

# **STATISTIKA INFERENSIAL**

**IM TIRTA**

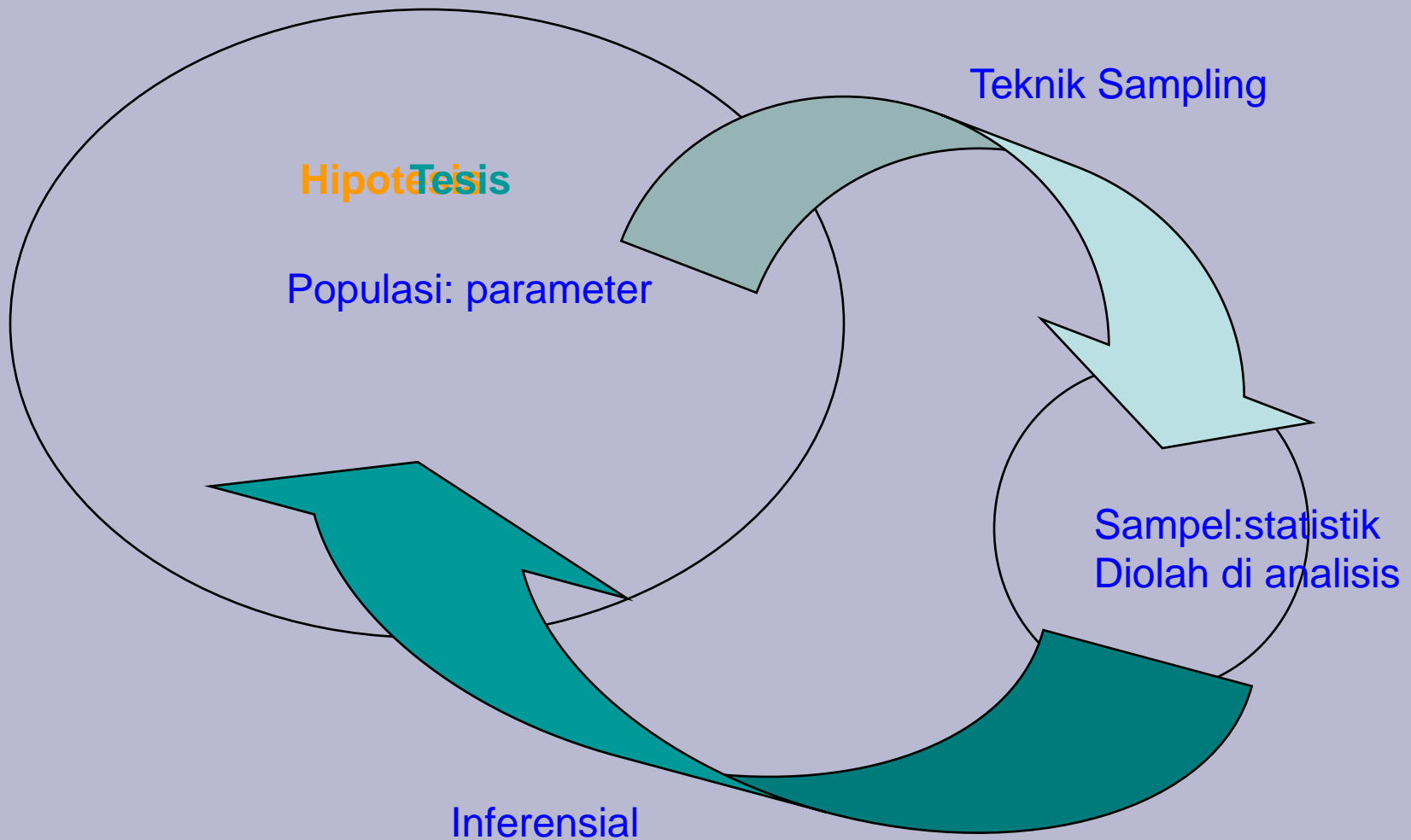
# RASIONAL

- Kondisi riil pengolahan informasi (Data): Karena **keterbatasan waktu, biaya dan tenaga** tidak memungkinkan mengumpulkan dan mengolah seluruh informasi yang ada di masyarakat (di lapangan)
- Implikasi: diperlukan **prosedur dan alat yang memungkinkan menarik kesimpulan yang valid yang berlaku bagi kelompok yang lebih luas, berdasarkan informasi yang diperoleh dari kelompok yang lebih kecil, dengan peluang kesalahan yang terukur**

# ISTILAH-ISTILAH UMUM

- Kelompok data yang diolah disebut **sampel**
- Kelompok data yang diwakili disebut **populasi**
- Prosedur statistika yang dimaksud (untuk menarik kesimpulan tentang populasi berdasarkan informasi pada sampel) disebut **statistika inferensial**
- Pernyataan dugaan sementara (dari kajian teori) tentang populasi disebut **hipotesis**. Hipotesis selanjutnya diuji dari informasi yang ada pada sampel

