



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



**Kampus
Merdeka**
INDONESIA JAYA

**MERDEKA
BELAJAR**

DATA SCIENCE DARMAJAYA
“YOUR BEST FUTURE IN DATA”

PERTEMUAN KE: [18]

[STATISTIKA UNTUK] SAINS DATA

OLEH: [EGI SAFITRI, S.MAT., M.SI]

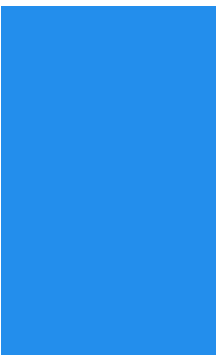


Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



**Kampus
Merdeka**
INDONESIA JAYA

**MERDEKA
BELAJAR**



Hipotesis



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



**Kampus
Merdeka**
INDONESIA JAYA

**MERDEKA
BELAJAR**

- Suatu pernyataan yang masih lemah kebenarannya dan perlu dibuktikan/ dugaan yang sifatnya masih sementara
- Hipotesis ini perlu untuk diuji untuk kemudian diterima/ ditolak



PENGUJIAN HIPOTESIS

Pengujian hipotesis statistik ialah prosedur yang memungkinkan keputusan dapat dibuat, yaitu keputusan untuk menolak atau menerima hipotesis yang sedang dipersoalkan atau diuji.

Hipotesis yang berupa anggapan didasarkan atas

Teori

Sebuah hipotesis dapat didasarkan pada teori-teori yang telah ada sebelumnya. Teori adalah suatu konsep atau rangkaian konsep yang telah dijelaskan dan diuji secara ilmiah. Hipotesis yang berasal dari teori seringkali mencoba menguji atau melengkapi pemahaman yang telah ada dalam teori tersebut.

Pengalaman

Pengalaman pribadi atau pengalaman kolektif juga dapat menjadi dasar untuk merumuskan hipotesis. Pengamat mungkin memiliki pengetahuan awal atau insight dari pengalaman sehari-hari yang mengarah pada suatu dugaan atau hipotesis tertentu.

Ketajaman berpikir

Sebuah hipotesis dapat muncul dari suatu kerangka berpikir atau model konseptual yang dihasilkan melalui pemikiran kritis. Kerangka berpikir ini bisa mencakup hubungan antar variabel atau konsep yang kemudian diuji dalam penelitian.

JENIS KESALAHAN

1. Kesalahan jenis 1.

Menolak hipotesis nol, yang seharusnya diterima

2. Kesalahan jenis 2.

Menerima hipotesis nol yang seharusnya ditolak

Keputusan- Situasi	H_0 Benar	H_0 Salah
Terima H_0	Keputusan tepat ($1-\alpha$)	Kesalahan Jenis II (β)
Tolak H_0	Kesalahan jenis I (α)	Keputusan Tepat ($1-\beta$)

DUA TIPE HIPOTESIS



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



**Kampus
Merdeka**
INDONESIA JAYA

**MERDEKA
BELAJAR**

- ▶ **Hipotesis korelatif** yaitu pernyataan tentang ada atau tidak adanya hubungan antara dua variabel atau lebih
- ▶ **Hipotesis komparatif** yaitu pernyataan tentang ada atau tidak adanya perbedaan antara dua kelompok atau lebih

PROSEDUR PENGUJIAN HIPOTESIS

- ▶ Menentukan formulasi hipotesis
- ▶ Menentukan taraf nyata (significant level)
- ▶ Menentukan kriteria pengujian
- ▶ Menentukan nilai uji statistik
- ▶ Membuat kesimpulan



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



**Kampus
Merdeka**
INDONESIA JAYA

**MERDEKA
BELAJAR**

PERUMUSAN HIPOTESIS



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



**Kampus
Merdeka**
INDONESIA JAYA

**MERDEKA
BELAJAR**

- Dinyatakan sebagai kalimat pernyataan (deklaratif)
- Melibatkan minimal dua variabel penelitian
- Mengandung suatu prediksi
- Harus dapat diuji (testable)

MENENTUKAN FORMULASI HIPOTESIS



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



Kampus
Merdeka
INDONESIA JAYA

**MERDEKA
BELAJAR**

Dibedakan 2 jenis :

- 1. Hipotesis nol** : suatu pernyataan yang akan diuji, hipotesis tersebut tidak memiliki perbedaan/perbedaannya nol dengan hipotesis sebenarnya.
- 2. Hipotesis alternatif** : segala hipotesis yang berbeda dengan hipotesis nol. Pemilihan hipotesis ini tergantung dari sifat masalah yang dihadapi

MENENTUKAN FORMULASI HIPOTESIS



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



Kampus
Merdeka
INDONESIA JAYA

MERDEKA
BELAJAR

- $H_0 : \mu = \mu_0$ dengan beberapa kemungkinan H_a
- $H_a : \mu < \mu_0$; $\mu > \mu_0$; ataukah $\mu \neq \mu_0$
satu sisi satu sisi dua sisi

MENENTUKAN FORMULASI HIPOTESIS



► Berdasarkan informasi yang dikemukakan pada sebuah media massa, bahwa harga beras jenis "A" di suatu wilayah adalah Rp. 3.200,-

