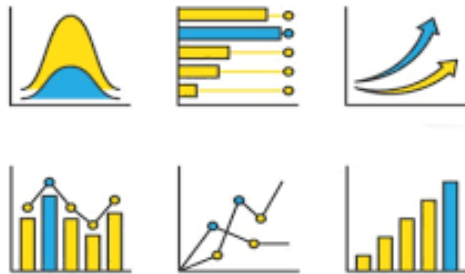


KUANTITATIF DAN KUALITATIF LINGKUNGAN (STATISTIKA LINGKUNGAN)

STATISTICS FOR DATA SCIENCE



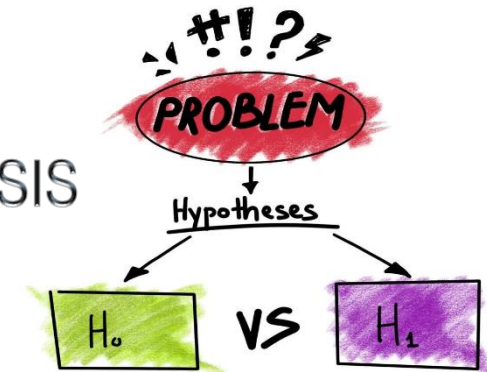
Prof. Dr. Ir. Zulkifli Alamsyah, M.Sc.

PROGRAM STUDI MAGISTER ILMU LINGKUNGAN
PASCASARJANA UNIVERSITAS JAMBI

KUANTITATIF DAN KUALITATIF LINGKUNGAN (STATISTIKA LINGKUNGAN)



ANALISIS DATA DAN PENGUJIAN HIPOTESIS (Hypothesis Testing)



Prof. Dr. Ir. Zulkifli Alamsyah, M.Sc.

PROGRAM STUDI MAGISTER ILMU LINGKUNGAN
PASCASARJANA UNIVERSITAS JAMBI

Pengertian Analisa data

Analisis data adalah *serangkaian kegiatan sistematis* untuk memeriksa, membersihkan, mentransformasi, dan memodelkan data dengan tujuan menemukan informasi yang berguna, mendukung kesimpulan, atau membantu pengambilan keputusan.

Definisi menurut beberapa ahli:

- **Miles & Huberman (1994):** analisis data adalah proses yang terdiri dari *data reduction, data display, dan conclusion drawing/verification*.
- **Creswell (2012):** analisis data adalah proses mengatur data, membaca keseluruhan informasi, membuat coding, dan menginterpretasikan makna dari temuan.
- **Cooper & Schindler (2014):** analisis data adalah transformasi data mentah menjadi informasi yang mampu memberikan insight.

Semua definisi menekankan: proses sistematis, pemberian makna, dan tujuan untuk menjawab pertanyaan penelitian.

Fungsi Analisis Data dalam Penelitian

- Mengubah data mentah menjadi informasi
Data mentah tidak punya makna tanpa analisis.
- Menjawab rumusan masalah atau tujuan penelitian
- Menguji hipotesis (kuantitatif)
- Membangun teori baru (kualitatif)
- Memberikan dasar ilmiah untuk rekomendasi kebijakan atau manajerial.

Ruang Lingkup Analisis Data

- Pengelolaan data
- Pengolahan statistik
- Interpretasi
- Verifikasi
- Visualisasi data
- Penarikan kesimpulan

Analisis data tidak hanya menghitung angka, tetapi juga memahami konteks, mengaitkan teori, dan menemukan makna.

