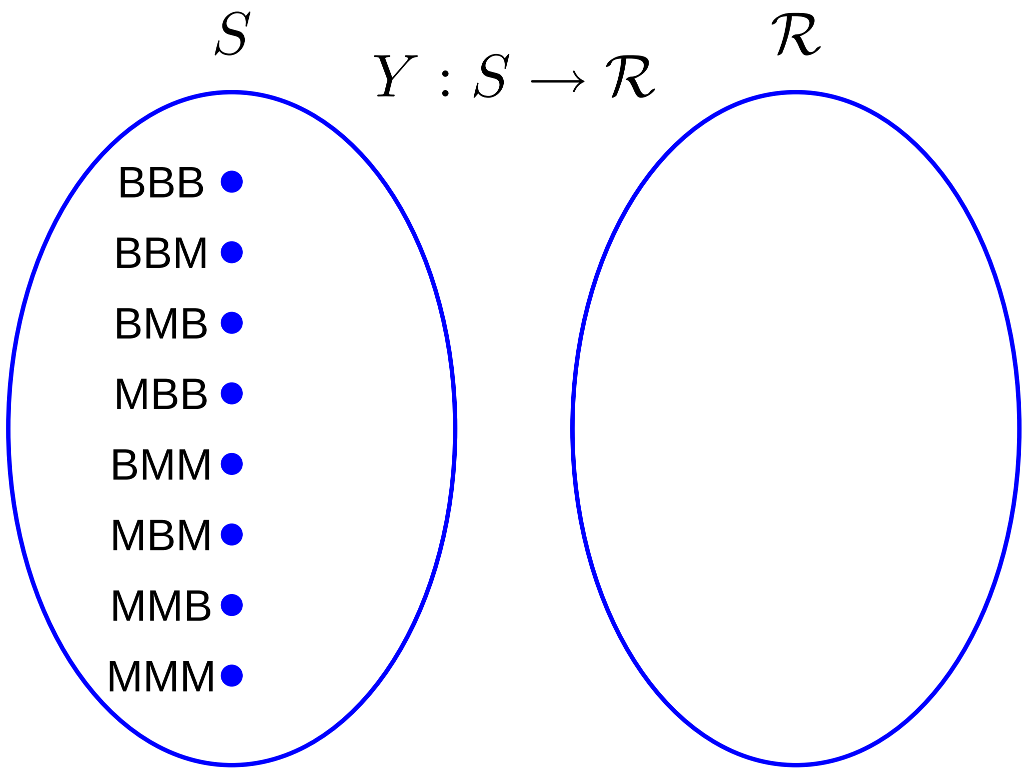
X : banyaknya M muncul dalam dua lemparan pertama

Peluang dan Variabel Random

Contoh

Distribusi peluang variabel random Y:

	y
$P(Y = y)$	



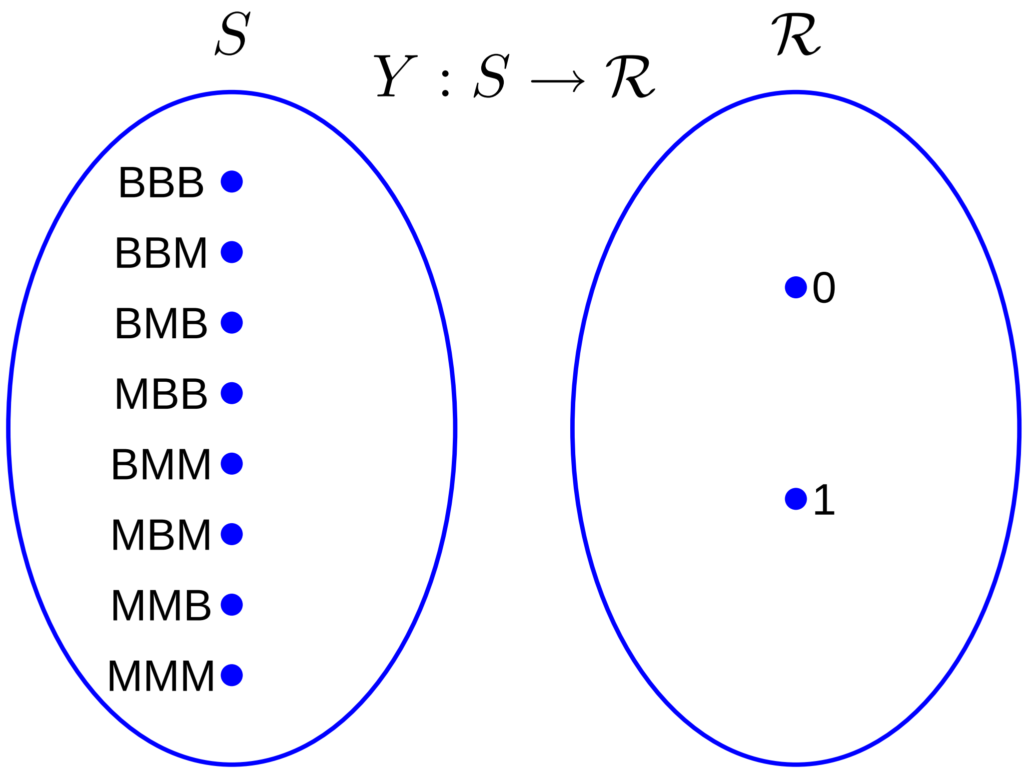
Y : banyaknya M muncul dalam lemparan ketiga

Peluang dan Variabel Random

Contoh

Distribusi peluang variabel random Y:

	y	
	0	1
$P(Y=y)$		



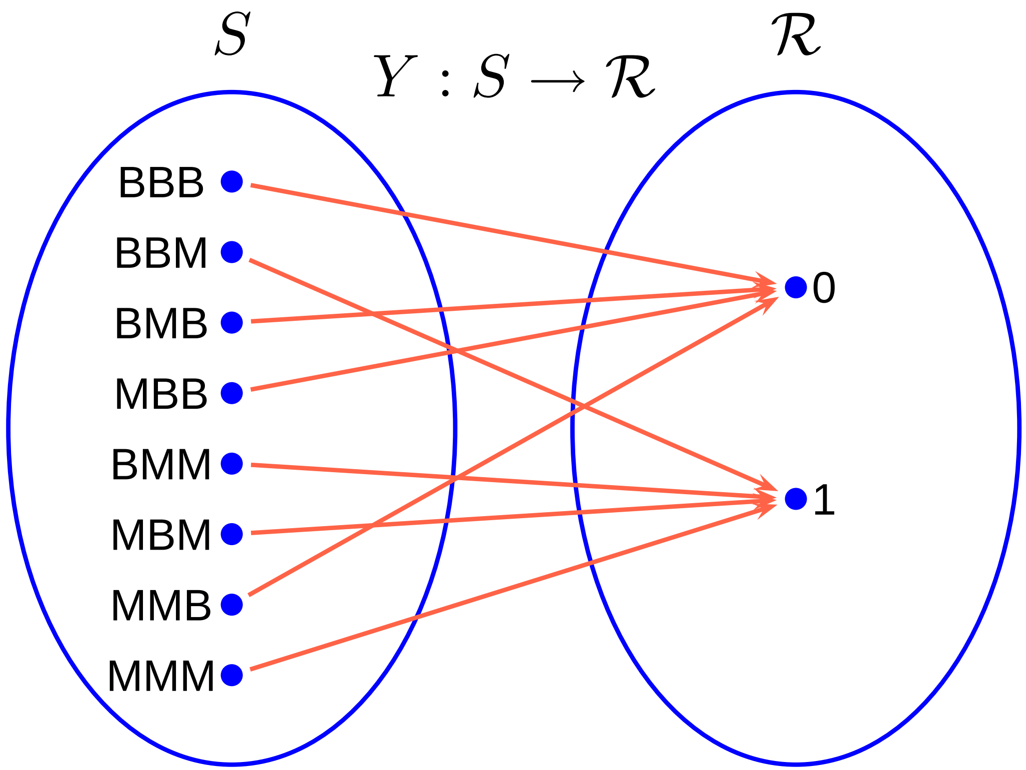
Y : banyaknya M muncul dalam lemparan ketiga

Peluang dan Variabel Random

Contoh

Distribusi peluang variabel random Y:

	y	
	0	1
$P(Y = y)$	1/2	1/2



Y : banyaknya M muncul dalam lemparan ketiga

Peluang dan Variabel Random

Contoh (dua variabel random)
Distribusi peluang bersama X dan Y, $P(X = x, Y = y)$::

x	y		$P(X = x)$
	0	1	
0			1/4
1			1/2
2			1/4
$P(Y = y)$	1/2	1/2	1

Peluang dan Variabel Random

Contoh (dua variabel random)
Distribusi peluang bersama X dan Y, $P(X = x, Y = y)$:

x	y		$P(X = x)$
	0	1	
0	$\{$BBB$\}$	$\{$BBM$\}$	1/4
1	$\{$BMB, MBB$\}$	$\{$BMM, MBM$\}$	1/2
2	$\{$MMB$\}$	$\{$MMM$\}$	1/4
$P(Y = y)$	1/2	1/2	1

Peluang dan Variabel Random

Contoh (dua variabel random)
Distribusi peluang bersama X dan Y, $P(X = x, Y = y)$:

x	y		$P(X = x)$
	0	1	
0	1/8	1/8	1/4
1	2/8	2/8	1/2
2	1/8	1/8	1/4
$P(Y = y)$	1/2	1/2	1

Peluang dan Variabel Random

Contoh (dua variabel random)
Distribusi peluang bersama X dan Y, $P(X = x, Y = y)$:

x	y		$P(X = x)$
	0	1	
0	1/8	1/8	1/4
1	2/8	2/8	1/2
2	1/8	1/8	1/4
$P(Y = y)$	1/2	1/2	1

Jika $P(X = x, Y = y) = P(X = x).P(Y = y)$ untuk setiap nilai dari X dan Y maka dua variabel random tersebut dikatakan independen

Peluang dan Variabel Random

Kovariansi

Ukuran numerik untuk variansi bersama dua variabel random

$$\begin{aligned}\mathrm{Kov}(X,Y) &= \mathrm{E}\left[(X-\mu_X)(Y-\mu_Y)\right] \\ &= \mathrm{E}(XY)-\mu_X\mu_Y\end{aligned}$$

Korelasi

Kovariansi dibagi dengan standar deviasi X dan standar deviasi Y

$$\mathrm{Kor}(X,Y) = \frac{\mathrm{Kov}(X,Y)}{\sigma_X.\sigma_Y}$$

Peluang dan Variabel Random

Harga harapan untuk penjumlahan dan pengurangan dua variabel random,

$$\begin{aligned} \mathrm{E}(X+Y) &= \mathrm{E}(X)+\mathrm{E}(Y) \\ \mathrm{E}(X-Y) &= \mathrm{E}(X)-\mathrm{E}(Y) \end{aligned}$$

Variansi untuk penjumlahan dan pengurangan dua variabel random,

$$\begin{aligned} \mathrm{Var}(X+Y) &= \mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Kov}(X,Y) \\ \mathrm{Var}(X-Y) &= \mathrm{Var}(X)+\mathrm{Var}(Y)-2\mathrm{Kov}(X,Y) \end{aligned}$$

Distribusi Variabel Random Diskret dan Kontinu

Eksperimen Bernoulli

Eksperimen dengan hanya dua hasil yang mungkin

Contoh

- melempar mata uang logam satu kali
- Mengamati telur ayam, apakah anak ayam itu jantan atau betina
- Mengamati kedelai yang ditanam, tumbuh atau tidak
- Reaksi obat pada tikus, positif atau negatif

Distribusi Variabel Random Diskret dan Kontinu

Sifat-sifat Eksperimen Bernoulli

- tiap usaha (*trial*) menghasilkan satu dari dua hasil yang mungkin, dinamakan sukses (S) dan gagal (G);
- peluang sukses, $P(S) = p$ dan peluang gagal $P(G) = 1 - p$, atau $P(G) = q$;
- usaha-usaha tersebut independen

Distribusi Variabel Random Diskret dan Kontinu

Distribusi Bernoulli

$$P(X = x; p) = p^x(1 - p)^{1-x},$$

dengan $x = 0, 1$ (gagal, sukses) dan p adalah peluang mendapatkan hasil sukses.

Distribusi Variabel Random Diskret dan Kontinu

Distribusi Binomial

Eksperimen Bernoulli dengan n usaha dan X : banyaknya sukses dalam n usaha tersebut.

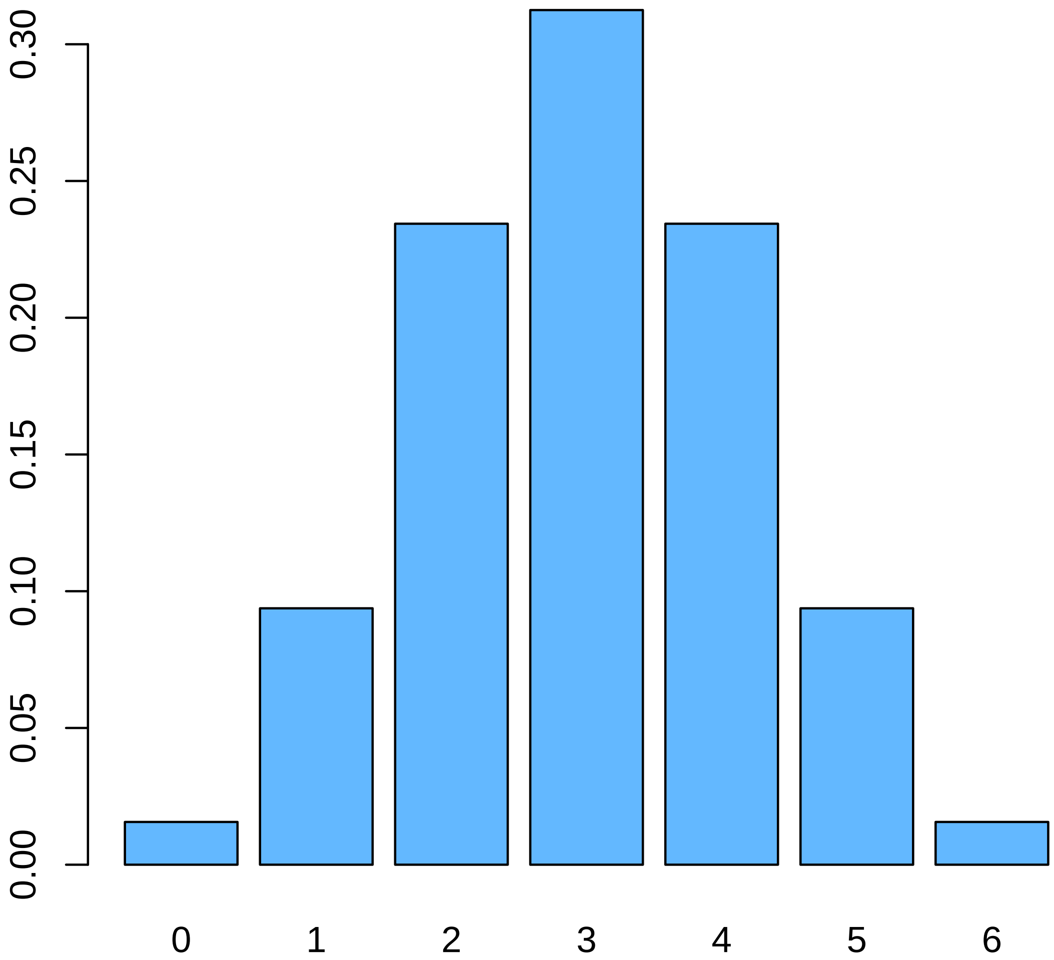
$$P(X = x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}, \qquad x = 0, 1, 2, \ldots, n$$

Mean dan variansi

$\mathrm{E}(X) = np$; $\mathrm{Var}(X) = np(1-p)$

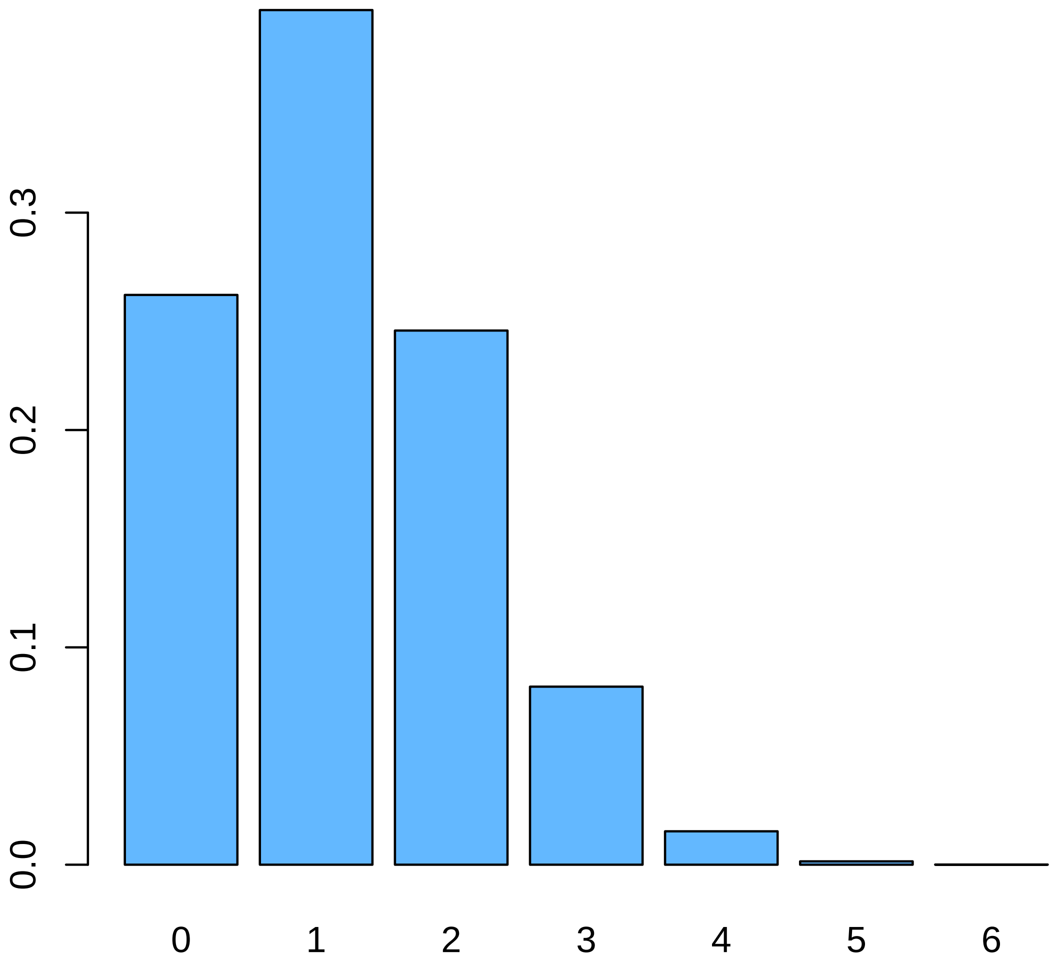
Distribusi Variabel Random Diskret dan Kontinu

Binomial dengan $n = 6$, $p = 0,5$



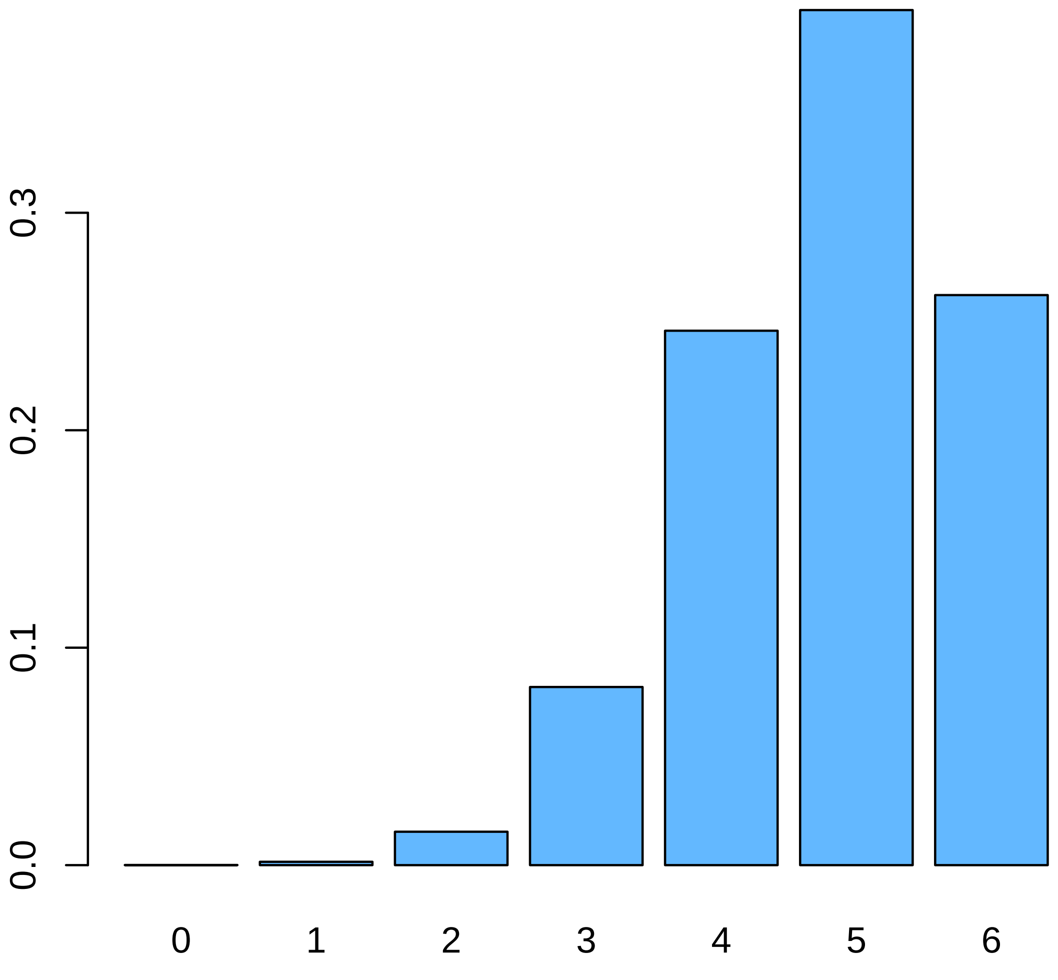
Distribusi Variabel Random Diskret dan Kontinu

Binomial dengan $n = 6$, $p = 0,2$



Distribusi Variabel Random Diskret dan Kontinu

Binomial dengan $n = 6$, $p = 0,8$



Distribusi Variabel Random Diskret dan Kontinu

Contoh (Distribusi Binomial)

Suatu uang logam yang baik (seimbang) dilempar 4 kali. X adalah banyaknya muka muncul dalam 4 kali pelemparan tersebut. Pelemparan dipandang sebagai usaha, dan sukses adalah muka muncul. X merupakan variabel random binomial dengan $n = 4$ dan $p = 1/2$ dengan distribusi peluang:

$$P(X = x; 4, \frac{1}{2}) = \binom{4}{x} \left(\frac{1}{2}\right)^x (1 - \frac{1}{2})^{4-x}, \qquad x = 0, 1, 2, 3, 4$$

Distribusi Variabel Random Diskret dan Kontinu

Contoh (Distribusi Binomial)

Suatu uang logam yang baik (seimbang) dilempar 4 kali. X adalah banyaknya muka muncul dalam 4 kali pelemparan tersebut. Pelemparan dipandang sebagai usaha, dan sukses adalah muka muncul. X merupakan variabel random binomial dengan $n = 4$ dan $p = 1/2$ dengan distribusi peluang:

$$P(X = x; 4, \frac{1}{2}) = \binom{4}{x} \left(\frac{1}{2}\right)^x (1 - \frac{1}{2})^{4-x}, \qquad x = 0, 1, 2, 3, 4$$

Peluang muka muncul dua kali, $X = 2$

$$\begin{aligned} P(X = 2; 4, \frac{1}{2}) &= \binom{4}{2} \left(\frac{1}{2}\right)^2 (1 - \frac{1}{2})^{4-2} \\ &= \frac{3}{8} \end{aligned}$$

Distribusi Variabel Random Diskret dan Kontinu

Contoh (Distribusi Binomial)

Suatu uang logam yang baik (seimbang) dilempar 4 kali. X adalah banyaknya muka muncul dalam 4 kali pelemparan tersebut. Pelemparan dipandang sebagai usaha, dan sukses adalah muka muncul. X merupakan variabel random binomial dengan $n = 4$ dan $p = 1/2$ dengan distribusi peluang:

$$P(X = x; 4, \frac{1}{2}) = \binom{4}{x} \left(\frac{1}{2}\right)^x (1 - \frac{1}{2})^{4-x}, \qquad x = 0, 1, 2, 3, 4$$

Peluang muka muncul paling tidak dua kali, $X \geq 2$

$$\begin{aligned} P(X \geq 2; 4, \frac{1}{2}) &= P(X = 2) + P(X = 3) + P(X = 4) \\ &= \frac{11}{16} \end{aligned}$$

Distribusi Variabel Random Diskret dan Kontinu

Distribusi Hipergeometrik

Eksperimen hipergeometrik:

- Dalam populasi berukuran N sebanyak k dinamakan sukses sedangkan sisanya $N - k$ dinamakan gagal
- sampel berukuran n diambil dari N benda
- Cara pengambilan sampel tanpa pengembalian

Distribusi Variabel Random Diskret dan Kontinu

Distribusi Hipergeometrik

Distribusi peluang:

$$P(X = x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}, \qquad x = 0, 1, 2, \ldots, \min(n, k)$$

Mean dan Variansi

$\mathrm{E}(X) = n\frac{k}{N}$; $\qquad \mathrm{Var}(X) = n\frac{k}{n}\frac{N-k}{N}\frac{N-n}{N-1}$

Distribusi Variabel Random Diskret dan Kontinu

Contoh (Distribusi Hipergeometrik)

Suatu kotak berisi 40 suku cadang dengan 3 rusak. Sampel berukuran 5 diambil sekaligus dari kotak. Pengambilan sampel ini adalah suatu eksperimen hipergeometrik dengan X adalah banyaknya suku cadang rusak, $N = 40$, $n = 5$ dan $k = 3$ dengan distribusi peluang:

$$P(X = x; 40, 5, 3) = \frac{\binom{3}{x}\binom{37}{5-x}}{\binom{40}{5}}, \qquad x = 0, 1, 2, 3$$

Distribusi Variabel Random Diskret dan Kontinu

Contoh (Distribusi Hipergeometrik)

Suatu kotak berisi 40 suku cadang dengan 3 rusak. Sampel berukuran 5 diambil sekaligus dari kotak. Pengambilan sampel ini adalah suatu eksperimen hipergeometrik dengan X adalah banyaknya suku cadang rusak, $N = 40$, $n = 5$ dan $k = 3$ dengan distribusi peluang:

$$P(X = x; 40, 5, 3) = \frac{\binom{3}{x}\binom{37}{5-x}}{\binom{40}{5}}, \qquad x = 0, 1, 2, 3$$

Peluang ditemukan satu suku cadang rusak dalam pengambilan sampel tersebut

$$P(X = 1; 40, 5, 3) \quad = \quad \frac{\binom{3}{1}\binom{37}{4}}{\binom{40}{5}} = 0,3011$$

Distribusi Variabel Random Diskret dan Kontinu

Contoh (Distribusi Hipergeometrik)

Suatu kotak berisi 40 suku cadang dengan 3 rusak. Sampel berukuran 5 diambil sekaligus dari kotak. Pengambilan sampel ini adalah suatu eksperimen hipergeometrik dengan X adalah banyaknya suku cadang rusak, $N = 40$, $n = 5$ dan $k = 3$ dengan distribusi peluang:

$$P(X = x; 40, 5, 3) = \frac{\binom{3}{x}\binom{37}{5-x}}{\binom{40}{5}}, \qquad x = 0, 1, 2, 3$$

Peluang ditemukan paling tidak satu suku cadang rusak dalam pengambilan sampel tersebut

$$\begin{aligned} P(X \geq 1; 40, 5, 3) &= P(X = 1) + P(X = 2) + P(X = 3) \\ &= 0,301 + 0,0354 + 0,0010 \\ &= 0,3376 \end{aligned}$$

Distribusi Variabel Random Diskret dan Kontinu

<u>Distribusi Poisson</u>

Sifat-sifat eksperimen Poisson:

- banyaknya sukses terjadi dalam suatu selang waktu atau daerah tertentu tidak terpengaruh (bebas) dari apa yang terjadi pada interval waktu atau daerah yang lain,
- peluang terjadinya sukses dalam interval waktu yang singkat atau daerah yang sempit sebanding dengan panjang interval waktu, atau luas daerah dan tidak tergantung pada banyaknya sukses yang terjadi di luar interval waktu atau daerah tersebut,
- peluang terjadinya lebih dari satu sukses dalam interval waktu yang singkat atau daerah yang sempit tersebut dapat diabaikan.

Distribusi Variabel Random Diskret dan Kontinu

Distribusi Poisson

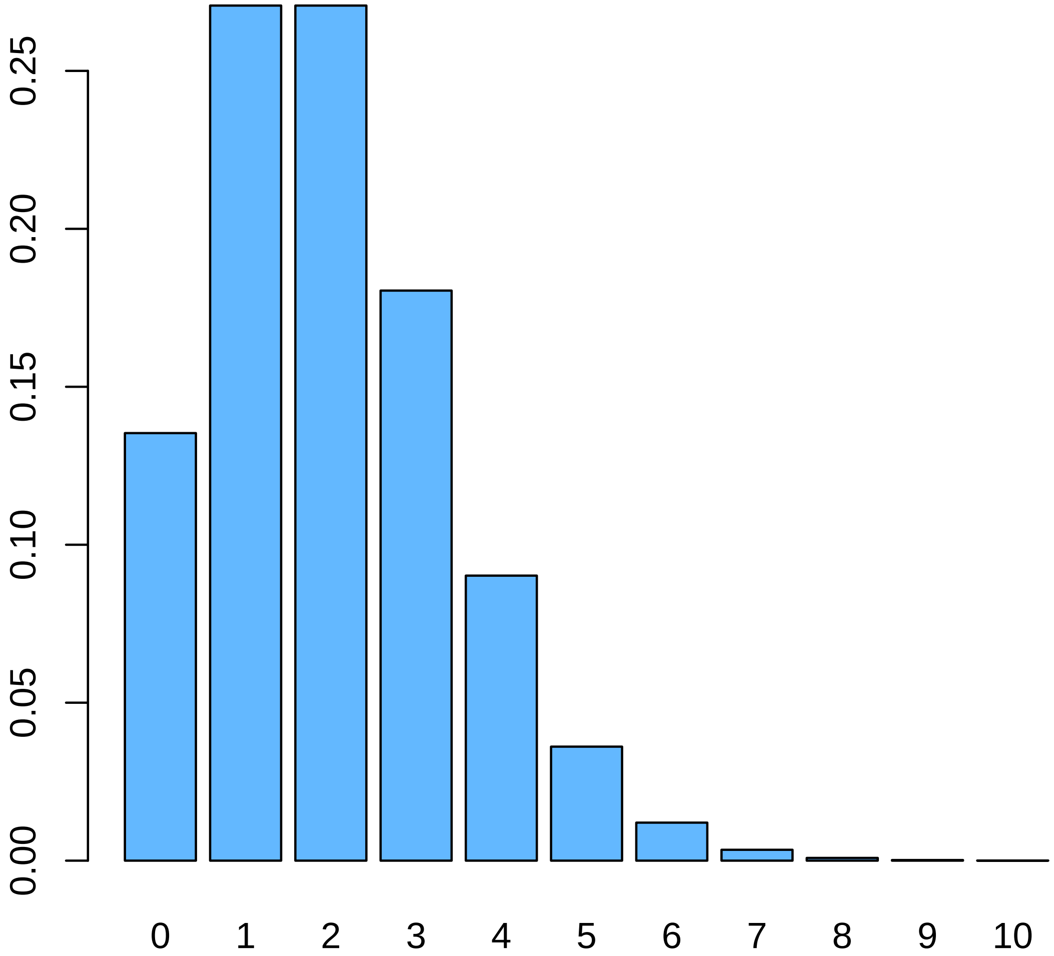
X adalah banyaknya sukses dalam eksperimen Poisson, yang mempunyai distribusi probabilitas

$$P(X = x; \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

Mean dan Variansi
$\mathrm{E}(X) = \lambda$; $\qquad \mathrm{Var}(X) = \lambda$

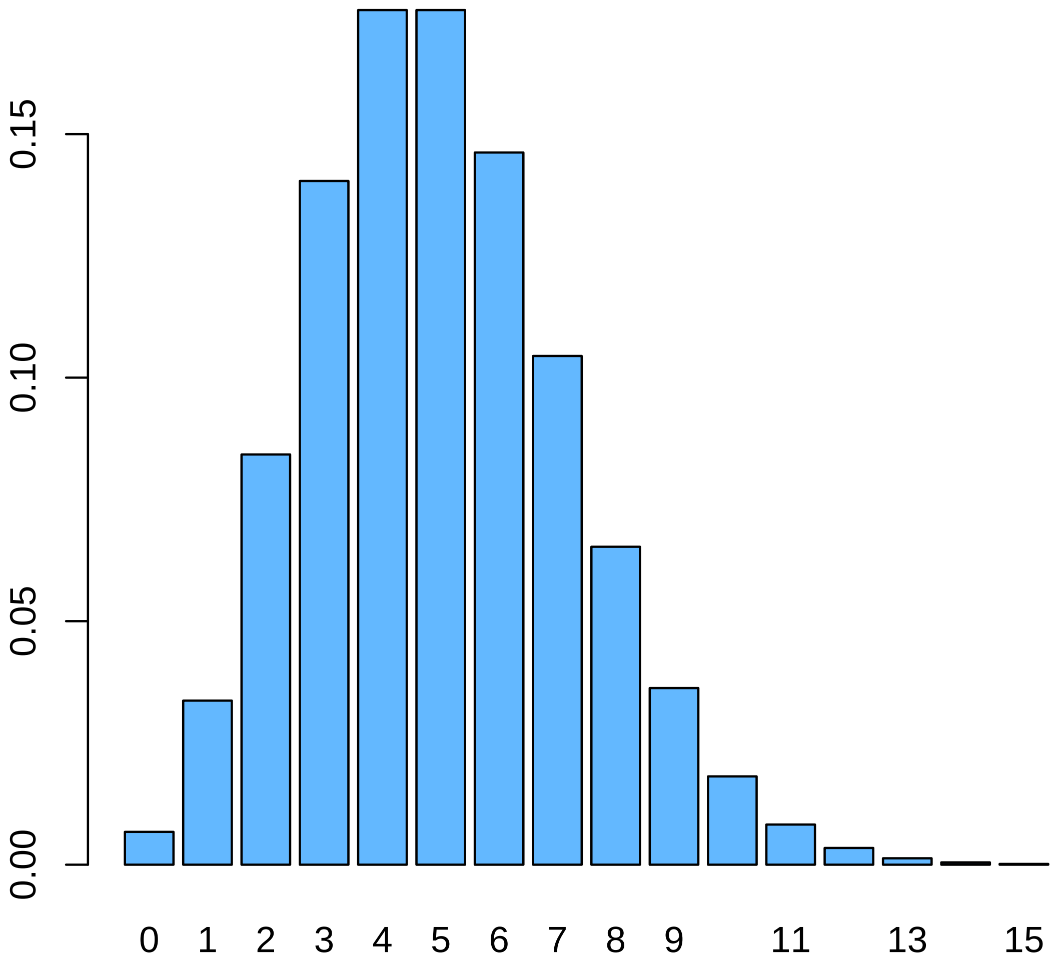
Distribusi Variabel Random Diskret dan Kontinu

Poisson dengan $\lambda = 2$



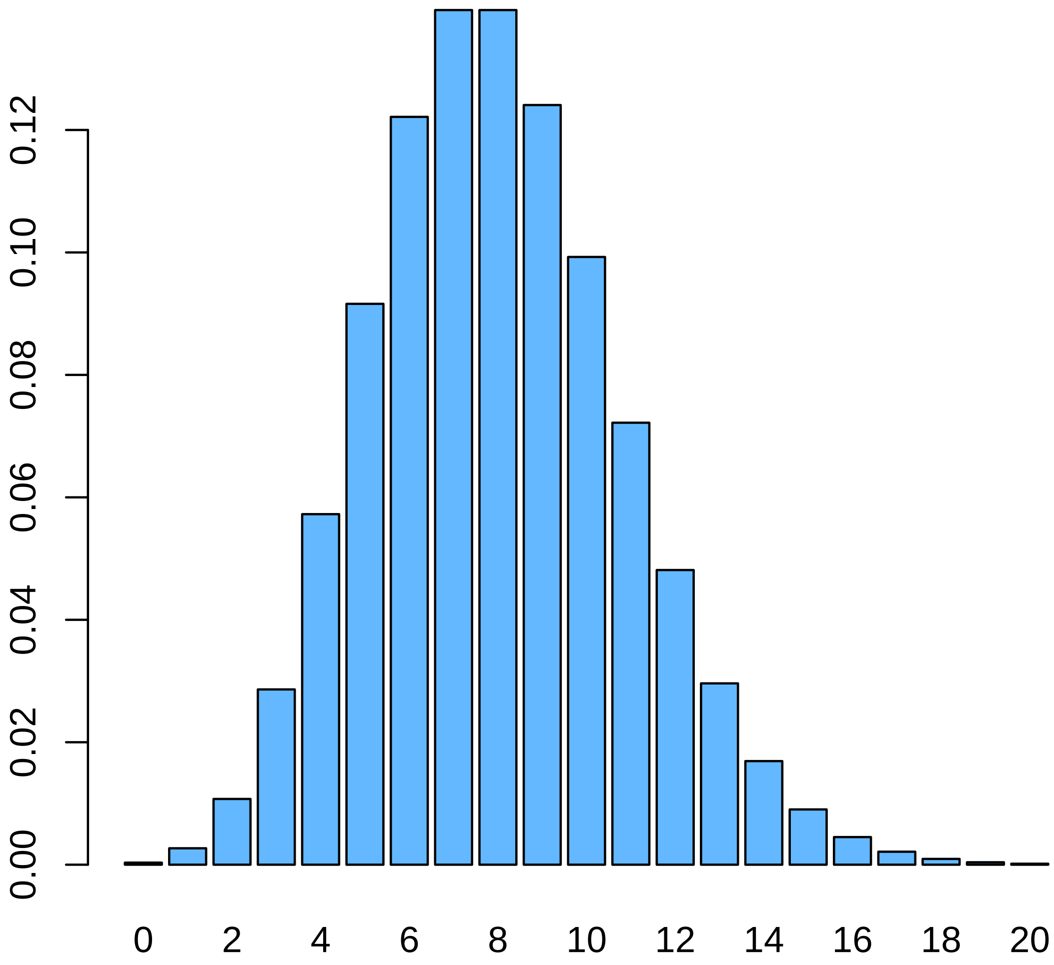
Distribusi Variabel Random Diskret dan Kontinu

Poisson dengan $\lambda = 5$



Distribusi Variabel Random Diskret dan Kontinu

Poisson dengan $\lambda = 8$



Distribusi Variabel Random Diskret dan Kontinu

Contoh (Distribusi Poisson)

Rata-rata banyaknya partikel radioaktif yang melewati suatu counter selama 1 milidetik dalam suatu percobaan di laboratorium adalah 4. Peluang 6 partikel melewati counter dalam suatu milidetik tertentu adalah

$$P(X = 6; \lambda = 4) = \frac{e^{-4}4^x}{6!} = 0,1042$$

Distribusi Variabel Random Diskret dan Kontinu

Distribusi Normal

Distribusi Normal dengan mean $\mathrm{E}(X) = \mu$ dan variansi $\mathrm{Var}(X) = \sigma^2$ mempunyai fungsi peluang,

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty < x < \infty$$

dan $-\infty < \mu < \infty$, $\sigma^2 > 0$

Distribusi Variabel Random Diskret dan Kontinu

Distribusi Normal

Distribusi Normal dengan mean $\mathrm{E}(X) = \mu$ dan variansi $\mathrm{Var}(X) = \sigma^2$ (ditulis $N(\mu, \sigma^2)$) mempunyai fungsi peluang,

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty < x < \infty$$

dengan $-\infty < \mu < \infty$, $\sigma^2 > 0$, $\pi = 3,141593\ldots$ dan $e = 2,718282\ldots$