(Pengujian Dua Pihak)

$$H_0 : \mu = \text{Rp. 3.200,-}$$

$$H_a : \mu \neq \text{Rp. 3.200,-}$$

► Berdasarkan informasi bahwa harga beras jenis "A" di suatu wilayah tidak kurang dari Rp. 3.200,- (Pengujian Satu Pihak – Kiri)

$$H_0 : \mu \geq \text{Rp. 3.200,-}$$

$$H_a : \mu < \text{Rp. 3.200,-}$$

MENENTUKAN FORMULASI HIPOTESIS



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



Kampus
Merdeka
INDONESIA JAYA

MERDEKA
BELAJAR

- Berdasarkan informasi bahwa harga beras jenis "A" di suatu wilayah tidak lebih dari Rp. 3.200,- (Pengujian Satu Pihak – Kanan)

$H_0 : \mu \leq \text{Rp. 3.200,-}$

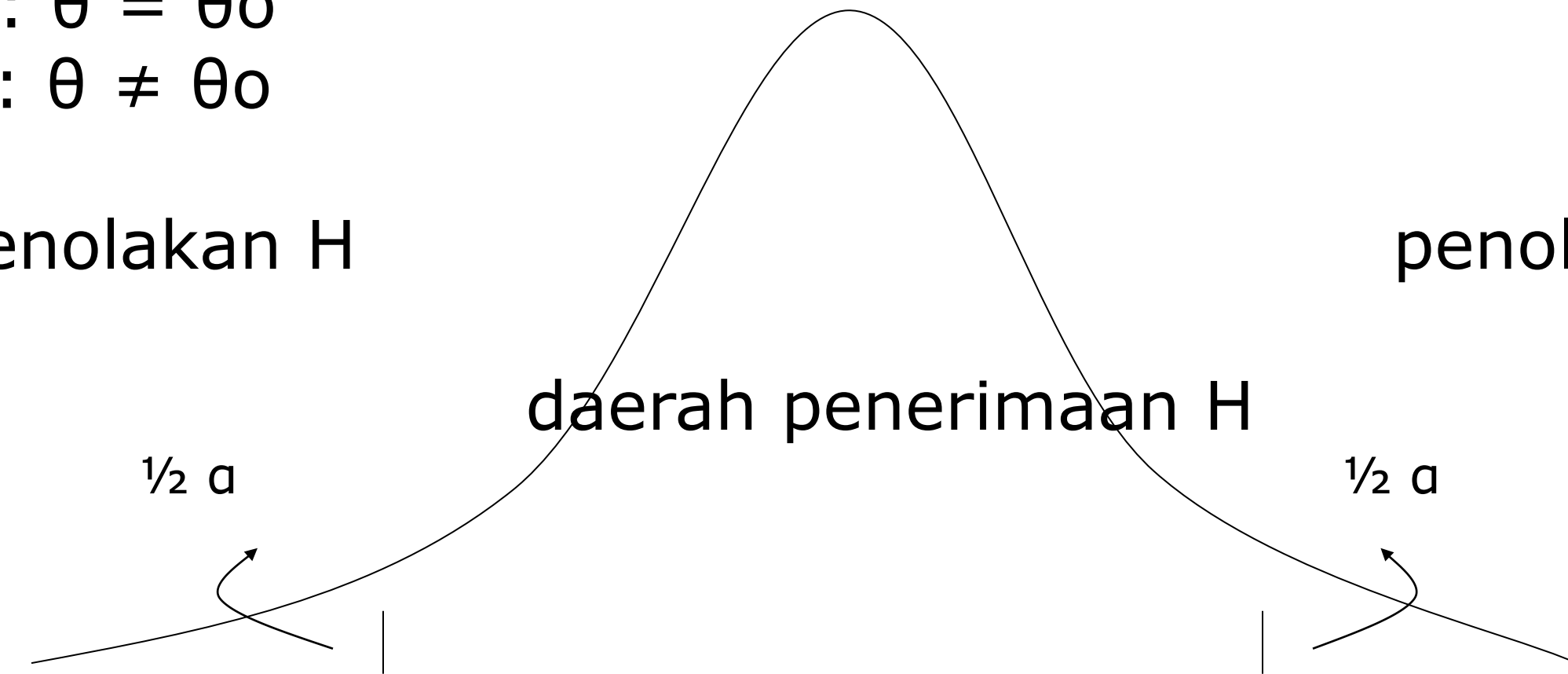
$H_a : \mu > \text{Rp. 3.200,-}$

UJI DUA PIHAK

- $H: \theta = \theta_0$
- $A: \theta \neq \theta_0$

penolakan H

penolakan H



Hipotesis H diterima jika: $-Z_{1/2(1-\alpha)} < Z < Z_{1/2(1-\alpha)}$

PENGUJIAN HIPOTESIS SATU RATA-RATA



Perumusan hipotesis

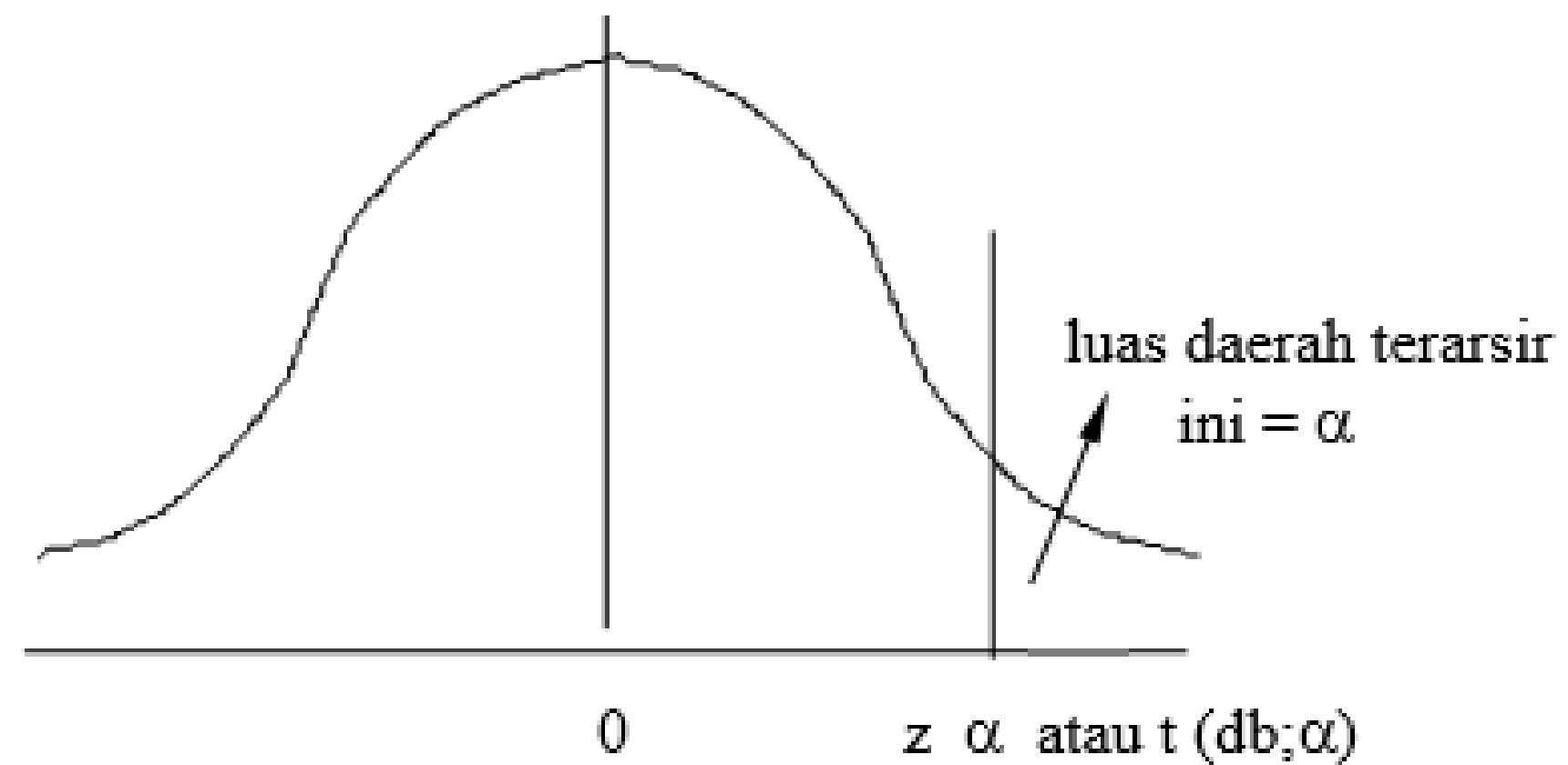
Menetapkan taraf nyata dan menghitung nilai Z_{tabel} atau t_{tabel}