Proses Analisis Data

Tahap 1 – Pengumpulan Data

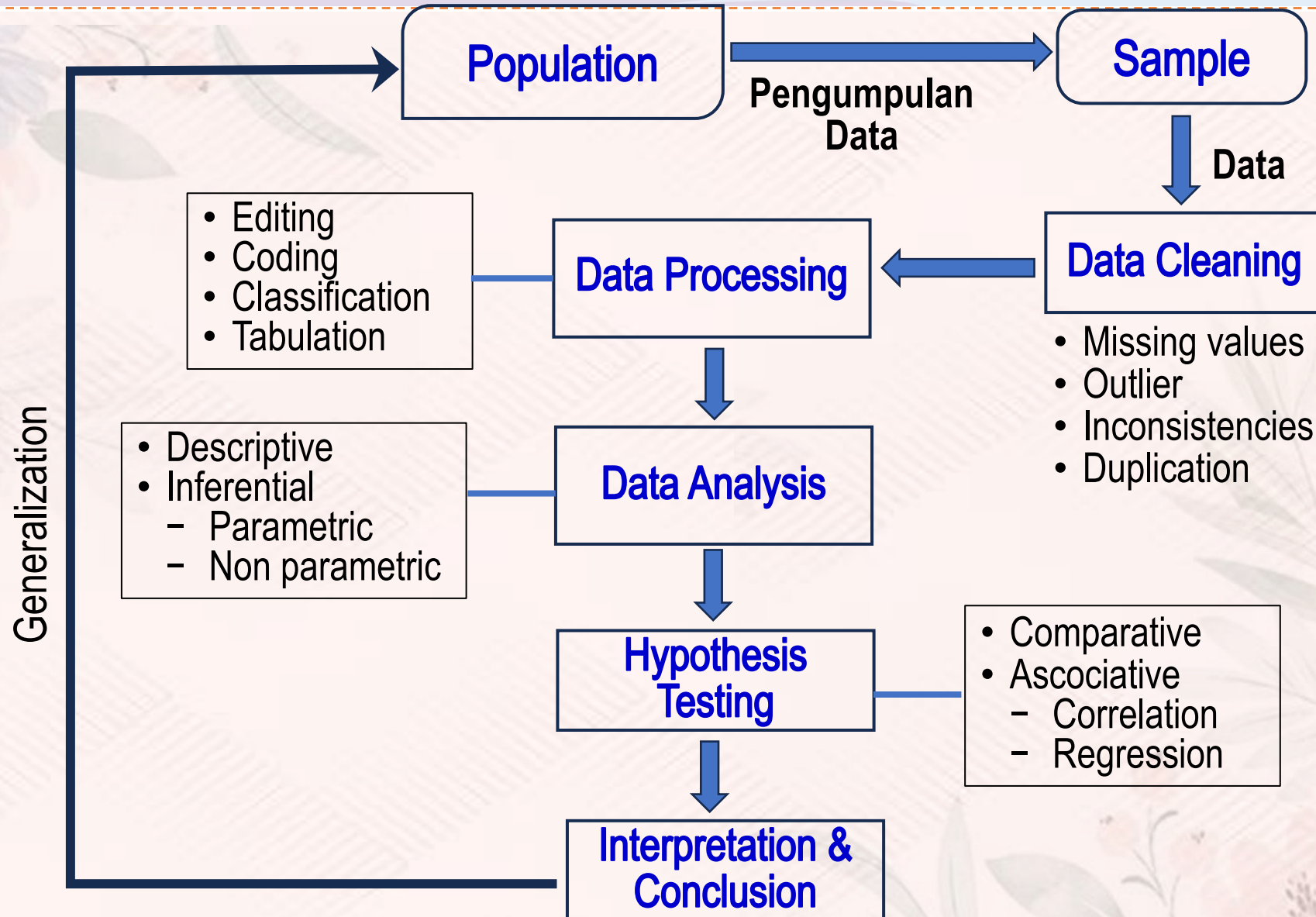
Memperoleh data sesuai tujuan penelitian.