Distribusi Normal standar: distribusi Normal dengan mean 0 dan variansi 1, ditulis $N(0, 1)$

Distribusi Variabel Random Diskret dan Kontinu

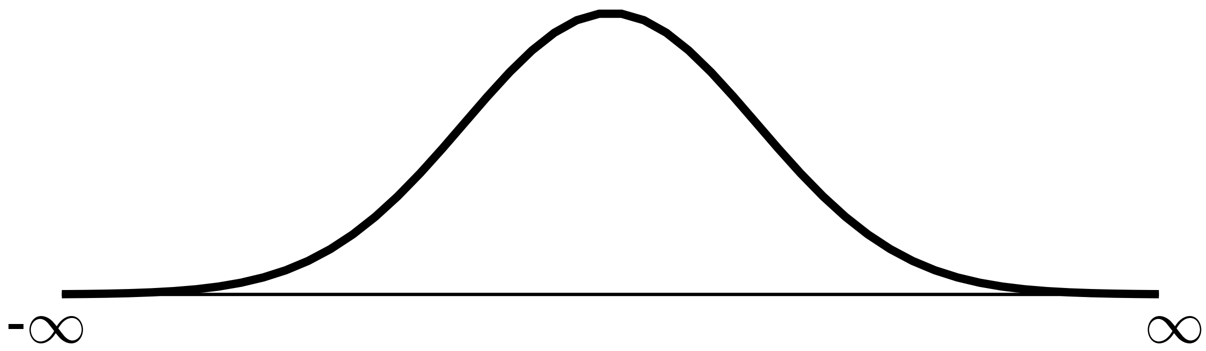
Kurva Normal

-∞ ∞

Sumbu x : $-\infty < x < \infty$

Distribusi Variabel Random Diskret dan Kontinu

Kurva Normal

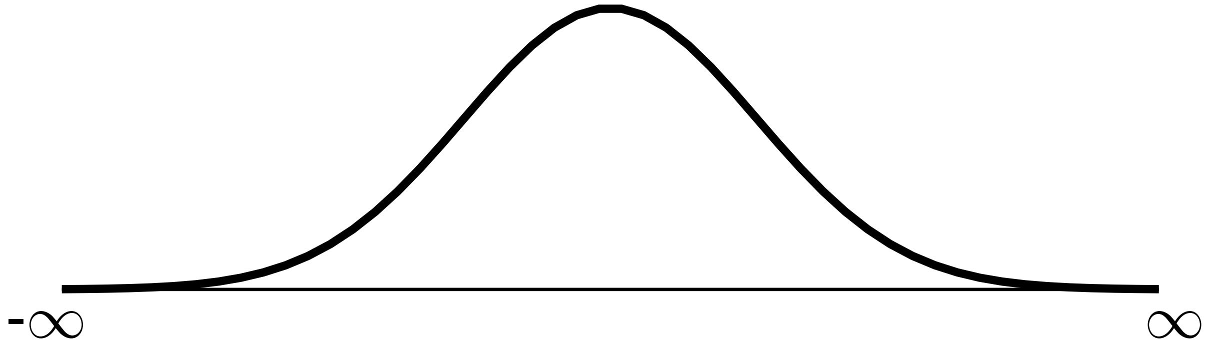


Sumbu x : $-\infty < x < \infty$
Fungsi peluang (sumbu y):

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad -\infty < x < \infty$$

Distribusi Variabel Random Diskret dan Kontinu

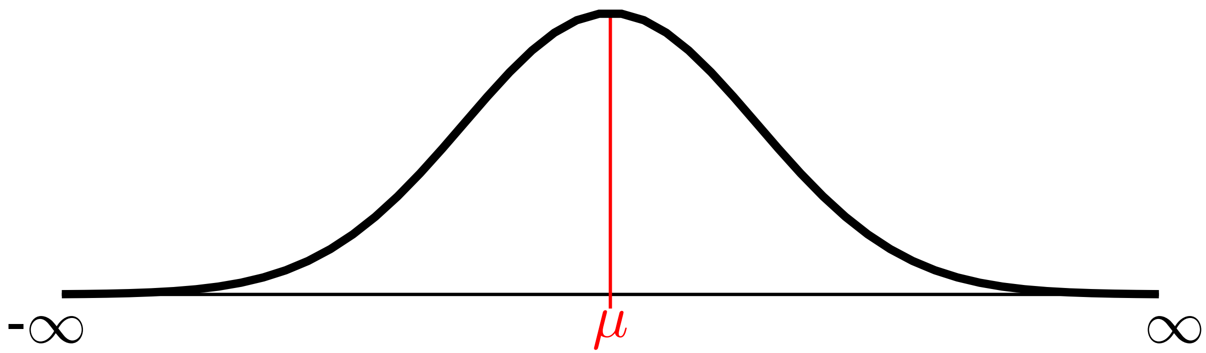
Kurva Normal



Sifat-sifat:

Distribusi Variabel Random Diskret dan Kontinu

Kurva Normal

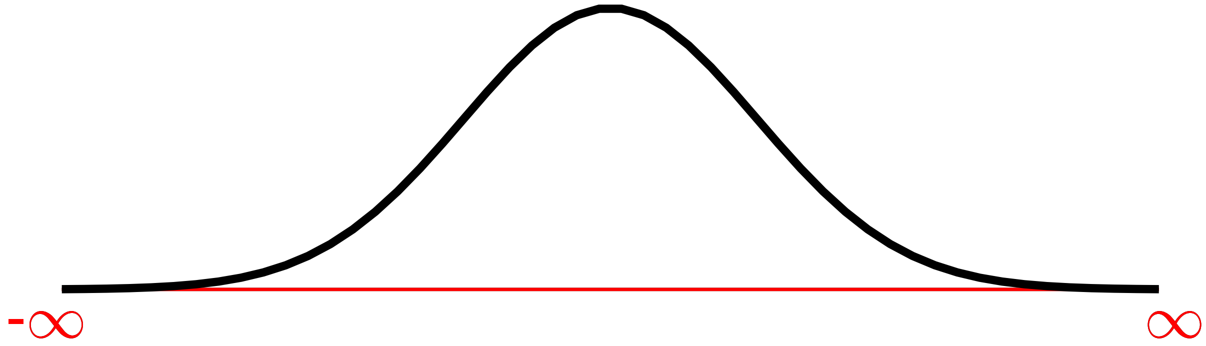


Sifat-sifat:

- simetris terhadap sumbu vertikal melalui μ,

Distribusi Variabel Random Diskret dan Kontinu

Kurva Normal

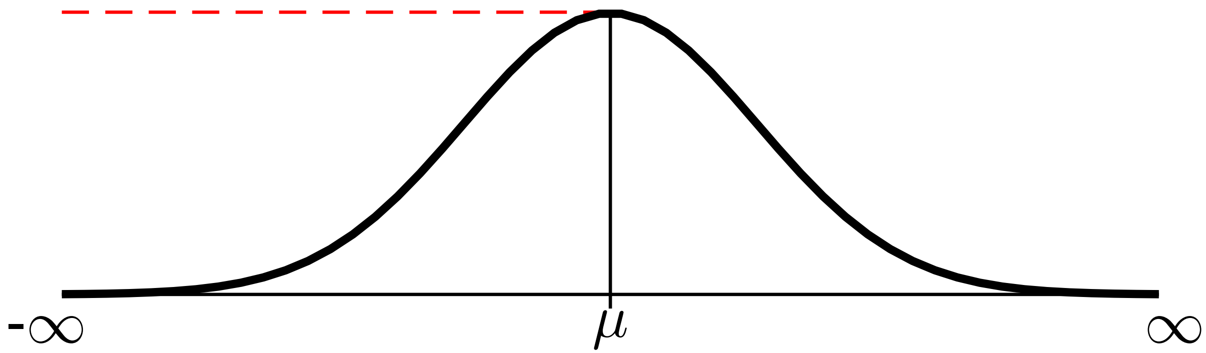


Sifat-sifat:

- simetris terhadap sumbu vertikal melalui μ,
- memotong sumbu mendatar (sumbu x) secara asimtotis,

Distribusi Variabel Random Diskret dan Kontinu

Kurva Normal

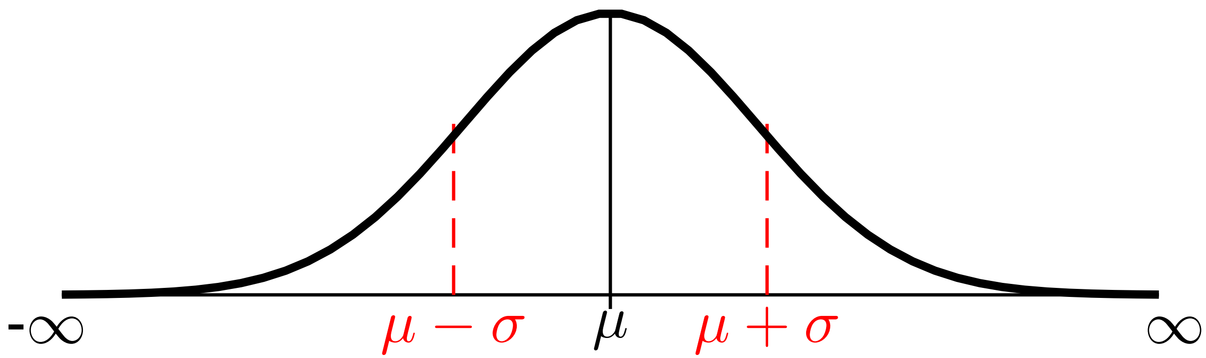


Sifat-sifat:

- simetris terhadap sumbu vertikal melalui μ,
- memotong sumbu mendatar (sumbu x) secara asimtotis,
- harga modus (maksimum) terletak pada $x = \mu$,

Distribusi Variabel Random Diskret dan Kontinu

Kurva Normal

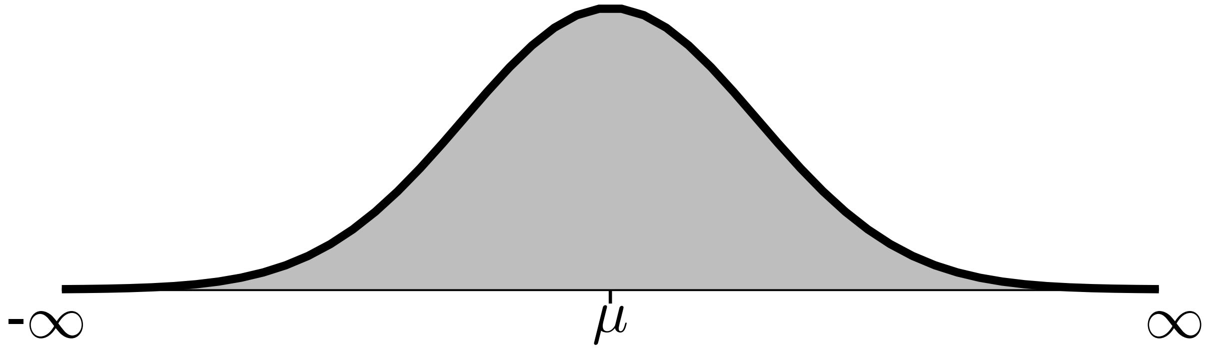


Sifat-sifat:

- simetris terhadap sumbu vertikal melalui μ,
- memotong sumbu mendatar (sumbu x) secara asimtotis,
- harga modus (maksimum) terletak pada $x = \mu$,
- mempunyai titik belok pada $x = \mu \pm \sigma$,

Distribusi Variabel Random Diskret dan Kontinu

Kurva Normal

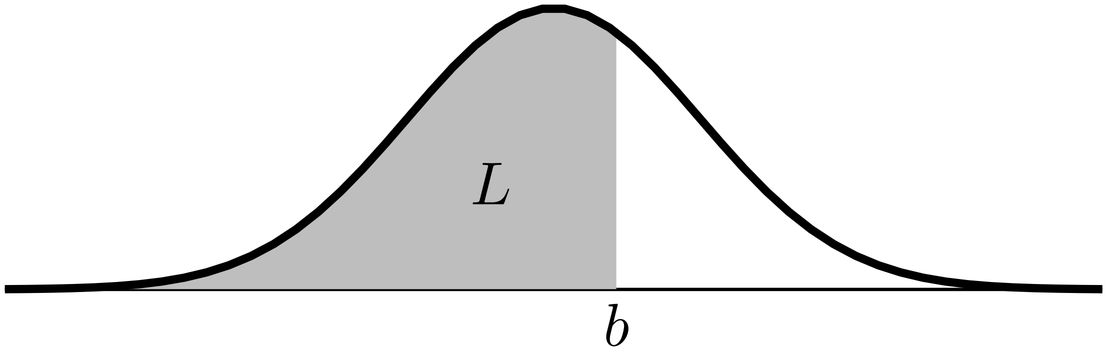


Sifat-sifat:

- simetris terhadap sumbu vertikal melalui μ,
- memotong sumbu mendatar (sumbu x) secara asimtotis,
- harga modus (maksimum) terletak pada $x = \mu$,
- mempunyai titik belok pada $x = \mu \pm \sigma$,
- luas kurva Normal sama dengan 1.

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal

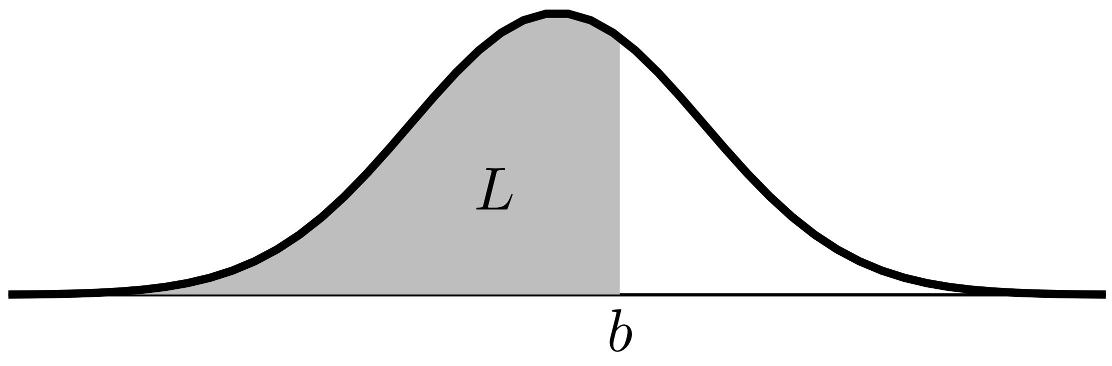


Luasan kurva di bawah kurva normal sampai batas b:

$$L = \int_{-\infty}^{b} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal

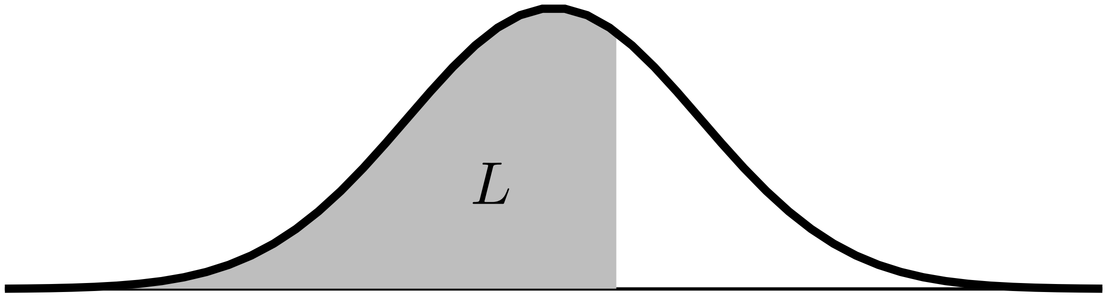


Dapat dihitung menggunakan tabel Normal Standar dengan terlebih dahulu mentransformasikan skala $X \sim N(\mu, \sigma^2)$ ke $Z \sim N(0, 1)$,

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

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal

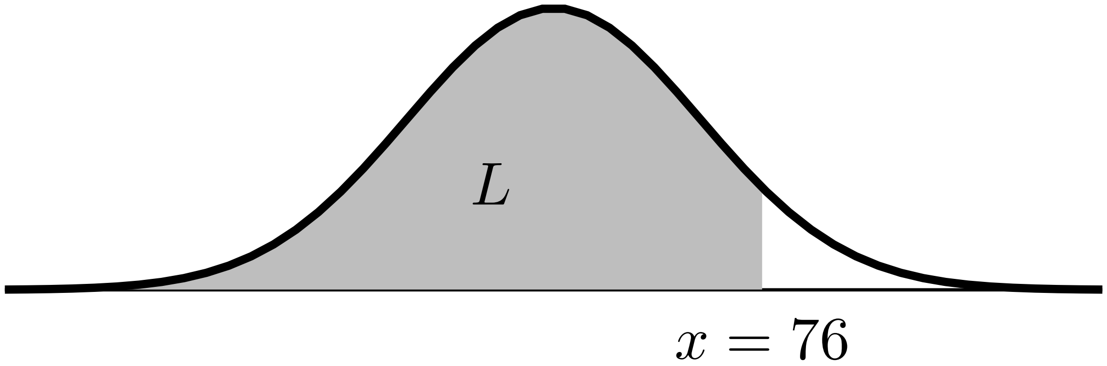


$$\frac{X-b}{\sigma}$$

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-3,4										
-3,3										
...										
0,0										
...										
3,3										
3,4										

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



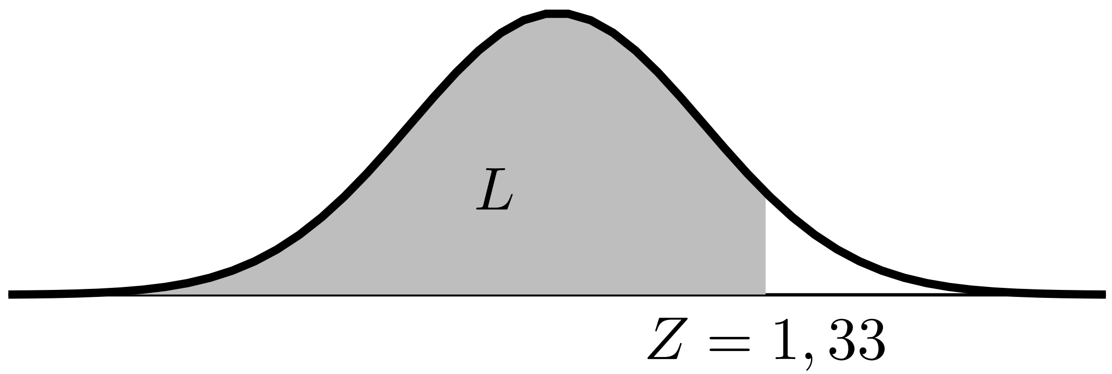
Contoh 1:

Distribusi Normal dengan mean $\mu = 60$ dan deviasi standar $\sigma = 12$, $N(60, 12^2)$

Hitunglah luas kurva Normal mulai ekor paling kiri ($-\infty$) sampai 76

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



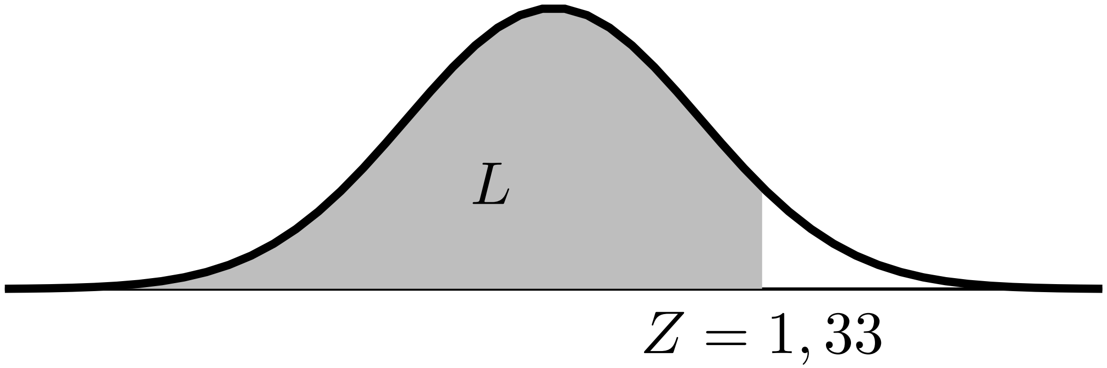
Contoh 1:

transformasi dari X ke Z,

$$\begin{aligned} Z &= \frac{X-\mu}{\sigma} \\ &= \frac{76-60}{12} \\ &= 1,33 \end{aligned}$$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal

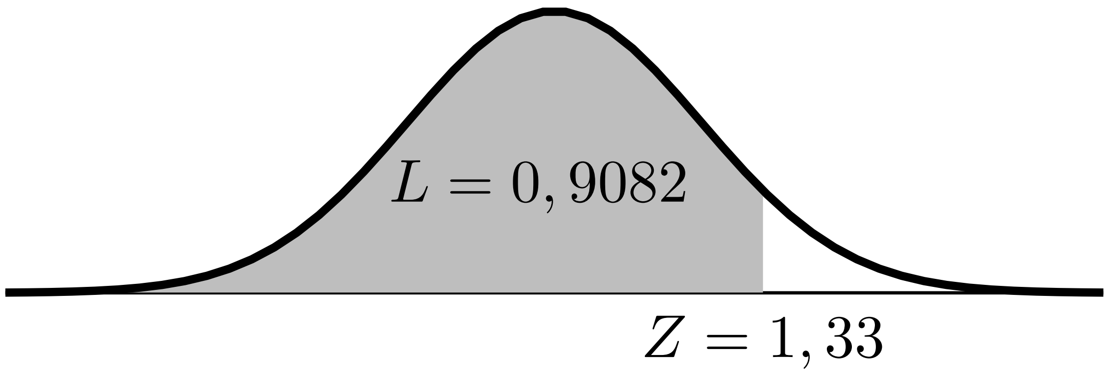


Contoh 1:

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
...										
0,0										
...										
1,3										

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal

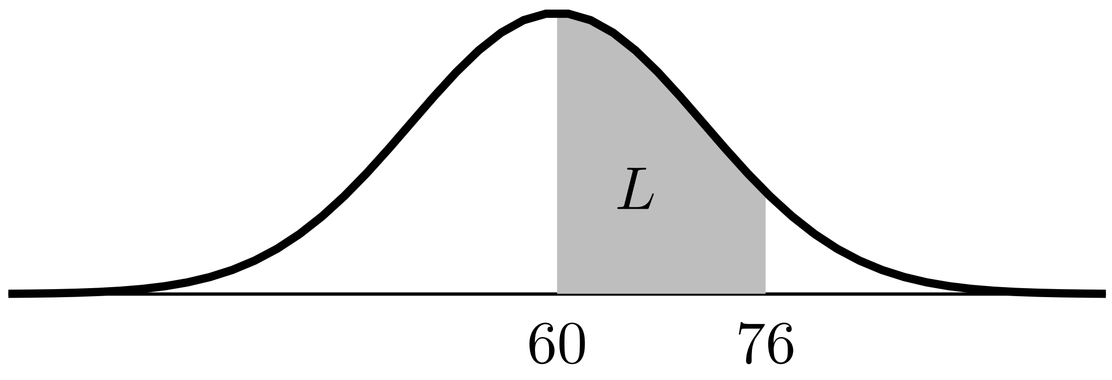


Contoh 1:

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
...										
0,0										
...										
1,3				0,9082						

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



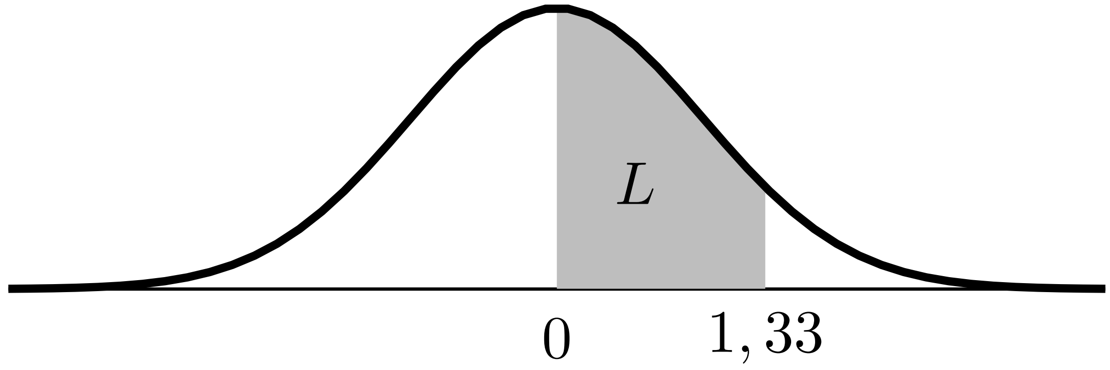
Contoh 2:

Distribusi Normal dengan mean $\mu = 60$ dan deviasi standar $\sigma = 12$, $N(60, 12^2)$

Hitunglah luas kurva Normal antara 60 sampai 76

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



Contoh 2:

transformasi dari $X = 60$ ke Z,

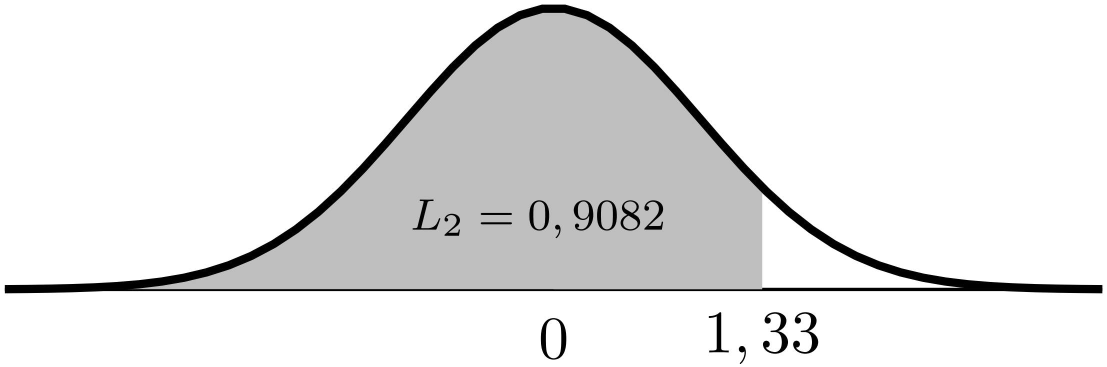
$$\begin{aligned} Z &= \frac{X-\mu}{\sigma} \\ &= \frac{60-60}{12} \\ &= 0 \end{aligned}$$

transformasi dari $X = 76$ ke Z,

$$\begin{aligned} Z &= \frac{X-\mu}{\sigma} \\ &= \frac{76-60}{12} \\ &= 1,33 \end{aligned}$$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal

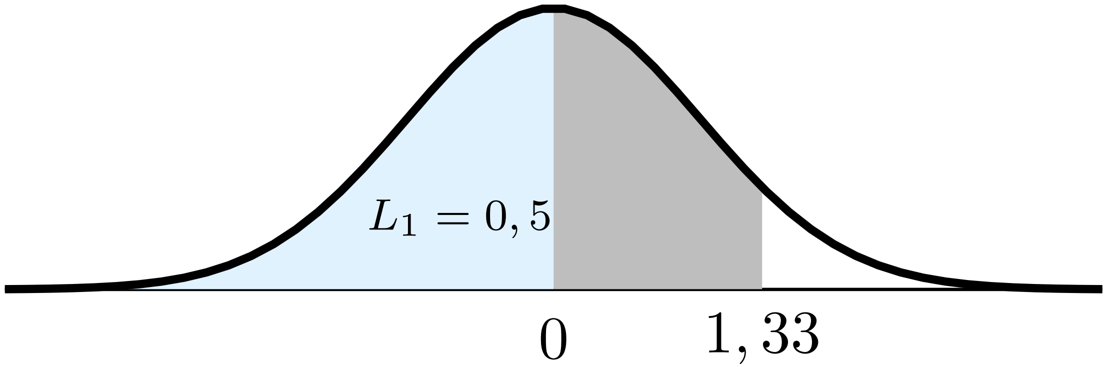


Contoh 2:

$$\begin{aligned} L &= L_2 - L_1 \\ &= 0,9082 - L_1 \end{aligned}$$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal

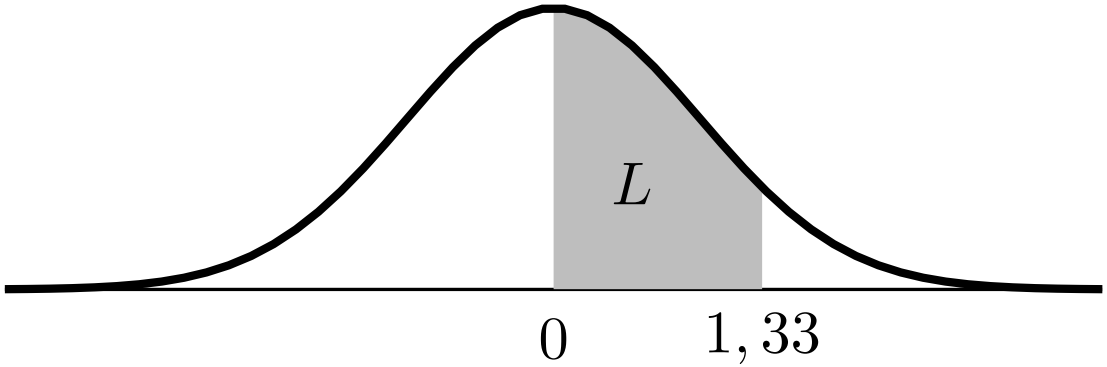


Contoh 2:

$$
\begin{aligned}
L &= L_2 - L_1 \\
&= 0,9082 - 0,5
\end{aligned}
$$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal

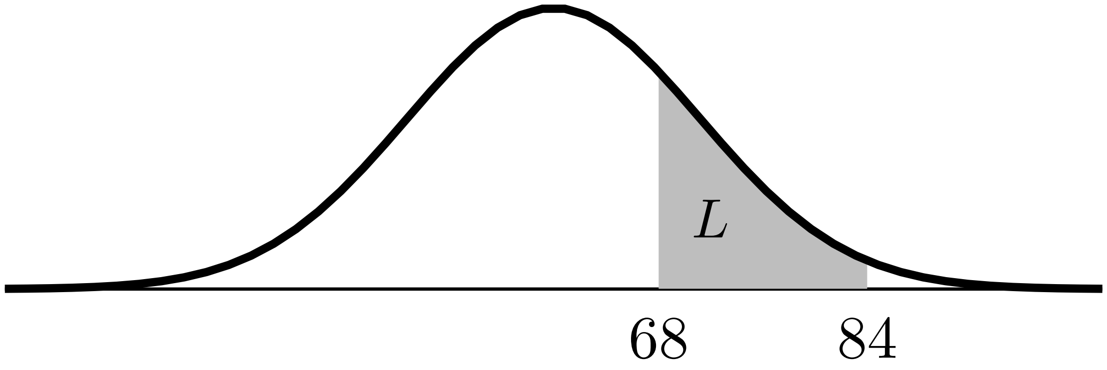


Contoh 2:

$$
\begin{aligned}
L &= L_2 - L_1 \\
&= 0,9082 - 0,5 \\
&= 0,4082
\end{aligned}
$$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal

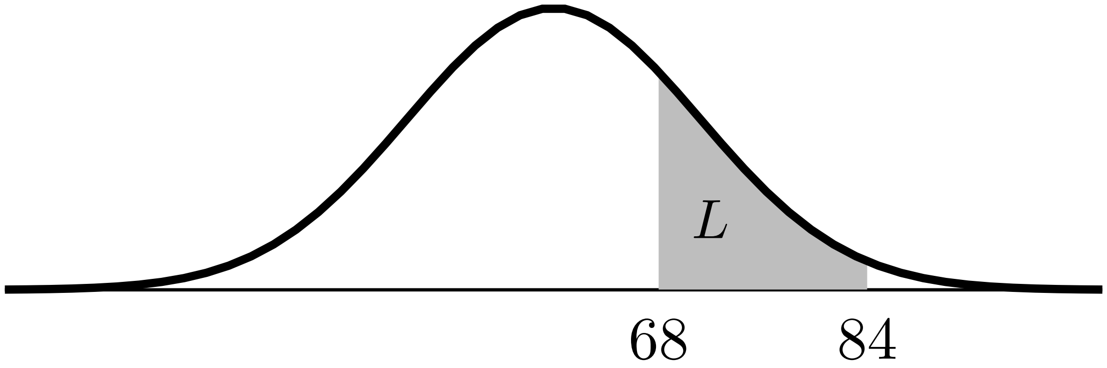


Contoh 3:

Distribusi Normal dengan mean $\mu = 60$ dan deviasi standar $\sigma = 12$, $N(60, 12^2)$. Hitunglah luas kurva Normal antara 68 sampai 84.

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



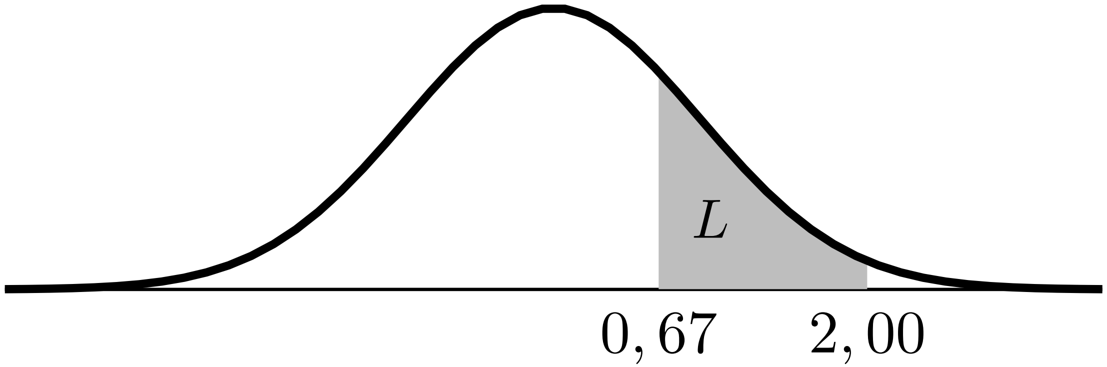
Contoh 3:

Distribusi Normal dengan mean $\mu = 60$ dan deviasi standar $\sigma = 12$, $N(60, 12^2)$. Hitunglah luas kurva Normal antara 68 sampai 84.

$L = P(68 \leq X \leq 84)$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



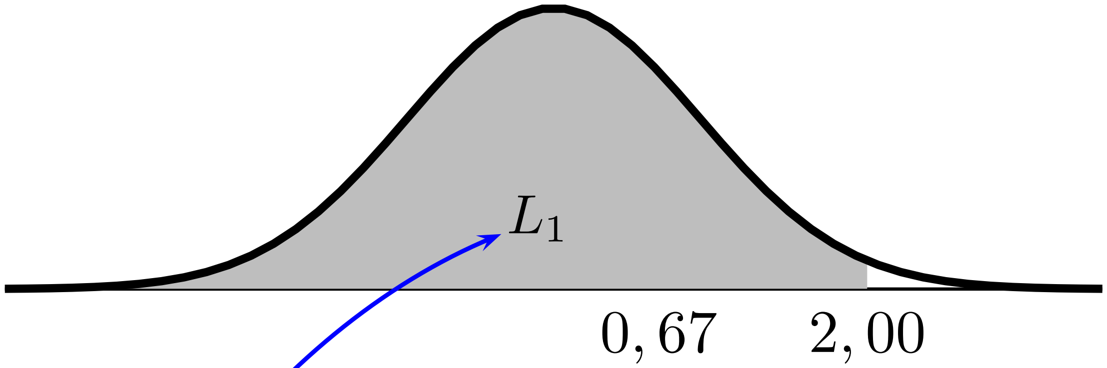
Contoh 3:

Distribusi Normal dengan mean $\mu = 60$ dan deviasi standar $\sigma = 12$, $N(60, 12^2)$. Hitunglah luas kurva Normal antara 68 sampai 84.

$$L = P(68 \le X \le 84)$$
$$= P(0,67 \le Z \le 2,00)$$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



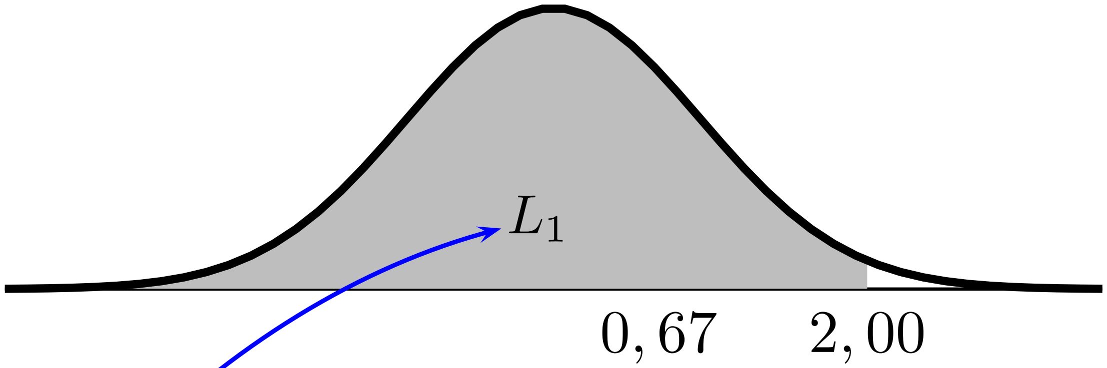
Contoh 3:

Distribusi Normal dengan mean $\mu = 60$ dan deviasi standar $\sigma = 12$, $N(60, 12^2)$. Hitunglah luas kurva Normal antara 68 sampai 84.

$$L = P(68 \leq X \leq 84)$$
$$= P(0,67 \leq Z \leq 2,00)$$
$$= P(-\infty < Z \leq 2,00) - P(-\infty < Z \leq 0,67)$$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



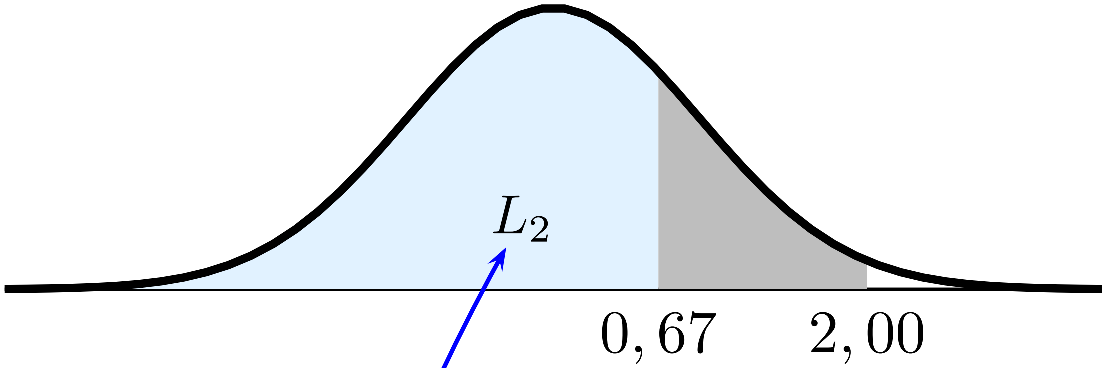
Contoh 3:

Distribusi Normal dengan mean $\mu = 60$ dan deviasi standar $\sigma = 12$, $N(60, 12^2)$. Hitunglah luas kurva Normal antara 68 sampai 84.

$$L = P(68 \leq X \leq 84)$$
$$= P(0,67 \leq Z \leq 2,00)$$
$$= 0,9772 - P(-\infty < Z \leq 0,67)$$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



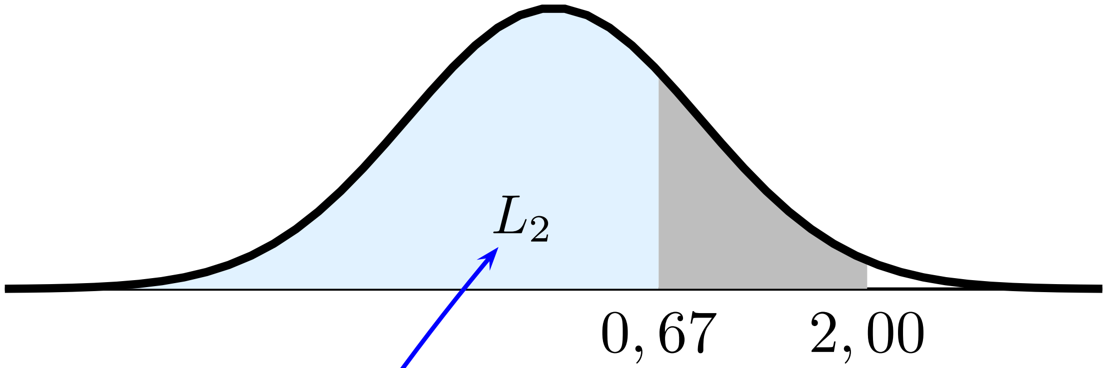
Contoh 3:

Distribusi Normal dengan mean $\mu = 60$ dan deviasi standar $\sigma = 12$, $N(60, 12^2)$. Hitunglah luas kurva Normal antara 68 sampai 84.

$$
\begin{aligned}
L &= P(68 \leq X \leq 84) \\
&= P(0,67 \leq Z \leq 2,00) \\
&= 0,9772 - P(-\infty < Z \leq 0,67)
\end{aligned}
$$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



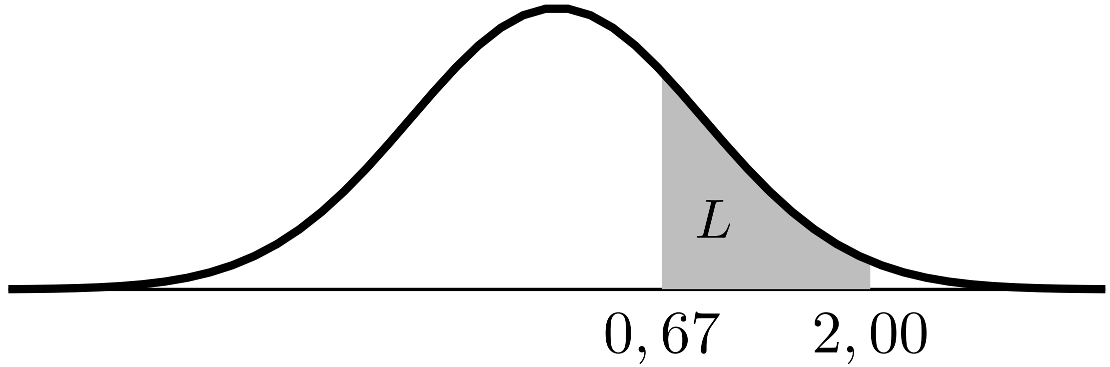
Contoh 3:

Distribusi Normal dengan mean $\mu = 60$ dan deviasi standar $\sigma = 12$, $N(60, 12^2)$. Hitunglah luas kurva Normal antara 68 sampai 84.

$$
\begin{aligned}
L &= P(68 \leq X \leq 84) \\
&= P(0,67 \leq Z \leq 2,00) \\
&= 0,9772 - 0,7486
\end{aligned}
$$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



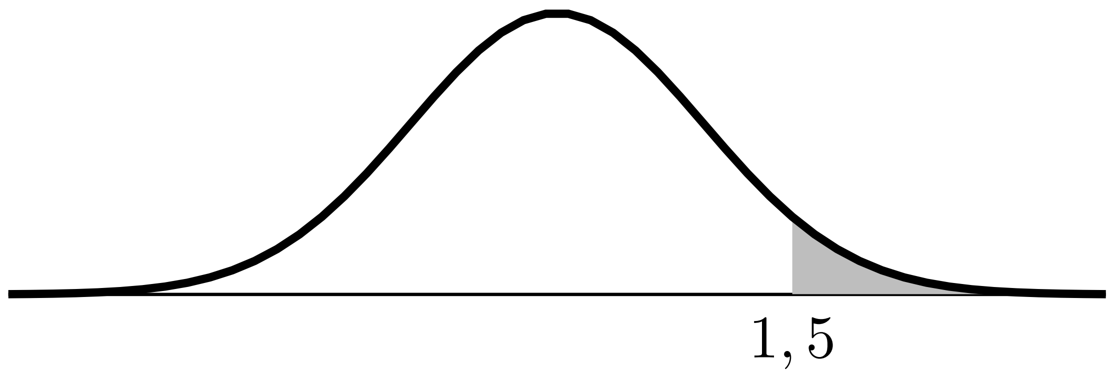
Contoh 3:

Distribusi Normal dengan mean $\mu = 60$ dan deviasi standar $\sigma = 12$, $N(60, 12^2)$. Hitunglah luas kurva Normal antara 68 sampai 84.

$$L = P(68 \leq X \leq 84)$$
$$= P(0,67 \leq Z \leq 2,00)$$
$$= 0,2286$$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal

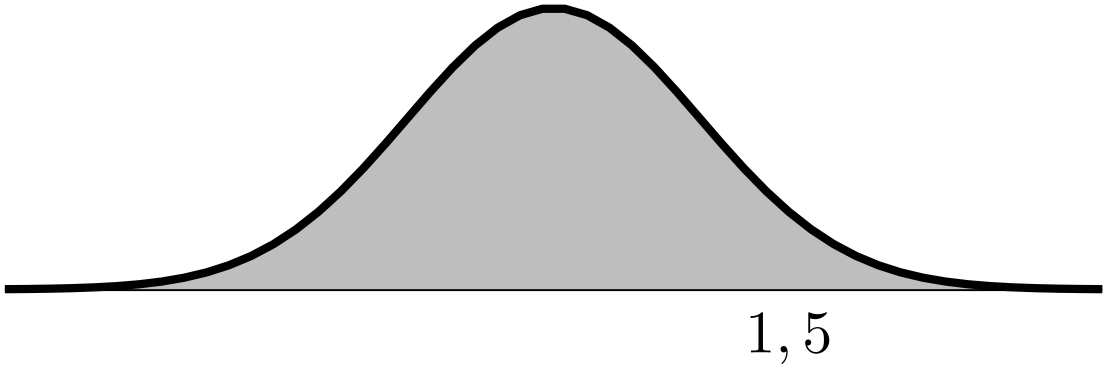


Contoh 4:

Diketahui $N(0,1)$, hitunglah $P(Z \geq 1,5)$.