Menghitung nilai z_0 atau t_0

Keputusan dan kesimpulan

$$H_0: \mu = \mu_0$$
$$H_0: \mu > \mu_0$$

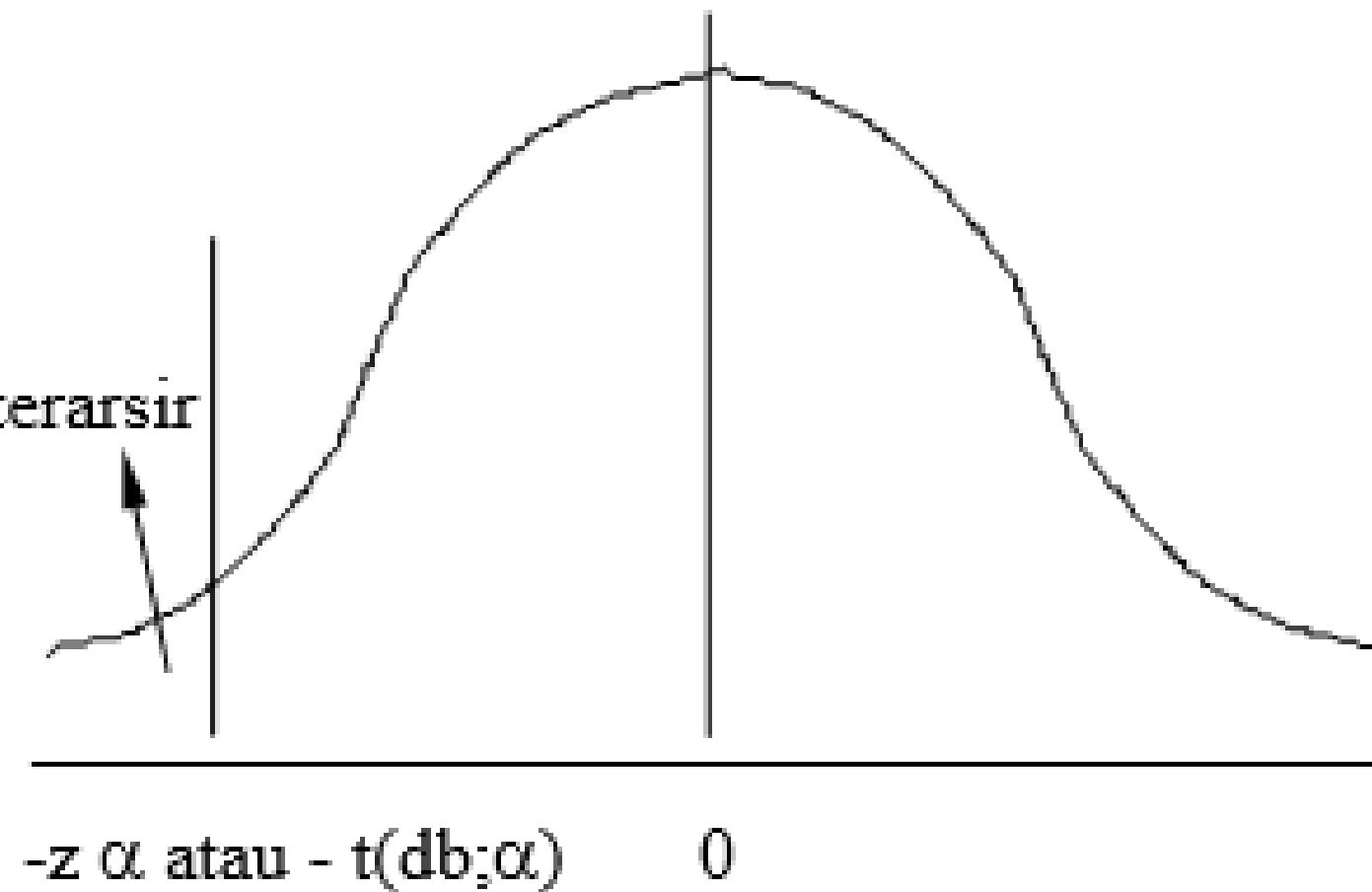
Uji satu arah kanan



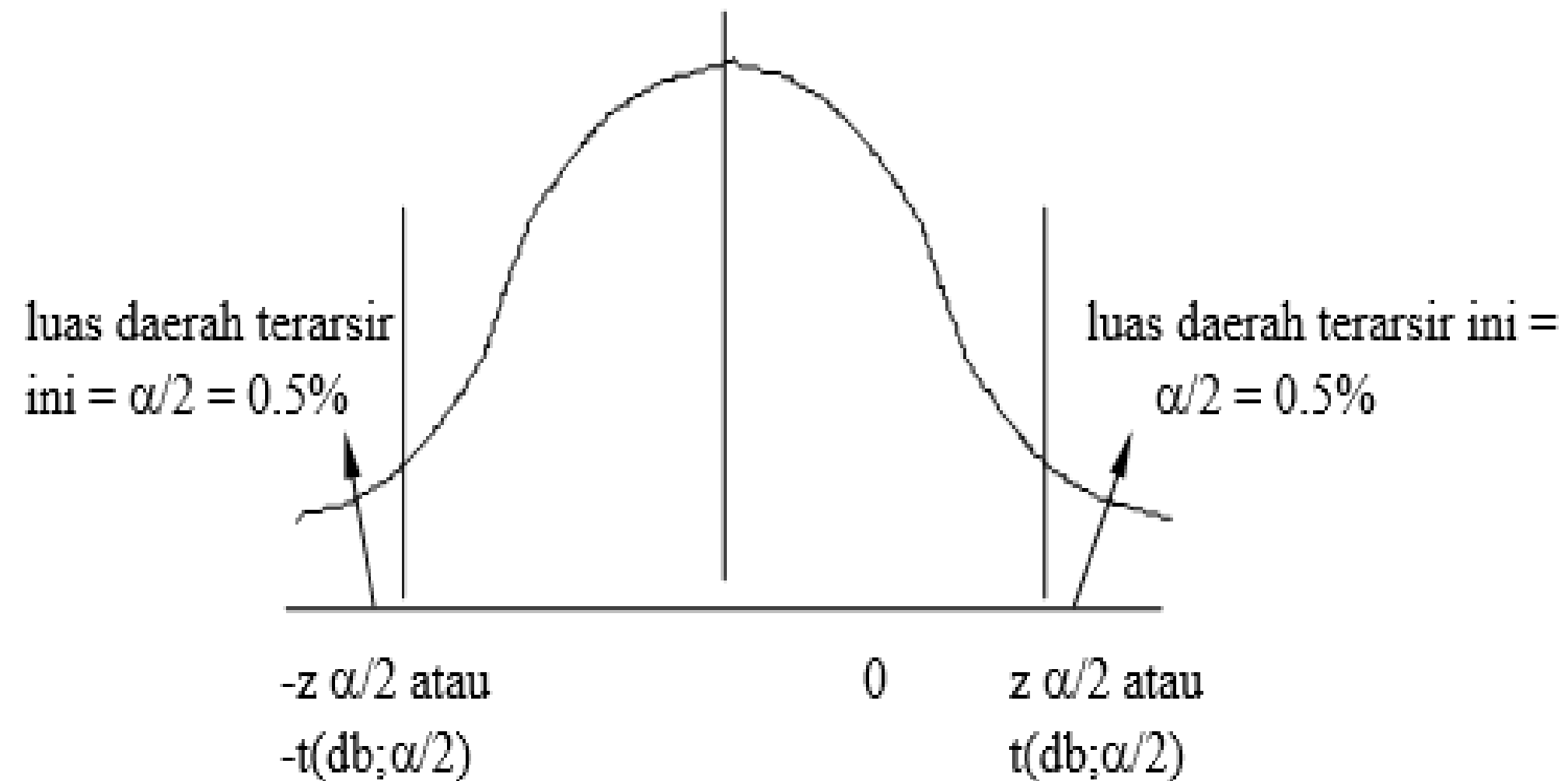
$$H_0: \mu = \mu_0$$
$$H_0: \mu < \mu_0$$