Sumber data:

- Primer: wawancara, survei, observasi, eksperimen
- Sekunder: data pemerintah, jurnal, laporan perusahaan, database (FAOSTAT, BPS, World Bank)

Pentingnya tahap ini:

- Data yang baik menghasilkan analisis yang benar
- Kesalahan di tahap pengumpulan → menghasilkan garbage in, garbage out



Pengertian Analisis data

- Secara sederhana analisa data dapat diartikan sebagai suatu proses pengolahan data yang sistematis untuk penarikan kesimpulan penelitian berdasarkan hasil pengolahan data
- .Kegiatan pengolahan data meliputi proses editing, coding/ recoding, klasifikasi atau pengelompokkan dan tabulating
- Analisis Data mencakup (1) Analisis Deskriptif dan (2) Analisis Statistik

ANALISIS DATA

Analisis Deskriptif

- Cara mengorganisasi dan menyimpulkan informasi secara numerik, dengan menelaah variabel penelitian satu persatu. Biasanya dengan menggunakan tabel, grafik atau diagram.

Analisis Statistik

- Cara untuk menarik suatu kesimpulan atau mengetahui suatu keadaan yang terjadi pada suatu populasi dengan menggunakan informasi yang didapatkan dan sampel yang merupakan bagian dari populasi tersebut. Merupakan pengambilan kesimpulan dengan melakukan pengujian hipotesis.

Analisis Deskriptif

- Penyajian hasil analisis deskriptif biasanya berupa frekuensi dan persentase, tabulasi silang, serta berbagai bentuk grafik dan Chart pada data yang bersifat kategorikal, serta berupa statistik-statistik kelompok (mean dan varians) pada data yang bukan kategorikal
- Analisis Statistik
 - Pengolahan data pada tingkat ini dimaksudkan untuk mengambil kesimpulan dengan pengujian hipotesis
 - Pada dasarnya pengujian hipotesis statistika terbagi menjadi :
 - Hipotesis tentang adanya hubungan (korelasi)
 - Hipotesis tentang adanya perbedaan kelompok
 - Hipotesis tentang adanya pengaruh

HIPOTESIS

- Hipotesis adalah pernyataan yang masih lemah tingkat kebenarannya sehingga masih harus diuji menggunakan teknik tertentu
- Hipotesis dirumuskan berdasarkan teori, dugaan, pengalaman pribadi/orang lain, kesan umum, kesimpulan yang masih sangat sementara
- Hipotesis adalah jawaban teoritik atau deduktif dan bersifat sementara.
- Hipotesis adalah pernyataan keadaan populasi yang akan diuji kebenarannya menggunakan data/informasi yang dikumpulkan melalui sampel.
- Jika pernyataan dibuat untuk menjelaskan nilai parameter populasi, maka disebut hipotesis statistik

PERUMUSAN HIPOTESIS

- Rumusan hipotesis pada dasarnya sudah dapat dibaca dari uraian masalah, tujuan penelitian, kajian teoritik, dan kerangka pikir sehingga rumusannya harus sejalan
- Rumusan hipotesis sebagai petunjuk arah dalam rancangan penelitian, teknik pengumpulan dan analisis data serta penyimpulan
- Dinyatakan sebagai kalimat pernyataan (deklaratif)
- Melibatkan minimal dua variabel penelitian
- Mengandung suatu prediksi
- Harus dapat diuji (testable)

TIPE HIPOTESIS

- **Hipotesis korelatif** yaitu pernyataan tentang ada atau tidak adanya hubungan antara dua variabel atau lebih
- **Hipotesis kausalitas** yaitu pernyataan tentang ada atau tidak adanya pengaruh antara satu atau lebih variabel bebas (independent) terhadap variabel terikat (dependent)
- **Hipotesis komparatif** yaitu pernyataan tentang ada atau tidak adanya perbedaan antara dua kelompok atau lebih
- **Hipotesis nihil/nol (H_0)** yaitu hipotesis yang menyatakan tidak adanya hubungan antara dua variabel atau lebih atau tidak adanya perbedaan antara dua kelompok atau lebih
- **Hipotesis alternatif (H_a)** yaitu hipotesis yang menyatakan adanya hubungan antara dua variabel atau lebih atau adanya perbedaan antara dua kelompok atau lebih