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



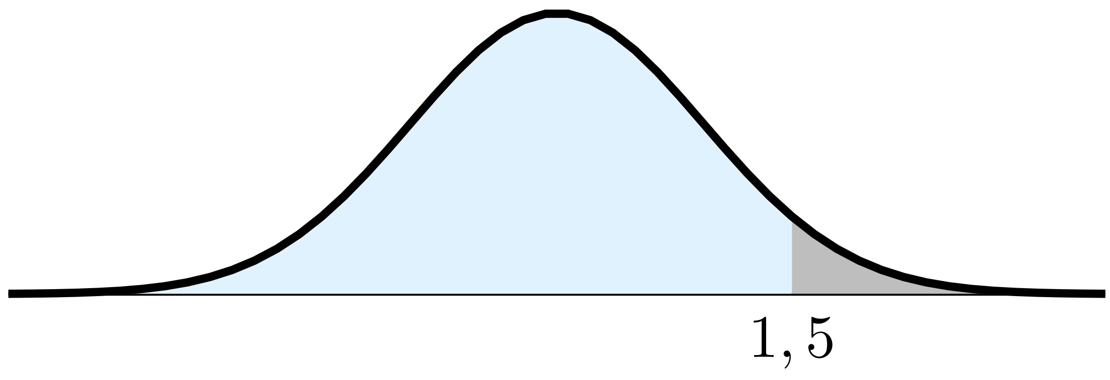
Contoh 4:

Diketahui $N(0,1)$, hitunglah $P(Z \geq 1,5)$.

$P(Z \geq 1,5) = 1 - P(-\infty \leq Z \leq 1,5)$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



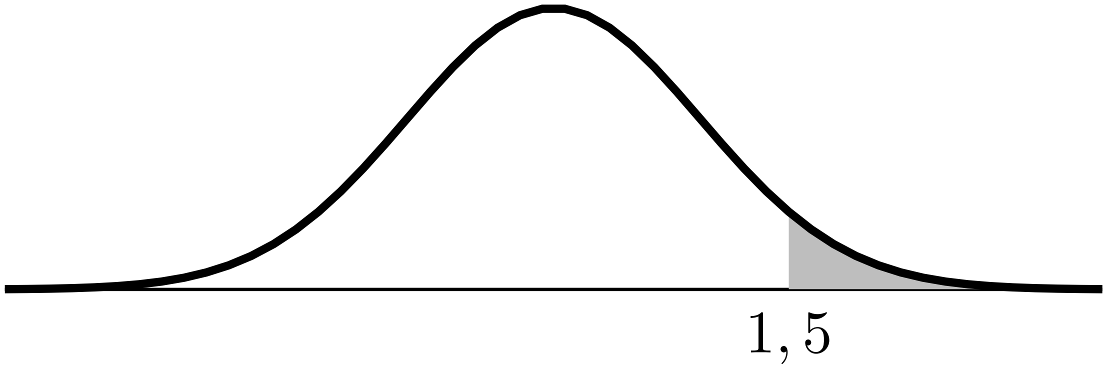
Contoh 4:

Diketahui $N(0,1)$, hitunglah $P(Z \geq 1,5)$.

$P(Z \geq 1,5) = 1 - P(-\infty \leq Z \leq 1,5)$

Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



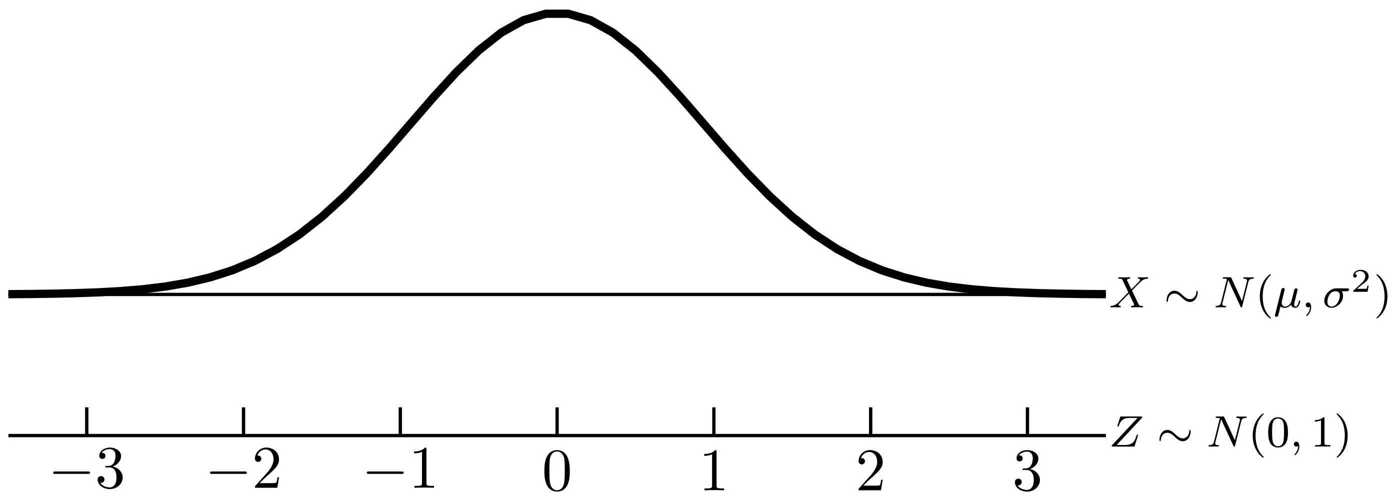
Contoh 4:

Diketahui $N(0,1)$, hitunglah $P(Z \geq 1,5)$.

$$\begin{aligned} P(Z \geq 1,5) &= 1 - P(-\infty \leq Z \leq 1,5) \\ &= 1 - 0,9332 \\ &= 0,0668 \end{aligned}$$

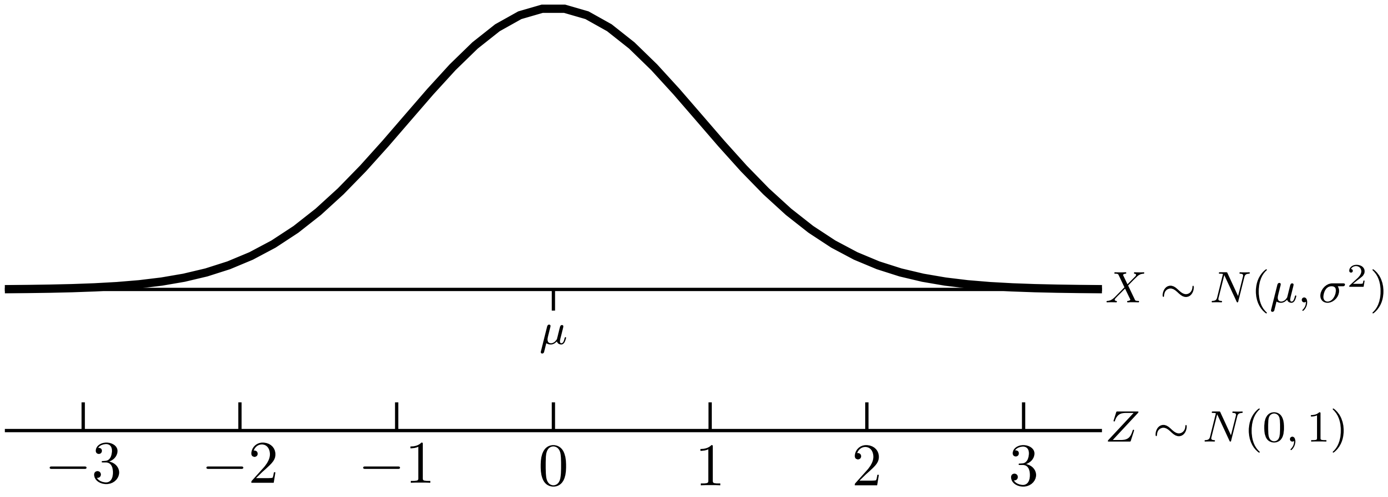
Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



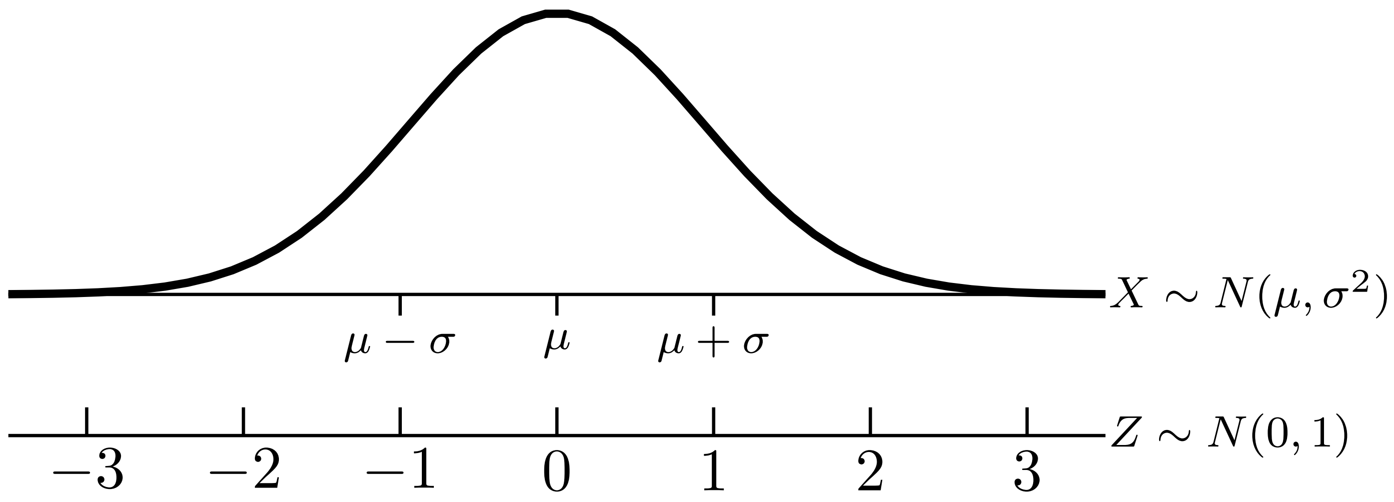
Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



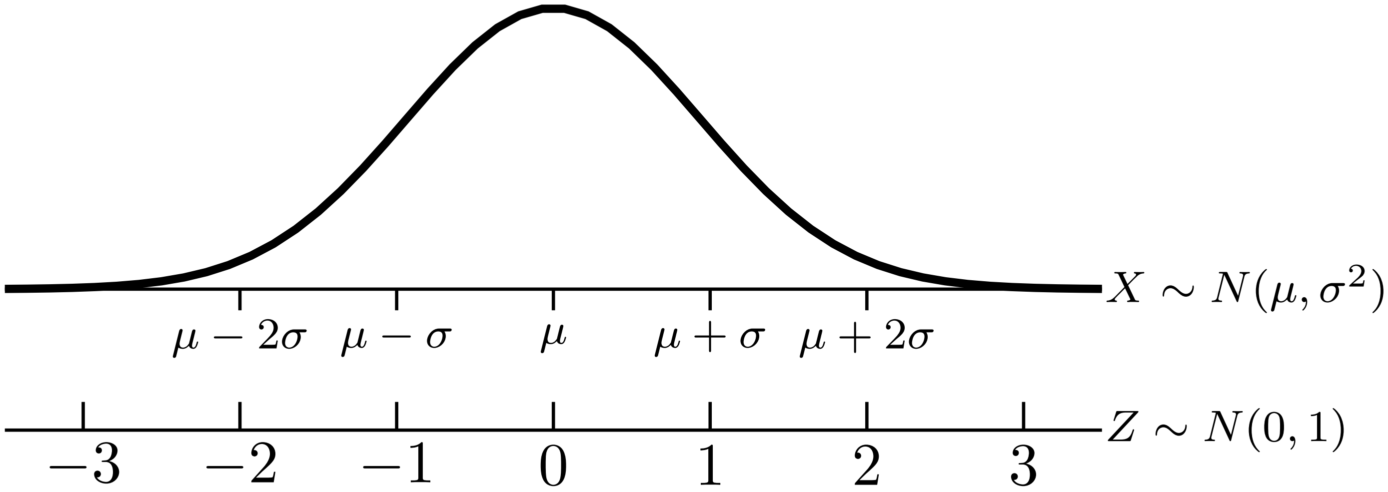
Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



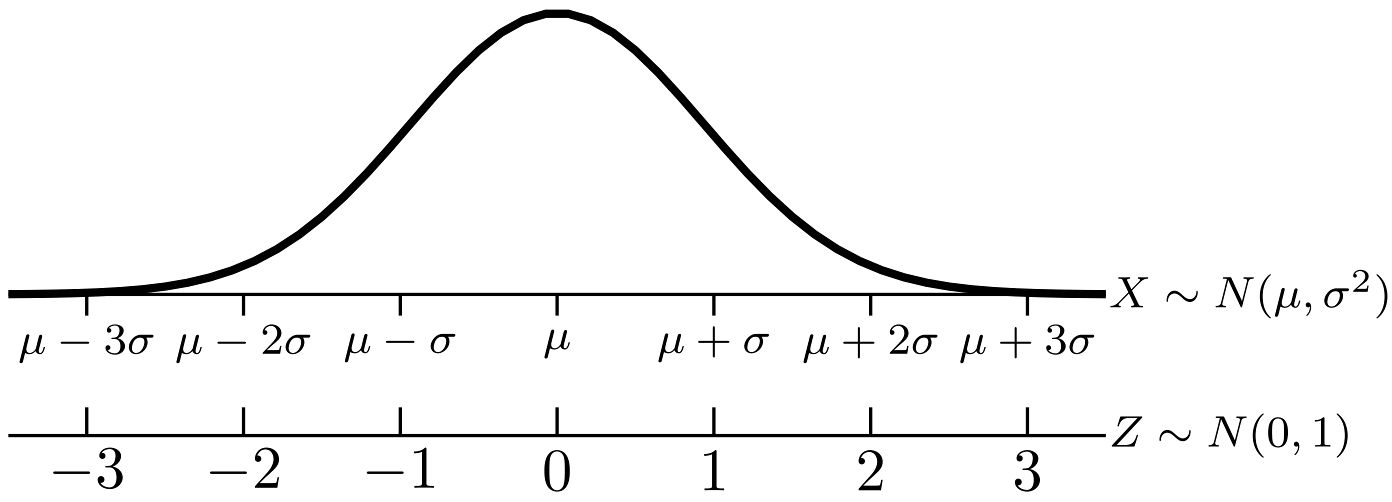
Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



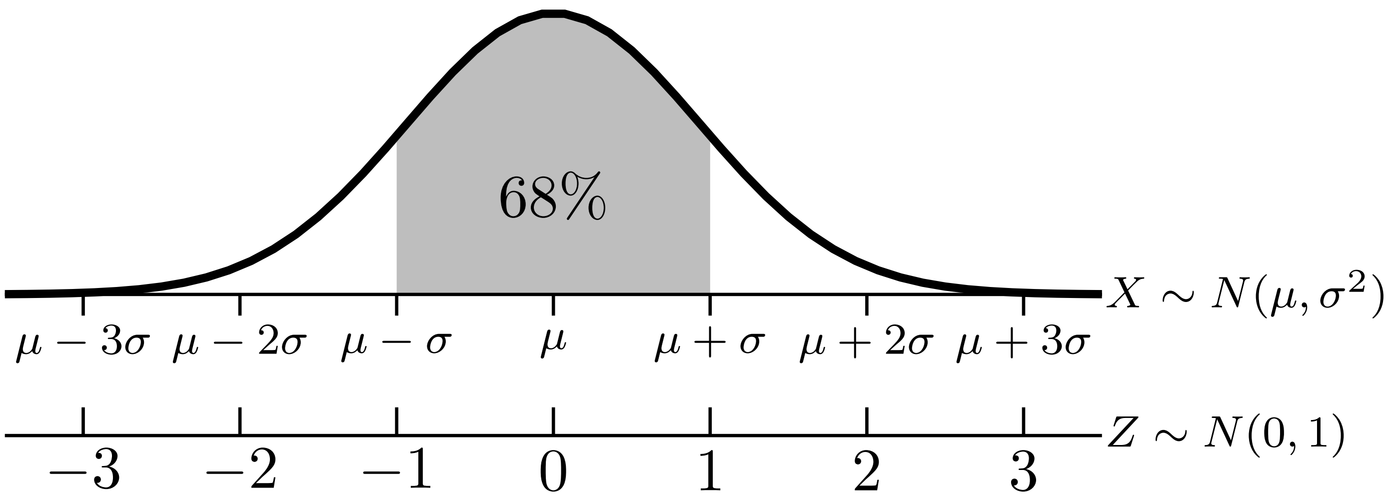
Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



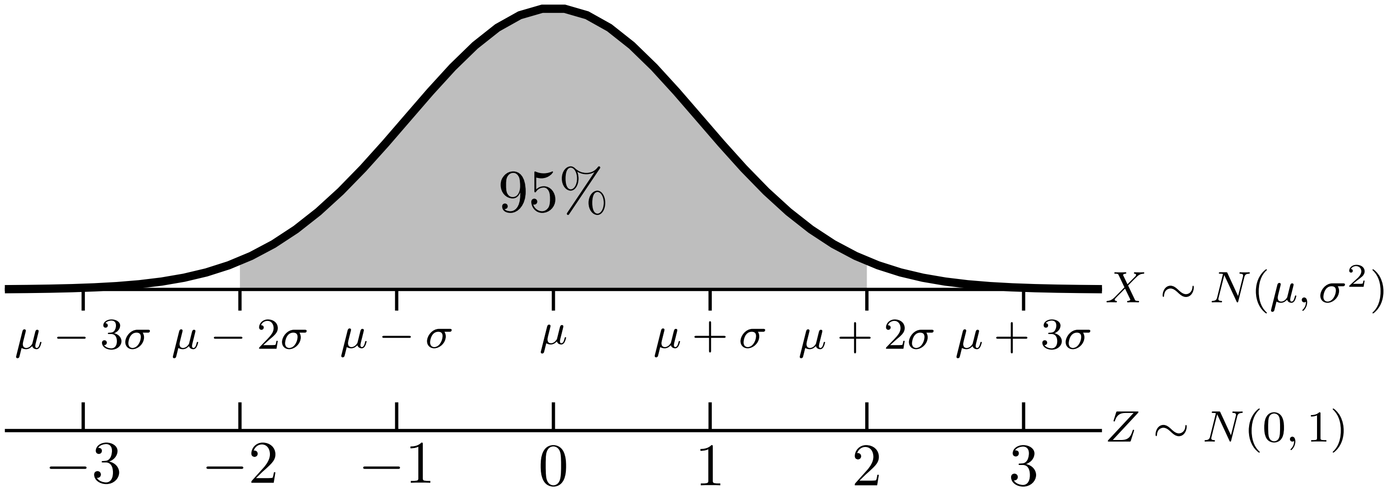
Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



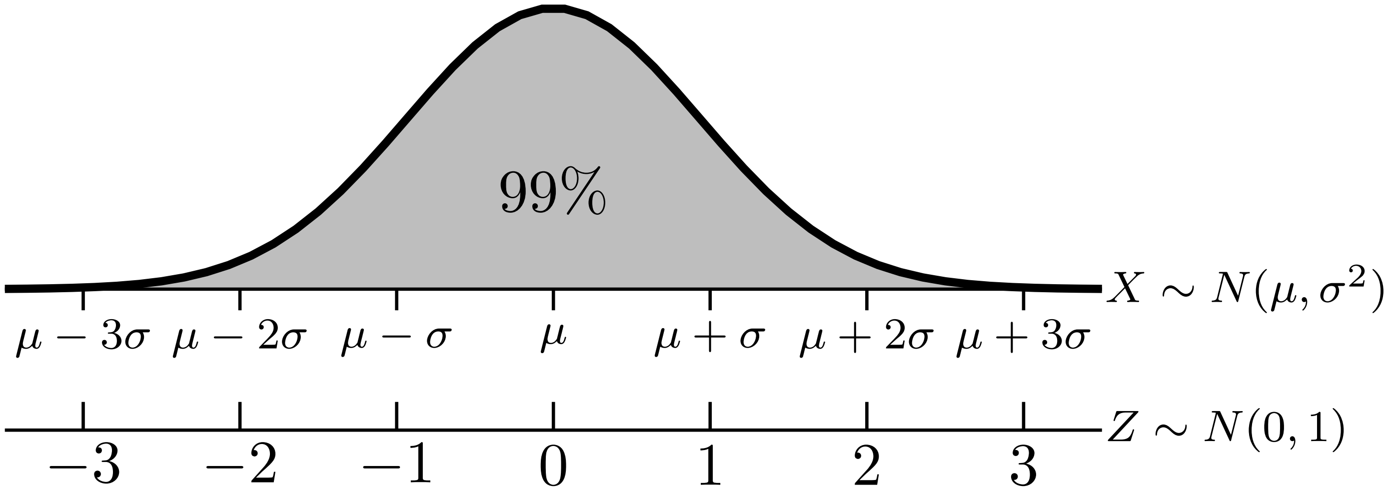
Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal



Distribusi Variabel Random Diskret dan Kontinu

Luasan di bawah Kurva Normal

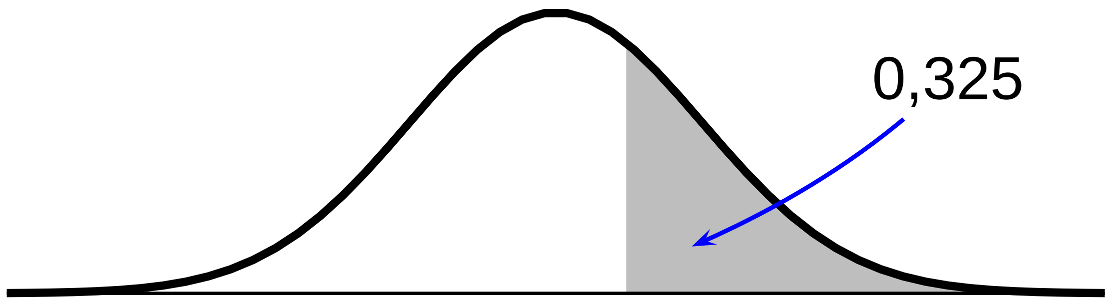


Distribusi Variabel Random Diskret dan Kontinu

Contoh 5 (Distribusi Normal)
Nilai-nilai ujian seleksi penerimaan mahasiswa baru secara nasional dianggap berdistribusi Normal dengan mean 45 dan deviasi standar 13. Jika hanya 32,5% calon mahasiswa yang akan diterima, berapakah nilai terendah calon mahasiswa yang diterima?

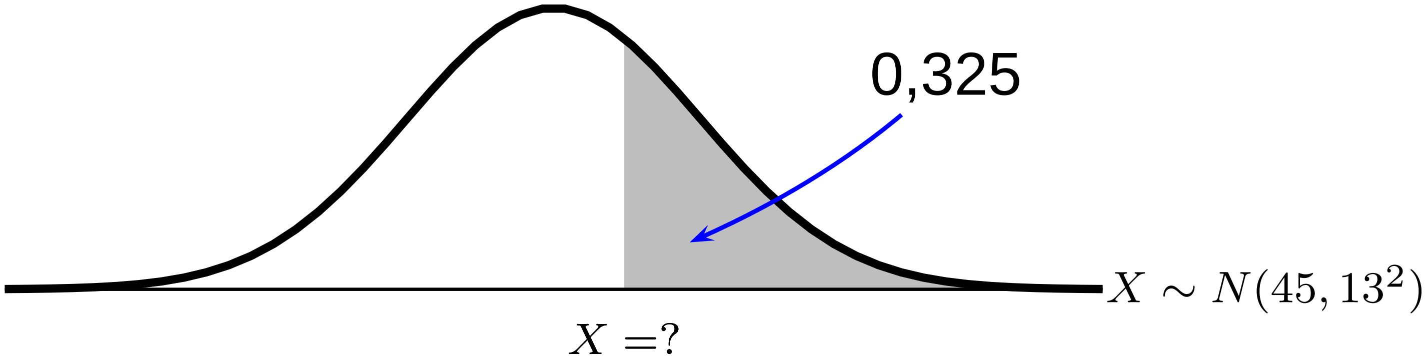
Distribusi Variabel Random Diskret dan Kontinu

Contoh 5 (Distribusi Normal)
Nilai-nilai ujian seleksi penerimaan mahasiswa baru secara nasional dianggap berdistribusi Normal dengan mean 45 dan deviasi standar 13. Jika hanya 32,5% calon mahasiswa yang akan diterima, berapakah nilai terendah calon mahasiswa yang diterima?



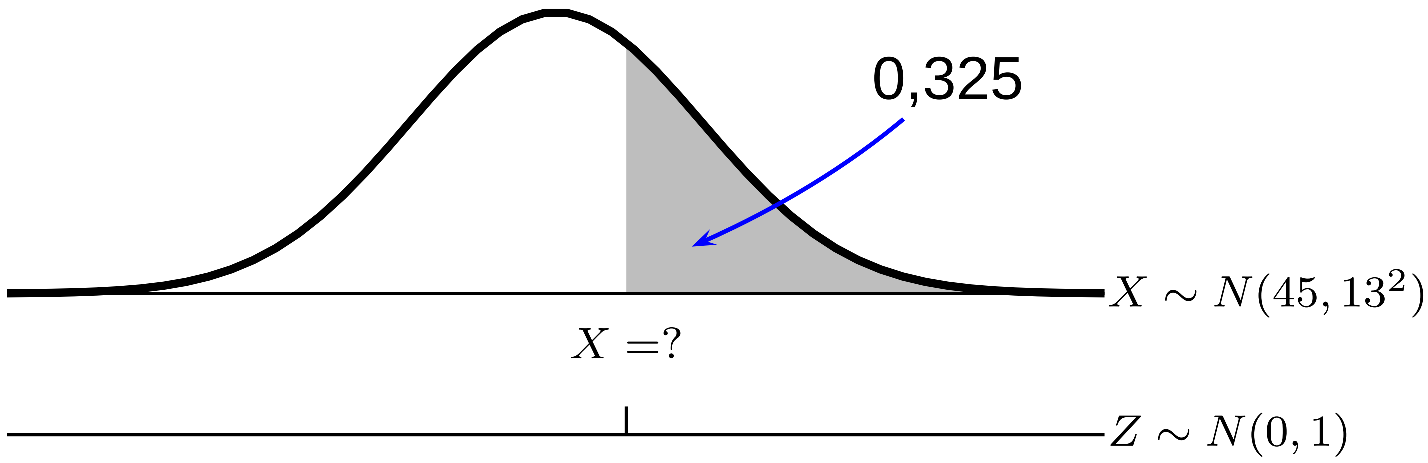
Distribusi Variabel Random Diskret dan Kontinu

Contoh 5 (Distribusi Normal)
Nilai-nilai ujian seleksi penerimaan mahasiswa baru secara nasional dianggap berdistribusi Normal dengan mean 45 dan deviasi standar 13. Jika hanya 32,5% calon mahasiswa yang akan diterima, berapakah nilai terendah calon mahasiswa yang diterima?



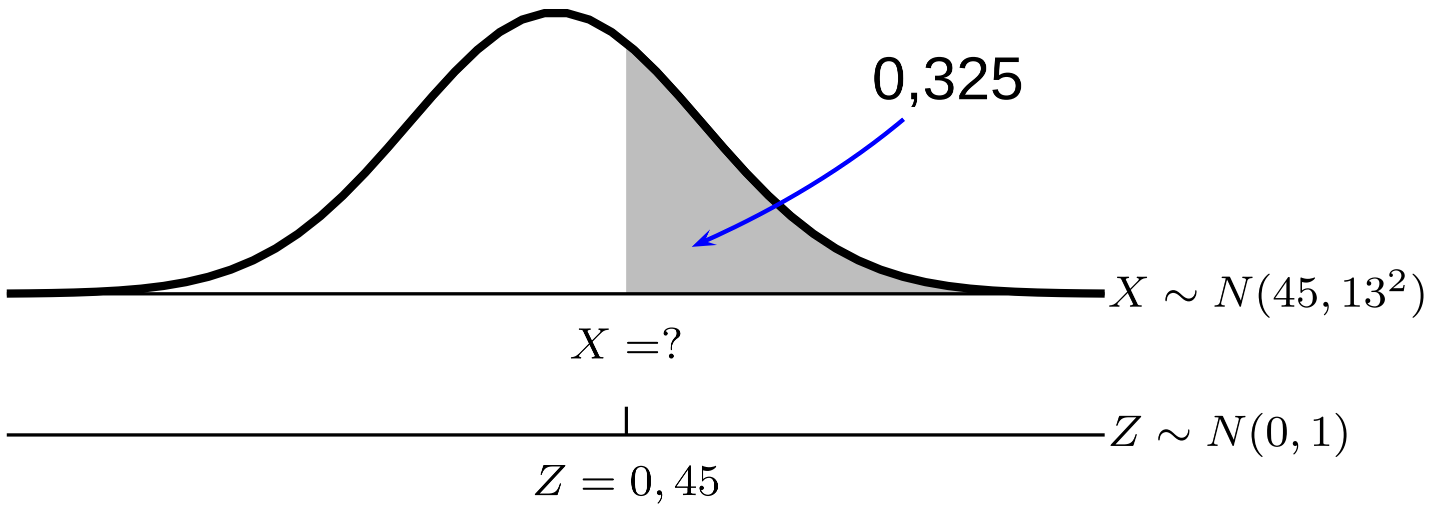
Distribusi Variabel Random Diskret dan Kontinu

Contoh 5 (Distribusi Normal)
Nilai-nilai ujian seleksi penerimaan mahasiswa baru secara nasional dianggap berdistribusi Normal dengan mean 45 dan deviasi standar 13. Jika hanya 32,5% calon mahasiswa yang akan diterima, berapakah nilai terendah calon mahasiswa yang diterima?



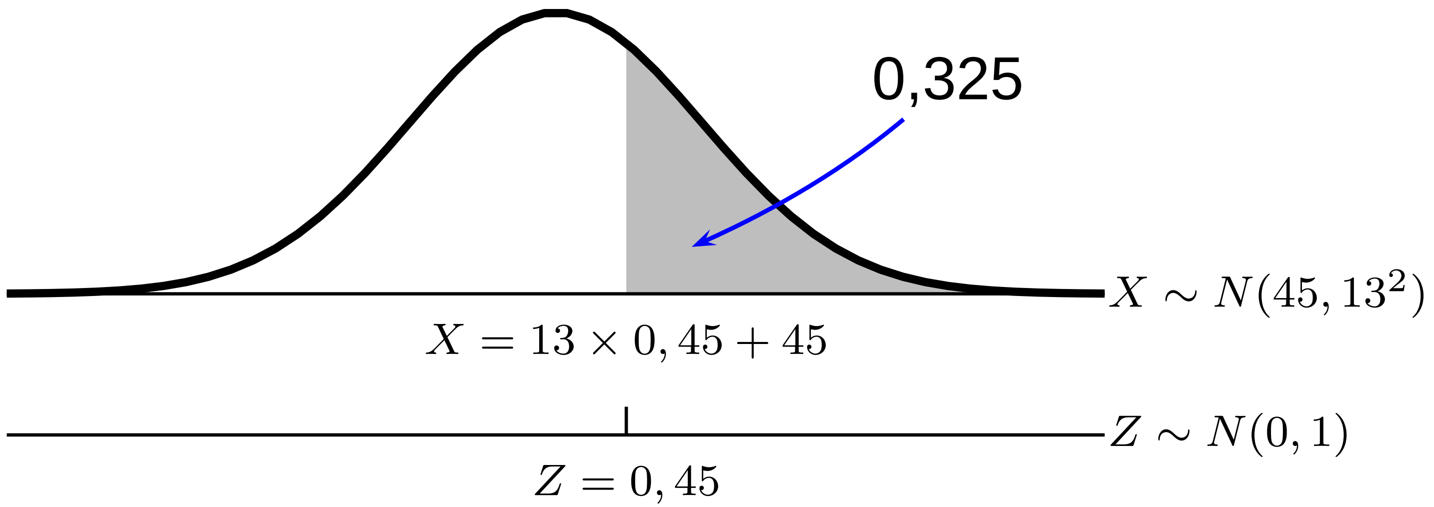
Distribusi Variabel Random Diskret dan Kontinu

Contoh 5 (Distribusi Normal)
Nilai-nilai ujian seleksi penerimaan mahasiswa baru secara nasional dianggap berdistribusi Normal dengan mean 45 dan deviasi standar 13. Jika hanya 32,5% calon mahasiswa yang akan diterima, berapakah nilai terendah calon mahasiswa yang diterima?



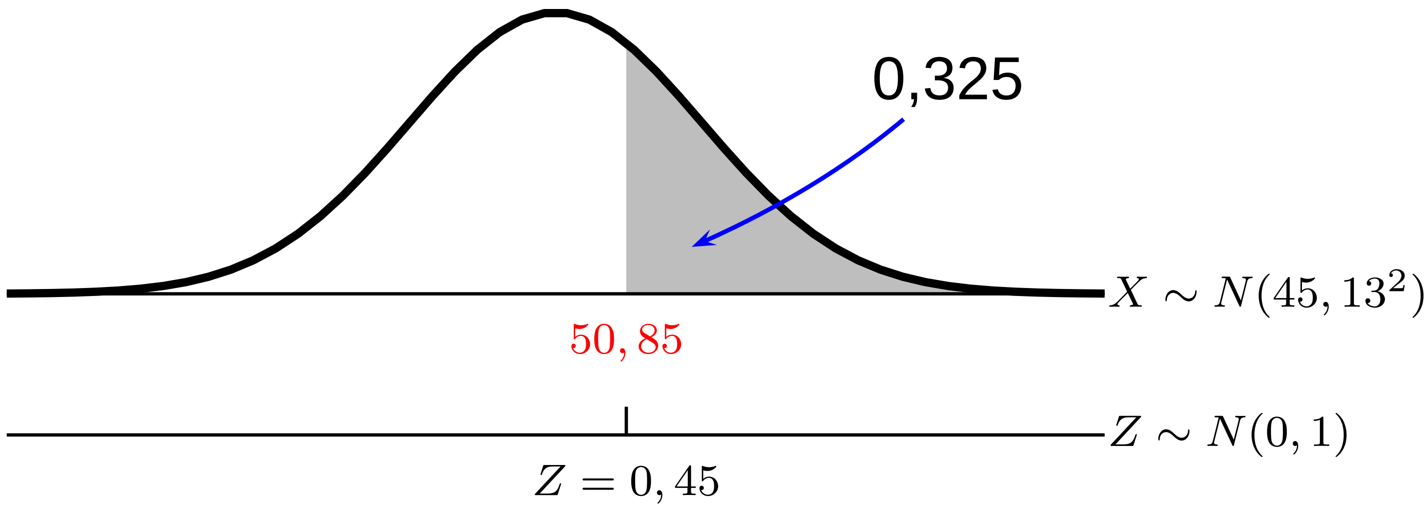
Distribusi Variabel Random Diskret dan Kontinu

Contoh 5 (Distribusi Normal)
Nilai-nilai ujian seleksi penerimaan mahasiswa baru secara nasional dianggap berdistribusi Normal dengan mean 45 dan deviasi standar 13. Jika hanya 32,5% calon mahasiswa yang akan diterima, berapakah nilai terendah calon mahasiswa yang diterima?