# POPULASI vs SAMPEL

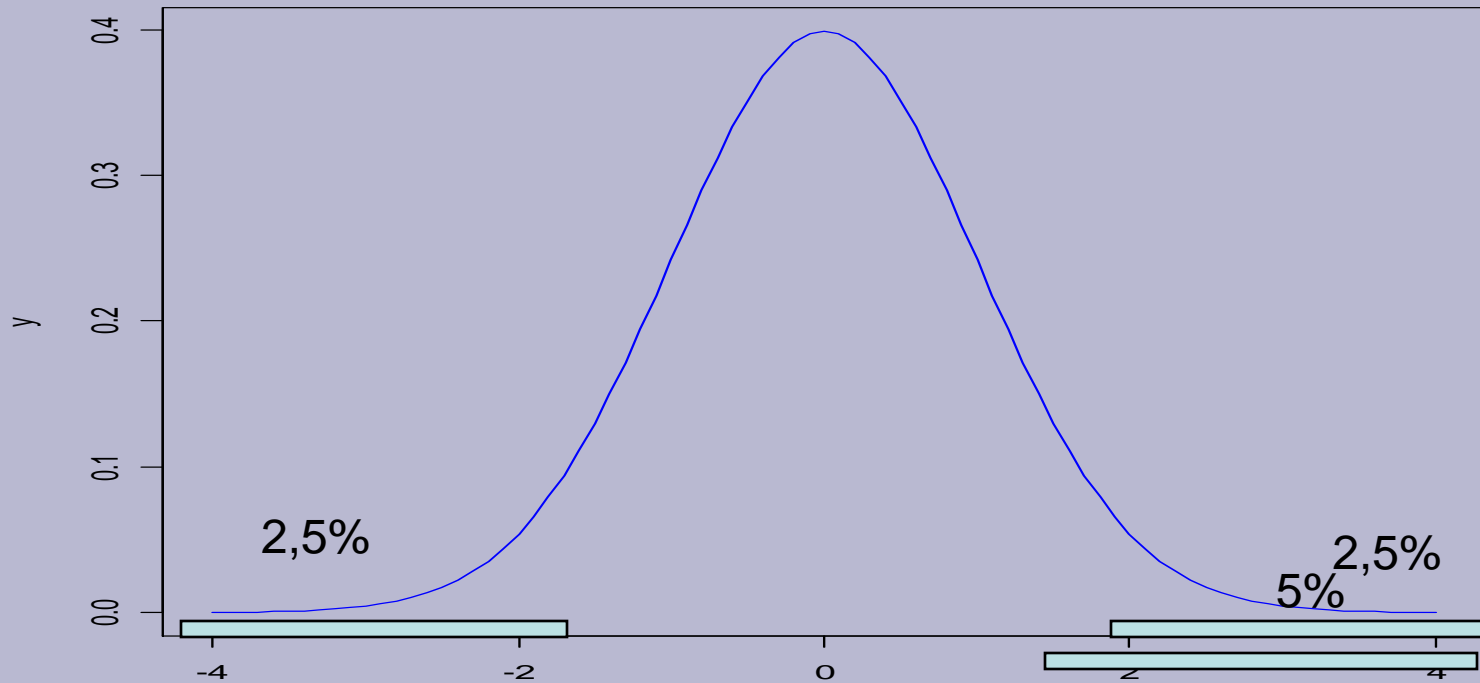


# HIPOTESIS

- **Hipotesis nol (nihil):**  $H_0$ , bersifat netral (tidak ada beda, tidak ada hubungan, ..tidak ...)
- **Hipotesis kerja (alternatif):**  $H_a$  ada dua macam yaitu
  - **Dua arah** (ada beda x dengany; ada hubungan, tanpa menentukan mana yang lebih baik )
  - **Satu arah** (menentukan kelompok yang lebih baik: x lebih baik y; lebih jelek, berhubungan positif)
  - Perumusan satu arah atau dua arah mempengaruhi pengambilan keputusan yang bersifat marjinal

# TARAF SIGNIFIKANSI

Peluang Normal



**Ha: dua arah (signifikansi 5%; 1,96)**

**Ha: satu arah (signifikansi 5%; 1,65)**

# ASUMSI UMUM SAMPEL

- Sampel yang diambil harus valid, representatif (ruang dan waktu) dan bersifat acak (wajib)
- Jumlah memenuhi syarat minimal ( $\geq 30$ )
- Memenuhi sebaran tertentu (umumnya sebaran normal: kontinu, simetrik)