Rumusan Hipotesis :

- Perbedaan antara kelompok A dan kelompok B

$$H_0 : \mu_A = \mu_B$$

$$H_a : \mu_A \neq \mu_B$$

- Kelompok A lebih tinggi dari kelompok B

$$H_0 : \mu_A = \mu_B$$

$$H_a : \mu_A > \mu_B$$

- Kelompok A lebih rendah dari kelompok B

$$H_0 : \mu_A = \mu_B$$

$$H_a : \mu_A < \mu_B$$

- Hipotesis Hubungan (asosiatif)
 - Suatu pertanyaan yang menunjukkan dugaan tentang hubungan antara dua variabel atau lebih
 - Contoh :
 - Adakah hubungan antara tingkat pendidikan seseorang dengan Produktivitas kerja?
 - Rumus hipotesis : Tidak ada hubungan antara tingkat pendidikan seseorang dengan Produktivitas kerja.
 - $H_0 : \rho = 0$
 - $H_a : \rho \neq 0$

Menentukan Jenis Metode Statistik yang akan Digunakan

- Kaitan antara jenis hipotesis, jumlah sampel dan skala pengukuran data menentukan jenis metode statistik yang digunakan
- untuk Pengujian hipotesis diperlukan pedoman penggunaan teknik statistik yang dikaitkan dengan jumlah sampel dan skala pengukuran

Metode Analisis Data Menurut Jenis Penelitian, Sampel dan Skala Pengukuran Data

Jenis Penelitian	Sampel	Skala pengukuran	Alat statistik
Deskriptif	Satu sampel	Nominal	Binomial atau Chi kuadrat
Deskriptif	Satu sampel	Ordinal	Run Test
Deskriptif	Satu sampel	Interval / Rasio	T-Test
Komparatif	Dua sampel berpasangan	Nominal	Mc. Nemar
Komparatif	Dua sampel berpasangan	Ordinal	Sign-Test / Wilcoxon Matched Pairs
Komparatif	Dua sampel berpasangan	Interval / rasio	T-Test
Komparatif	Dua sampel independen	Nominal	Fisher / Chi kuadrat
Komparatif	Dua sampel independen	Ordinal	Median-Test, Mann Whitney, U-Test, Kolmogorov Smirnov, Run Wald-Wolvowitz
Komparatif	Dua sampel independen	Interval / rasio	T-Test

Jenis Penelitian	Sampel	Skala pengukuran	Alat statistik
Komparatif	K sampel berpasangan	Nominal	Cochran Q
Komparatif	K sampel berpasangan	Ordinal	Friedman
Komparatif	K sampel berpasangan	Interval / rasio	Anova
Komparatif	K sampel independen	Nominal	Chi Kuadrat
Komparatif	K sampel independen	Ordinal	Median Extention, Kruskal walls
Komparatif	K sampel independen	Interval / rasio	Anova
Asosiatif		Nominal	Koefisien kontigensi
Asosiatif		Ordinal	Koefisien spearman Rank
Asosiatif		Interval / rasio	Korelasi pearson product moment, korelasi parsial, korelasi berganda, regresi