Distribusi Variabel Random Diskret dan Kontinu

Contoh 5 (Distribusi Normal)
Nilai-nilai ujian seleksi penerimaan mahasiswa baru secara nasional dianggap berdistribusi Normal dengan mean 45 dan deviasi standar 13. Jika hanya 32,5% calon mahasiswa yang akan diterima, berapakah nilai terendah calon mahasiswa yang diterima?



Distribusi Sampling Statistik

Populasi: himpunan keseluruhan obyek yang diamati.

Sampel: himpunan bagian dari populasi.

Sampel Random: sampel yang diperoleh dengan cara pengambilan sampel sedemikian sehingga setiap elemen populasi mempunyai kemungkinan yang sama untuk terambil.

Parameter: suatu harga (numerik) yang dihitung dari populasi, memberi deskripsi/karakteristik pada populasi.

Statistik: suatu harga (numerik) yang dihitung dari sampel.

Distribusi sampling statistik: distribusi peluang suatu statistik.

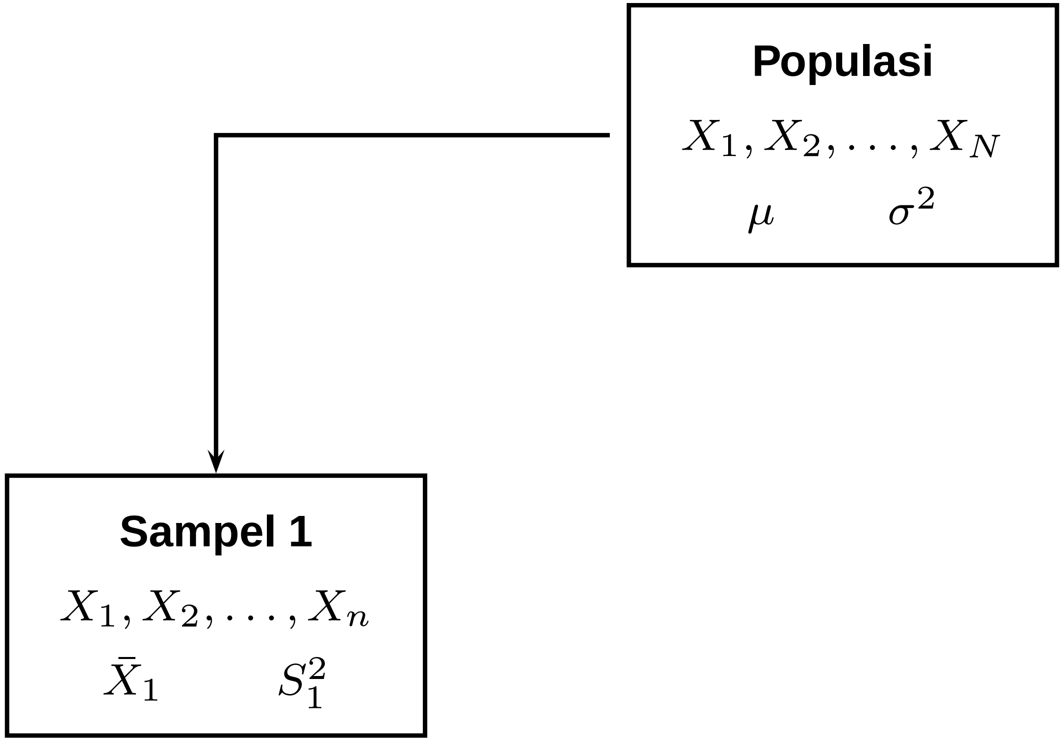
Distribusi Sampling Statistik

Populasi

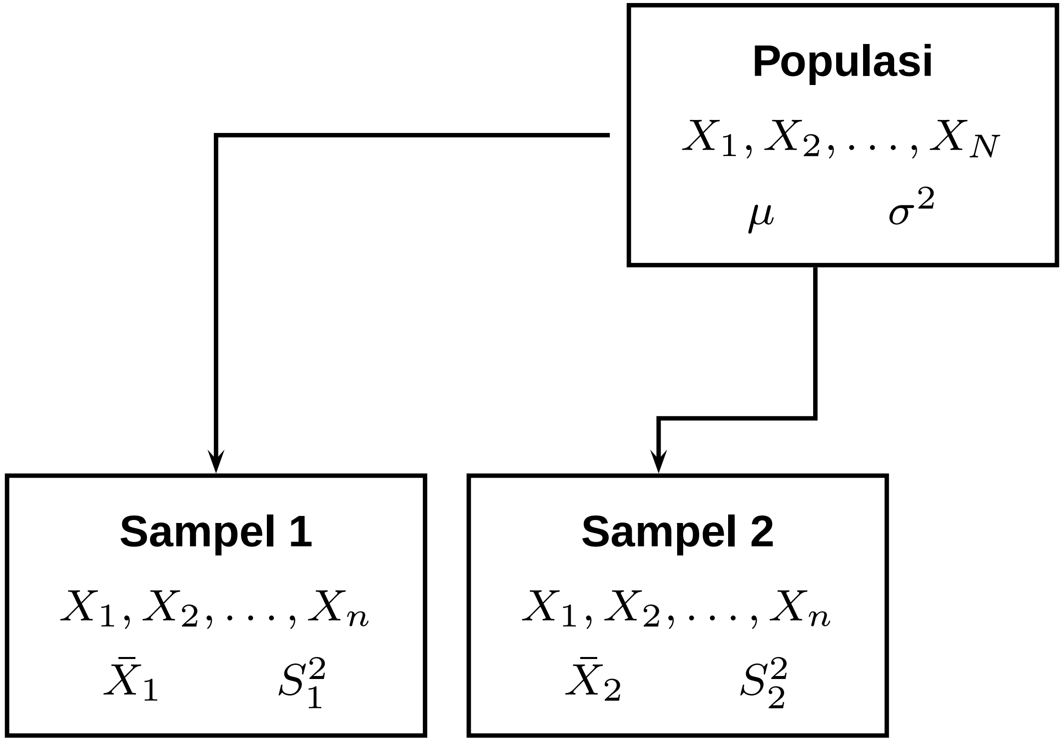
$$X_1, X_2, \ldots, X_N$$

$$\mu \qquad \sigma^2$$

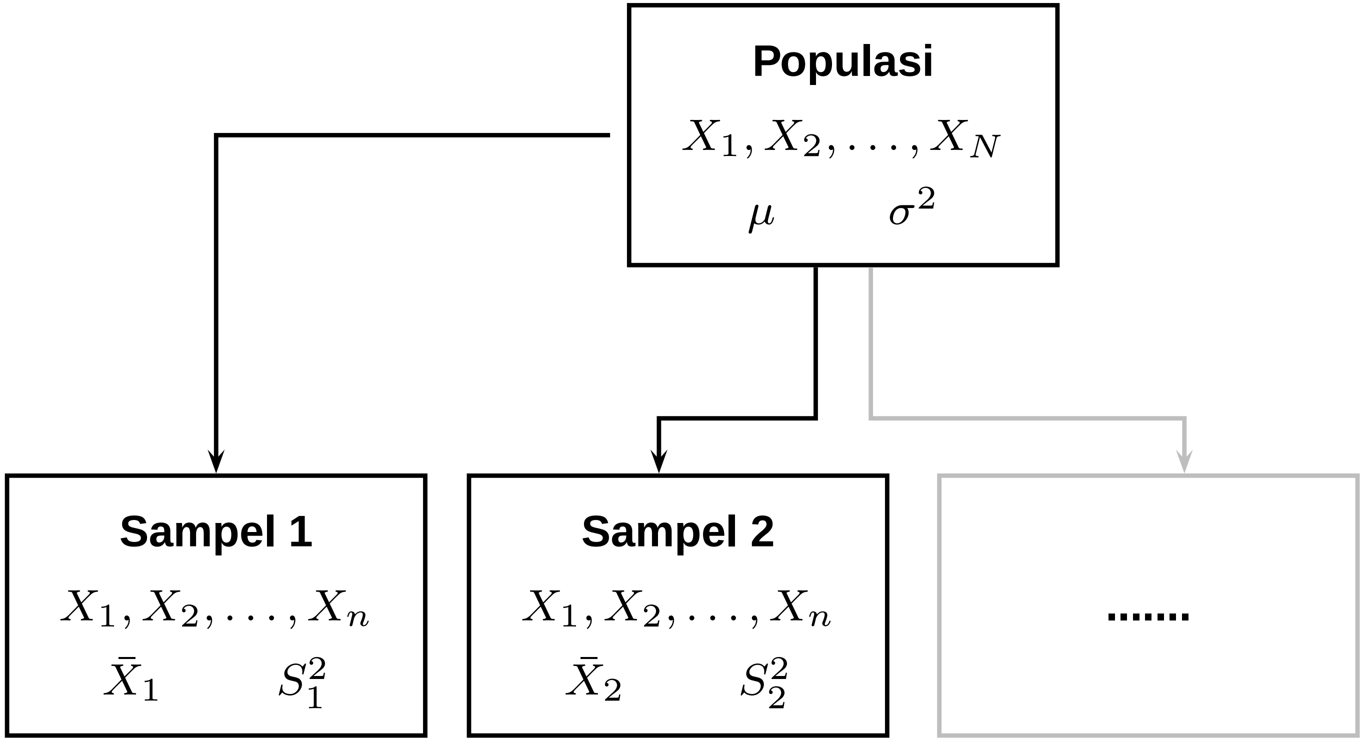
Distribusi Sampling Statistik



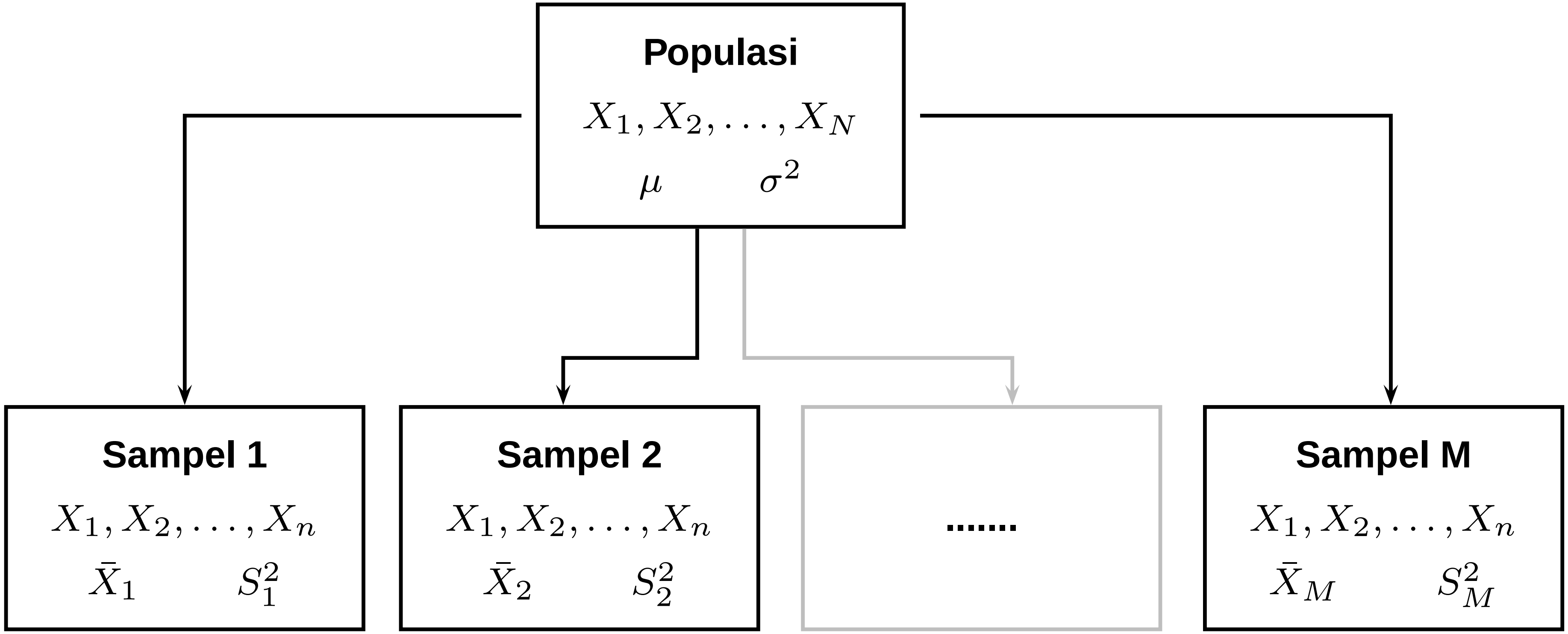
Distribusi Sampling Statistik



Distribusi Sampling Statistik



Distribusi Sampling Statistik



Distribusi Sampling Statistik

Contoh:

Populasi
$\{2, 4, 3\}$, $N = 3$

Sampling dengan pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

Distribusi peluang

x	$P(X = x)$
2	1/3
3	1/3
4	1/3

$$\mathrm{E}(X) = (2 + 3 + 4)\tfrac{1}{3} = 3$$

$$\mathrm{Var}(X) = (2^2 + 3^2 + 4^2)\tfrac{1}{3} - 3^2 = 2/3$$

Sampling dengan pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1

$\{2, 2\}$, $n = 2$

$\bar{X}_1 = 2$

Sampling dengan pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1

$\{2, 2\}$, $n = 2$

$\bar{X}_1 = 2$

Sampel 2

$\{2, 3\}$, $n = 2$

$\bar{X}_2 = 2,5$

Sampling dengan pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1

$\{2, 2\}$, $n = 2$

$\bar{X}_1 = 2$

Sampel 2

$\{2, 3\}$, $n = 2$

$\bar{X}_2 = 2,5$

Sampel 3

$\{2, 4\}$, $n = 2$

$\bar{X}_3 = 3$

Sampling dengan pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1	**Sampel 2**	**Sampel 3**	**Sampel 4**
$\{2, 2\}$, $n = 2$	$\{2, 3\}$, $n = 2$	$\{2, 4\}$, $n = 2$	$\{3, 2\}$, $n = 2$
$\bar{X}_1 = 2$	$\bar{X}_2 = 2,5$	$\bar{X}_3 = 3$	$\bar{X}_4 = 2,5$

Sampling dengan pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1	Sampel 2	Sampel 3	Sampel 4	Sampel 5
$\{2, 2\}$, $n = 2$	$\{2, 3\}$, $n = 2$	$\{2, 4\}$, $n = 2$	$\{3, 2\}$, $n = 2$	$\{3, 3\}$, $n = 2$
$\bar{X}_1 = 2$	$\bar{X}_2 = 2,5$	$\bar{X}_3 = 3$	$\bar{X}_4 = 2,5$	$\bar{X}_5 = 3$

Sampling dengan pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1
$\{2, 2\}$, $n = 2$
$\bar{X}_1 = 2$

Sampel 2
$\{2, 3\}$, $n = 2$
$\bar{X}_2 = 2,5$

Sampel 3
$\{2, 4\}$, $n = 2$
$\bar{X}_3 = 3$

Sampel 4
$\{3, 2\}$, $n = 2$
$\bar{X}_4 = 2,5$

Sampel 5
$\{3, 3\}$, $n = 2$
$\bar{X}_5 = 3$

Sampel 6
$\{3, 4\}$, $n = 2$
$\bar{X}_6 = 3,5$

Sampling dengan pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1
$\{2, 2\}$, $n = 2$
$\bar{X}_1 = 2$

Sampel 2
$\{2, 3\}$, $n = 2$
$\bar{X}_2 = 2,5$

Sampel 3
$\{2, 4\}$, $n = 2$
$\bar{X}_3 = 3$

Sampel 4
$\{3, 2\}$, $n = 2$
$\bar{X}_4 = 2,5$

Sampel 5
$\{3, 3\}$, $n = 2$
$\bar{X}_5 = 3$

Sampel 6
$\{3, 4\}$, $n = 2$
$\bar{X}_6 = 3,5$

Sampel 7
$\{4, 2\}$, $n = 2$
$\bar{X}_7 = 3$

Sampling dengan pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1
$\{2, 2\}$, $n = 2$
$\bar{X}_1 = 2$

Sampel 2
$\{2, 3\}$, $n = 2$
$\bar{X}_2 = 2,5$

Sampel 3
$\{2, 4\}$, $n = 2$
$\bar{X}_3 = 3$

Sampel 4
$\{3, 2\}$, $n = 2$
$\bar{X}_4 = 2,5$

Sampel 5
$\{3, 3\}$, $n = 2$
$\bar{X}_5 = 3$

Sampel 6
$\{3, 4\}$, $n = 2$
$\bar{X}_6 = 3,5$

Sampel 7
$\{4, 2\}$, $n = 2$
$\bar{X}_7 = 3$

Sampel 8
$\{4, 3\}$, $n = 2$
$\bar{X}_8 = 3,5$

Sampling dengan pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi
$\{2, 4, 3\}$, $N = 3$
$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1 $\{2, 2\}$, $n = 2$ $\bar{X}_1 = 2$	**Sampel 2** $\{2, 3\}$, $n = 2$ $\bar{X}_2 = 2,5$	**Sampel 3** $\{2, 4\}$, $n = 2$ $\bar{X}_3 = 3$	**Sampel 4** $\{3, 2\}$, $n = 2$ $\bar{X}_4 = 2,5$	**Sampel 5** $\{3, 3\}$, $n = 2$ $\bar{X}_5 = 3$
Sampel 6 $\{3, 4\}$, $n = 2$ $\bar{X}_6 = 3,5$	**Sampel 7** $\{4, 2\}$, $n = 2$ $\bar{X}_7 = 3$	**Sampel 8** $\{4, 3\}$, $n = 2$ $\bar{X}_8 = 3,5$	**Sampel 9** $\{4, 4\}$, $n = 2$ $\bar{X}_9 = 4$	

Sampling dengan pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2,4,3\}$, $N=3$

$\mu=3$, $\sigma^2=2/3$

Sampel 1 $\{2,2\}$, $n=2$ $\bar{X}_1=2$	**Sampel 2** $\{2,3\}$, $n=2$ $\bar{X}_2=2,5$	**Sampel 3** $\{2,4\}$, $n=2$ $\bar{X}_3=3$	**Sampel 4** $\{3,2\}$, $n=2$ $\bar{X}_4=2,5$	**Sampel 5** $\{3,3\}$, $n=2$ $\bar{X}_5=3$
Sampel 6 $\{3,4\}$, $n=2$ $\bar{X}_6=3,5$	**Sampel 7** $\{4,2\}$, $n=2$ $\bar{X}_7=3$	**Sampel 8** $\{4,3\}$, $n=2$ $\bar{X}_8=3,5$	**Sampel 9** $\{4,4\}$, $n=2$ $\bar{X}_9=4$	

Sampling dengan pengembalian $\Rightarrow M=N^n=3^2=9$

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

$\bar{x}$	$P(\bar{X} = \bar{x})$
2,0	1/9
2,5	2/9
3,0	3/9
3,5	2/9
4,0	1/9

Sampling dengan pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

$\bar{x}$	$P(\bar{X} = \bar{x})$
2,0	1/9
2,5	2/9
3,0	3/9
3,5	2/9
4,0	1/9

$$\mu_{\bar{X}} = \mathrm{E}(\bar{X}) = 2(\tfrac{1}{9}) + 2,5(\tfrac{2}{9}) + 3(\tfrac{3}{9}) + 3,5(\tfrac{2}{9}) + 4(\tfrac{1}{9}) = 3$$

$$\sigma^2_{\bar{X}} = \mathrm{Var}(\bar{X}) = 2^2(\tfrac{1}{9}) + 2,5^2(\tfrac{2}{9}) + 3^2(\tfrac{3}{9}) + 3,5^2(\tfrac{2}{9}) + 4^2(\tfrac{1}{9}) - 3^2 = 1/3$$

Sampling dengan pengembalian

Distribusi Sampling Statistik

Sampling dengan pengembalian

Untuk sampel berukuran n dari populasi berukuran N dengan mean μ dan variansi σ^2, mean dan variansi dari statistik $\bar{X}$:

$$\mu_{\bar{X}} = \mathrm{E}(\bar{X}) = \mu$$
$$\sigma^2_{\bar{X}} = \mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}$$

Distribusi Sampling Statistik

Contoh:

Populasi
$\{2, 4, 3\}$, $N = 3$
$\mu = 3$, $\sigma^2 = 2/3$

Sampling tanpa pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1

$\{2, 3\}$, $n = 2$

$\bar{X}_2 = 2,5$

Sampling tanpa pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1

$\{2, 3\}$, $n = 2$

$\bar{X}_2 = 2,5$

Sampel 2

$\{2, 4\}$, $n = 2$

$\bar{X}_2 = 3$

Sampling tanpa pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1

$\{2, 3\}$, $n = 2$

$\bar{X}_2 = 2,5$

Sampel 2

$\{2, 4\}$, $n = 2$

$\bar{X}_2 = 3$

Sampel 3

$\{3, 4\}$, $n = 2$

$\bar{X}_3 = 3,5$

Sampling tanpa pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

Sampel 1

$\{2, 3\}$, $n = 2$

$\bar{X}_2 = 2, 5$

Sampel 2

$\{2, 4\}$, $n = 2$

$\bar{X}_2 = 3$

Sampel 3

$\{3, 4\}$, $n = 2$

$\bar{X}_3 = 3, 5$

Sampling tanpa pengembalian $\Rightarrow M = \binom{N}{n} = \binom{3}{2} = 3$

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

$\bar{x}$	$P(\bar{X} = \bar{x})$
2,5	1/3
3,0	1/3
3,5	1/3

Sampling tanpa pengembalian

Distribusi Sampling Statistik

Contoh:

Populasi

$\{2, 4, 3\}$, $N = 3$

$\mu = 3$, $\sigma^2 = 2/3$

$\bar{x}$	$P(\bar{X} = \bar{x})$
2,5	1/3
3,0	1/3
3,5	1/3

$$\mu_{\bar{X}} = \mathrm{E}(\bar{X}) =) + 2,5(\tfrac{1}{3}) + 3(\tfrac{1}{3}) + 3,5(\tfrac{1}{3}) = 3$$

$$\mu_{\bar{X}} = \mathrm{Var}(\bar{X}) =) + 2,5^2(\tfrac{1}{3}) + 3^2(\tfrac{1}{3}) + 3,5^2(\tfrac{1}{3}) - 3^2 = 1/6$$

Sampling tanpa pengembalian

Distribusi Sampling Statistik

Sampling tanpa pengembalian

Untuk sampel berukuran n dari populasi berukuran N dengan mean μ dan variansi σ^2, mean dan variansi dari statistik $\bar{X}$:

$$\mu_{\bar{X}} = \mathrm{E}(\bar{X}) = \mu$$

$$\sigma^2_{\bar{X}} = \mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}\frac{N-n}{N-1}$$

Distribusi Sampling Statistik

Sifat-sifat Distribusi Sampling untuk Mean

Sifat 1: Apabila sampel-sampel random dengan n elemen masing-masing diambil dari suatu populasi yang mempunyai mean μ dan variansi σ^2 , maka distribusi sampling mean akan mempunyai mean $\mu_{\bar{X}} = \mu$ dan variansi $\sigma^2_{\bar{X}} = \sigma^2/n$.

Sifat 2: Apabila populasi (dalam sifat 1) berdistribusi Normal, maka distribusi sampling untuk mean juga berdistribusi Normal.

Distribusi Sampling Statistik

Sifat-sifat Distribusi Sampling untuk Mean

Sifat 3 (Teorema Limit Pusat): Apabila sampel-sampel random diambil dari suatu populasi yang berdistribusi sembarang, yang mempunyai mean μ dan variansi σ^2, maka untuk n besar, distribusi sampling untuk mean dapat dianggap mendekati Normal dengan $\mu_{\bar{X}} = \mu$ dan variansi $\sigma^2_{\bar{X}} = \sigma^2/n$, sehingga

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

mendekati Normal Standar.

Distribusi Sampling Statistik

Contoh 1:

Seorang peneliti di bidang pertanian akan meneliti hasil dari suatu varietas padi di Indonesia. Akan diteliti 5 tanah pertanian tersebar di seluruh Indonesia yang dapat ditanami padi tersebut.

Distribusi Sampling Statistik

Contoh 1:

Seorang peneliti di bidang pertanian akan meneliti hasil dari suatu varietas padi di Indonesia. Akan diteliti 5 tanah pertanian tersebar di seluruh Indonesia yang dapat ditanami padi tersebut.

Populasi untuk masalah ini adalah hasil padi jenis tersebut yang diperoleh dari seluruh tanah pertanian di Indonesia.

Distribusi Sampling Statistik

Contoh 1:

Seorang peneliti di bidang pertanian akan meneliti hasil dari suatu varietas padi di Indonesia. Akan diteliti 5 tanah pertanian tersebar di seluruh Indonesia yang dapat ditanami padi tersebut.

Sampel untuk masalah ini adalah hasil padi yang diperoleh dari 5 tanah pertanian yang terpilih. Sampel ini akan merupakan sampel random jika, setiap tanah pertanian di Indonesia mempunyai peluang yang sama untuk terpilih ; dan pemilihan satu tanah pertanian tidak mempengaruhi atau dipengaruhi pemilihan tanah yang lain.

Distribusi Sampling Statistik

Contoh 1:
Seorang peneliti di bidang pertanian akan meneliti hasil dari suatu varietas padi di Indonesia. Akan diteliti 5 tanah pertanian tersebar di seluruh Indonesia yang dapat ditanami padi tersebut.

Sampel untuk masalah ini adalah hasil padi yang diperoleh dari 5 tanah pertanian yang terpilih. Sampel ini akan merupakan sampel random jika, setiap tanah pertanian di Indonesia mempunyai peluang yang sama untuk terpilih ; dan pemilihan satu tanah pertanian tidak mempengaruhi atau dipengaruhi pemilihan tanah yang lain.

Hal ini dapat dilakukan dengan mendaftar terlebih dahulu semua tanah pertanian di Indonesia dan diberi nomor identitas, kemudian dipilih 5 tanah pertanian secara random berdasarkan nomor identitas (misalnya dengan tabel bilangan random).

Distribusi Sampling Statistik

Contoh 2:

Suatu sampel random berukuran 40 diambil dari suatu populasi dengan mean 41,4 dan variansi 84,64. Hitung peluang bahwa mean sampel itu terletak antara 40 dan 45. Anggap ukuran populasinya sangat besar relatif terhadap ukuran sampel.

Distribusi Sampling Statistik

Contoh 2:

Suatu sampel random berukuran 40 diambil dari suatu populasi dengan mean 41,4 dan variansi 84,64. Hitung peluang bahwa mean sampel itu terletak antara 40 dan 45. Anggap ukuran populasinya sangat besar relatif terhadap ukuran sampel.

Berdasarkan Sifat 1, distribusi sampling untuk mean ($\bar{X}$) mempunyai mean ($\mathrm{E}(\bar{X})$, harga harapan): $\mu_{\bar{X}} = \mu = 41,4$ dan variansi ($\mathrm{Var}(\bar{X})$): $\sigma_{\bar{X}} = \sigma^2/n = 84,64/40 = 2,116$.

Distribusi Sampling Statistik

Contoh 2:

Suatu sampel random berukuran 40 diambil dari suatu populasi dengan mean 41,4 dan variansi 84,64. Hitung peluang bahwa mean sampel itu terletak antara 40 dan 45. Anggap ukuran populasinya sangat besar relatif terhadap ukuran sampel.

Berdasarkan Sifat 1, distribusi sampling untuk mean ($\bar{X}$) mempunyai mean ($\mathrm{E}(\bar{X})$, harga harapan): $\mu_{\bar{X}} = \mu = 41,4$ dan variansi ($\mathrm{Var}(\bar{X})$): $\sigma_{\bar{X}} = \sigma^2/n = 84,64/40 = 2,116$.

Ukuran sampel $n = 40$ cukup besar untuk berlakunya Sifat 3,

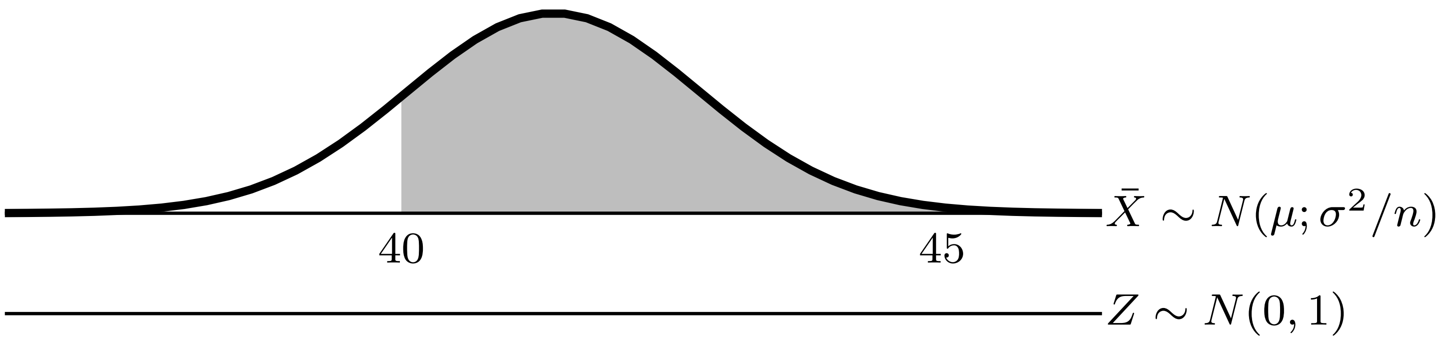
Distribusi Sampling Statistik

Contoh 2:

Suatu sampel random berukuran 40 diambil dari suatu populasi dengan mean 41,4 dan variansi 84,64. Hitung peluang bahwa mean sampel itu terletak antara 40 dan 45. Anggap ukuran populasinya sangat besar relatif terhadap ukuran sampel.

Berdasarkan Sifat 1, distribusi sampling untuk mean ($\bar{X}$) mempunyai mean ($\mathrm{E}(\bar{X})$, harga harapan): $\mu_{\bar{X}} = \mu = 41,4$ dan variansi ($\mathrm{Var}(\bar{X})$): $\sigma_{\bar{X}} = \sigma^2/n = 84,64/40 = 2,116$.

Ukuran sampel $n = 40$ cukup besar untuk berlakunya Sifat 3,



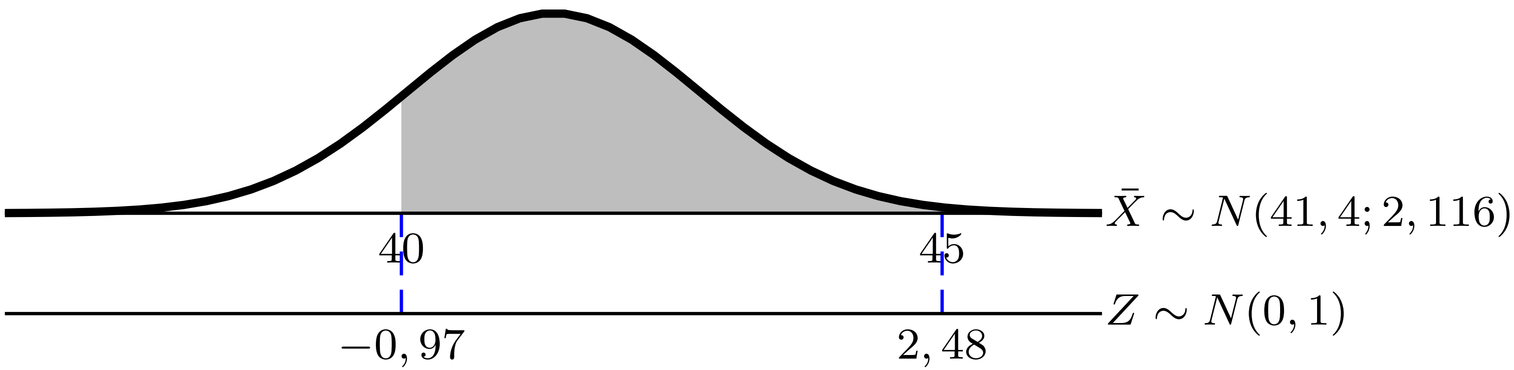
Distribusi Sampling Statistik

Contoh 2:

Suatu sampel random berukuran 40 diambil dari suatu populasi dengan mean 41,4 dan variansi 84,64. Hitung peluang bahwa mean sampel itu terletak antara 40 dan 45. Anggap ukuran populasinya sangat besar relatif terhadap ukuran sampel.

Berdasarkan Sifat 1, distribusi sampling untuk mean ($\bar{X}$) mempunyai mean ($\mathrm{E}(\bar{X})$, harga harapan): $\mu_{\bar{X}} = \mu = 41,4$ dan variansi ($\mathrm{Var}(\bar{X})$): $\sigma_{\bar{X}} = \sigma^2/n = 84,64/40 = 2,116$.

Ukuran sampel $n = 40$ cukup besar untuk berlakunya Sifat 3,



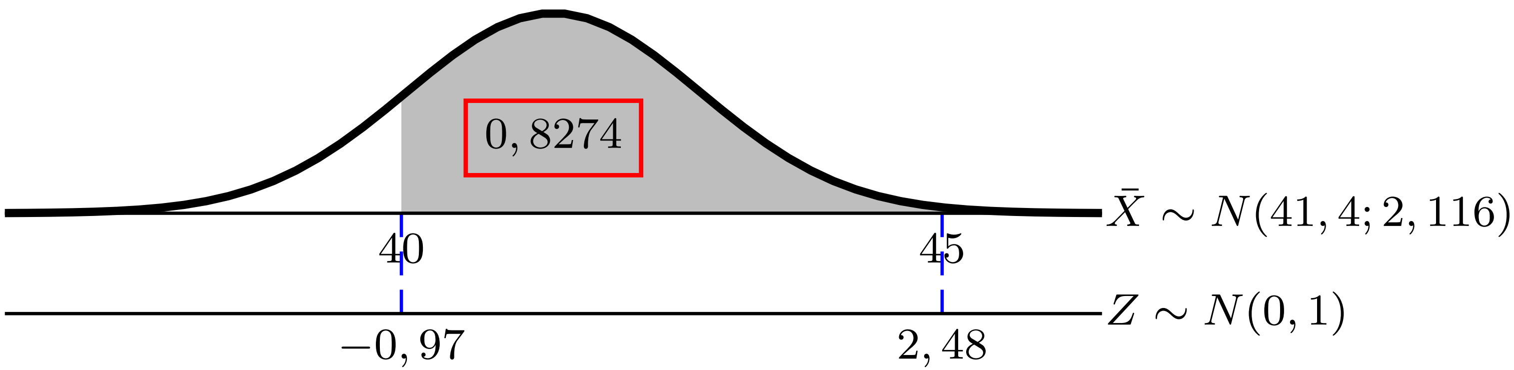
Distribusi Sampling Statistik

Contoh 2:

Suatu sampel random berukuran 40 diambil dari suatu populasi dengan mean 41,4 dan variansi 84,64. Hitung peluang bahwa mean sampel itu terletak antara 40 dan 45. Anggap ukuran populasinya sangat besar relatif terhadap ukuran sampel.

Berdasarkan Sifat 1, distribusi sampling untuk mean ($\bar{X}$) mempunyai mean ($\mathrm{E}(\bar{X})$, harga harapan): $\mu_{\bar{X}} = \mu = 41,4$ dan variansi ($\mathrm{Var}(\bar{X})$): $\sigma_{\bar{X}} = \sigma^2/n = 84,64/40 = 2,116$.

Ukuran sampel $n = 40$ cukup besar untuk berlakunya Sifat 3,



Distribusi Sampling Statistik

Contoh 3:

Diketahui suatu populasi dengan mean 82 dan deviasi standar 12.

a. Jika suatu sampel random berukuran 64 diambil, berapa peluang bahwa mean sampel akan terletak antara 80,8 dan 83,2 ?

Distribusi Sampling Statistik

Contoh 3:

Diketahui suatu populasi dengan mean 82 dan deviasi standar 12.

a. Jika suatu sampel random berukuran 64 diambil, berapa peluang bahwa mean sampel akan terletak antara 80,8 dan 83,2 ?
Karena $n = 64$ cukup besar, dapat digunakan Teorema limit pusat (sifat 3). Distribusi $\bar{X}$ akan mendekati normal dengan mean $\mu_{\bar{X}} = 82$ dan deviasi standar $\sigma_{\bar{X}} = 12/\sqrt{64} = 1,5$
$P(80,8 \leq \bar{X} \leq 83,2)$ dapat dihitung melalui $Z = \frac{\bar{X}-82}{1,5}$

$$\begin{aligned} P(80,8 \leq \bar{X} \leq 83,2) &= P(\frac{80,8-82}{1,5} \leq Z \leq \frac{83,2-82}{1,5}) \\ &= P(-0,8 \leq Z \leq 0,8) \\ &= 0,5762 \end{aligned}$$

Distribusi Sampling Statistik

Contoh 3:

Diketahui suatu populasi dengan mean 82 dan deviasi standar 12.

b. Berapa probabilitasnya jika ukuran sampel random 100 ?

Distribusi Sampling Statistik

Contoh 3:

Diketahui suatu populasi dengan mean 82 dan deviasi standar 12.

b. Berapa probabilitasnya jika ukuran sampel random 100 ? Untuk $n = 100$, $\sigma_{\bar{X}} = 12/\sqrt{100} = 1,2$

$$
\begin{aligned}
P(80,8 \leq \bar{X} \leq 83,2) &= P(\frac{80,8-82}{1,2} \leq Z \leq \frac{83,2-82}{1,2}) \\
&= P(-1,0 \leq Z \leq 1,0) \\
&= 0,6826
\end{aligned}
$$

Pendekatan Normal untuk Binomial

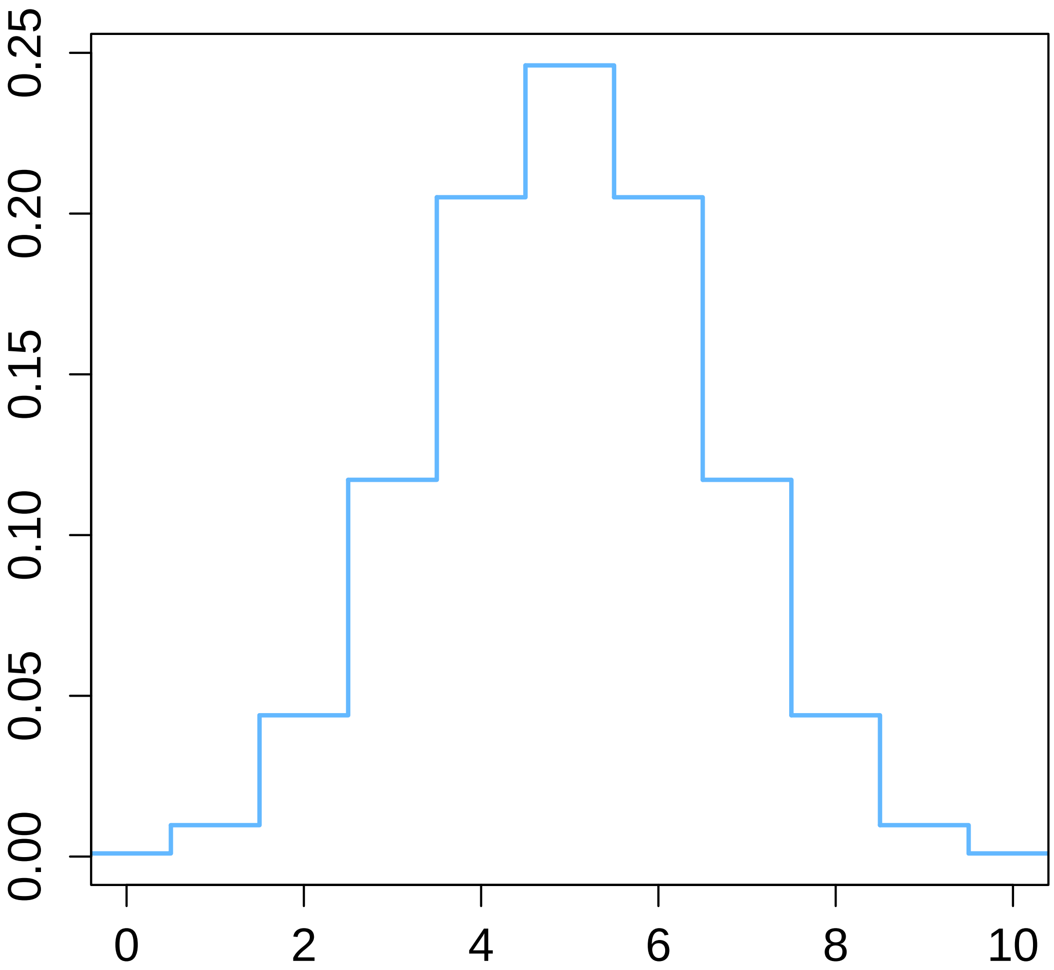
Teorema
Bila X adalah variabel random binomial dengan mean $\mu = np$ dan variansi $\sigma^2 = npq$, maka untuk n besar

$$Z = \frac{X - np}{\sqrt{npq}}$$

merupakan variabel random normal standar.