Uji satu arah kiri

luas daerah terarsir
ini = α



$H_0: \mu = \mu_0$
 $H_0: \mu \neq \mu_0$ Uji dua arah



MENENTUKAN TARAF NYATA



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



**Kampus
Merdeka**
INDONESIA JAYA

**MERDEKA
BELAJAR**

- Besarnya batas toleransi dalam menerima kesalahan hasil hipotesis terhadap nilai parameter populasinya
- Besarnya taraf nyata bergantung pada keberanian pembuat keputusan yang dalam hal ini berapa besarnya kesalahan yang akan ditolerir
- Besarnya kesalahan tersebut disebut sebagai daerah kritis pengujian/ daerah penolakan

MENENTUKAN KRITERIA PENGUJIAN



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



**Kampus
Merdeka**
INDONESIA JAYA

**MERDEKA
BELAJAR**

- Bentuk pembuatan keputusan dalam menerima/ menolak hipotesis nol dengan cara membandingkan nilai α tabel distribusinya dgn nilai statistiknya sesuai dgn btk pengujiannya
- Penerimaan H_0 : nilai uji statistiknya berada di luar nilai kritis
- Penolakan H_0 : nilai uji statistiknya berada dalam nilai kritis

MENGHITUNG NILAI UJI STATISTIK



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



Kampus
Merdeka
INDONESIA JAYA

**MERDEKA
BELAJAR**

Untuk $n > 30$

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

Untuk $n \leq 30$

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

Dimana:

n = banyaknya elemen sample ($n > 30$)

$$\bar{X} = \frac{1}{n} \sum X_i$$

σ_x = kesalahan baku $\bar{X} = \sigma / \sqrt{n}$

μ_0 = nilai μ sesuai dengan H_0

MENGHITUNG NILAI UJI STATISTIK



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



Kampus
Merdeka
INDONESIA JAYA

**MERDEKA
BELAJAR**

Jika simpangan baku populasi diketahui,

$$Z_o = \frac{\bar{X} - \mu_o}{\sigma_{\bar{x}}} = \frac{\bar{X} - \mu_o}{\frac{\sigma}{\sqrt{n}}}$$

Jika simpangan baku populasi tidak diketahui,

$$Z_o = \frac{\bar{X} - \mu_o}{s_{\bar{x}}} = \frac{\bar{X} - \mu_o}{\frac{s}{\sqrt{n}}}$$

KEPUTUSAN DAN KESIMPULAN

- I. $H_0 : \mu = \mu_0$ apabila $Z_0 > Z_\alpha$, H_0 ditolak
 $H_0 : \mu > \mu_0$ apabila $Z_0 \leq Z_\alpha$, H_0 diterima
- II. $H_0 : \mu = \mu_0$ apabila $Z_0 < -Z_\alpha$, H_0 ditolak
 $H_0 : \mu < \mu_0$ apabila $Z_0 \geq -Z_\alpha$, H_0 diterima
- III. $H_0 : \mu = \mu_0$ apabila $Z_0 > Z_\alpha$, H_0 atau $Z_0 < -Z_{\alpha/2}$, H_0 ditolak.
 $H_0 : \mu \neq \mu_0$ apabila $-Z_{\alpha/2} \leq Z_0 < Z_{\alpha/2}$, H_0 diterima



CONTOH SOAL

1. Dari 100 nasabah bank rata-rata melakukan penarikan \$495 per bulan melalui ATM, dengan standar deviasi = \$45. Dengan taraf nyata 5%, ujilah Apakah rata-rata nasabah menarik uang melalui ATM sebesar \$500 per bulan dengan alternatif lebih besar dari itu.