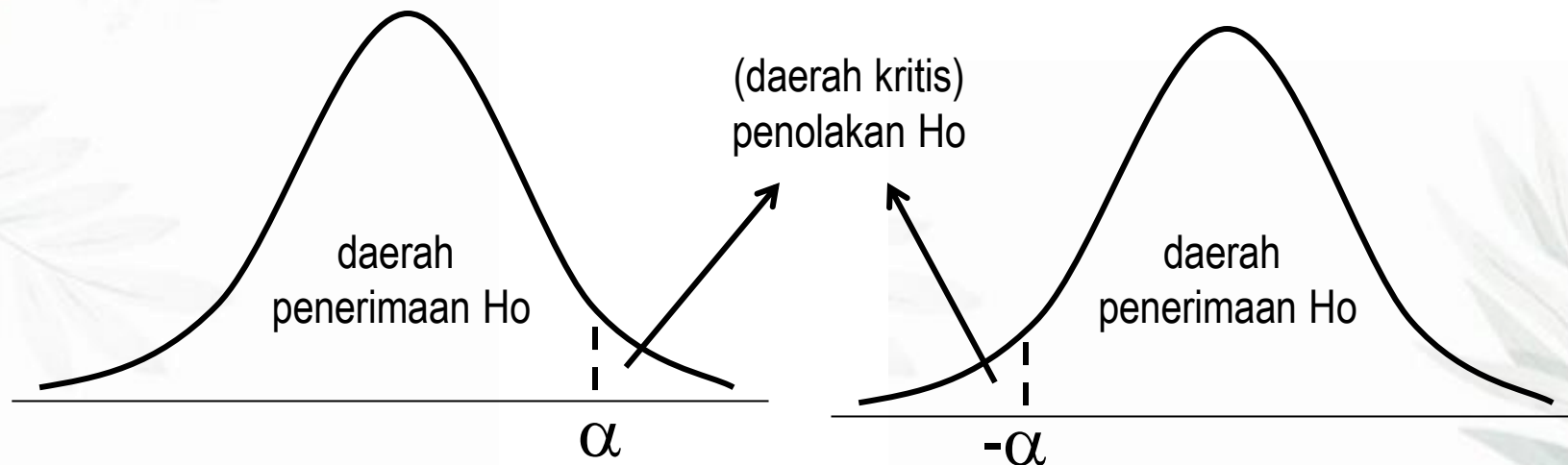
KESALAHAN DALAM PENGAMBILAN KEPUTUSAN

KEPUTUSAN	Ho benar	Ho salah
Terima Ho	Tepat	Kesalahan Tipe II (β)
Tolak Ho	Kesalahan Tipe I (α)	Tepat

Kesalahan Tipe I adalah kesalahan yg dibuat pd waktu menguji hipotesis di mana kita menolak Ho pd hal sesungguhnya Ho itu benar. \Rightarrow peluang menolak Ho yg benar

Kesalahan Tipe II adalah kesalahan yg dibuat pd waktu menguji hipotesis di mana kita menerima Ho pd hal sesungguhnya Ho itu salah. \Rightarrow peluang menolak Ho yg salah