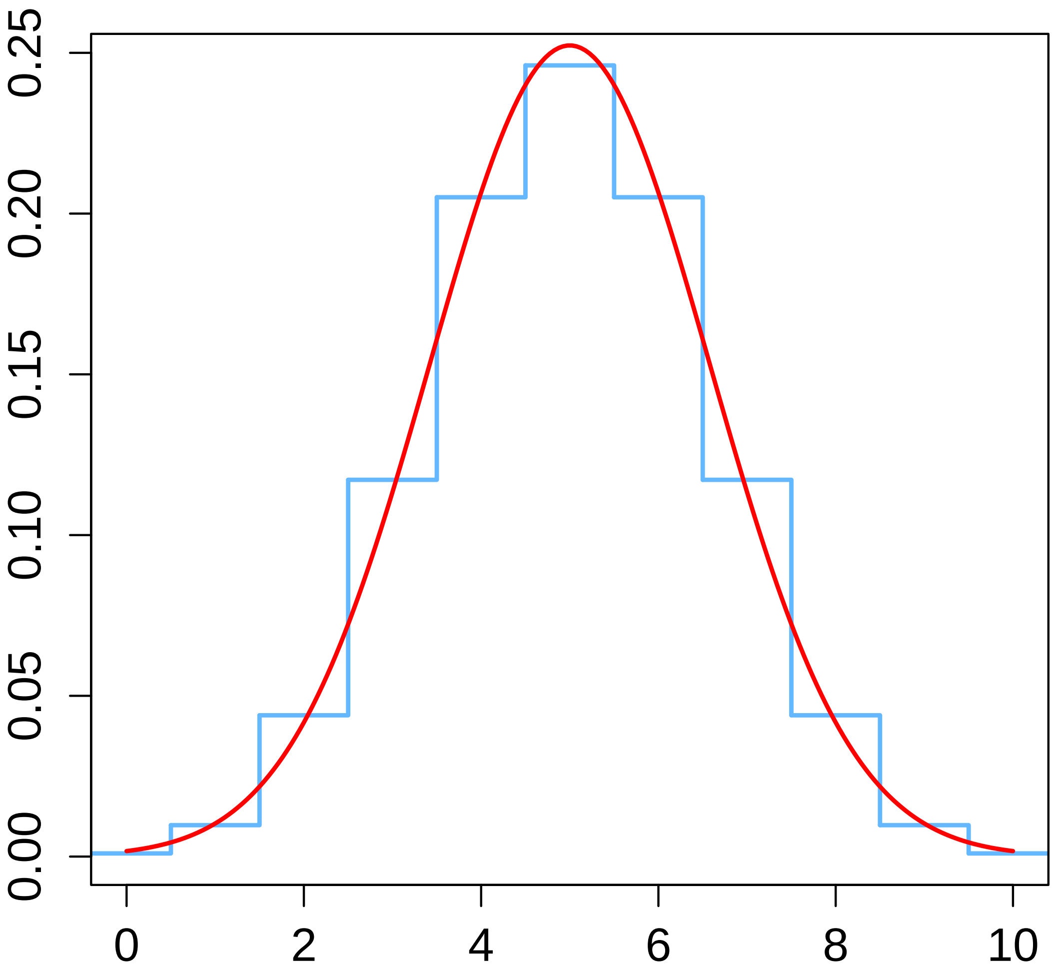
Pendekatan Normal untuk Binomial

Binomial$(n = 10, p = 0,5) \rightarrow$ Normal



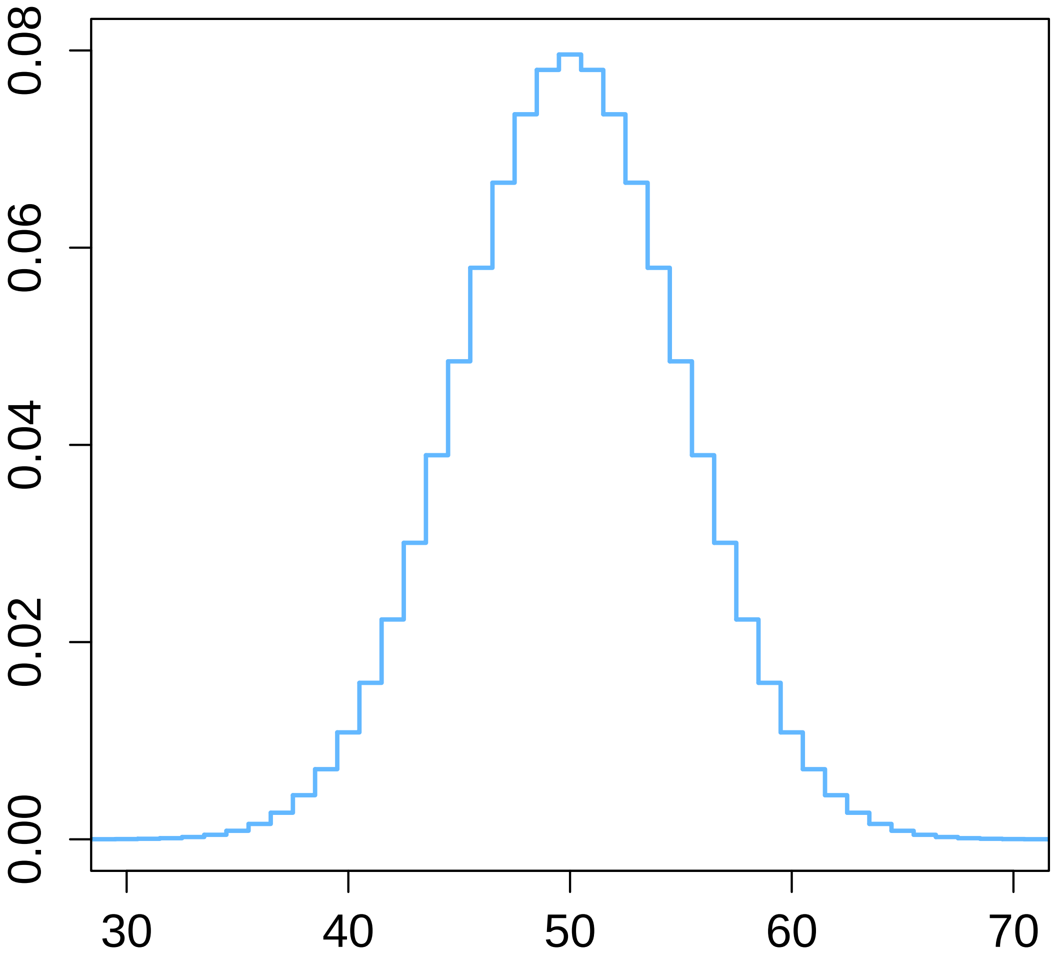
Pendekatan Normal untuk Binomial

Binomial$(n = 10, p = 0,5) \rightarrow$ Normal



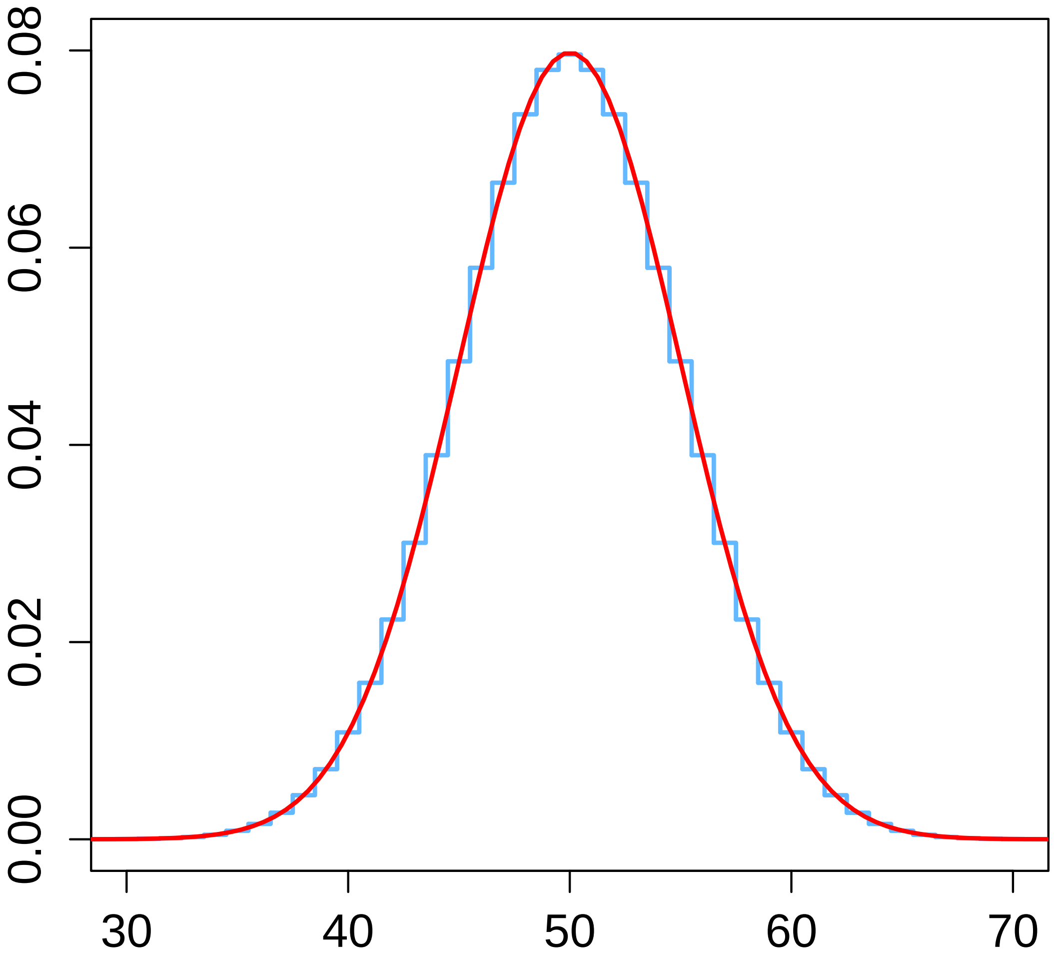
Pendekatan Normal untuk Binomial

Binomial$(n = 100, p = 0,5) \rightarrow$ Normal



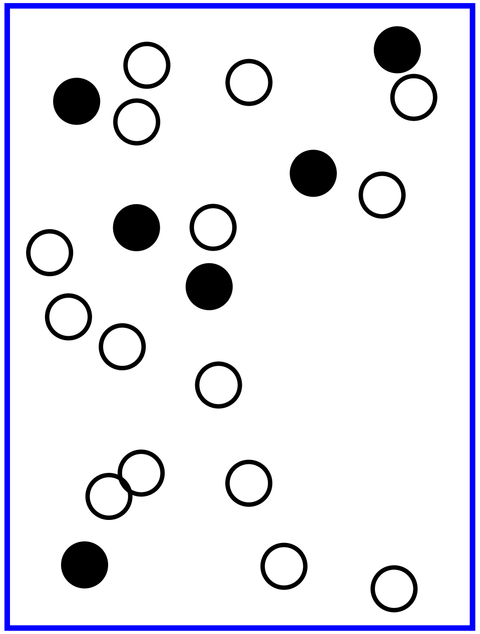
Pendekatan Normal untuk Binomial

Binomial$(n = 100, p = 0,5) \rightarrow$ Normal



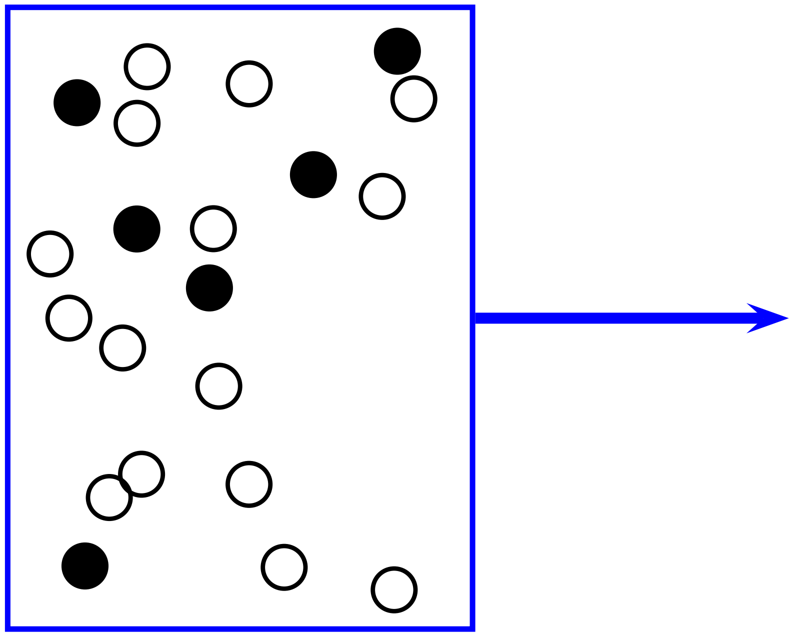
Inferensi Statistik

Permasalahan dalam peluang



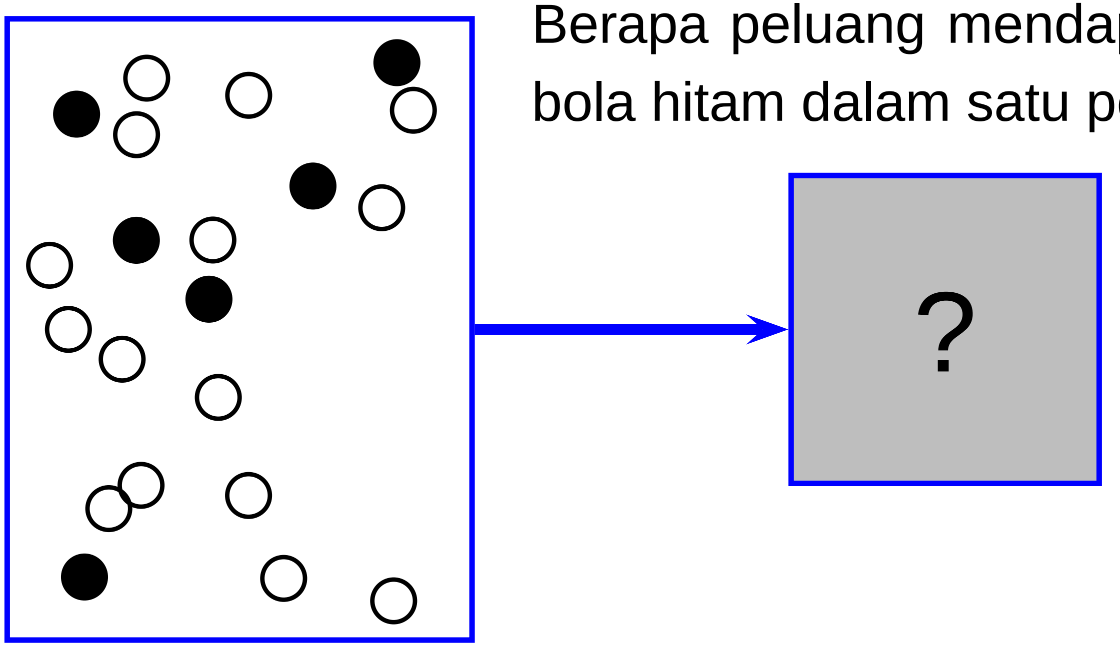
Inferensi Statistik

Permasalahan dalam peluang



Inferensi Statistik

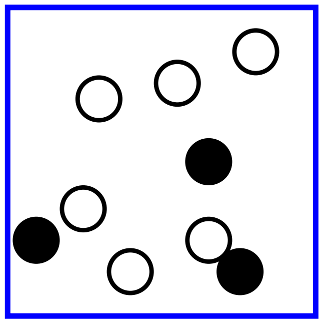
Permasalahan dalam peluang



Berapa peluang mendapatkan satu bola hitam dalam satu pengambilan

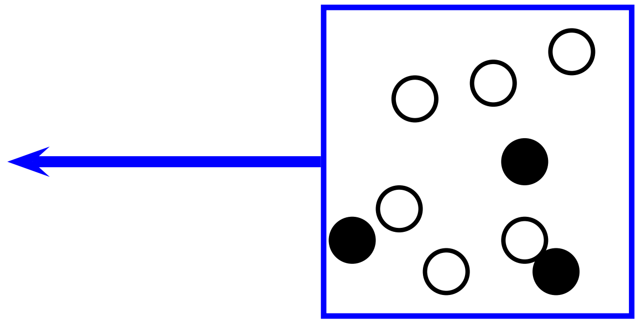
Inferensi Statistik

Permasalahan dalam inferensi



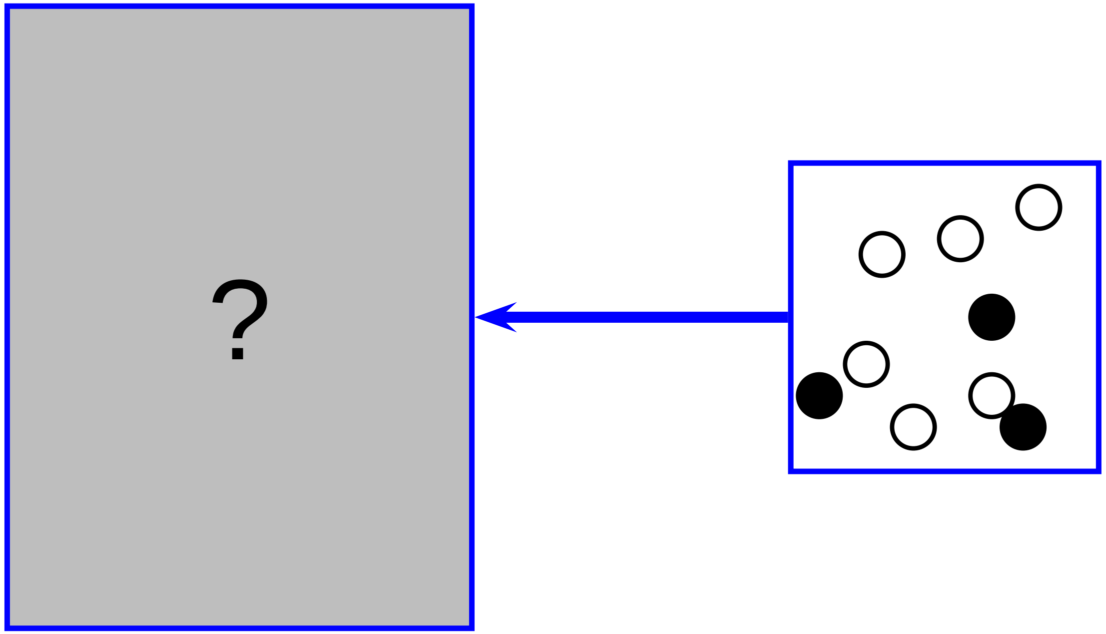
Inferensi Statistik

Permasalahan dalam inferensi



Inferensi Statistik

Permasalahan dalam inferensi



Bagaimana karakteristik populasi berdasarkan sampel

Inferensi Statistik

Inferensi statistik: pengambilan kesimpulan tentang parameter populasi berdasarkan analisis pada sampel

Konsep-konsep inferensi statistik: *estimasi titik*, *estimasi interval* dan *uji hipotesis*

Estimasi parameter: Menduga nilai parameter populasi berdasarkan data/statistik.

Estimasi titik: Menduga nilai tunggal parameter populasi. Misalnya parameter μ diduga dengan statistik $\bar{X}$

Estimasi interval: Menduga nilai parameter populasi dalam bentuk interval. Misalnya diduga dengan suatu interval $A \leq \mu \leq B$

Inferensi Statistik

Contoh: estimator titik untuk mean μ

- rata-rata

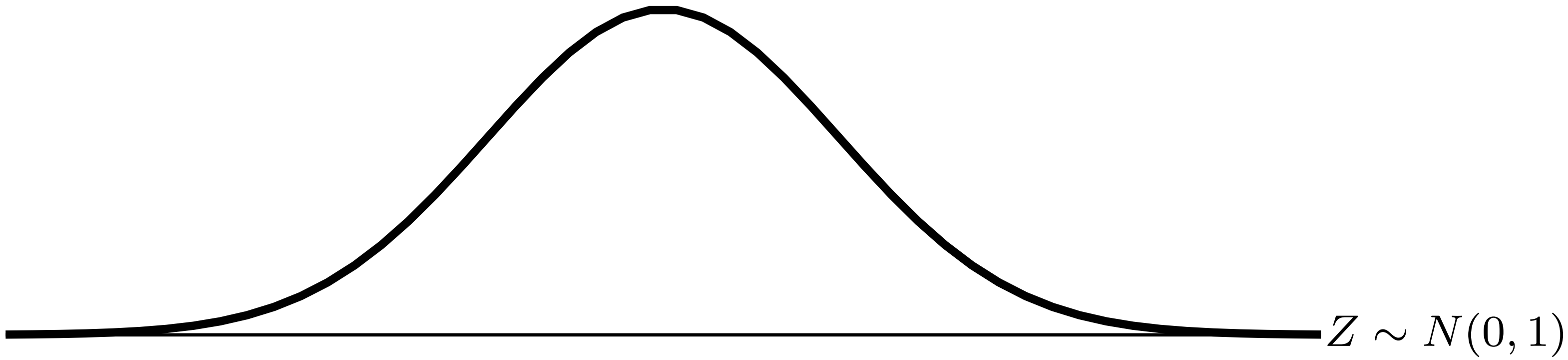
$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

- Median
- rata-rata dua harga ekstrim

$$\frac{X_{\text{min}} + X_{\text{maks}}}{2}$$

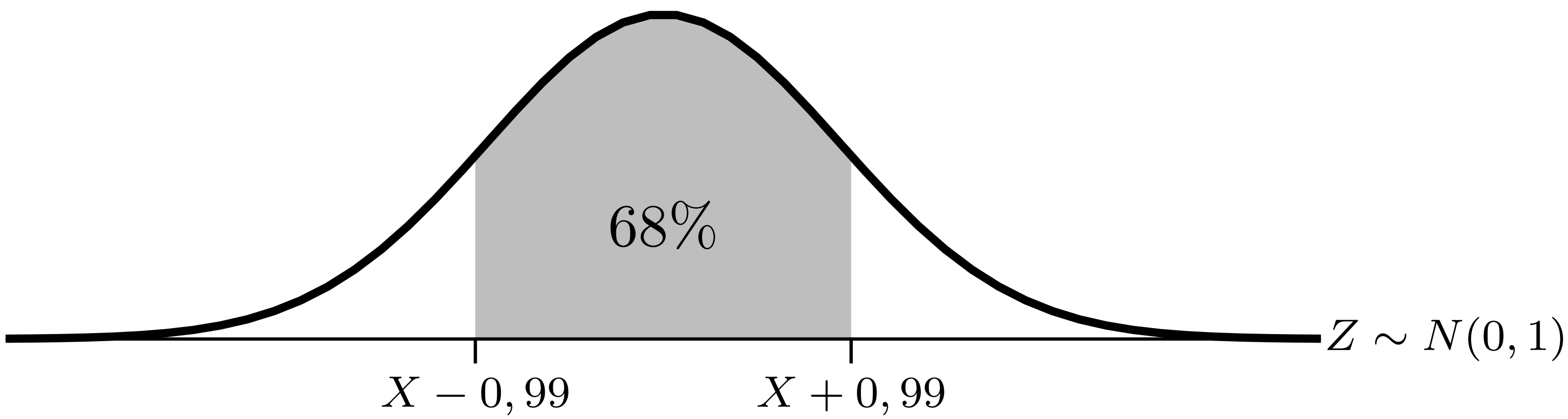
Inferensi Statistik

Contoh: Estimasi Interval
Diketahui variabel random Normal X dengan mean $\mathrm{E}(X) = \mu$ dan $\mathrm{Var}(X) = 1$. Maka $(X - \mu)$ akan berdistribusi Normal standar.



Inferensi Statistik

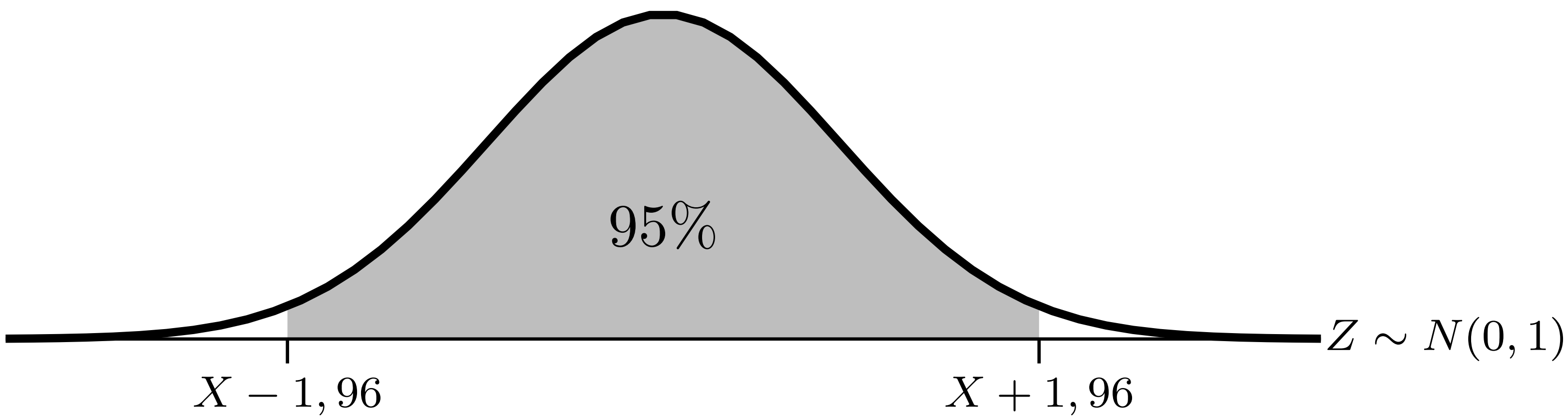
Contoh: Estimasi Interval
Diketahui variabel random Normal X dengan mean $\mathrm{E}(X) = \mu$ dan $\mathrm{Var}(X) = 1$. Maka $(X - \mu)$ akan berdistribusi Normal standar.



Interval Konfidensi (estimasi interval) 68%

Inferensi Statistik

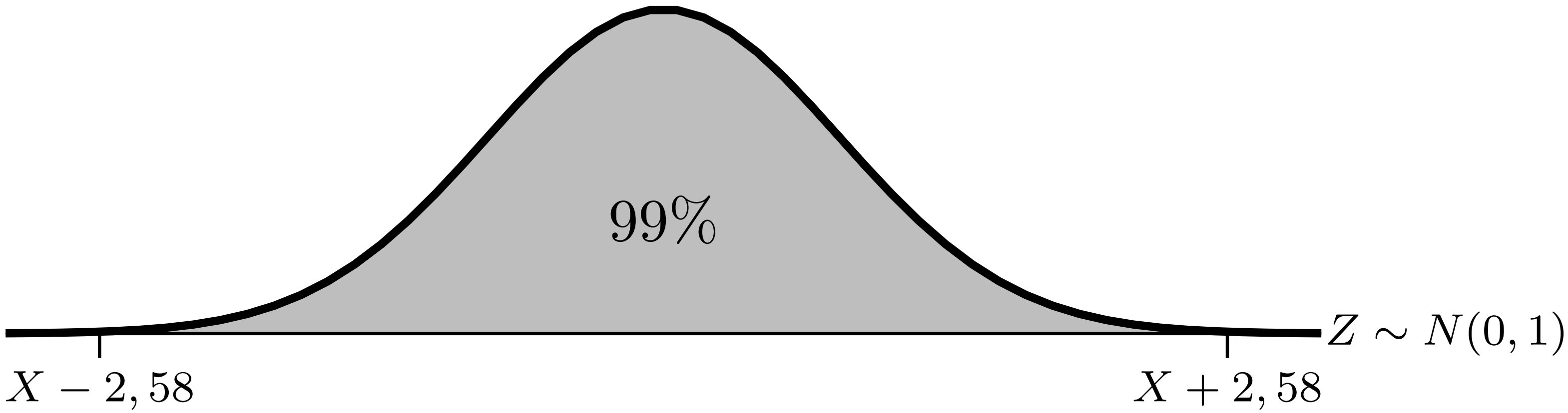
Contoh: Estimasi Interval
Diketahui variabel random Normal X dengan mean $\mathrm{E}(X) = \mu$ dan $\mathrm{Var}(X) = 1$. Maka $(X - \mu)$ akan berdistribusi Normal standar.



Interval Konfidensi (estimasi interval) 95%

Inferensi Statistik

Contoh: Estimasi Interval
Diketahui variabel random Normal X dengan mean $\mathrm{E}(X) = \mu$ dan $\mathrm{Var}(X) = 1$. Maka $(X - \mu)$ akan berdistribusi Normal standar.



Interval Konfidensi (estimasi interval) 99%

Inferensi Statistik

Uji hipotesis: suatu proses untuk menentukan apakah dugaan tentang nilai parameter/karakteristik populasi didukung kuat oleh data sampel atau tidak

Hipotesis penelitian: hipotesis tentang pernyataan dari hasil penelitian yang akan dilakukan

Hipotesis Statistik: suatu pernyataan tentang parameter populasi

Inferensi Statistik

Hipotesis nol (H_0). Hipotesis yang akan diuji oleh suatu prosedur statistik, biasanya berupa suatu pernyataan tidak adanya perbedaan atau tidak adanya hubungan. Pernyataan nol dapat diartikan bahwa pernyataan tetang parameter tidak didukung secara kuat oleh data.

Hipotesis alternatif (H_1). Hipotesis yang merupakan lawan dari H_0, biasanya berupa pernyataan tentang adanya perbedaan atau adanya hubungan. H_1 digunakan untuk menunjukkan bahwa pernyataan mendapat dukungan kuat dari data.

Logika Uji Hipotesis. Tidak dapat dibuktikan bahwa suatu hipotesis itu benar, tapi dapat dibuktikan bahwa suatu hipotesis itu salah.

Inferensi Statistik

Tipe Kesalahan dalam Uji Hipotesis

Keputusan Uji	Kenyataan	
	H_0 benar	H_0 salah
H_0 tidak ditolak	benar	salah (Tipe II)
H_0 ditolak	salah (Tipe I)	benar

Inferensi Statistik

Tipe Kesalahan dalam Uji Hipotesis

Keputusan Uji	Kenyataan	
	H_0 benar	H_0 salah
H_0 tidak ditolak	benar	salah (Tipe II)
H_0 ditolak	salah (Tipe I)	benar

Peluang melakukan kesalahan tipe I

$P(\text{menolak } \mathrm{H}_0 \ \text{ yang benar}) = \alpha$

Inferensi Statistik

Tipe Kesalahan dalam Uji Hipotesis

Keputusan Uji	Kenyataan	
	H_0 benar	H_0 salah
H_0 tidak ditolak	benar	salah (Tipe II)
H_0 ditolak	salah (Tipe I)	benar

Peluang melakukan kesalahan tipe I
$P(\text{menolak } \mathrm{H}_0 \ \text{ yang benar}) = \alpha$

Peluang melakukan kesalahan tipe II
$P(\text{tidak menolak } \mathrm{H}_0 \ \text{ yang salah}) = \beta$

Inferensi Statistik

Contoh (Hipotesis statistik dan statistik penguji)

Ingin diuji secara statistik pernyataan : *suatu obat baru lebih baik dari obat yang selama ini digunakan.*

Misalkan p adalah proporsi (prosentase) orang yang sembuh setelah minum obat tersebut, dan obat dikatakan baik jika proporsi orang yang sembuh lebih dari 60 %.

Pernyataan H_0 dan H_1 adalah sebagai berikut :

H_0 : $p \leq 0,6$ (obat baru tidak lebih baik)

H_1 : $p > 0,6$ (obat baru lebih baik)

Inferensi Statistik

Contoh (Hipotesis statistik dan statistik penguji)

Ingin diuji secara statistik pernyataan : *suatu obat baru lebih baik dari obat yang selama ini digunakan.*

H_0 : $p \leq 0,6$ (obat baru tidak lebih baik)

H_1 : $p > 0,6$ (obat baru lebih baik)

Dilakukan eksperimen terhadap 20 pasien.
X : banyak pasien yang sembuh
$X \sim \mathrm{Binomial}(n = 20, p = 0,6)$

Inferensi Statistik

Contoh (Hipotesis statistik dan statistik penguji)

Ingin diuji secara statistik pernyataan : *suatu obat baru lebih baik dari obat yang selama ini digunakan.*

H_0 : $p \leq 0,6$ (obat baru tidak lebih baik)

H_1 : $p > 0,6$ (obat baru lebih baik)

Dilakukan eksperimen terhadap 20 pasien.
X : banyak pasien yang sembuh
$X \sim \mathrm{Binomial}(n = 20, p = 0,6)$

X besar (banyak yang sembuh) $\Rightarrow$ menolak H_0,
X kecil (banyak yang tidak sembuh) $\Rightarrow$ mendukung H_0

Inferensi Statistik

Daerah penolakan (Daerah kritik): himpunan (daerah) harga-harga dimana H0 ditolak

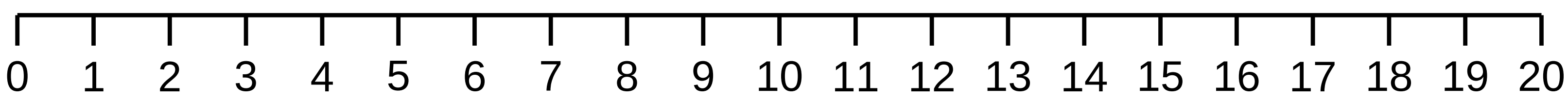
Statistik Penguji: statistik atau variabel random yang digunakan untuk menentukan apakah H_0 ditolak atau tidak ditolak. Bila statistik penguji masuk dalam daerah penolakan maka H_0 ditolak, sebaliknya jika tidak maka H_0 tidak ditolak.

Inferensi Statistik

Daerah penolakan (Daerah kritik): himpunan (daerah) harga-harga dimana H0 ditolak

Statistik Penguji: statistik atau variabel random yang digunakan untuk menentukan apakah H_0 ditolak atau tidak ditolak. Bila statistik penguji masuk dalam daerah penolakan maka H_0 ditolak, sebaliknya jika tidak maka H_0 tidak ditolak.

Contoh (lanjutan):
Daerah penolakan:

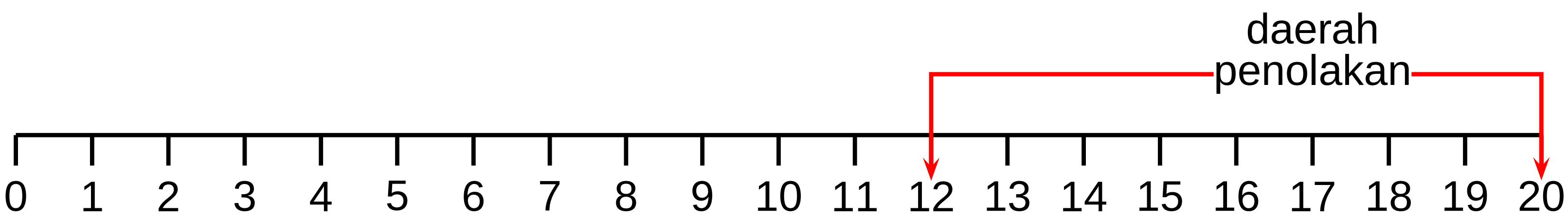


Inferensi Statistik

Daerah penolakan (Daerah kritik): himpunan (daerah) harga-harga dimana H0 ditolak

Statistik Penguji: statistik atau variabel random yang digunakan untuk menentukan apakah H_0 ditolak atau tidak ditolak. Bila statistik penguji masuk dalam daerah penolakan maka H_0 ditolak, sebaliknya jika tidak maka H_0 tidak ditolak.

Contoh (lanjutan):
Daerah penolakan: $X \geq 12$

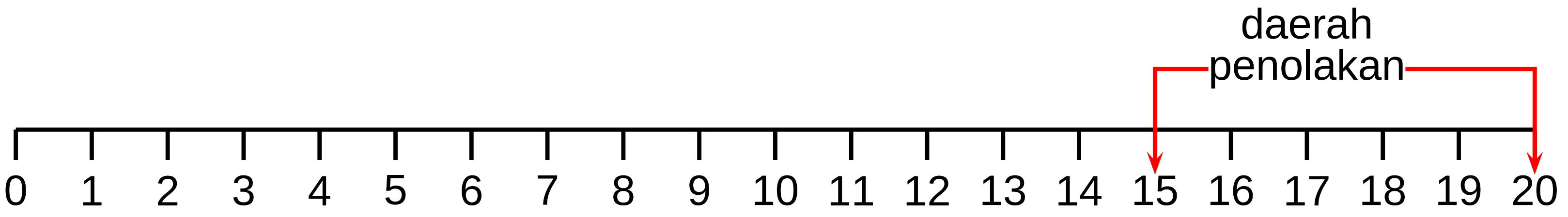


Inferensi Statistik

Daerah penolakan (Daerah kritik): himpunan (daerah) harga-harga dimana H0 ditolak

Statistik Penguji: statistik atau variabel random yang digunakan untuk menentukan apakah H_0 ditolak atau tidak ditolak. Bila statistik penguji masuk dalam daerah penolakan maka H_0 ditolak, sebaliknya jika tidak maka H_0 tidak ditolak.

Contoh (lanjutan):
Daerah penolakan: $X \geq 15$



Inferensi Statistik

$P(\text{Tipe I}) = \alpha$ untuk beberapa nilai p dengan menganggap H_0 benar ($p \leq 0,6$) dan daerah penolakan $X \geq 12$

	p di bawah H_0			
$P(\text{Tipe I}) = \alpha$	0,2	0,3	0,4	0,6
$P(X \geq 12)$	0,00	0,005	0,057	0,596

Inferensi Statistik

Harga peluang untuk $p = 0,6$ untuk beberapa kriteria penolakan

	$X \geq 12$	$X \geq 14$	$X \geq 16$	$X \geq 18$
Peluang	0,596	0,25	0,051	0,004

p-*value*: nilai α yang terkecil.