Penyelesaian:

Diketahui:

$$\bar{X} = 495$$

$$s = 45$$

$$n = 100$$

$$\mu_0 = 500$$



Prosedur pengujian hipotesis:

1. Perumusan hipotesis

$$H_0: \mu = 500$$
$$H_a: \mu > 500$$

2. Menetapkan taraf nyata

$$\alpha = 5\% = 0.05.$$

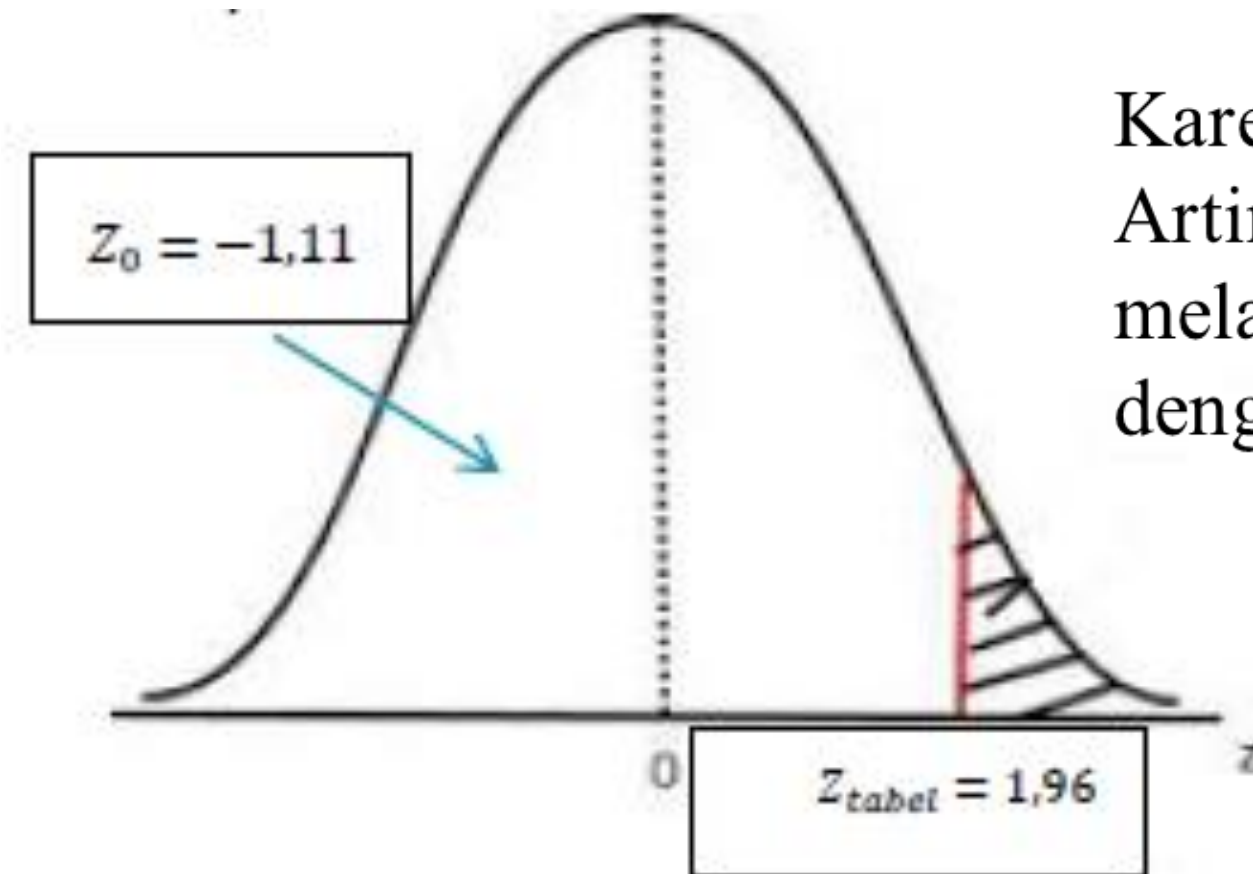
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

$$Z_{tabel} = 1,96$$

3. Menghitung nilai Z_0

$$Z_0 = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{495 - 500}{45/\sqrt{100}} = -\frac{5}{4,5} = -1,11$$

4. Keputusan dan kesimpulan



Karena $Z_0 < Z_{tabel}$ maka H_0 diterima.
Artinya rata-rata nasabah menarik uang melalui ATM adalah sebesar \$500 per bulan dengan tingkat keyakinan 95%.



2. Seorang *job-specialist* menguji 25 karyawan dan mendapatkan bahwa rata-rata penguasaan pekerjaan kesekretarisan adalah 22 bulan dengan standar deviasi = 4 bulan. Dengan taraf nyata 5%, ujilah apakah rata-rata penguasaan kerja kesekretarisan tidak sama dengan 20 bulan?