UJI SATU ARAH

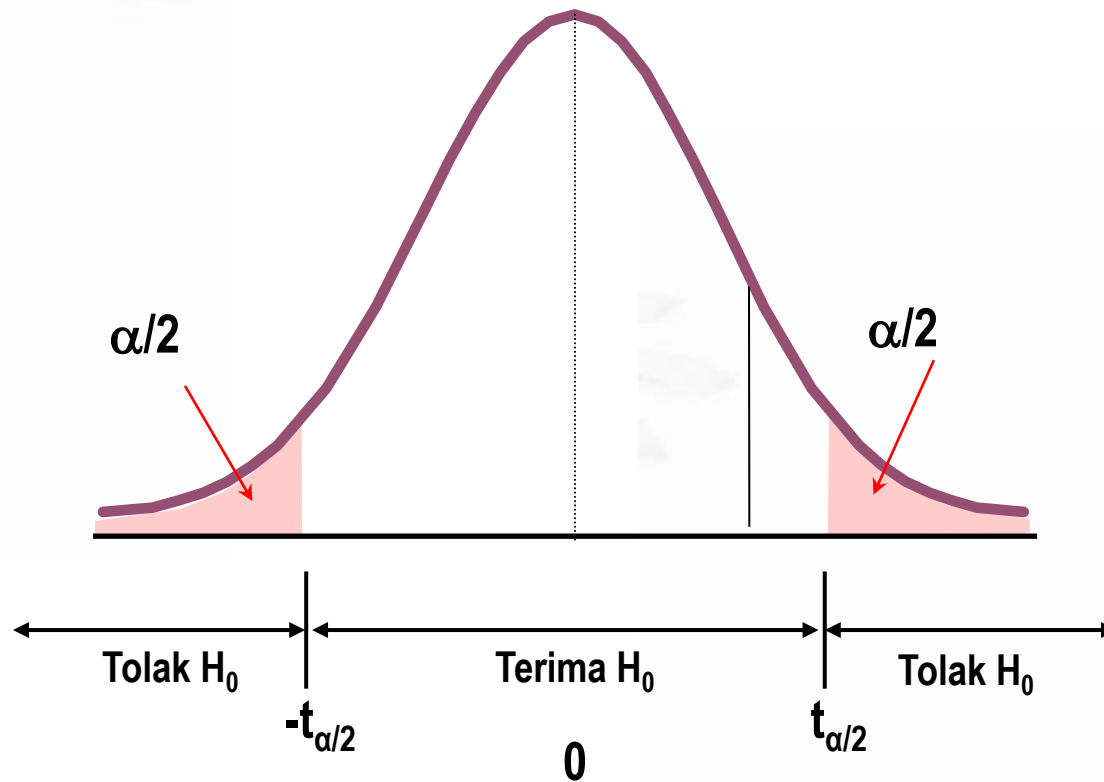


Catatan:

α = tingkat signifikansi

x = suatu bilangan tertentu

UJI DUA ARAH



Proses Pengambilan Keputusan Menggunakan Statistika:

- ❁ Identifikasi Permasalahan
- ❁ Perumusan Hipotesis
- ❁ Perencanaan dan Pelaksanaan Studi
- ❁ Pengujian Hipotesis
- ❁ Penarikan Kesimpulan

Identifikasi Permasalahan

- Apa pertanyaan mendasarkan yang ingin diketahui?
- Permasalahan akan menggiring kita kepada perumusan hipotesis dan penggunaan prosedur statistik.

Contoh:

Pembangunan regional pada dasarnya bertujuan untuk meningkatkan perekonomian wilayah sehingga dapat meningkatkan kesempatan kerja, pemerataan pendapatan dan kesejahteraan masyarakat. Meskipun upaya-upaya pembangunan sudah dilaksanakan secara merata, masih ditemukan ketimpangan pendapatan antar wilayah.

Perumusan Hipotesis

Hipotesis adalah kesimpulan sementara yang akan diuji kebenarannya

Scientific hypothesis atau sering disebut **hipotesis penelitian** merupakan pernyataan verbal terhadap jawaban permasalahan penelitian.

Contoh: Perbedaan karakteristik wilayah menyebabkan perbedaan pendapatan masyarakat

Statistical hypothesis is a statement or claim about a **population parameter** that can be tested using sample data.

→ Statistical hypothesis dinyatakan dalam bentuk **parameter** yang akan diuji.

$H_0: \mu_1 = \mu_2$ atau $\mu_1 - \mu_2 = 0$ (null hypothesis)

$H_a: \mu_1 \neq \mu_2$ atau $\mu_1 - \mu_2 \neq 0$ (alternative hyp.)

Untuk contoh hipotesis tersebut, Alat Uji statistic yang digunakan adalah Uji-t (t-test), dengan formula:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sum X_1^2 - (\sum X_1)^2/n_1 + \sum X_2^2 - (\sum X_2)^2/n_2}{n_1 + n_2 - 2} \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

- **Tentukan Tingkat signifikansi (significance level)**

Contoh: $\alpha = 0.01$ pada tingkat keyakinan 99%

atau $\alpha = 0.05$ pada tingkat keyakinan 95%

- **Cek nilai t-tabel pada Tabel Statistik**

- **Kaedah keputusan alat uji statistik**

Jika $t_{hitung} > t_{table}$ maka tolak H_0

$t_{hitung} \leq t_{table}$ maka terima H_a

Uji t untuk kasus satu sampel (One-sample t test)

- Pengujian rata-rata populasi
- Lebih praktis dan realistis – dapat digunakan bila variance populasi (σ^2) tidak diketahui

Asumsi-asumsi mengenai Distribusi Sampel:

- 1) Nilai Rata-rata (mean) = μ
- 2) Nilai varians = σ^2/n
- 3) Bentuk distribusi sampel = normal

Contoh kasus:

Dengan adanya pembangunan di suatu wilayah, diprediksi pendapatan perkapita masyarakat meningkat dari pendapatan tahun lalu sebesar Rp. 3.000.000.