Inferensi Statistik

Tahap-tahap Uji Hipotesis Secara umum

1. Tentukan model probabilitas yang cocok dari data
2. Tentukan Hipotesis H_0 dan H_1
3. Tentukan Statistik Penguji, yang harus merupakan fungsi dari data dan tidak memuat parameter yang tidak diketahui
4. Tentukan tingkat signifikansi
5. Tentukan daerah kritik berdasarkan tingkat signifikansi
6. Hitung Statistik Penguji, apakah masuk daerah kritik atau tidak
7. Alternatif: Hitung p-*value* berdasarkan statistik penguji
8. Ambil kesimpulan berdasarkan 6 atau 7

Inferensi Statistik

- Satu Populasi
  - Populasi sembarang
    - μ
    - p
  - Populasi Normal
    - μ
    - σ^2
- Dua Populasi
  - Populasi sembarang
    - $\mu_1^2,\ \mu_2^2$
    - $p_1^2,\ p_2^2$
  - Populasi Normal
    - $\mu_1^2,\ \mu_2^2$
    - $\sigma_1^2,\ \sigma_2^2$
- $k > 2$ Populasi — Analisis Variansi (ANAVA)

Inferensi Statistik Satu Populasi Sembarang

Estimasi interval mean (μ) suatu populasi

Teorema Limit Pusat

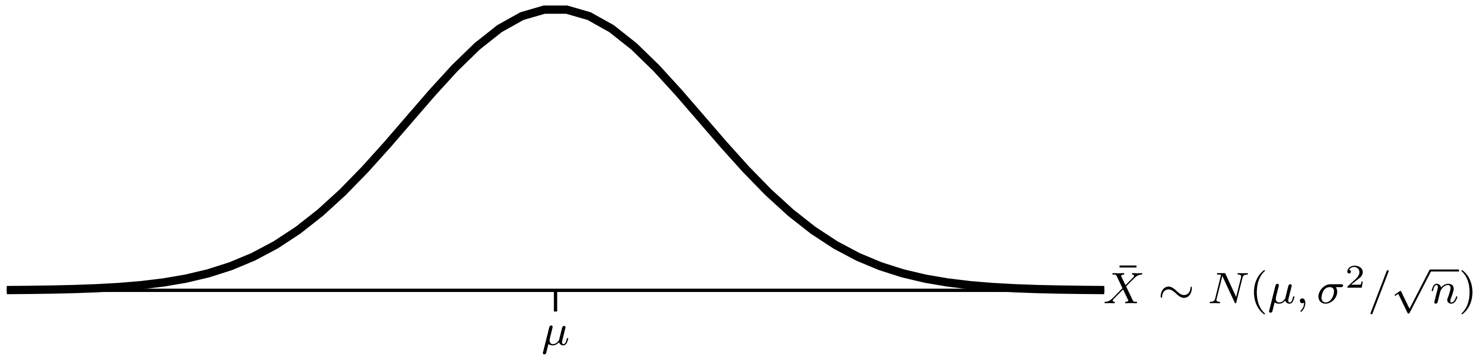
Apabila sampel-sampel random diambil dari suatu populasi yang berdistribusi sembarang, yang mempunyai mean μ dan variansi σ^2, maka untuk n besar, distribusi sampling untuk mean dapat dianggap mendekati Normal dengan $\mu_{\bar{X}} = \mu$ dan variansi $\sigma^2_{\bar{X}} = \sigma^2/n$, sehingga

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

mendekati Normal Standar.

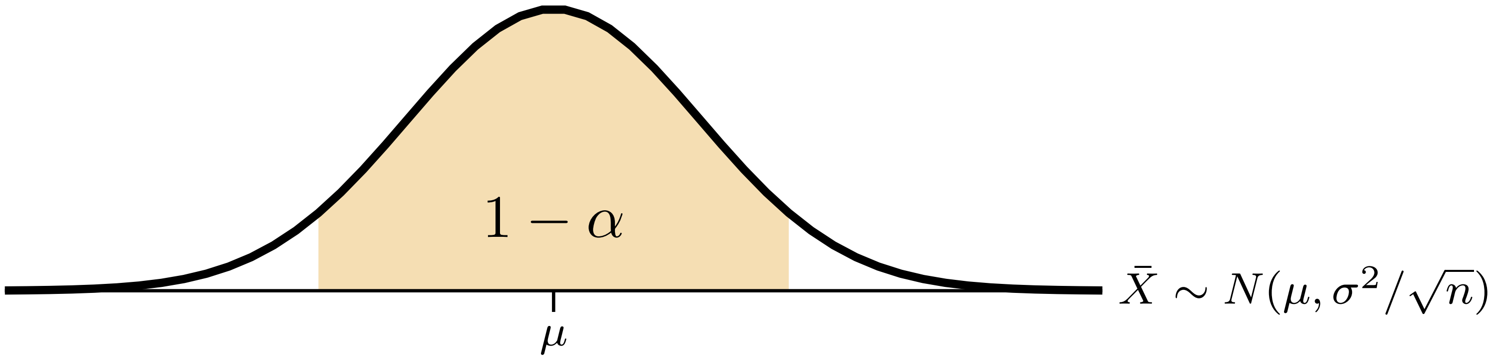
Inferensi Statistik Satu Populasi Sembarang

Estimasi interval mean (μ) suatu populasi



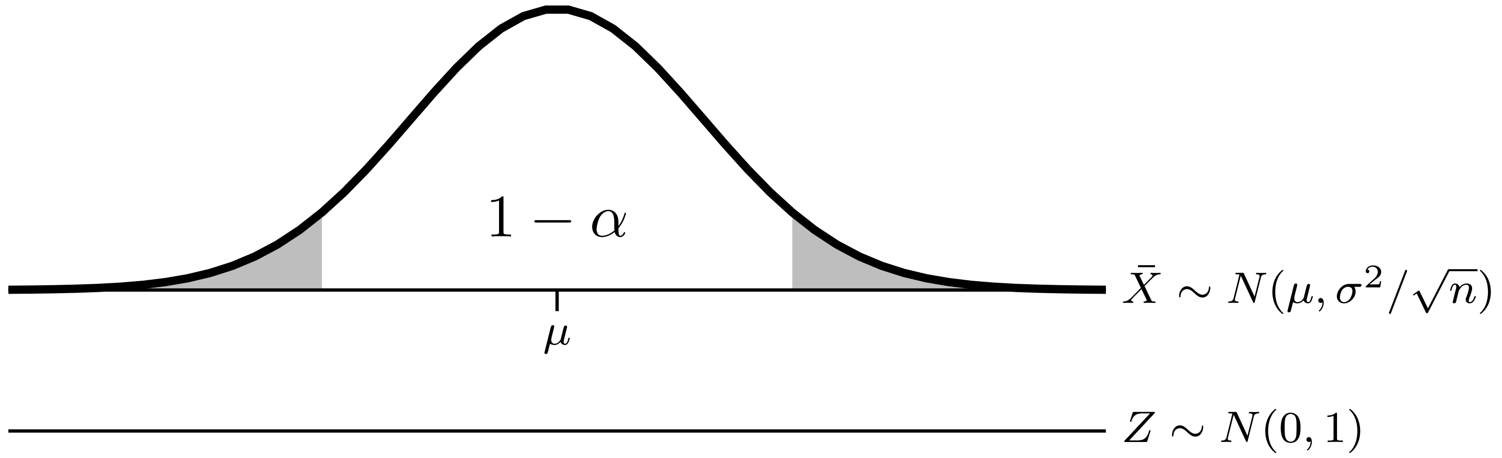
Inferensi Statistik Satu Populasi Sembarang

Estimasi interval mean (μ) suatu populasi



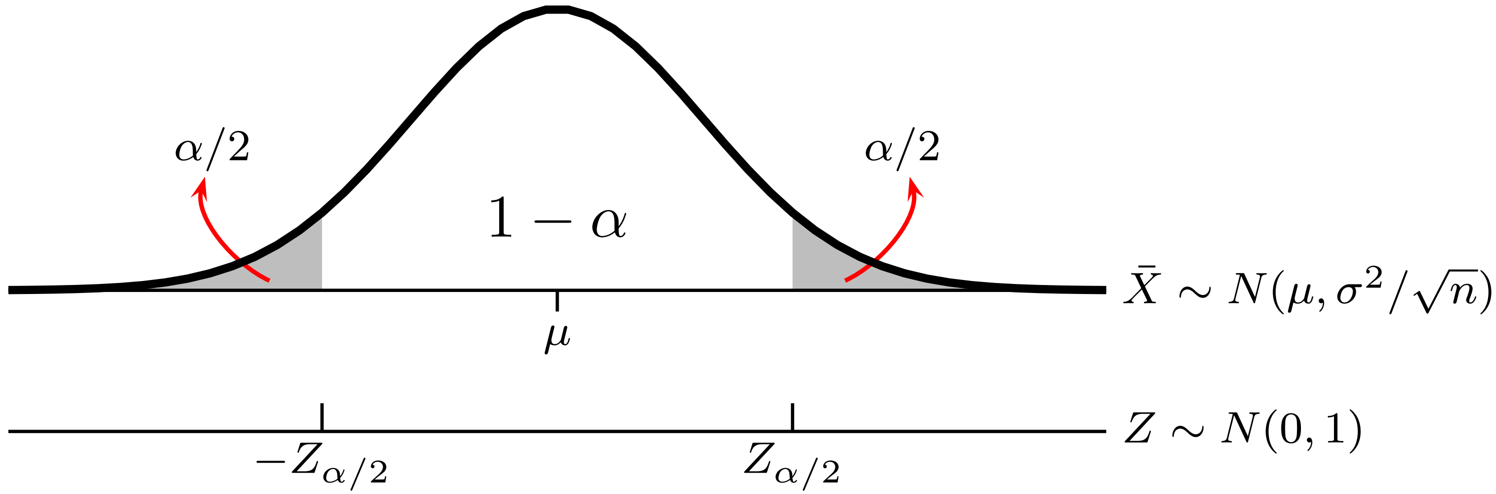
Inferensi Statistik Satu Populasi Sembarang

Estimasi interval mean (μ) suatu populasi



Inferensi Statistik Satu Populasi Sembarang

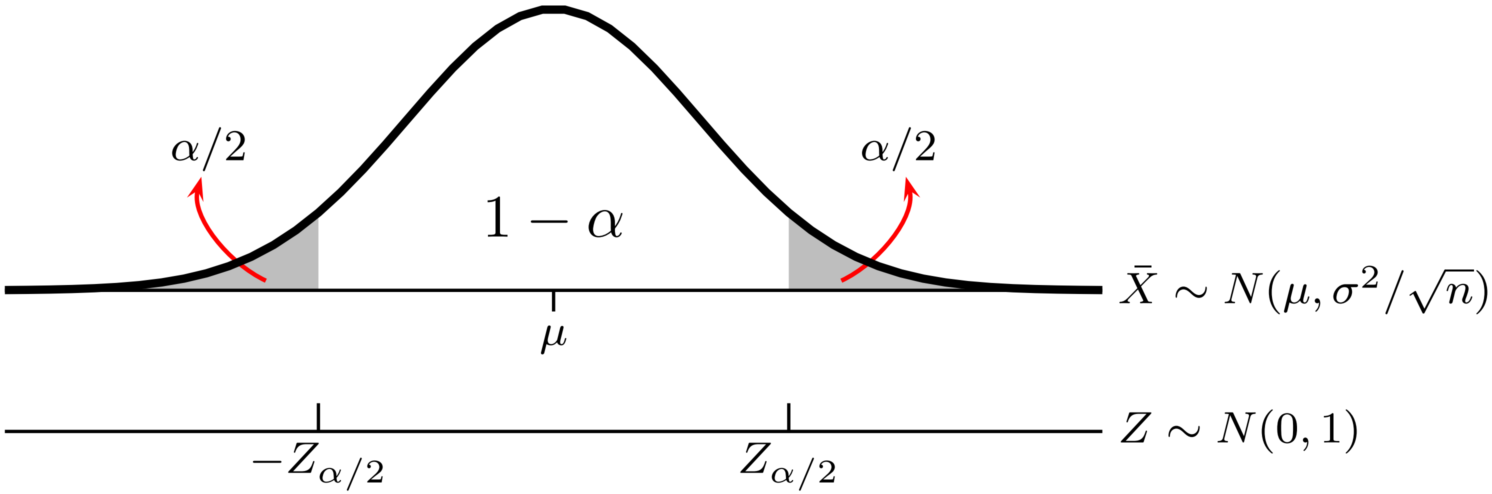
Estimasi interval mean (μ) suatu populasi



$$P(-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}) \approx 1-\alpha$$

Inferensi Statistik Satu Populasi Sembarang

Estimasi interval mean (μ) suatu populasi

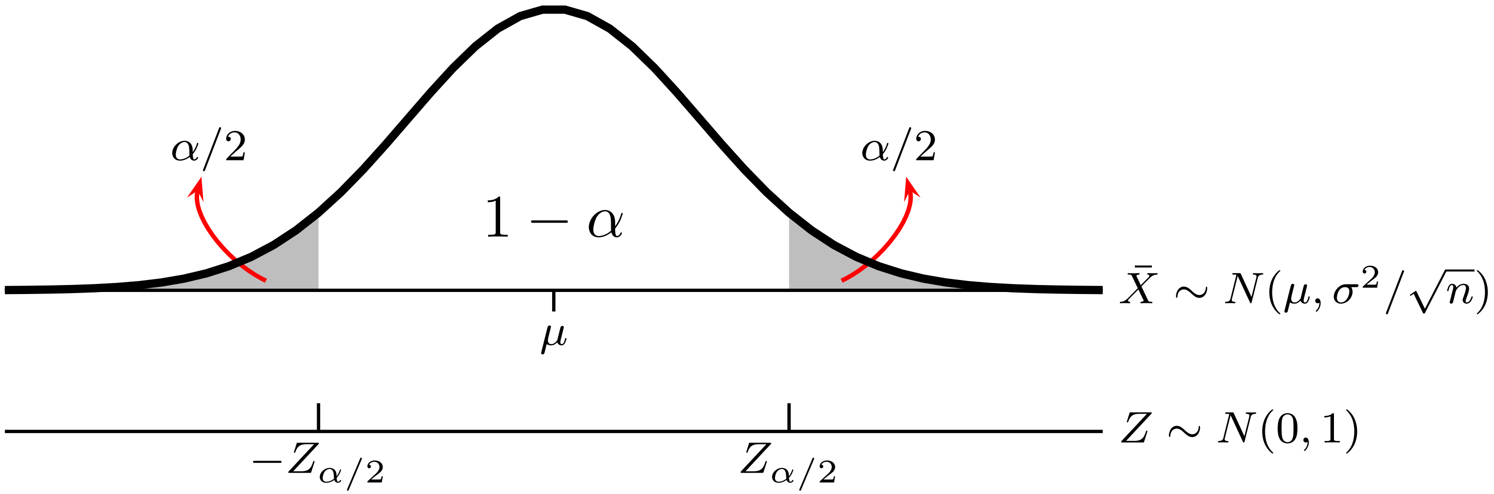


$$P(-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}) \approx 1-\alpha$$

$$P(-Z_{\alpha/2} \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq Z_{\alpha/2}) \approx 1-\alpha$$

Inferensi Statistik Satu Populasi Sembarang

Estimasi interval mean (μ) suatu populasi



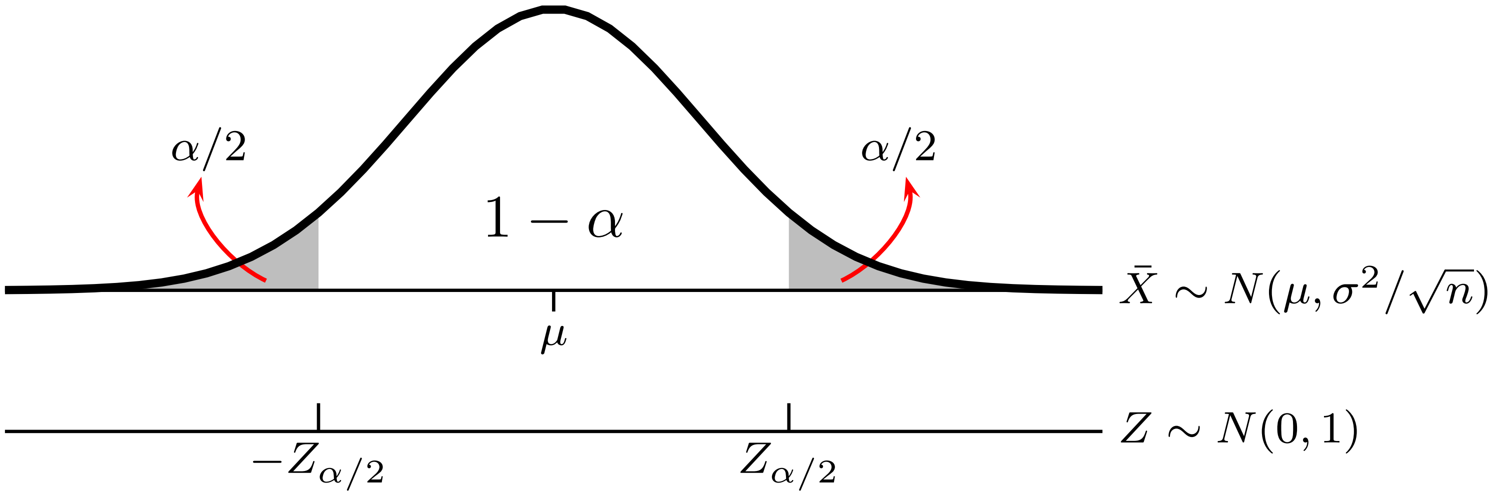
$$P(-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}) \approx 1-\alpha$$

$$P(-Z_{\alpha/2} \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq Z_{\alpha/2}) \approx 1-\alpha$$

$$P(\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) \approx 1-\alpha$$

Inferensi Statistik Satu Populasi Sembarang

Estimasi interval mean (μ) suatu populasi



Interval Konfidensi $(1-\alpha)100\%$ untuk mean μ

$B \leq \mu \leq A$

$$B = \bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$A = \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Inferensi Statistik Satu Populasi Sembarang

Contoh:
Suatu sampel random dengan 150 keluarga di suatu kota menunjukkan penghasilan bulanan rata-rata Rp 325 000,00 dengan deviasi standar Rp 25 000,00. Hitung interval konfidensi 95% untuk rata-rata penghasilan bulanan seluruh keluarga di kota tersebut.

Inferensi Statistik Satu Populasi Sembarang

Contoh:
Suatu sampel random dengan 150 keluarga di suatu kota menunjukkan penghasilan bulanan rata-rata Rp 325 000,00 dengan deviasi standar Rp 25 000,00. Hitung interval konfidensi 95% untuk rata-rata penghasilan bulanan seluruh keluarga di kota tersebut.

Jawab:
X : penghasilan bulanan di kota tersebut
$\bar{X} = 325.000$; $s = 25.000$; $n = 150$.
Interval konfidensi 95% untuk rata-rata penghasilan bulanan (μ):

$$B = \bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 325.000 - 1{,}96\frac{25}{\sqrt{150}} = 324.996$$

$$A = \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} = 325.000 + 1{,}96\frac{25}{\sqrt{150}} = 325.004$$

Interval konfidensi 95%: $324.996 \leq \mu \leq 325.004$

σ dapat diganti s

Inferensi Statistik Satu Populasi Sembarang

Uji Hipotesis Mean (μ) Populasi

1. Hipotesis
   A. $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$
   B. $H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$
   C. $H_0 : \mu \geq \mu_0$ vs. $H_1 : \mu < \mu_0$
2. Tingkat signifikansi α
3. Statistik Penguji

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

atau

$$Z = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

jika σ tidak diketahui diganti s. Distribusi dari Z adalah Normal Standar.

Inferensi Statistik Satu Populasi Sembarang

Uji Hipotesis Mean (μ) Populasi

4. Daerah penolakan (berdasarkan α dan Hipotesis)

   A. H_0 ditolak apabila $Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$

   B. H_0 ditolak apabila $Z > Z_{\alpha}$

   C. H_0 ditolak apabila $Z < -Z_{\alpha}$

Inferensi Statistik Satu Populasi Sembarang

Contoh:
Ujian standar intelegensia telah diadakan beberapa tahun dengan nilai rata-rata 70 dengan deviasi standar 8. Sekelompok mahasiswa terdiri dari 100 orang mahasiswa, diberi pelajaran dengan mengutamakan bidang Matematika. Apabila dari 100 mahasiswa ini diperoleh hasil ujian dengan nilai rata-rata 75, apakah cukup alasan untuk mempecayai bahwa pengutamaan bidang Matematika menaikkan hasil ujian standar?

Inferensi Statistik Satu Populasi Sembarang

Estimasi interval proporsi (p) suatu populasi

Jika $X \sim \mathrm{Binomial}(n, p)$, maka variabel random $\frac{x}{n}$ mempunyai mean p dan variansi $\frac{p(1-p)}{n}$

Untuk n besar

$$Z = \frac{\frac{x}{n} - p}{\sqrt{\frac{\frac{x}{n}(1-\frac{x}{n})}{n}}}$$

mendekati Normal Standar (Teorema Limit Pusat)

Inferensi Statistik Satu Populasi Sembarang

Estimasi interval proporsi (p) suatu populasi

Interval Konfidensi $(1-\alpha)100\%$ untuk p
$B \le p \le A$

$B = \hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

$A = \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

dengan $\hat{p} = \frac{x}{n}$

Inferensi Statistik Satu Populasi Sembarang

Contoh:
Jika 610 dari 900 sampel random petani di suatu daerah adalah buruh tani, hitunglah interval konfidensi 90% untuk proporsi buruh tani di daerah itu.

Inferensi Statistik Satu Populasi Sembarang

Uji Hipotesis proporsi (p) Populasi

1. Hipotesis
   A. $H_0 : p = p_0$ vs. $H_1 : p \neq p_0$
   B. $H_0 : p \leq p_0$ vs. $H_1 : p > p_0$
   C. $H_0 : p \geq p_0$ vs. $H_1 : p < p_0$
2. Tingkat signifikansi α
3. Statistik Penguji

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

   Distribusi dari Z adalah Normal Standar.

Inferensi Statistik Satu Populasi Sembarang

Uji Hipotesis proporsi (p) Populasi

4. Daerah penolakan (berdasarkan α dan Hipotesis)

   A. H_0 ditolak apabila $Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$

   B. H_0 ditolak apabila $Z > Z_{\alpha}$

   C. H_0 ditolak apabila $Z < -Z_{\alpha}$

Inferensi Statistik Satu Populasi Sembarang

Hubungan antara Interval Konfidensi dan Uji Hipotesis

Interval Konfidensi $(1-\alpha)100\%$ untuk mean μ

$$\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Daerah penolakan dengan tingkat signifikansi α untuk uji hipotesis $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$

$$Z > Z_{\alpha/2} \text{ atau } Z < -Z_{\alpha/2}$$

Daerah penerimaan

$$-Z_{\alpha/2} \leq Z \leq Z_{\alpha/2}$$

Inferensi Statistik Satu Populasi Sembarang

Hubungan antara Interval Konfidensi dan Uji Hipotesis

Interval Konfidensi $(1-\alpha)100\%$ untuk mean μ

$$\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Daerah penolakan dengan tingkat signifikansi α untuk uji hipotesis $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$

$$Z > Z_{\alpha/2} \text{ atau } Z < -Z_{\alpha/2}$$

Daerah penerimaan

$$-Z_{\alpha/2} \leq \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \leq Z_{\alpha/2}$$

Inferensi Statistik Satu Populasi Sembarang

Hubungan antara Interval Konfidensi dan Uji Hipotesis

Interval Konfidensi $(1-\alpha)100\%$ untuk mean μ

$$\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Daerah penolakan dengan tingkat signifikansi α untuk uji hipotesis $H_0 : \mu = \mu_0$ vs. $H_1 : \mu \neq \mu_0$

$$Z > Z_{\alpha/2} \text{ atau } Z < -Z_{\alpha/2}$$

Daerah penerimaan

$$\bar{X} - Z_{\alpha/2}\tfrac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \bar{X} + Z_{\alpha/2\frac{\sigma}{\sqrt{n}}}$$

Inferensi Statistik Satu Populasi Sembarang

Ringkasan

Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
μ mean	$Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ $Z \sim N(0,1)$	$B \leq \mu \leq A$ $B = \bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ $A = \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$	$H_1 : \mu \neq \mu_0$ $H_1 : \mu > \mu_0$ $H_1 : \mu < \mu_0$	$Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$ $Z > Z_\alpha$ $Z < -Z_\alpha$
p proporsi	$Z = \dfrac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ $Z \sim N(0,1)$	$B \leq p \leq A$ $B = \hat{p} - Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $A = \hat{p} + Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$H_1 : p \neq p_0$ $H_1 : p > p_0$ $H_1 : p < p_0$	$Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$ $Z > Z_\alpha$ $Z < -Z_\alpha$

Inferensi Statistik Satu Populasi Normal

- Data dianggap berdistribusi Normal
- Ukuran sampel tidak harus besar
- Jenis parameter:
  - mean μ
  - variansi σ^2
- Distribusi Sampling
  - Normal
  - t
  - Chi-kuadrat (*Chi-square*)

Inferensi Statistik Satu Populasi Normal

Normal Standar

Jika $X_1, \ldots, X_n$ adalah sampel random berasal dari populasi Normal dengan mean μ dan variansi σ^2 maka variabel random

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

berdistribusi Normal Standar $N(0, 1)$

Inferensi Statistik Satu Populasi Normal

<u>Distribusi t</u>

Jika $X_1, \ldots, X_n$ adalah sampel random berasal dari populasi Normal dengan mean μ dan variansi σ^2 maka variabel random

$$t = \frac{\bar{X} - \mu}{s/\sqrt{n}}$$

berdistribusi t dengan derajad bebas $n - 1$.

Untuk n yang semakin besar, distribusi t akan mendekati distribusi Normal.

Inferensi Statistik Satu Populasi Normal

Distribusi Chi-kuadrat $2k$

Diketahui $X_1, \ldots, X_k$ adalah variabel random yang berdistribusi Normal yang independen satu dengan yang lain. Distribusi variabel random

$$\chi^2 = X_1^2 + \ldots + X_k^2$$

berdistribusi Chi-kuadrat berderajad bebas k dengan mean $\mathrm{E}(\chi^2) = k$ dan variansi $\mathrm{Var}(\chi^2) = 2k$

Inferensi Statistik Satu Populasi Normal

Distribusi Chi-kuadrat $n-1$

Diketahui $X_1, \ldots, X_n$ adalah variabel random yang berdistribusi Normal dengan mean μ dan variansi σ^2 maka variabel random

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

berdistribusi Chi-kuadrat dengan derajad bebas $n-1$

Inferensi Statistik Satu Populasi Normal

<u>Distribusi Normal Standar</u>

Apabila sampel random berukuran n diambil dari suatu populasi yang berdistribusi Normal dengan mean μ dan variansi σ^2, maka variabel random

$$Z = \frac{s^2 - \sigma^2}{\sigma^2 \sqrt{\frac{2}{n-1}}}$$

berdistribusi $N(0,1)$ untuk n besar.

Inferensi Statistik Satu Populasi Normal

Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
μ mean	Bila σ^2 diketahui $Z = \dfrac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$ $Z \sim N(0,1)$	$B \le \mu \le A$ $B = \bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ $A = \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$	$H_1 : \mu \neq \mu_0$ $H_1 : \mu > \mu_0$ $H_1 : \mu < \mu_0$	$Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$ $Z > Z_\alpha$ $Z < -Z_\alpha$
	Bila σ^2 tidak diketahui $t = \dfrac{\bar{X} - \mu_0}{s/\sqrt{n}}$ $t \sim$ distribusi t dgn. derajad bebas $n-1$	$B \le \mu \le A$ $B = \bar{X} - t_{(n-1,\alpha/2)}\frac{s}{\sqrt{n}}$ $A = \bar{X} + t_{(n-1,\alpha/2)}\frac{s}{\sqrt{n}}$	$H_1 : \mu \neq \mu_0$ $H_1 : \mu > \mu_0$ $H_1 : \mu < \mu_0$	$t > t_{(n-1,\alpha/2)}$ atau $t < -t_{(n-1,\alpha/2)}$ $t > t_{(n-1,\alpha)}$ $t < -t_{(n-1,\alpha)}$

Inferensi Statistik Satu Populasi Normal

Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
σ^2 variansi	$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$ $\chi^2 \sim$ chi-square dgn. derajad bebas $k = n-1$	$B \leq \sigma^2 \leq A$ $B = \frac{(n-1)s^2}{\chi^2_{(n-1,\alpha/2)}}$ $A = \frac{(n-1)s^2}{\chi^2_{(n-1,1-\alpha/2)}}$	$H_1 : \sigma^2 \neq \sigma_0^2$ $H_1 : \sigma^2 > \sigma_0^2$ $H_1 : \sigma^2 < \sigma_0^2$	$\chi^2 > \chi^2_{(k,\alpha/2)}$ atau $\chi^2 < \chi^2_{(k,1-\alpha/2)}$ $\chi^2 > \chi^2_{(k,\alpha)}$ $\chi^2 < \chi^2_{(k,1-\alpha)}$
	Untuk n besar, $$Z = \frac{s^2 - \sigma^2}{\sigma^2\sqrt{\frac{2}{n-1}}}$$ $Z \sim N(0,1)$	$B \leq \sigma^2 \leq A$ $B = \frac{s^2}{1+Z_{\alpha/2}\sqrt{\frac{2}{n-1}}}$ $A = \frac{s^2}{1-Z_{\alpha/2}\sqrt{\frac{2}{n-1}}}$	$H_1 : \sigma^2 \neq \sigma_0^2$ $H_1 : \sigma^2 > \sigma_0^2$ $H_1 : \sigma^2 < \sigma_0^2$	$Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$ $Z > Z_\alpha$ $Z < -Z_\alpha$

Inferensi Statistik Dua Populasi Sembarang

Distribusi sampling selisih dua mean

Misalkan $X_{11}, X_{12}, \ldots, X_{1n_1}$ dan $X_{21}, X_{22}, \ldots, X_{2n_2}$ adalah dua sampel random independen satu sama lain yang diambil dari populasi yang mempunyai mean μ_1 dan μ_2 serta variansi σ_1^2 dan σ_2^2, maka untuk n_1 dan n_2 besar, variabel random

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

berdistribusi Normal Standar, dengan

$$\bar{X}_1 = \sum_{i=1}^{n_1} \frac{X_{1i}}{n_1} \qquad \bar{X}_2 = \sum_{i=1}^{n_2} \frac{X_{2i}}{n_2}$$

Inferensi Statistik Dua Populasi Sembarang

Distribusi sampling selisih dua mean

Jika σ_1^2 dan σ_2^2 tidak diketahui, dan diasumsikan $\sigma_1^2 \neq \sigma_2^2$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

berdistribusi Normal Standar dengan s_1^2 dan s_2^2 adalah variansi sampel

Inferensi Statistik Dua Populasi Sembarang

Distribusi sampling selisih dua mean

Jika σ_1^2 dan σ_2^2 tidak diketahui, dan diasumsikan $\sigma_1^2 = \sigma_2^2$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})}}$$

berdistribusi Normal Standar dengan

$$s_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

yang disebut sebagai *pooled variance*

Inferensi Statistik Dua Populasi Sembarang

Distribusi sampling selisih dua proporsi

Misalkan $X_{11}, X_{12}, \ldots, X_{1n_1}$ dan $X_{21}, X_{22}, \ldots, X_{2n_2}$ adalah dua sampel random independen satu sama lain yang diambil dari populasi yang berdistribusi binomial. Untuk n_1 dan n_2 besar, variabel random

$$Z = \frac{\left(\frac{X_1}{n_1} - \frac{X_2}{n_2}\right) - (p_1 - p_2)}{\sqrt{\frac{\frac{X_1}{n_1}(1-\frac{X_1}{n_1})}{n_1} + \frac{\frac{X_2}{n_2}(1-\frac{X_2}{n_2})}{n_2}}}$$

berdistribusi Normal Standar.

Inferensi Statistik Dua Populasi Sembarang

Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
$\mu_1 - \mu_2$ selisih dua mean	σ_1^2 dan σ_2^2 diketahui $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ $Z \sim N(0,1)$	$B \leq \mu_1 - \mu_2 \leq A$ $B = (\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ $A = (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$H_1 : \mu_1 - \mu_2 \neq \mu_0$ $H_1 : \mu_1 - \mu_2 > \mu_0$ $H_1 : \mu_1 - \mu_2 < \mu_0$	$Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$ $Z > Z_\alpha$ $Z < -Z_\alpha$
	σ_1^2 dan σ_2^2 tdk diketahui, $\sigma_1^2 \neq \sigma_2^2$ $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ $Z \sim N(0,1)$	$B \leq \mu_1 - \mu_2 \leq A$ $B = (\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ $A = (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$H_1 : \mu_1 - \mu_2 \neq \mu_0$ $H_1 : \mu_1 - \mu_2 > \mu_0$ $H_1 : \mu_1 - \mu_2 < \mu_0$	$Z > Z_{\alpha/2}$ atau $Z < -Z_{\alpha/2}$ $Z > Z_\alpha$ $Z < -Z_\alpha$

Inferensi Statistik Dua Populasi Sembarang

Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
	σ_1^2 dan σ_2^2 tdk diketahui, $\sigma_1^2 = \sigma_2^2$ $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})}}$ $Z \sim N(0,1)$ $s_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}$	$B \leq \mu_1 - \mu_2 \leq A$ $B = (\bar{X}_1 - \bar{X}_2) -$ $Z_{\frac{\alpha}{2}} \sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}$ $A = (\bar{X}_1 - \bar{X}_2) +$ $Z_{\frac{\alpha}{2}} \sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}$	$H_1 : \mu_1 - \mu_2 \neq \mu_0$ $H_1 : \mu_1 - \mu_2 > \mu_0$ $H_1 : \mu_1 - \mu_2 < \mu_0$	$Z > Z_{\frac{\alpha}{2}}$ atau $Z < -Z_{\frac{\alpha}{2}}$ $Z > Z_\alpha$ $Z < -Z_\alpha$
$p_1 - p_2$ Selisih dua proporsi	$Z =$ $\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}}$ $Z \sim N(0,1)$ $\hat{p}_1 = \frac{X_1}{n_1}$; $\hat{p}_2 = \frac{X_2}{n_2}$	$B \leq p_1 - p_2 \leq A$ $B = (\hat{p}_1 - \hat{p}_2) -$ $Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}$ $A = (\hat{p}_1 - \hat{p}_2) +$ $Z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}$	$H_1 : p_1 - p_2 \neq p_0$ $H_1 : p_1 - p_2 > p_0$ $H_1 : p_1 - p_2 < p_0$	$Z > Z_{\frac{\alpha}{2}}$ atau $Z < -Z_{\frac{\alpha}{2}}$ $Z > Z_\alpha$ $Z < -Z_\alpha$

Inferensi Statistik Dua Populasi Normal

Distribusi sampling selisih dua mean

Misalkan $X_{11}, X_{12}, \ldots, X_{1n_1}$ dan $X_{21}, X_{22}, \ldots, X_{2n_2}$ adalah dua sampel random independen satu sama lain yang berdistribusi Normal dengan mean μ_1 dan μ_2 serta variansi σ_1^2 dan σ_2^2, maka variabel random

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

berdistribusi Normal Standar, dengan

$$\bar{X}_1 = \sum_{i=1}^{n_1} \frac{X_{1i}}{n_1} \qquad \bar{X}_2 = \sum_{i=1}^{n_2} \frac{X_{2i}}{n_2}$$

Inferensi Statistik Dua Populasi Normal

Distribusi sampling selisih dua mean

Jika σ_1^2 dan σ_2^2 tidak diketahui, dan diasumsikan $\sigma_1^2 \neq \sigma_2^2$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

berdistribusi t dengan derajad bebas

$$k = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1+1} + \frac{(s_2^2/n_2)^2}{n_2+1}} - 2, \quad \text{atau} \quad k = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$$

dengan s_1^2 dan s_2^2 adalah variansi sampel

Inferensi Statistik Dua Populasi Normal

Distribusi sampling selisih dua mean

Jika σ_1^2 dan σ_2^2 tidak diketahui, dan diasumsikan $\sigma_1^2 = \sigma_2^2$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})}}$$

berdistribusi t dengan derajad bebas $n_1 + n_2 - 2$ dan

$$s_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

yang disebut sebagai *pooled variance*

Inferensi Statistik Dua Populasi Normal

Distribusi sampling Perbandingan dua variansi

Misalkan $X_{11}, X_{12}, \ldots, X_{1n_1}$ dan $X_{21}, X_{22}, \ldots, X_{2n_2}$ adalah dua sampel random independen satu sama lain yang berdistribusi Normal dengan mean μ_1 dan μ_2 serta variansi σ_1^2 dan σ_2^2, maka variabel random

$$F = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}$$

berdistribusi F dengan derajad bebas pembilang $n_1 - 1$, derajad bebas penyebut $n_2 - 1$

Inferensi Statistik Dua Populasi Normal

Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
$\mu_1 - \mu_2$ Selisih dua mean	σ_1^2 dan σ_2^2 diketahui $Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ $Z \sim N(0,1)$	$B \leq \mu_1 - \mu_2 \leq A$ $B = (\bar{X}_1 - \bar{X}_2) - Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$ $A = (\bar{X}_1 - \bar{X}_2) + Z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	$H_1 : \mu_1 - \mu_2 \neq \mu_0$ $H_1 : \mu_1 - \mu_2 > \mu_0$ $H_1 : \mu_1 - \mu_2 < \mu_0$	$Z > Z_{\frac{\alpha}{2}}$ atau $Z < -Z_{\frac{\alpha}{2}}$ $Z > Z_\alpha$ $Z < -Z_\alpha$
	σ_1^2 dan σ_2^2 tdk diketahui dan $\sigma_1^2 \neq \sigma_2^2$ $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ $t \sim t_k$ dgn $k = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1+1} + \frac{(s_2^2/n_2)^2}{n_2+1}} - 2$ atau $k = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$	$B \leq \mu_1 - \mu_2 \leq A$ $B = (\bar{X}_1 - \bar{X}_2) - t_{\frac{\alpha}{2},k} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$ $A = (\bar{X}_1 - \bar{X}_2) + t_{\frac{\alpha}{2},k} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$H_1 : \mu_1 - \mu_2 \neq \mu_0$ $H_1 : \mu_1 - \mu_2 > \mu_0$ $H_1 : \mu_1 - \mu_2 < \mu_0$	$t > t_{\frac{\alpha}{2},k}$ atau $t < -t_{\frac{\alpha}{2},k}$ $t > t_{\alpha,k}$ $t < -t_{\alpha,k}$

Inferensi Statistik Dua Populasi Normal

Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
	σ_1^2 dan σ_2^2 tdk diketahui dan $\sigma_1^2 = \sigma_2^2$ $t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}}$ $t \sim t_k$ dgn. $k = n_1 + n_2 - 2$ $S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$	$B \le \mu_1 - \mu_2 \le A$ $B = (\bar{X}_1 - \bar{X}_2) - t_{\frac{\alpha}{2},k}\sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}$ $A = (\bar{X}_1 - \bar{X}_2) + t_{\frac{\alpha}{2},k}\sqrt{S_p^2(\frac{1}{n_1} + \frac{1}{n_2})}$	$H_1: \mu_1 - \mu_2 \neq \mu_0$ $H_1: \mu_1 - \mu_2 > \mu_0$ $H_1: \mu_1 - \mu_2 < \mu_0$	$t > t_{\frac{\alpha}{2},k}$ atau $t < -t_{\frac{\alpha}{2},k}$ $t > t_{\alpha,k}$ $t < -t_{\alpha,k}$
σ_1^2 / σ_2^2 Perban-dingan dua variansi	$F = s_1^2 / s_2^2$ dengan $F \sim F_{\frac{\alpha}{2},k_1,k_2}$ $k_1 = n_1 - 1, k_2 = n_2 - 1$	$B \le \sigma_1^2/\sigma_2^2 \le A$ $B = \frac{s_1^2/s_2^2}{F_{(k_1,k_2,\frac{\alpha}{2})}}$ $A = \frac{s_1^2}{s_2^2} F_{(k_1,k_2,\frac{\alpha}{2})}$ catatan: $F(1 - \frac{\alpha}{2}, k_1, k_2) = 1/F(\frac{\alpha}{2}, k_2, k_1)$	$H_1: \sigma_1 \neq \sigma_2$ $H_1: \sigma_1 > \sigma_2$ $H_1: \sigma_1 < \sigma_2$	$F > F_{\frac{\alpha}{2},k_1,k_2}$ atau $F < 1/F_{\frac{\alpha}{2},k_2,k_1}$ $F > F_{\alpha,k_1,k_2}$ $F < 1/F_{\alpha,k_2,k_1}$