Penyelesaian:

Diketahui:

$$\bar{X} = 22$$

$$s = 4$$

$$n = 25$$

$$\mu_0 = 20$$

$$\alpha = 5\%$$



Prosedur pengujian hipotesis:

1. Perumusan hipotesis

$$H_0: \mu = 20$$

$$H_a: \mu \neq 20$$

2. Menetapkan taraf nyata

$$\alpha = 5\% = 0,05$$

Statistics UEU 2019
Derajat kebebasan = $n - 1 = 25 - 1 = 24$

$$t_{tabel} = 2,064$$

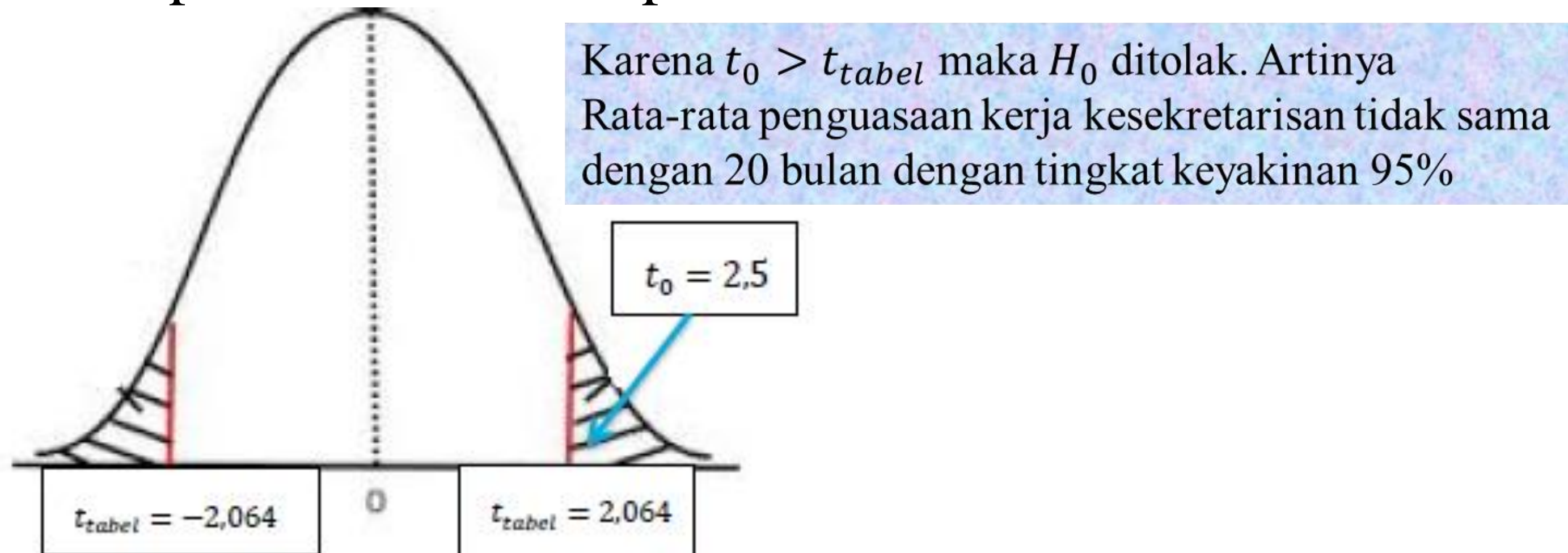


cum. prob	t _{.50}	t _{.75}	t _{.80}	t _{.85}	t _{.90}	t _{.95}	t _{.975}	t _{.99}	t _{.995}	t _{.999}	t _{.9995}
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725

3. Menghitung nilai t_0

$$t_0 = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{22 - 20}{4/\sqrt{25}} = \frac{2}{0,8} = 2,5$$

4. Keputusan dan kesimpulan



PENGUJIAN HIPOTESIS DUA KAIA-KAIA (DATA INDEPENDEN)

Untuk $n > 30$:

$$Z_0 = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1 - \bar{X}_2}}$$

dimana:

$$\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Apabila σ_1^2 dan σ_2^2 tidak diketahui, dapat diestimasi dengan:

$$s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Untuk $n \leq 30$:

$$t_o = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}$$

t_o mempunyai distribusi t dengan derajat kebebasan sebesar $n_1 + n_2 - 2$. Cara pengujiannya seperti yang sebelumnya, artinya $Z_o(t_o)$ dibandingkan dengan $Z_\alpha, Z_{\alpha/2}, -Z_{\frac{\alpha}{2}}(t_\alpha, t_{\frac{\alpha}{2}}, -t_{\alpha/2})$.

Keterangan:

\bar{X}_1 = rata-rata sampel 1

\bar{X}_2 = rata-rata sampel 2

σ_1^2 = varians dari sampel 1

σ_2^2 = varians dari sampel 2

n_1 = banyaknya sampel 1

n_2 = banyaknya sampel 2



CONTOH SOAL

Seorang pemilik toko yang menjual dua macam bola lampu merek A dan B, berpendapat bahwa tak ada perbedaan rata-rata lamanya menyala bola lampu kedua merek tersebut dengan alternatif ada perbedaan (tak sama). Guna menguji pendapatnya itu, kemudian dilakukan eksperimen dengan jalan menjalankan 100 buah bola lampu merek A dan 50 buah bola lampu merek B, sebagai random sampling.

Dari hasil random sampling didapatkan bahwa bola lampu merek A dapat menyala rata-rata selama 952 jam, sedangkan merek B 987 jam, dengan standar deviasi 85 jam dan 92 jam. Dengan menggunakan $\alpha = 5\%$, ujilah pendapat tersebut.