Seorang peneliti ingin menguji apakah prediksi tersebut dapat diterima.

Hipotesis yang diajukan adalah: pendapatan per kapita masyarakat saat ini di wilayah tersebut lebih besar dari Rp.3.000.000.

Langkah-langkah yang harus dilakukan:

- Mengumpulkan data dari sejumlah sampel
- Menghitung nilai rata-rata pendapatan per kapita
- Menguji hipotesis

Langkah-langkah pengujian hipotesis:

- Mengajukan hipotesis statistik:
Ho: $\mu = 3.000.000$
Ha: $\mu > 3.000.000$
- Menggunakan alat uji t-statistik dengan formula:

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

X = Rata-rata sampel

μ = Dugaan rata-rata populasi

s = Standar deviasi sampel

n = Jumlah sampel

Karena varians populasi tidak diketahui, maka digunakan pendekatan varians sampel, dengan rumus:

$$s^2 = \frac{\sum(X_i - \bar{X})^2}{n - 1} \quad \text{dan} \quad s = s^{1/2}$$

Dimana:

s^2 = varians sampel; s = standar deviasi sampel
 X_i = nilai pengamatan (sampel) ke- i , utk $i = 1, \dots, n$

- Kaedah keputusan uji statistik

Jika $t_{\text{hitung}} > t_{\text{table}(\alpha, n-1)}$ maka tolak H_0

$t_{\text{hitung}} \leq t_{\text{table}(\alpha, n-1)}$ maka terima H_0

Contoh:

Dari 36 sampel, diperoleh data rata-rata pendapatan per kapita sebesar Rp. 3.150.000 per tahun, dengan standar deviasi Rp. 600.000.

Apakah secara statistik nilai ini kecil dari pendapatan perkapita tahun lalu?

Pengujian:

$$t = \left| \frac{X - \mu}{s / \sqrt{n}} \right| = \left| \frac{3.150.000 - 3.000.000}{600.000 / \sqrt{36}} \right|$$

$$t = \left| \frac{150.000}{600.000 / 6} \right| = 1,50$$

Kesimpulan pengujian hipotesis:

- ❓ Nilai t-hitung = 1.50
- ❓ Nilai t-tabel $\alpha(0.1, 39) = 1.3036$
- ❓ Nilai t-tabel $\alpha(0.05, 39) = 1.6849$

Sesuai dengan kaedah keputusan:

Secara statistik, 90% dapat diyakini bahwa pendapatan per kapita masyarakat saat ini lebih dari Rp. 3.000.000 per tahun.

Kasus Dua Sampel

- ✘ Independet
- ✘ Dependent
 - Terjadi berpasangan secara alami
 - Pengukuran dari subjek yang sama
 - Dua subjek yang dipasangkan atas suatu variabel

Fokus perhatian pada kasus dua sampel adalah:

- Perbandingan antara kedua kelompok sampel
- Perbedaan rata-rata kedua kelompok sampel

Asumsi-asumsi mengenai Distribusi selisih dua Rata-rata Sampel ($X_1 - X_2$) :

1) Nilai Rata-rata (mean) = $\mu_1 - \mu_2$

2) Nilai varians:

$$\sigma^2_{X_1-X_2} = \sigma_1^2/n_1 + \sigma_2^2/n_2 \text{ (independent)}$$

$$= \sigma_1^2/n_1 + \sigma_2^2/n_2 - 2 \rho(\sigma_1/\sqrt{n_1})(\sigma_2/\sqrt{n_2}); \text{ (dep.)}$$

3) Bentuk distribusi sampel = normal

Contoh kasus:

Upaya-upaya pembangunan di semua wilayah telah dilakukan secara adil dan merata. Dalam perkembangannya, terlihat pertumbuhan ekonomi yang cukup signifikan antar wilayah.