Inferensi Statistik Dua Populasi Normal

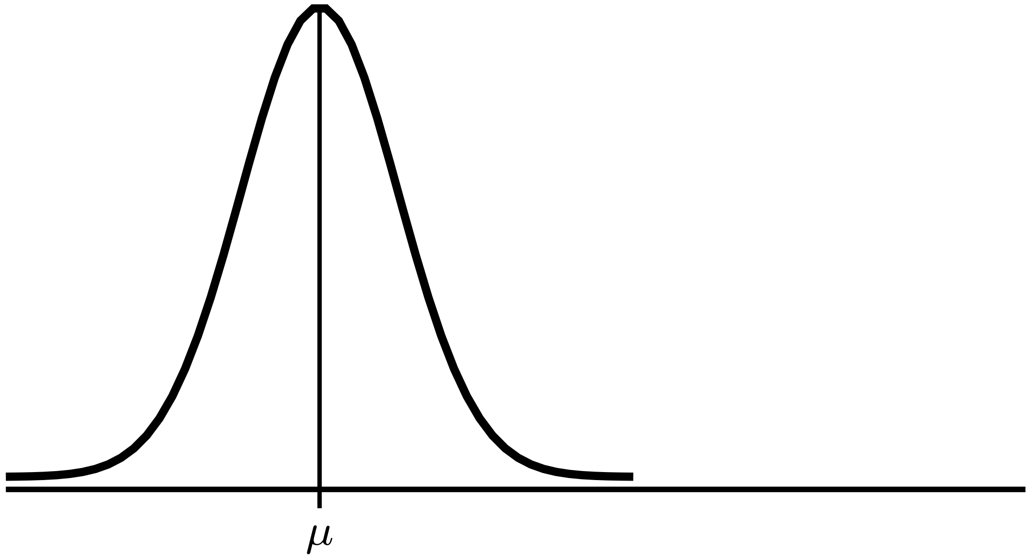
Parameter	Statistik	Interval Konfidensi (1-α)100%	Hipotesis alternatif	Daerah Kritik
μ_d mean selisih data berpasangan	$t = \frac{\bar{D}-\mu_D}{s_D/\sqrt{n}}$ dengan $t \sim$ distribusi t dgn derajad bebas $k = n-1$	$B \leq \mu \leq A$ $B = \bar{X} - t_{(n-1,\alpha/2)}\frac{s}{\sqrt{n}}$ $A = \bar{X} + t_{(n-1,\alpha/2)}\frac{s}{\sqrt{n}}$	$H_1 : \mu_D \neq \mu_0$ $H_1 : \mu_D > \mu_0$ $H_1 : \mu_D < \mu_0$	$t > t_{(n-1,\alpha/2)}$ atau $t < -t_{(n-1,\alpha/2)}$ $t > t_{(n-1,\alpha)}$ $t < -t_{(n-1,\alpha)}$

Analisis Variansi Satu Arah

- Perluasan dari uji mean dua populasi Normal (data berasal dari populasi Normal)
- Ada k mean populasi yang dibandingkan
- Berdasarkan pada pemecahan variansi

Analisis Variansi Satu Arah

Inferensi mean populasi Normal

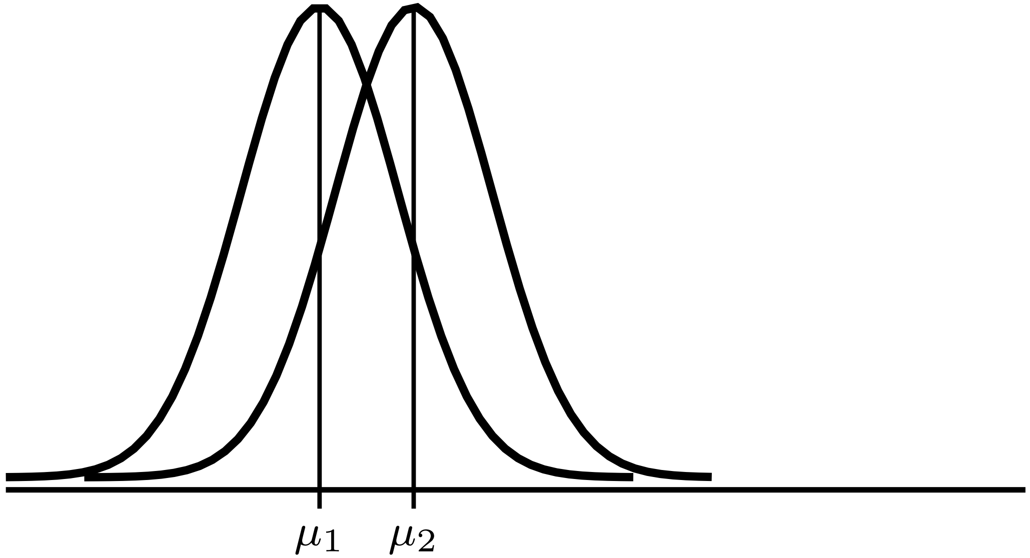


Uji mean satu populasi

$\mathrm{H}_0 : \mu = \mu_0$

Analisis Variansi Satu Arah

Inferensi mean populasi Normal



Uji mean satu populasi

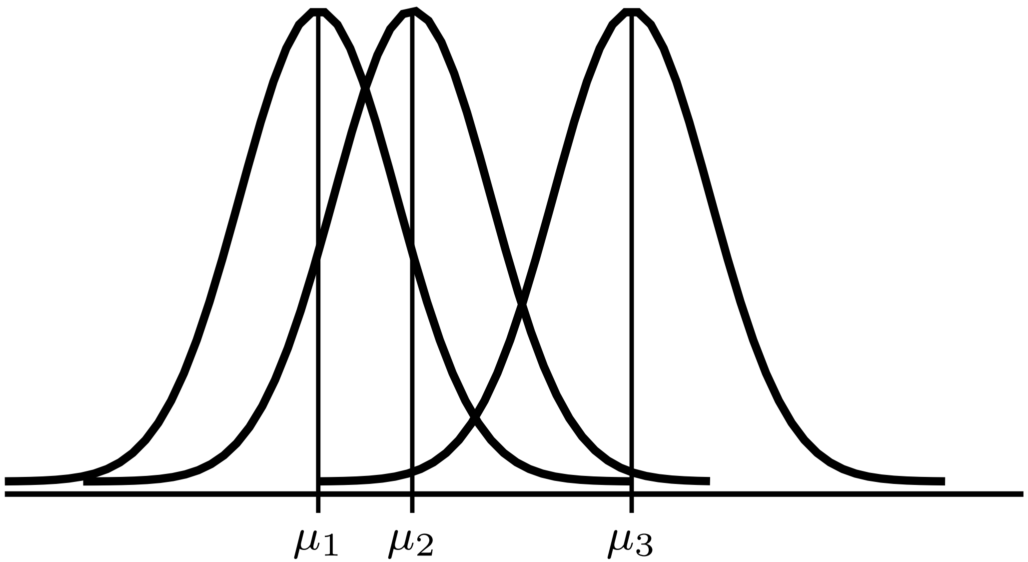
$\mathrm{H}_0 : \mu = \mu_0$

Uji mean dua populasi

$\mathrm{H}_0 : \mu_1 = \mu_2$

Analisis Variansi Satu Arah

Inferensi mean populasi Normal



Uji mean satu populasi

$\mathrm{H}_0 : \mu = \mu_0$

Uji mean dua populasi

$\mathrm{H}_0 : \mu_1 = \mu_2$

Uji mean k populasi

$\mathrm{H}_0 : \mu_1 = \mu_2 = \mu_3$

Analisis Variansi Satu Arah

Uji Hipotesis

$\mathrm{H}_0 : \mu_1 = \mu_2 = \ldots = \mu_k$
H_1 : minimal ada dua mean yang tidak sama

Statistik Penguji

$$F = \frac{\mathrm{MST}}{\mathrm{MSE}}$$

dimana $F \sim F_{(k-1,n-k)}$
MST: *mean square treatment* (kuadrat rata-rata perlakuan)
MSE: *mean square error* (kuadrat rata-rata sesatan)
yang diperoleh dari Tabel **Anava** (Analisis Variansi)

Daerah Kritis

H_0 ditolak jika $F > F_{(k-1,n-k)}$

Analisis Variansi Satu Arah

Tabel Anava

Sumber Variansi	derajad bebas	Jumlah Kuadrat (SS)	Rata-rata Jumlah Kuadrat (MS)	rasio F
Perlakuan	$k-1$	$\mathrm{SST} = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$	$\mathrm{MST} = \frac{SST}{k-1}$	$F = \frac{\mathrm{MST}}{\mathrm{MSE}}$
Sesatan	$N-k$	$\mathrm{SSE} = \sum_{i=1}^{k}(n_i - 1)S_i^2$	$\mathrm{MSE} = \frac{SSE}{N-k}$	

$$\bar{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$$

$$S_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2$$

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}, \quad N = \sum n_i$$

Analisis Variansi Satu Arah

Contoh:

Dipunyai empat varitas padi yang akan kita uji produktivitasnya. Dua puluh empat petak tanah yang kira-kira mempunyai kesuburan yang sama dipilih. Kemudian 24 petak itu dibagi secara random menjadi empat kelompok, masing-masing 6 petak yang selanjutnya tiap kelompok ditanami satu varitas padi. Apakah rata-rata produktivitas 4 varitas padi tersebut sama?

varitas			
A	B	C	D
24	13	21	27
13	21	13	30
18	11	26	24
24	23	23	29
16	28	16	26
23	18	12	34

Analisis Variansi Satu Arah

	varitas		
A	B	C	D
24	13	21	27
13	21	13	30
18	11	26	24
24	23	23	29
16	28	16	26
23	18	12	34

Analisis Variansi Satu Arah

	varitas			
	A	B	C	D
	24	13	21	27
	13	21	13	30
	18	11	26	24
	24	23	23	29
	16	28	16	26
	23	18	12	34
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33
S_i^2	21,87	40,40	32,30	12,27

Analisis Variansi Satu Arah

	varitas			
	A	B	C	D
	24	13	21	27
	13	21	13	30
	18	11	26	24
	24	23	23	29
	16	28	16	26
	23	18	12	34
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33
S_i^2	21,87	40,40	32,30	12,27

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij} = \frac{513}{24} = 21,38$$

Analisis Variansi Satu Arah

	varitas			
	A	B	C	D
	24	13	21	27
	13	21	13	30
	18	11	26	24
	24	23	23	29
	16	28	16	26
	23	18	12	34
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33
S_i^2	21,87	40,40	32,30	12,27

$\bar{X} = \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij} = \frac{513}{24} = 21,38$

$\text{SST} = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$

$= 6(19{,}67-21{,}38)^2+6(19{,}00-21{,}38)^2+6(18{,}50-21{,}38)^2+6(28{,}33-21{,}38)^2$

Analisis Variansi Satu Arah

	varitas			
	A	B	C	D
	24	13	21	27
	13	21	13	30
	18	11	26	24
	24	23	23	29
	16	28	16	26
	23	18	12	34
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33
S_i^2	21,87	40,40	32,30	12,27

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij} = \frac{513}{24} = 21,38$$

$$\text{SST} = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$$
$$= 391,46$$

Analisis Variansi Satu Arah

	varitas			
	A	B	C	D
	24	13	21	27
	13	21	13	30
	18	11	26	24
	24	23	23	29
	16	28	16	26
	23	18	12	34
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33
S_i^2	21,87	40,40	32,30	12,27

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij} = \frac{513}{24} = 21,38$$

$$\text{SST} = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$$

$$= 391,46$$

$$\text{SSE} = \sum_{i=1}^{k}(n_i - 1)S_i^2$$

$$= (6-1)21{,}87 + (6-1)40{,}40 + (6-1)32{,}30 + (6-1)12{,}27$$

Analisis Variansi Satu Arah

	varitas			
	A	B	C	D
	24	13	21	27
	13	21	13	30
	18	11	26	24
	24	23	23	29
	16	28	16	26
	23	18	12	34
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33
S_i^2	21,87	40,40	32,30	12,27

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij} = \frac{513}{24} = 21,38$$

$$\begin{aligned}\text{SST} &= \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2 \\ &= 391,46 \\ \text{SSE} &= \sum_{i=1}^{k}(n_i - 1)S_i^2 \\ &= 534,17\end{aligned}$$

Analisis Variansi Satu Arah

Tabel Anava Hasil 4 Varitas Padi

Sumber Variansi	derajad bebas	Jumlah Kuadrat (SS)	Rata-rata Jumlah Kuadrat (MS)	rasio F
Perlakuan	$k-1$	$\mathrm{SST} = \sum_{i=1}^{k} n_i(\bar{X}_i - \bar{X})^2$	$\mathrm{MST} = \frac{SST}{k-1}$	$F = \frac{\mathrm{MST}}{\mathrm{MSE}}$
Sesatan	$N-k$	$\mathrm{SSE} = \sum_{i=1}^{k}(n_i - 1)S_i^2$	$\mathrm{MSE} = \frac{SSE}{N-k}$	

Analisis Variansi Satu Arah

Tabel Anava Hasil 4 Varitas Padi

Sumber Variansi	derajad bebas	Jumlah Kuadrat (SS)	Rata-rata Jumlah Kuadrat (MS)	rasio F
Varitas	3	$\mathrm{SST} = 391,46$	$\mathrm{MST} = \frac{SST}{k-1}$	$F = \frac{\mathrm{MST}}{\mathrm{MSE}}$
Sesatan	20	$\mathrm{SSE} = 534,17$	$\mathrm{MSE} = \frac{SSE}{N-k}$	

Analisis Variansi Satu Arah

Tabel Anava Hasil 4 Varitas Padi

Sumber Variansi	derajad bebas	Jumlah Kuadrat (SS)	Rata-rata Jumlah Kuadrat (MS)	rasio F
Varitas	3	$\mathrm{SST} = 391,46$	$\mathrm{MST} = \frac{391,46}{3}$	$F = \frac{\mathrm{MST}}{\mathrm{MSE}}$
Sesatan	20	$\mathrm{SSE} = 534,17$	$\mathrm{MSE} = \frac{434,17}{20}$	

Analisis Variansi Satu Arah

Tabel Anava Hasil 4 Varitas Padi

Sumber Variansi	derajad bebas	Jumlah Kuadrat (SS)	Rata-rata Jumlah Kuadrat (MS)	rasio F
Varitas	3	$\mathrm{SST} = 391,46$	$\mathrm{MST} = 130,49$	$\frac{130,49}{26,71}$
Sesatan	20	$\mathrm{SSE} = 534,17$	$\mathrm{MSE} = 26,71$	

Analisis Variansi Satu Arah

Tabel Anava Hasil 4 Varitas Padi

Sumber Variansi	derajad bebas	Jumlah Kuadrat (SS)	Rata-rata Jumlah Kuadrat (MS)	rasio F
Varitas	3	$\mathrm{SST} = 391,46$	$\mathrm{MST} = 130,49$	$4,8856$
Sesatan	20	$\mathrm{SSE} = 534,17$	$\mathrm{MSE} = 26,71$	

Analisis Variansi Satu Arah

Tabel Anava Hasil 4 Varitas Padi

Sumber Variansi	derajad bebas	Jumlah Kuadrat (SS)	Rata-rata Jumlah Kuadrat (MS)	rasio F
Varitas	3	$\mathrm{SST} = 391,46$	$\mathrm{MST} = 130,49$	$4,8856$
Sesatan	20	$\mathrm{SSE} = 534,17$	$\mathrm{MSE} = 26,71$	

Statistik Penguji

$F = 4,8856$

Daerah Kritik ($\alpha = 0,05$)

H_0 ditolak jika $F > F_{(3,20)} = 3,10$

Analisis Variansi Satu Arah

Tabel Anava Hasil 4 Varitas Padi

Sumber Variansi	derajad bebas	Jumlah Kuadrat (SS)	Rata-rata Jumlah Kuadrat (MS)	rasio F
Varitas	3	$\mathrm{SST} = 391,46$	$\mathrm{MST} = 130,49$	$4,8856$
Sesatan	20	$\mathrm{SSE} = 534,17$	$\mathrm{MSE} = 26,71$	

Statistik Penguji

$F = 4,8856$

Daerah Kritik ($\alpha = 0,05$)

H_0 ditolak jika $F > F_{(3,20)} = 3,10$

Kesimpulan

$F = 4,8856 > 3,10 \Rightarrow \mathrm{H}_0$ ditolak, paling tidak ada dua mean yang tidak sama

Analisis Variansi Satu Arah

Pembandingan Ganda (*Multiple Comparisons*)

Merupakan analisis lanjutan bila H_0 ditolak dalam Anava.

Metode:

- Tukey
- Scheffé
- Bonferroni
- Newman - Keuls

Analisis Variansi Satu Arah

<u>Metode Scheffé</u>

- Mempunyai asumsi sama seperti Anava
- Menggunakan tabel F

Hipotesis

$\mathrm{H}_0 : \mu_i = \mu_j$
$\mathrm{H}_1 : \mu_i \neq \mu_j$
untuk $i \neq j$ dan $i, j = 1, 2, \ldots, k$

Daerah Kritik (Keputusan)

H_0 ditolak, jika $\mid \bar{X}_i - \bar{X}_j \mid > \sqrt{(k-1)S^2 F_{\alpha,k-1,N-k}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$

($S^2 = \mathrm{MSE}$ dalam Anava satu arah)

Analisis Variansi Satu Arah

<u>Contoh</u>

Pembandingan ganda untuk varitas padi contoh di muka.

Diketahui (dari hitungan di muka):

	varitas			
	A	B	C	D
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33

$$\text{MSE} = 26,71$$

Analisis Variansi Satu Arah

Contoh

Pembandingan ganda untuk varitas padi contoh di muka.

Diketahui (dari hitungan di muka):

	varitas			
	A	B	C	D
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33

$$\text{MSE} = 26,71$$

Pembandingan ganda Scheffé:

Pembandingan	$\mid \bar{X}_i - \bar{X}_j \mid$	$\sqrt{(k-1)\text{MSE}F_{\alpha,k-1,N-k}\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$	Kesimpulan
μ_A vs. μ_B			
μ_A vs. μ_C			
μ_A vs. μ_D			
μ_B vs. μ_C			
μ_B vs. μ_D			
μ_C vs. μ_D			

Analisis Variansi Satu Arah

Contoh

Pembandingan ganda untuk varitas padi contoh di muka.

Diketahui (dari hitungan di muka):

	varitas			
	A	B	C	D
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33

$$\text{MSE} = 26,71$$

Pembandingan ganda Scheffé:

Pembandingan	$\mid \bar{X}_i - \bar{X}_j \mid$	$\sqrt{3(26,71)(3,10)\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$	Kesimpulan
μ_A vs. μ_B			
μ_A vs. μ_C			
μ_A vs. μ_D			
μ_B vs. μ_C			
μ_B vs. μ_D			
μ_C vs. μ_D			

Analisis Variansi Satu Arah

Contoh

Pembandingan ganda untuk varitas padi contoh di muka.

Diketahui (dari hitungan di muka):

	varitas			
	A	B	C	D
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33

$\text{MSE} = 26,71$

Pembandingan ganda Scheffé:

Pembandingan	$\mid \bar{X}_i - \bar{X}_j \mid$	$\sqrt{248,403 \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}$	Kesimpulan
μ_A vs. μ_B	0,67		
μ_A vs. μ_C	1,17		
μ_A vs. μ_D	8,66		
μ_B vs. μ_C	0,50		
μ_B vs. μ_D	9,33		
μ_C vs. μ_D	9,83		

Analisis Variansi Satu Arah

Contoh

Pembandingan ganda untuk varitas padi contoh di muka.

Diketahui (dari hitungan di muka):

	varitas			
	A	B	C	D
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33

$\text{MSE} = 26,71$

Pembandingan ganda Scheffé:

Pembandingan	$\mid \bar{X}_i - \bar{X}_j \mid$	$\sqrt{248,403\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$	Kesimpulan
μ_A vs. μ_B	0,67	$\sqrt{248,403\left(\frac{1}{6}+\frac{1}{6}\right)}$	
μ_A vs. μ_C	1,17		
μ_A vs. μ_D	8,66		
μ_B vs. μ_C	0,50		
μ_B vs. μ_D	9,33		
μ_C vs. μ_D	9,83		

Analisis Variansi Satu Arah

Contoh

Pembandingan ganda untuk varitas padi contoh di muka.

Diketahui (dari hitungan di muka):

	varitas			
	A	B	C	D
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33

$$\text{MSE} = 26,71$$

Pembandingan ganda Scheffé:

Pembandingan	$\mid \bar{X}_i - \bar{X}_j \mid$	$\sqrt{248,403\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$	Kesimpulan
μ_A vs. μ_B	0,67	9,1	
μ_A vs. μ_C	1,17	9,1	
μ_A vs. μ_D	8,66	9,1	
μ_B vs. μ_C	0,50	9,1	
μ_B vs. μ_D	9,33	9,1	
μ_C vs. μ_D	9,83	9,1	

Analisis Variansi Satu Arah

Contoh

Pembandingan ganda untuk varitas padi contoh di muka.

Diketahui (dari hitungan di muka):

	varitas			
	A	B	C	D
n_i	6	6	6	6
$\bar{X}_i$	19,67	19,00	18,50	28,33

$$\mathrm{MSE} = 26,71$$

Pembandingan ganda Scheffé:

Pembandingan	$\mid \bar{X}_i - \bar{X}_j \mid$	$\sqrt{248,403\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$	Kesimpulan
μ_A vs. μ_B	0,67	9,1	H_0 diterima
μ_A vs. μ_C	1,17	9,1	H_0 diterima
μ_A vs. μ_D	8,66	9,1	H_0 diterima
μ_B vs. μ_C	0,50	9,1	H_0 diterima
μ_B vs. μ_D	9,33	9,1	H_0 ditolak
μ_C vs. μ_D	9,83	9,1	H_0 ditolak

Regresi Linear Sederhana

Analisis Regresi digunakan untuk menyelidiki hubungan antara variabel dependen (respon) Y dengan variabel independen (variabel penjelas, prediktor) X.

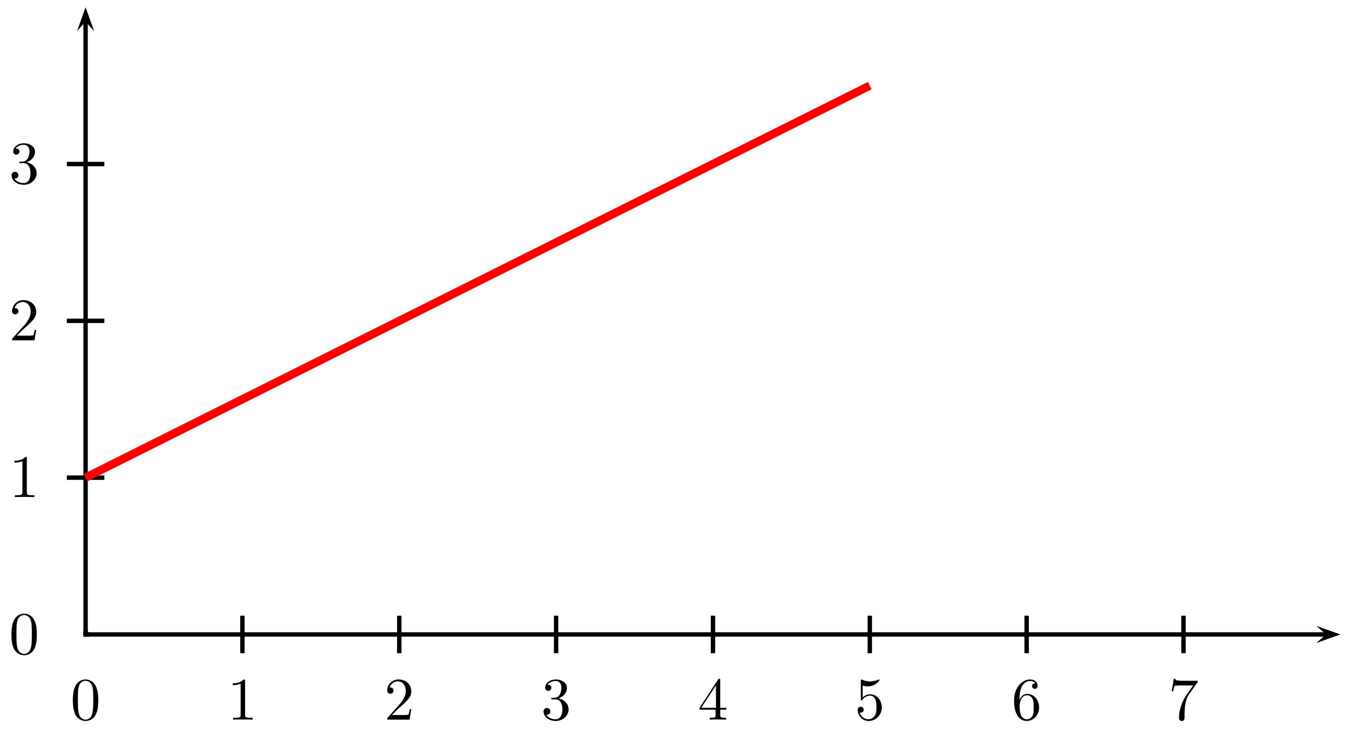
Hubungan antara Y dan X:

- fungsional
- statistik

Regresi Linear Sederhana

Contoh Hubungan Fungsional

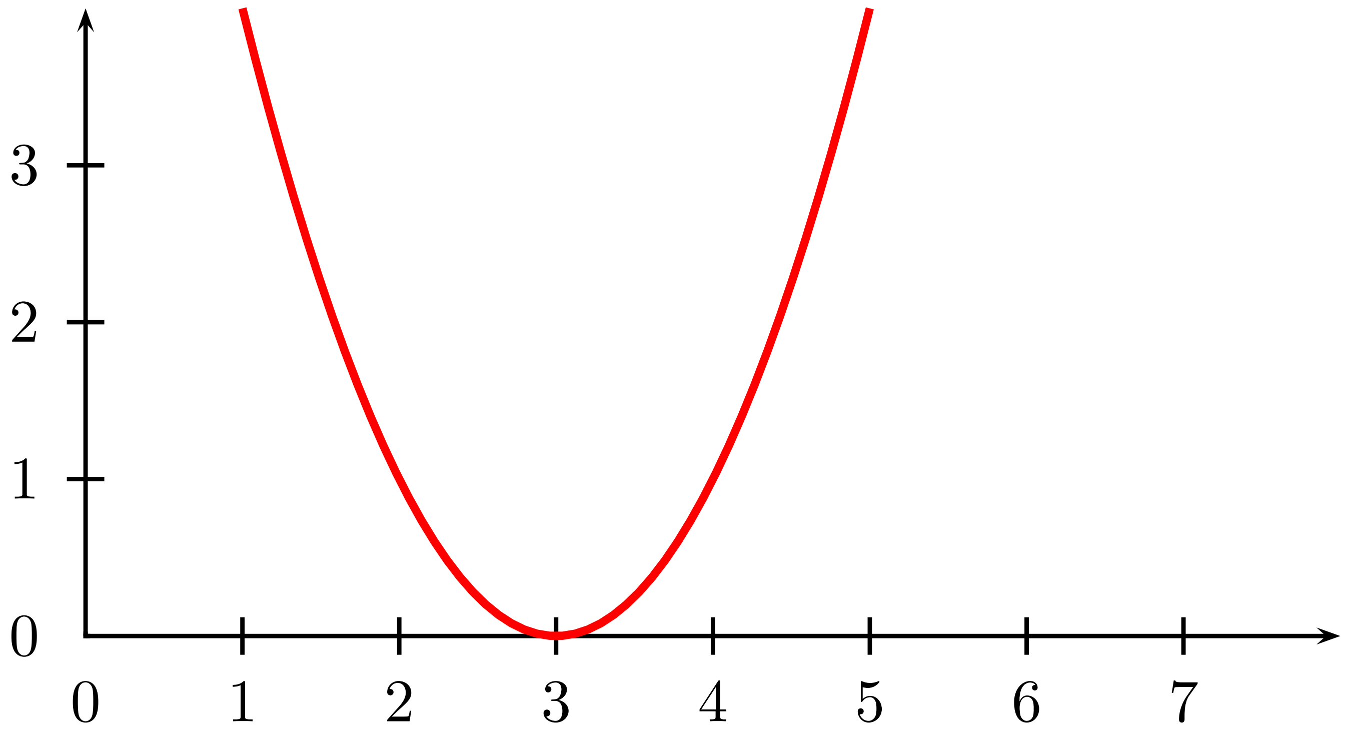
$$Y = 1 + \frac{1}{2}X$$



Regresi Linear Sederhana

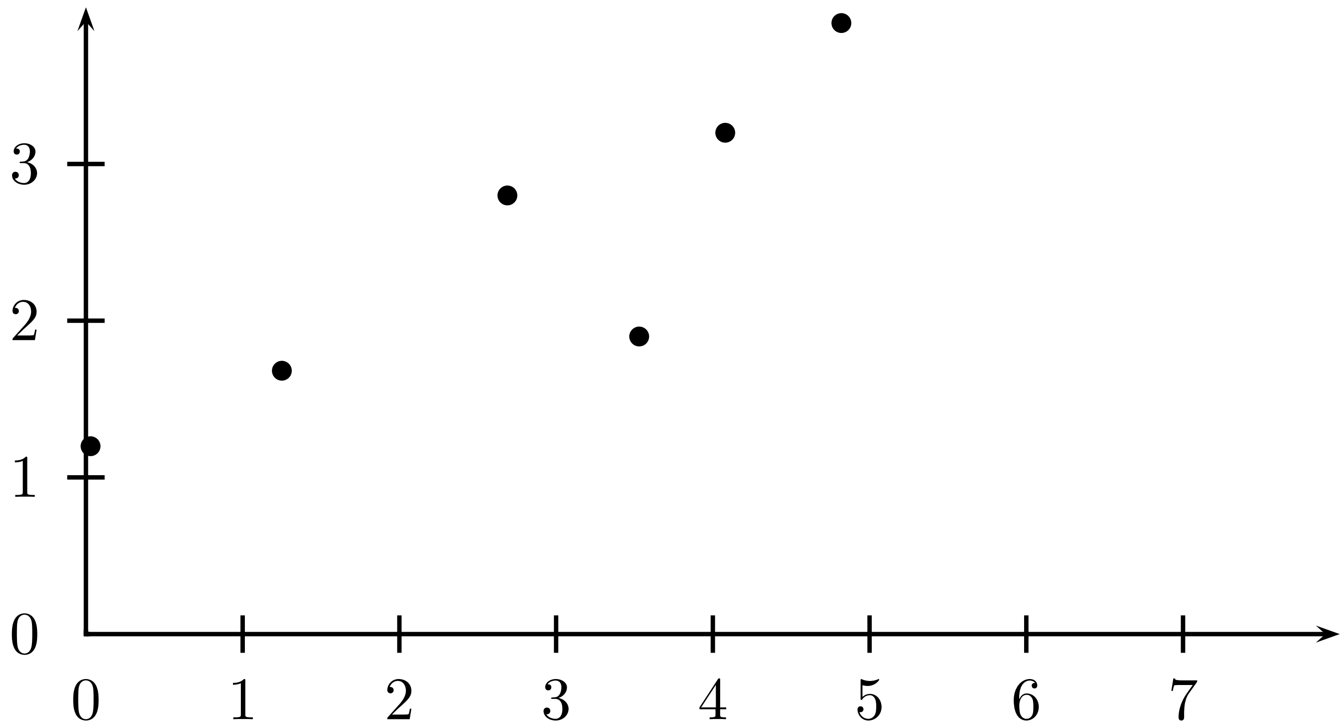
Contoh Hubungan Fungsional

$$Y = (X - 3)^2$$



Regresi Linear Sederhana

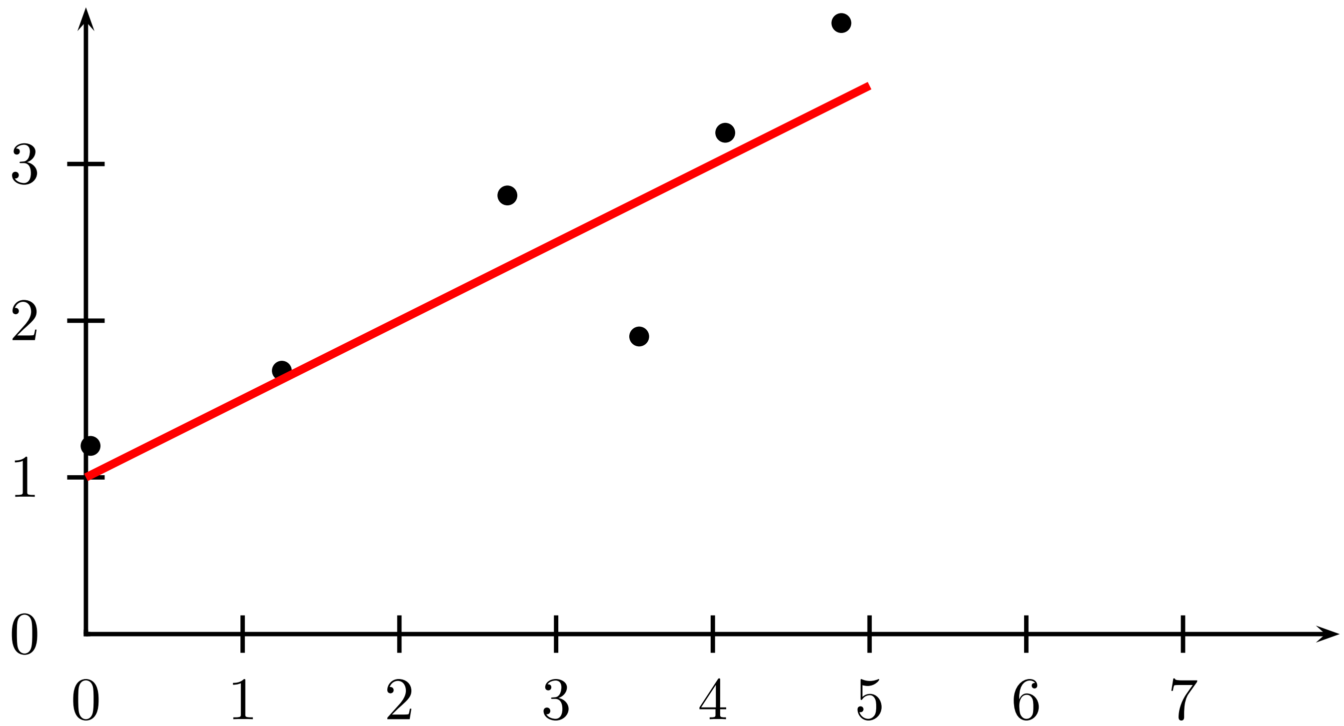
Contoh Hubungan Statistik



Regresi Linear Sederhana

Contoh Hubungan Statistik

$$Y = 1 + \frac{1}{2}X$$



Regresi Linear Sederhana

Contoh

Dipunyai data umur dan tinggi dari sampel 8 buah pohon jenis tertentu sbb.:

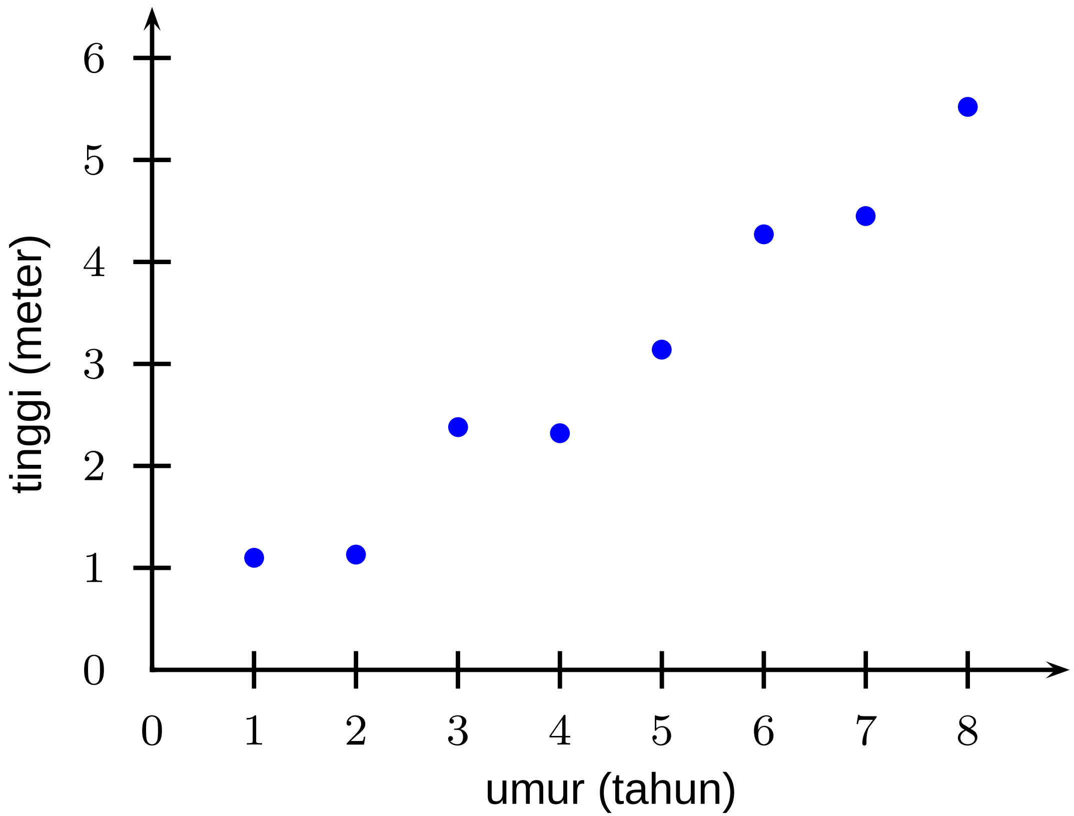
umur (tahun):	1	2	3	4	5	6	7	8
tinggi (meter):	1,10	1,13	2,38	2,32	3,14	4,27	4,45	5,52

Regresi Linear Sederhana

Contoh

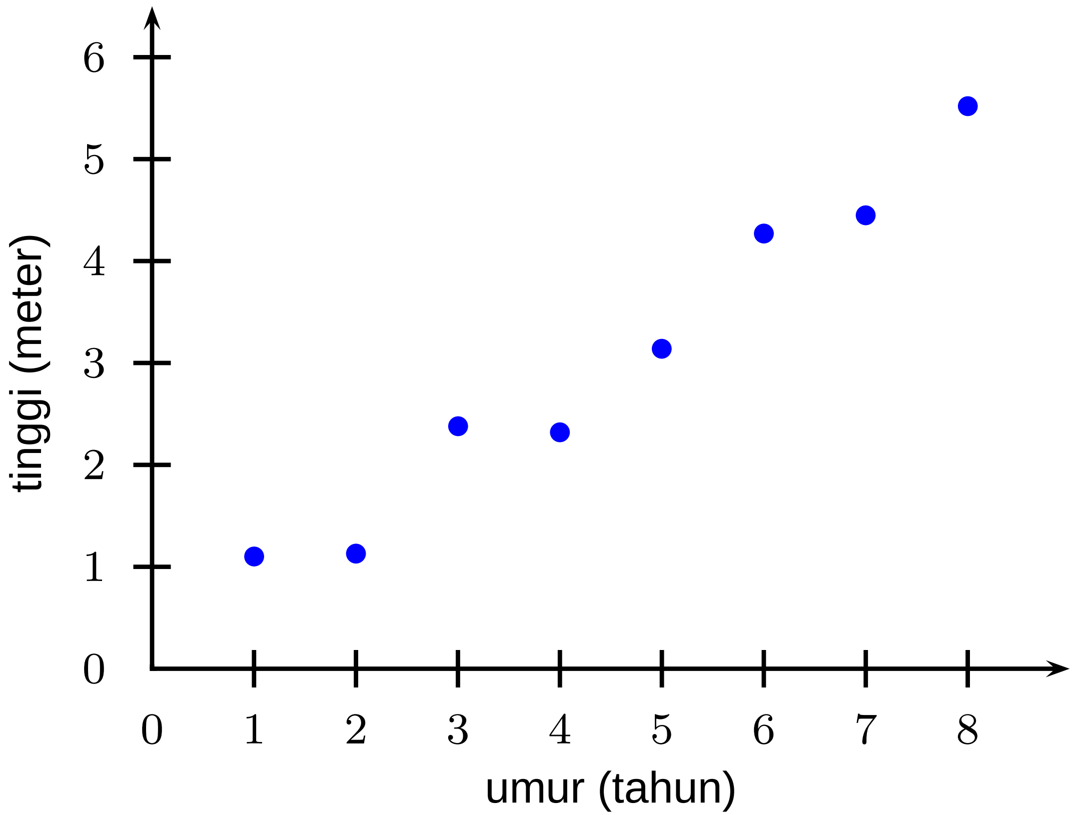
Dipunyai data umur dan tinggi dari sampel 8 buah pohon jenis tertentu sbb.:

umur (tahun):	1	2	3	4	5	6	7	8
tinggi (meter):	1,10	1,13	2,38	2,32	3,14	4,27	4,45	5,52