Penyelesaian:

Diketahui

$$n_1 = 100$$

$$n_2 = 50$$

$$\bar{X}_1 = 952$$

$$\bar{X}_2 = 987$$

$$\sigma_1 = 85$$

$$\sigma_2 = 92$$

$$\alpha = 5\%$$

Statistics UEU 2019

Prosedur pengujian hipotesis

1. Perumusan hipotesis

$$H_0: \mu_1 - \mu_2 = 0$$

$$H_0: \mu_1 - \mu_2 \neq 0$$

2. Menetapkan taraf nyata

$$\alpha = 5\% = 0,05$$

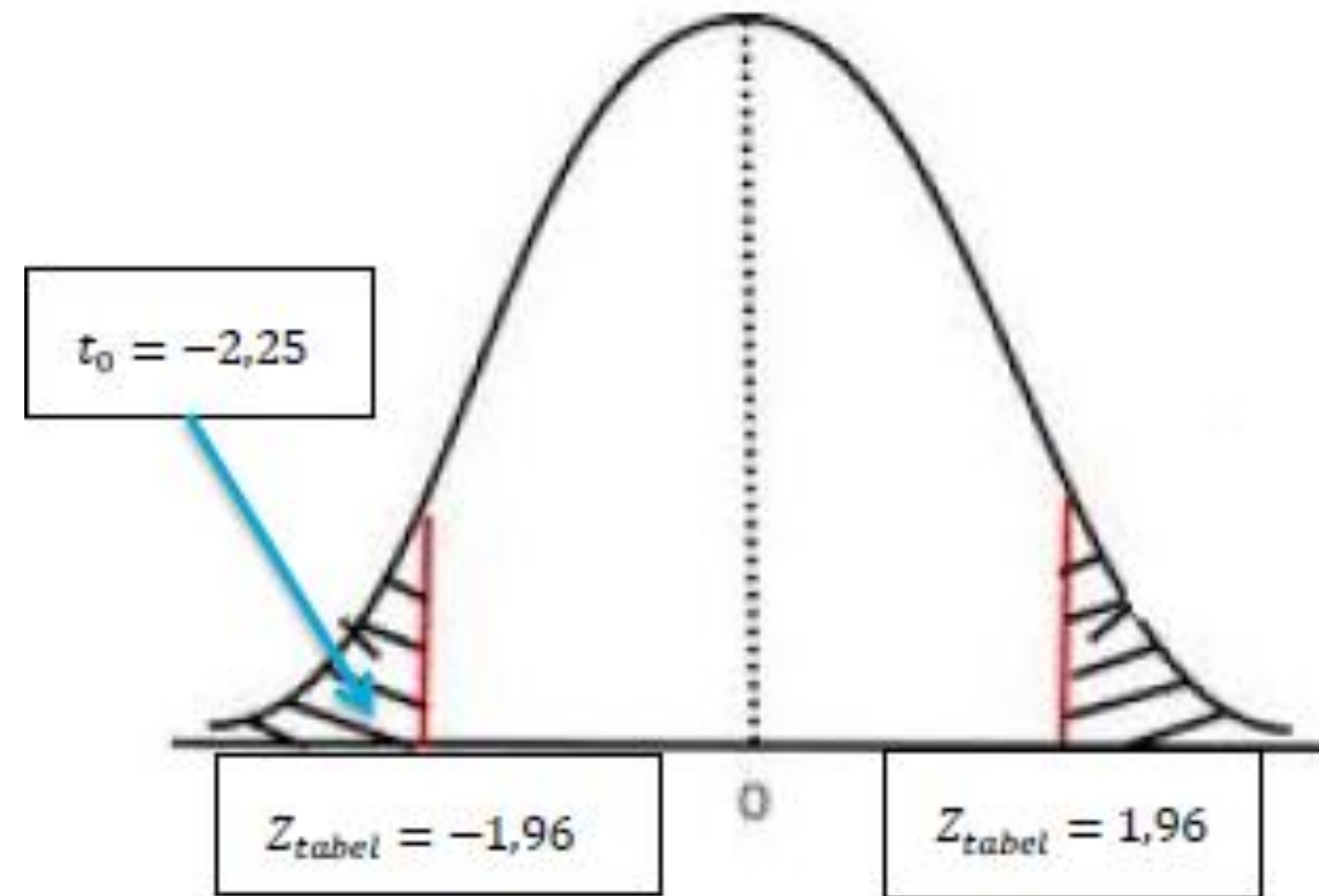
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

$$Z_{tabel} = 1,96$$

3. Menghitung nilai Z_0

$$\text{Statisti } Z_0 = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{952 - 987}{\sqrt{\frac{85^2}{100} + \frac{92^2}{50}}} = -2,25$$

4. Keputusan dan kesimpulan



Statistics UEU 2019

Karena $Z_0 < -Z_{tabel}$ maka H_0 ditolak. Artinya, rata-rata lamanya menyala dari bola lampu kedua merek tersebut tidak sama dengan tingkat keyakinan 95%.



Institut Informatika & Bisnis
DARMAJAYA
Yayasan Alfian Husin



**Kampus
Merdeka**
INDONESIA JAYA

**MERDEKA
BELAJAR**

THANK YOU!!

DATA SCIENCE DARMAJAYA "YOUR BEST FUTURE IN DATA"