Seorang peneliti tertarik mempelajari apakah terdapat perbedaan pertumbuhan ekonomi antara suatu wilayah dengan wilayah lainnya.

Hipotesis yang diajukan adalah: karena perbedaan karakteristik wilayah, terdapat perbedaan pertumbuhan ekonomi (yang diukur dari pendapatan per kapita masyarakat) antara wilayah yang satu dengan wilayah lainnya.

Langkah-langkah yang harus dilakukan:

- Mengumpulkan data dari sejumlah sampel pada kedua wilayah
- Menghitung nilai rata-rata pendapatan per kapita pada kedua wilayah
- Menguji hipotesis

Langkah-langkah pengujian hipotesis:

- Mengajukan hipotesis statistik:

$$H_0: \mu_1 - \mu_2 = 0 \text{ atau } \mu_1 = \mu_2$$

$$H_a: \mu_1 - \mu_2 \neq 0 \text{ atau } \mu_1 \neq \mu_2$$

- Menggunakan alat uji t-statistik dgn formula:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sum X_1^2 - (\sum X_1)^2/n_1 + \sum X_2^2 - (\sum X_2)^2/n_2}{n_1 + n_2 - 2} \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

- Kaedah keputusan uji statistik

Jika $t_{hitung} > t_{table(\alpha/2, n1+n2-2)}$ maka tolak H_0

$t_{hitung} \leq t_{table(\alpha/2, n1+n2-2)}$ maka terima H_0

Misal,

Dari masing-masing wilayah diperoleh sampel sebanyak 36 rumah tangga. Data yang diperoleh adalah:

$$\Sigma X_1 = 376$$

$$\Sigma X_1^2 = 4030$$

$$X_1 = 10.44$$

$$s_1^2 = 2.94$$

$$\Sigma X_2 = 262$$

$$\Sigma X_2^2 = 2076$$

$$X_2 = 7.28$$

$$s_2^2 = 4.83$$

Apakah secara statistik terdapat perbedaan pendapatan yang signifikan pada kedua wilayah tersebut?

Bentuk umum formula uji-t untuk dua beda rata-rata:

$$t = \frac{X_1 - X_2}{S_{x_1 - x_2}}$$

Dimana:

$X_1 - X_2$ = beda dua rata-rata

$S_{x_1 - x_2}$ = Standar deviasi dari beda dua rata-rata

$$S_{x_1 - x_2} = \sqrt{\left[\frac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2}{n_1 + n_2 - 2} \right] \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}$$

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{(n_1 - 1) s_1^2 + (n_2 - 1) s_2^2}{n_1 + n_2 - 2} \right] \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}}$$

Nilai t-tabel: Critical Values of t (Degree of freedom = $n_1 + n_2 - 2$)

Degree of Freedom	Upper Tail Areas				
	0.1	0.05	0.025	0.01	0.005
66	1.2945	1.6683	1.9966	2.3820	2.6524
67	1.2943	1.6879	1.9960	2.3833	2.6512
68	1.2941	1.6876	1.9955	2.3824	2.6501
69	1.2939	1.6872	1.9949	2.3816	2.6490
70	1.2938	1.6669	1.9944	2.3808	2.6479
71	1.2936	1.6666	1.9939	2.3800	2.6469
72	1.2934	1.6663	1.9935	2.3793	2.6459
73	1.2933	1.6660	1.9930	2.3785	2.6449
74	1.2931	1.6657	1.9925	2.3778	2.6439
75	1.2929	1.6654	1.9921	2.3771	2.6430