Regresi Linear Sederhana

Contoh
Akan dicari garis linear $\hat{y} = a + bx$ yang paling "mewakili" hubungan antara x (umur) dan y (tinggi)



Regresi Linear Sederhana

Estimasi (Penduga) garis regresi

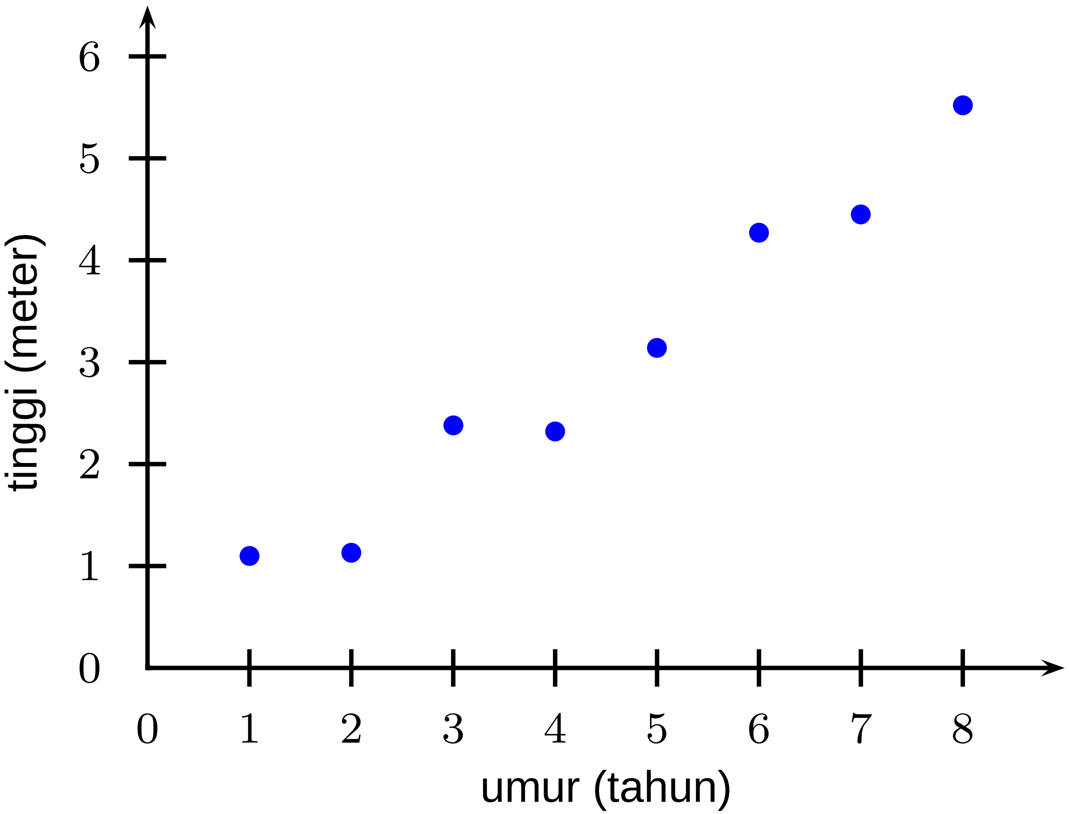
$$\hat{y} = a + bx$$

Menggunakan **Metode Kuadrat Terkecil (MKT)**
Prinsip MKT: Dicari nilai a dan b yang meminimalkan jumlah kuadrat residu (JKR), $\mathrm{JKR} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

Regresi Linear Sederhana

Contoh

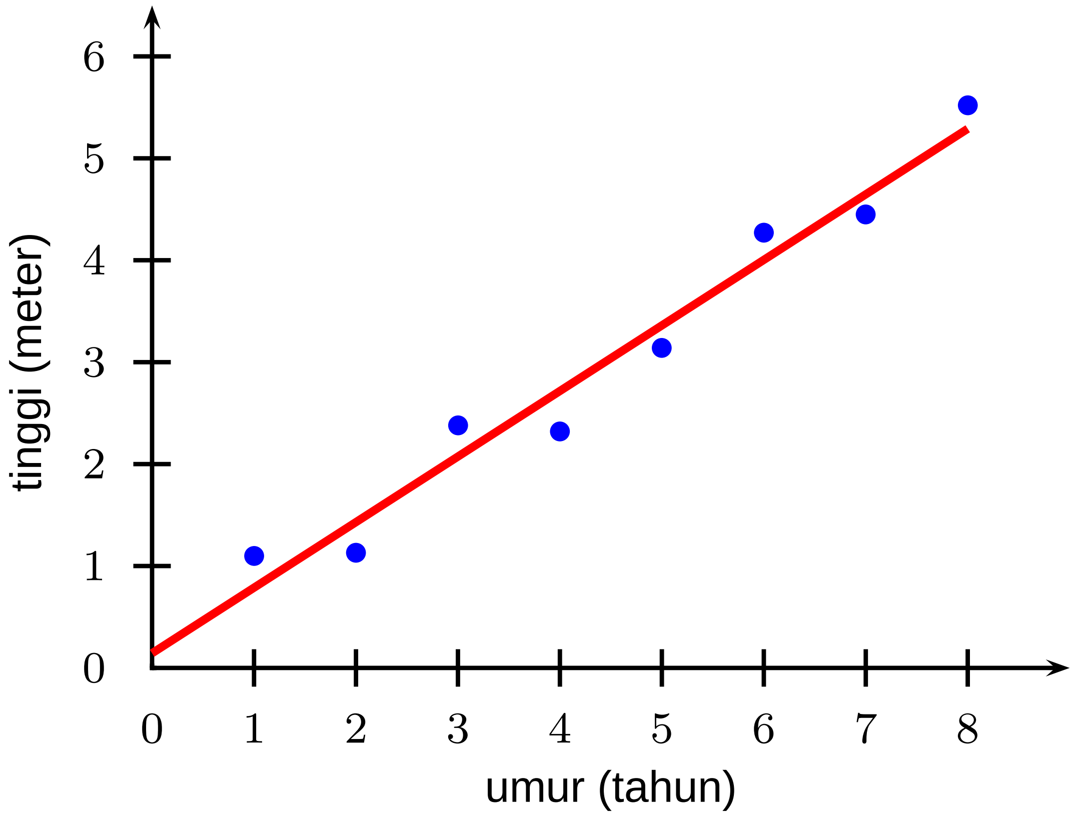
Data umur dan tinggi pohon



Regresi Linear Sederhana

Contoh

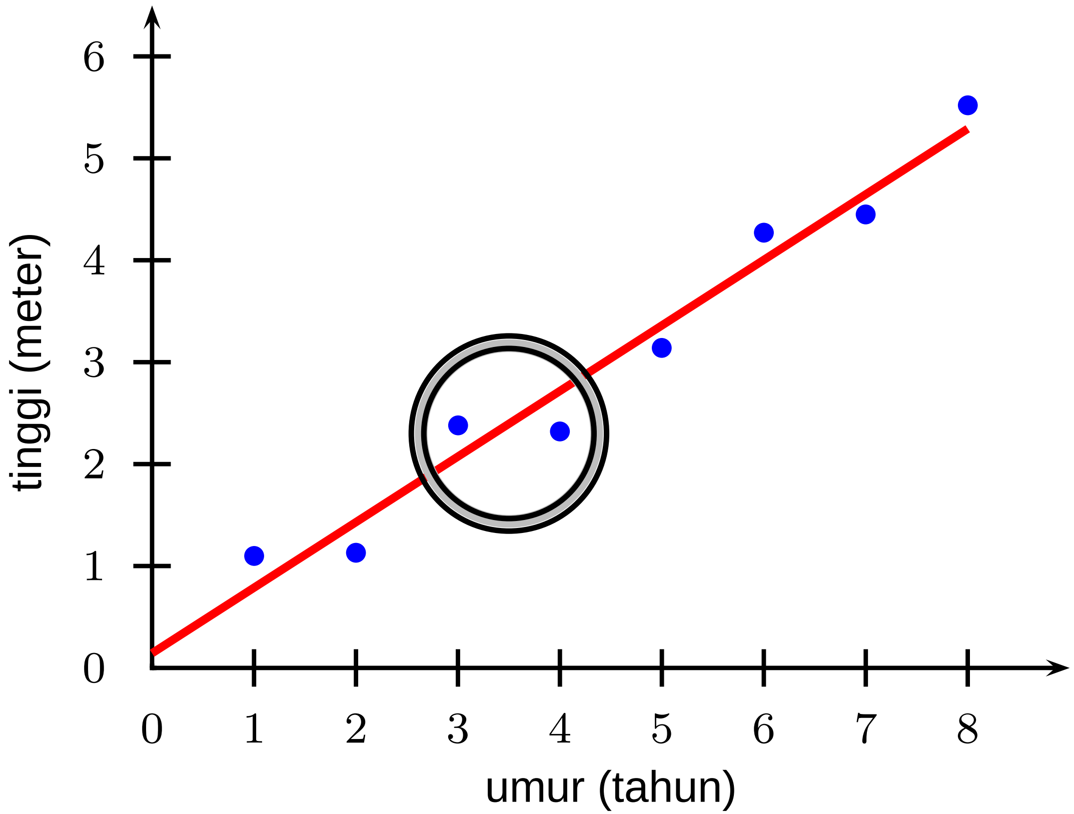
Data umur dan tinggi pohon



Regresi Linear Sederhana

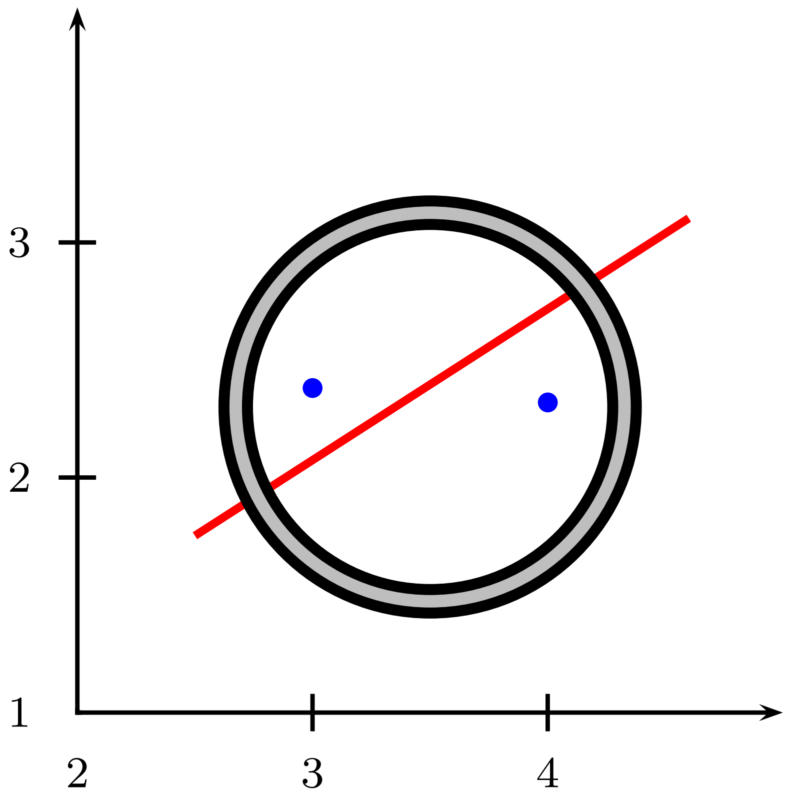
Contoh

Data umur dan tinggi pohon



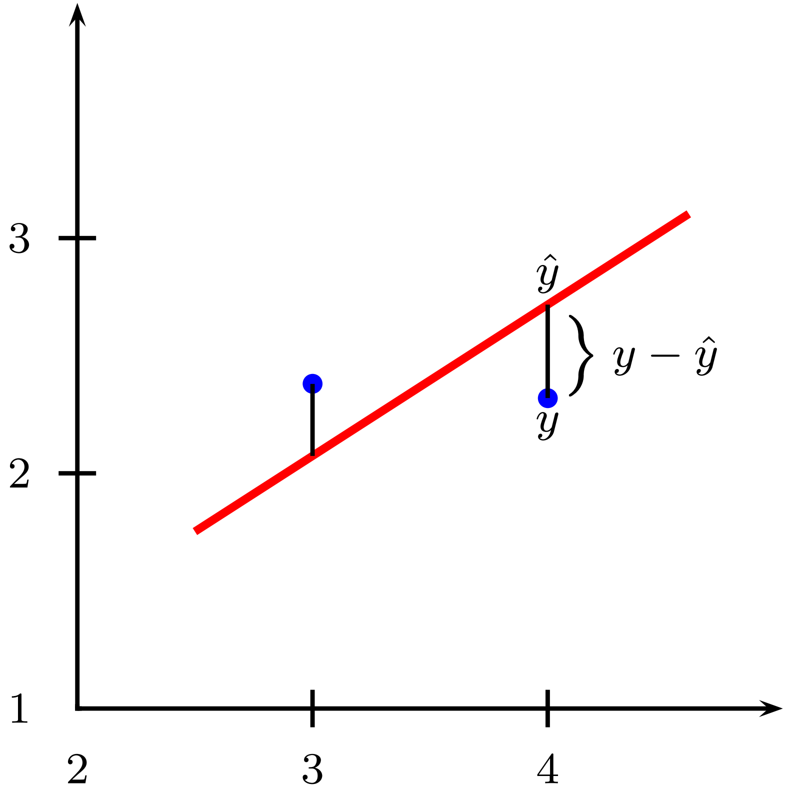
Regresi Linear Sederhana

Contoh
Data umur dan tinggi pohon



Regresi Linear Sederhana

Contoh
Data umur dan tinggi pohon



Regresi Linear Sederhana

Persamaan normal

Penyelesaian yang diperoleh dari MKT:

$$\begin{aligned} \sum_{i=1}^{n} y_i &= na + b \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i y_i &= a \sum_{i=1}^{n} x_i + b \sum_{i=1}^{n} x_i^2 \end{aligned}$$

Regresi Linear Sederhana

Persamaan normal

Penyelesaian yang diperoleh dari MKT:

$$
\begin{aligned}
a &= \frac{\sum y_i \sum x_i^2 - \sum x_i \sum x_i y_i}{n \sum x_i^2 - (\sum x_i)^2} \\
b &= \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}
\end{aligned}
$$

Jika b dihitung terlebih dahulu, a bisa dihitung dengan

$$
a = \frac{\sum y_i - b \sum x_i}{n} = \bar{y} - b\bar{x}
$$

Regresi Linear Sederhana

Persamaan normal

Penyelesaian yang diperoleh dari MKT:

$$a = \frac{(24,31)(204) - (36)(136,41)}{8(204) - (36)^2} = 0,1443$$

$$b = \frac{8(136,41) - (36)(24,31)}{8(204) - (36)^2} = 0,6432$$

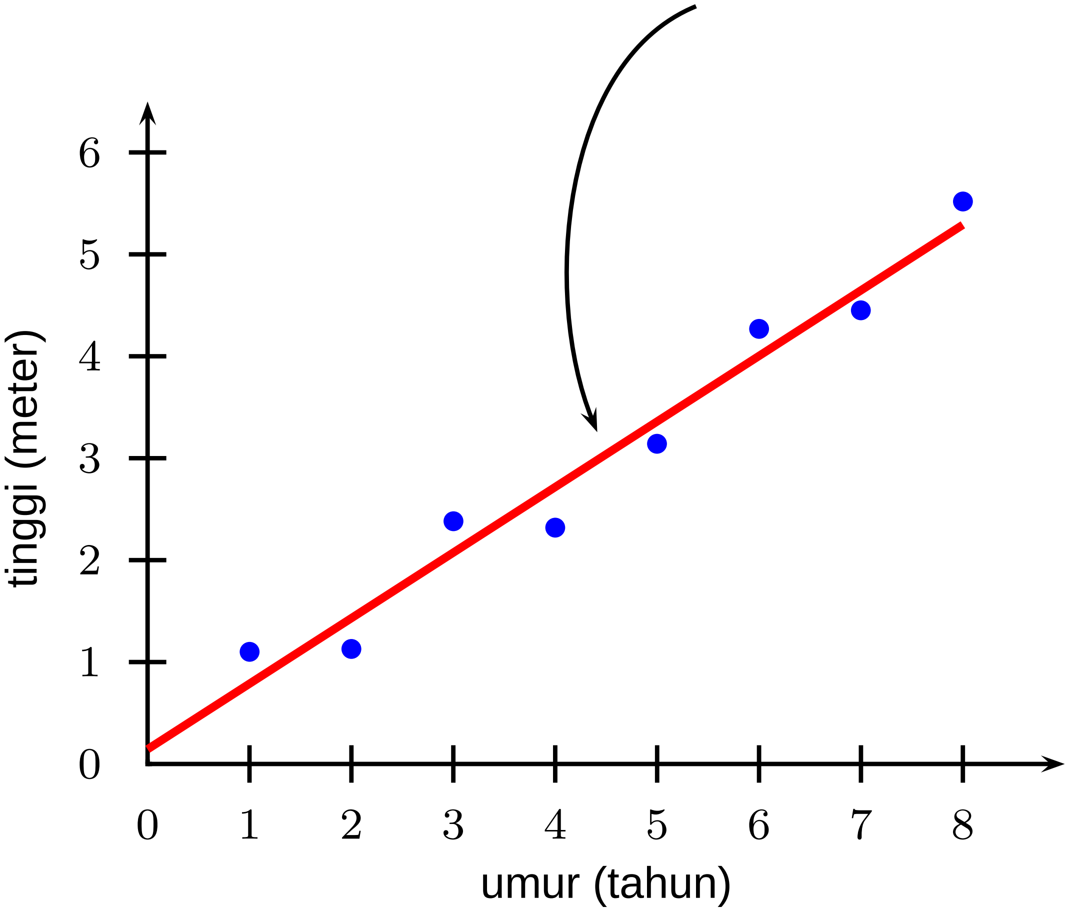
Jika $b = 0,6432$ dihitung terlebih dahulu, a bisa dihitung dengan

$$a = \frac{24,31 - (0,6432)(36)}{8} = 0,1443$$

Regresi Linear Sederhana

Contoh

Data umur dan tinggi pohon: $\hat{y} = 0,1443 + 0,6432x$



Regresi Linear Sederhana

Korelasi

Koefisien korelasi antara Y dan X adalah

$$\rho = \frac{\mathrm{Kov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}, \quad -1 \leq \rho \leq 1$$

dimana Y dan X dianggap variabel random berdistribusi bersama tertentu (lihat kembali bagian Probabilitas di muka)

Regresi Linear Sederhana

Korelasi

ρ menunjukkan tingkat hubungan linear antara kedua variabel.

Estimasi titik untuk ρ adalah

$$
\begin{aligned}
r &= \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 (y_i - \bar{y})^2}} \\
&= \frac{n \sum xy - \sum x \sum y}{\sqrt{n \sum x^2 - (\sum x)^2} \sqrt{n \sum y^2 - (\sum y)^2}}
\end{aligned}
$$

Regresi Linear Sederhana

Contoh Korelasi

Data umur dan tinggi pohon:

$$
\begin{aligned}
r &= \frac{n\sum xy - \sum x \sum y}{\sqrt{n\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}} \\
&= \frac{8(136,41) - (36)(24,31)}{\sqrt{8(204) - (36)^2}\sqrt{8(91,8991) - (24,31)^2}} \\
&= 0.982
\end{aligned}
$$

Regresi Linear Sederhana

<u>Koefisien Determinasi r^2</u>
Digunakan untuk mengukur keeratan hubungan linear antara x dan y dalam regresi linear

$$r^2 = 100\left(\frac{\mathrm{SSR}}{\mathrm{SST}}\right)\%$$

dengan
$\mathrm{SST} = \sum(y_i - \bar{y})^2$ (jumlah kuadrat total),
$\mathrm{SSE} = \sum(y - \hat{y})^2$ (jumlah kuadrat sesatan/residu/*error*),
$\mathrm{SSR} = \mathrm{SST} - \mathrm{SSE}$ (jumlah kuadrat regresi)

Menunjukkan berapa persen variasi dari y yang dapat diterangkan oleh x

Regresi Linear Sederhana

Contoh Koefisien Determinasi

Data umur dan tinggi pohon:

$\mathrm{SST} = \sum(y_i - \bar{y})^2 = 18,02709$

$\mathrm{SSE} = \sum(y - \hat{y})^2 = 0,6506536$

$\mathrm{SSR} = \mathrm{SST} - \mathrm{SSE} = 18,02709 - 0,6506536 = 17,37644$

$$\begin{aligned} r^2 &= 100\left(\frac{\mathrm{SSR}}{\mathrm{SST}}\right)\% \\ &= 100\left(\frac{17,37644}{18,02709}\right)\% \\ &= 96,39\% \end{aligned}$$

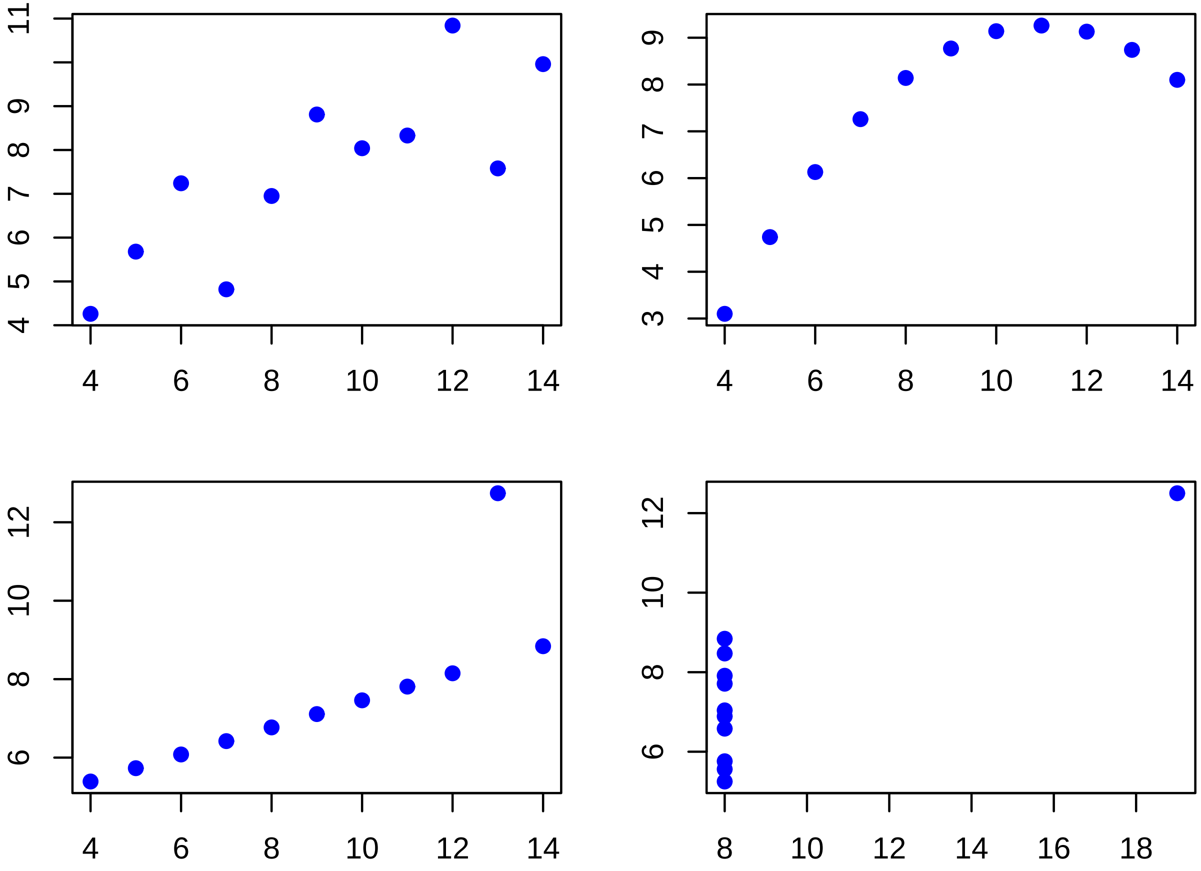
Untuk regresi sederhana, kuadrat dari koefisien determinasi sama dengan korelasi

Regresi Linear Sederhana

<u>Korelasi</u>

Korelasi hanya untuk hubungan linear

r=0,82



Regresi Linear Sederhana

Inferensi dalam Regresi

Diperoleh observasi berupa pasangan data (Y, X),
$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$

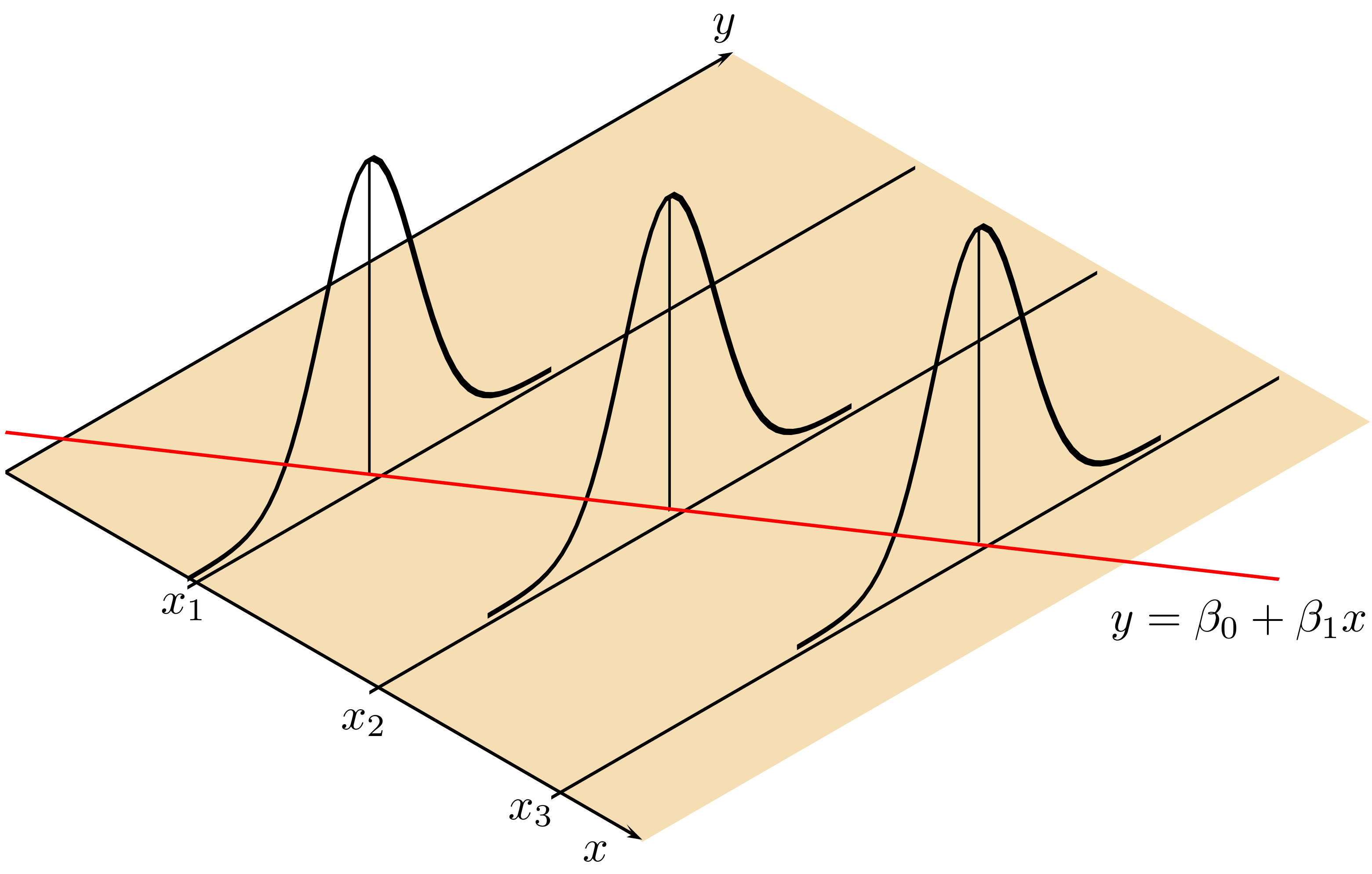
Model

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

Asumsi: suku eror ϵ_i independen dan berdistribusi normal $N(0, \sigma^2)$
Parameter β_0 dan β_1 dinamakan koefisien regresi.

Regresi Linear Sederhana

Inferensi dalam Regresi



Regresi Linear Sederhana

Inferensi dalam Regresi

- Penduga untuk β_0 dan β_1 adalah b_0 dan b_1:

$$
\begin{aligned}
b_1 &= \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \\
b_0 &= \frac{\sum y_i - b_1 \sum x_i}{n}
\end{aligned}
$$

- Penduga untuk σ^2

$$
\begin{aligned}
s^2 &= \sum_{i=1}^{n} \frac{(y_i - \hat{y})^2}{n-2} \\
&= \frac{\sum y_i^2 - b_0 \sum y_i - b_1 \sum x_i y_i}{n-2}
\end{aligned}
$$

Regresi Linear Sederhana

Inferensi Kemiringan Garis Regresi (β_1)

Statistik Penguji

$$t = \frac{b_1 - \beta_1}{s\sqrt{\frac{n}{n\sum x^2 - (\sum x)^2}}}$$

yang berdistribusi t dengan derajad bebas $n - 2$

Regresi Linear Sederhana

Inferensi Perpotongan Garis Regresi (β_0)

Statistik Penguji

$$t = \frac{b_0 - \beta_0}{s\sqrt{\frac{\sum x^2}{n\sum x^2 - (\sum x)^2}}}$$

yang berdistribusi t dengan derajad bebas $n-2$

Regresi Linear Sederhana

Contoh

Hitunglah interval konfidensi 95% untuk kemiringan garis regresi (β_1) dari data umur dan tinggi pohon di muka.

Telah dihitung $\beta_0 = 0,1443$ dan $\beta_1 = 0,6432$ (atau a dan b dalam notasi di muka)

$$
\begin{aligned}
s^2 &= \frac{\sum y_i^2 - b_0 \sum y_i - b_1 \sum x_i y_i}{n-2} \\
&= \frac{91,8991 - (0,1443)(24,31) - (0,6432)(136,41)}{6} \\
&= 0,1087092
\end{aligned}
$$

Interval konfidensi 95% untuk β_1: $\beta_1 \pm t_{(\alpha/2,n-2)} s \sqrt{\frac{n}{n\sum x^2 - (\sum x)^2}}$

$0,6432 \pm 0,0508$ atau $(0,592; 0,694)$

Regresi Linear Sederhana

Contoh

Hitunglah interval konfidensi 95% untuk perpotongan garis regresi (β_0) dari data umur dan tinggi pohon di muka.

Telah dihitung $\beta_0 = 0,1443$ dan $\beta_1 = 0,6432$ (atau a dan b dalam notasi di muka) serta $s^2 = 0,1087092$

Interval konfidensi 95% untuk β_0: $\beta_0 \pm t_{(\alpha/2,n-2)} s \sqrt{\frac{\sum x^2}{n\sum x^2 - (\sum x)^2}}$

$0,1443 \pm 0,2569$ atau $(-0,1126; 0,4012)$

Latihan

1. Dalam suatu eksperimen *plant breeding* dengan dua tipe bunga A dan B. Probabilitas terjadinya tipe A diharapkan lebih besar dari 7/16. Seorang ahli melakukan eksperimen dengan 100 kuntum bunga dan mendapatkan bahwa separuhnya adalah tipe A, dengan menggunakan $\alpha = 0{,}01$; kesimpulan apakah yang dapat kita tarik?

Latihan

2. Suatu jenis tikus tertentu yang mendpatkan makanan biasa menunjukkan kenaikan rata-rata 65 gram selama tiga bulan pertama dari hidupnya. Suatu sampel random dengan 40 ekor tikus seperti itu diberi makanan dengan protein tinggi dan menunjukkan kenaikan berat rata-rata 82 gram dengan deviasi standar 17,6 gram selama tiga bulan pertama hidupnya. Apakah fakta cukup mendukung dugaan bahwa makanan yang berprotein tinggi akan memperbesar kenaikan berat tikus?

Latihan

3. Suatu perusahaan alat elektronik ingin menguji dua macam kualitas hasil produksinya. Untuk ini diadakan percobaan-percobaan dan diperoleh hasil sebagai berikut: 10 produk kualitas A mempunyai tahan hidup rata-rata 2600 jam dengan deviasi standar 300 jam. Sedangkan 15 produk kualitas B mempunyai tahan hidup rata-rata 2400 jam dengan deviasi standar 250 jam. Berdasarkan hasil percobaan di atas, apakah kita percaya bahwa kedua kualitas produk elektronik itu berbeda tahan hidupnya? (Anggap distribusi kedua populasi normal dengan variansi sama).

Latihan

4. Seorang Zoologist ingin menggunakan tikus yang berat waktu lahirnya mempunyai variabilitas yang rendah. Tersedia dua jenis tikus yang berbeda. Dia mengambil sampel random dengan 10 jenis pertama dan 16 jenis kedua. Diperoleh $S_1^2 = 0{,}36$ gram dan $S_2^2 = 0{,}87$ gram. Apakah variabilitas dua jenis tersebut berbeda? ($\alpha = 0{,}02$)

Latihan

5. Suatu stimulan akan diuji akibatnya terhadap tekanan darah. Dua belas orang laki-laki diambil secara random dari laki-laki dalam kelompok umur 30 - 40 tahun. Tekanan darah mereka diukur sebelum dan sesudah diberi stimulan. Hasilnya adalah sbb.:

   Tekanan darah sebelum dan sesudah (mmHg)

orang ke-	1	2	3	4	5	6	7	8	9	10	11	12
sebelum	120	124	130	118	140	128	140	135	126	130	126	127
sesudah	128	130	131	127	132	125	141	137	118	134	129	130

   Hitung interval konfidensi 95% untuk rata-rata selisih tekanan darah sesudah dan sebelum stimulan untuk semua orang laki-laki dalam kelompok 30 - 40 tahun.

Latihan

6. Ada hipotesis yang menyatakan bahwa untuk sepasang bayi kembar, berat badan bayi yang lahir kemudian lebih berat dari bayi yang lahir sebelumnya. Apabila kita ingin menguji pernyataan tersebut, uji statistik apa yang digunakan?

Latihan

7. Suatu survei menyatakan bahwa dalam suatu daerah tertentu 20 % rumah tangga berada di bawah garis kemiskinan. Suatu program pengentasan kemiskinan dilaksanakan pada daerah tersebut. Untuk mengetahui apakah program tersebut berhasil, sampel sebesar 400 rumah tangga diambil dari daerah tersebut, 68 rumah tangga dinyatakan berada di bawah garis kemiskinan. Berhasilkah program ini ? ($\alpha = 0.05$)

Latihan

8. Sebuah program *diet* untuk mengurangi berat badan diterapkan pada 12 pria dan 14 wanita. Diperoleh hasilnya sebagai berikut (dalam kg):

Wanita	X_1	109	135	88	118	132	154	121	146	129	94	104	116	136	142
	X_2	85	105	54	85	105	123	98	115	97	64	69	89	115	106
Pria	Y_1	137	127	106	127	122	109	121	115	93	118	139	113		
	Y_2	118	99	79	109	99	83	105	98	75	95	117	92		

(X_1, X_2 adalah berat wanita sebelum dan sesudah melakukan diet ; Y_1, Y_2 adalah berat pria sebelum dan sesudah melakukan *diet*).

a. Apakah program *diet* tersebut berhasil secara umum (tanpa memandang pria atau wanita)? ($\alpha = 0,05$)

b. Bila ingin diketahui program *diet* tersebut lebih baik untuk wanita atau pria, inferensi statistik apakah yang dapat digunakan?

Latihan

9. Dari sampel random $n = 25$ bola lampu, diperoleh tahan hidup rata-rata 1,85 tahun dan standar deviasi 0,5 tahun.

   a. Hitunglah interval konfidensi 95% untuk rata-rata tahan hidup bola lampu

   b. Hitunglah interval konfidensi 90% untuk variansi tahan hidup bola lampu

   c. Apabila $n = 64$, hitunglah interval konfidensi 90% untuk variansi tahan hidup bola lampu

Latihan

10. Ingin diketahui *mean* berapa lama seorang mahasiswa melakukan *chatting* di internet. Untuk itu itu akan dilakukan survei di beberapa warung internet di kampus. Penelitian pendahuluan menunjukkan bahwa standar deviasi dari lama *chatting* adalah 67 menit dan berdistribusi Normal. Bila kesalahan estimasi interval survei ini tidak boleh lebih dari 10 menit dengan tingkat konfidensi 95%, berapa ukuran sampel yang harus digunakan?

Latihan

11. Ingin diketahui apakah suatu metode pembiakan tanaman pisang yang menggunakan cara modern menghasilkan pisang dengan berat yang lebih besar daripada pisang yang dikembangkan dengan cara tradisional. Diperoleh informasi sebagai berikut:

Jenis pisang	cara tradisional	cara modern
banyak sampel	100	120
rata-rata pertandan	4,2 kg	4,8 kg
deviasi standar	1,2 kg	0,9 kg

a. Ujilah apakah terdapat perbedaan nyata dari hasil kedua metode pembiakan tersebut ($\alpha = 3\%$). Anggap kedua variansinya sama.

b. Jika variansinya tidak diketahui apakah sama atau tidak, uji apakah yang dapat saudara gunakan untuk menguji kesamaan dua variansi? Dengan menganggap kedua populasi berdistribusi normal, tulislah hipotesis dan uji statistiknya

Latihan

12. Apakah cukup bukti yang menyatakan bahwa lebih dari Is there sufficient evidence at 1% level (α=0.01) that more than 30% mahasiswa baru gagal memenuhi standar pengethauna dan pemahaman matematika jika 60 mahasiswa baru dari sampel 130 mahasiswa gagal memenuhi standar?

Latihan

13. Suatu alat pengukur tekanan darah elektronik akan diuji ketepatan hasil pengukurannya. Bila hasil pengukuran tekanan darah sama atau mendekati hasil pengukuran alat ukur standar maka alat pengukur elektronik ini dinyatakan dapat dipakai. Dari 15 orang yang terpilih sebagai sampel dilakukan dua kali pengukuran masing-masing dengan alat ukur tekanan darah standar dan dengan alat ukur elektronik. Diperoleh hasil pengukuran tekanan darah diastolik sebagai berikut:

alat standar	68	82	94	106	92	80	76	74	119	93	86	65	74	84	100
alat elektronik	72	84	89	100	97	88	84	70	103	84	86	63	69	87	93

Apakah alat pengukur tekanan darah elektronik ini dapat dipakai? (α=0,05)

Latihan

14. Diketahui data gizi dan berat badan 50 anak usia 4-5 tahun di suatu desa seperti pada tabel berikut

n	status gizi	berat badan (kg)	
		rata-rata	deviasi std.
35	baik	13,5	2,5
15	buruk	7,5	1,5

a. Hitunglah interval konfidensi 95% untuk proporsi anak dengan gizi buruk!

b. Hitunglah interval konfidensi 95% untuk mean berat anak dengan status gizi baik!

c. Statistik penguji apakah yang dapat digunakan untuk inferensi mean berat anak dengan status gizi buruk? Jelaskan!

Latihan

15. Dengan menggunakan tabel Normal standar hitunglah:

   a. $P(-2 \leq Z \leq 1.5)$

   b. $P(Z \geq 1)$

   c. k, jika diketahui $P(0 \leq Z \leq k) = 0{,}4236$

   d. $P(\mu - \sigma \leq X \leq \mu + \sigma)$

   e. k yang memenuhi $P(X \leq k) = 0{,}05$

dengan Z adalah variabel random normal standar dan X adalah variabel random dengan mean μ dan variansi σ^2

Latihan

16. Suatu desain mobil diperkirakan akan menurunkan konsumsi bahan bakar sekaligus variabilitasnya. Sampel random dengan 16 mobil biasa diperoleh deviasi standar untuk konsumsi bahan bakar (liter per 100 km) sebesar 3,1. Sedangkan sampel random dengan 12 mobil desain ini diperoleh deviasi standar 1,8. Dengan asumsi sampel berasal dari distribusi normal ujilah bahwa desain mobil baru tersebut memang dapat menurunkan variabilitas konsumsi bahan bakar ($\alpha = 0{,}05$).