# PELANGGARAN ASUMSI

- Keterwakilan, validitas dan keacakan data **bersifat wajib**
- Jumlah dan sebaran menentukan jenis uji statistika yang dapat dipergunakan (misalnya parametrik vs non parametrik)
- Jumlah sampel yang relatif besar (dan interval) memungkinkan lebih leluasa memilih metode analisis



# STATISTIKA INFERENSIAL

- Secara umum mempelajari **hubungan beberapa variabel dengan berbagai skala pengukuran** atau pencacahan
- Dapat pula merupakan persoalan sederhana hanya **menduga (menaksir) pemusatan populasi (parameter) berdasarkan pemusatan sampel (statistik)**
- Hubungan antara **variabel nominal dapat ditafsirkan sebagai perbedaan kelompok**. Misalnya mempelajari hubungan jenis kelamin (laki-perempuan) dengan prestasi belajar statistika ekuivalen dengan mempelajari perbedaan prestasi belajar statistika antara kelompok laki-laki dan perempuan.

# STATISTIKA INFERENSIAL

- Taksiran parameter dapat berupa **taksiran titik dan interval (dengan tingkat keyakinan tertentu berdasarkan sebaran data)**
- Variabel (objek yang diamati) dibedakan menjadi **variabel bebas (penjelas/ *explanatory*) dan variabel respon (terikat/*response*)**
- Jenis kelamin (penjelas) vs prestasi belajar (respons)
- Tingkat pendidikan (penjelas) vs jumlah anak (respons) atau penghasilan
- Wilayah (penjelas) vs persentase buta huruf (respons)
- Penghasilan (respon) vs tingkat pendidikan, jenis kelamin

# UJI BEDA KELOMPOK

Ada berbagai kondisi dalam uji beda kelompok atau hubungan antara variabel nominal dengan lainnya

1. Kelompok pada variabel bebas ada 2 atau lebih (banyaknya kategori dalam salah satu variabel nominal 2 atau lebih)
2. Skala variabel respon bisa kontinu (beda rata-rata) atau diskrit (beda frekwensi atau proporsi)

# SIFAT STATISTIK SAMPEL

- Statistik sebagai informasi (numerik) dari sampel memiliki sifat: **Stokastik (probabilistik)** karena diambil dari sampel acak karena itu harus selalu ada ukuran **pemusatan** (rata-rata) dan ukuran **penyebaran** (deviasi baku), keduanya tidak bisa dipisahkan dan bergantung pada jenis sebaran data (normal dll).
- Perbedaan dalam sampel yang diyakini juga terjadi pada populasi disebut **signifikan** bukan disebabkan karena faktor kebetulan.

# KOMPONEN STATISTIK

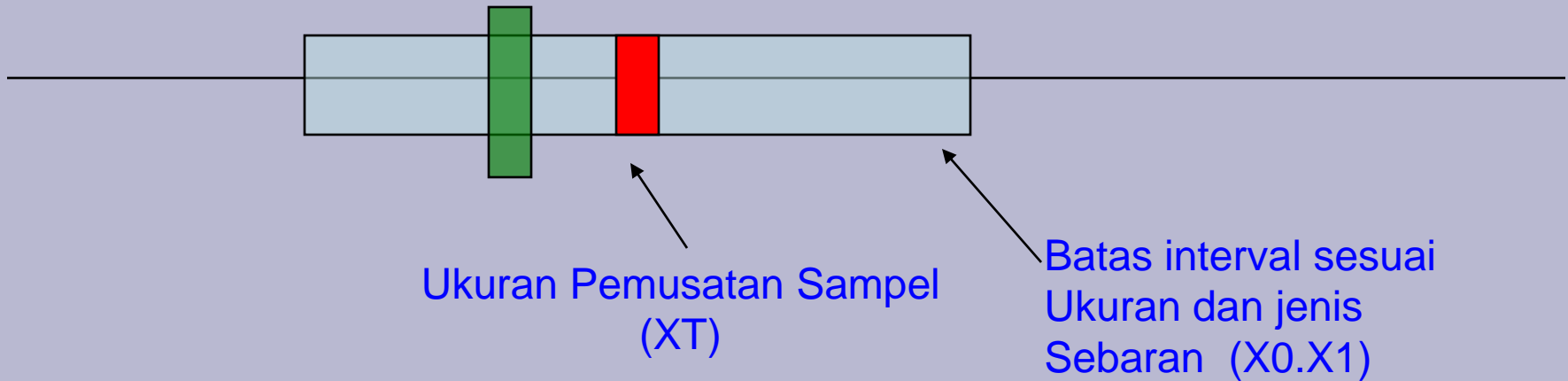
- Rata-rata sampel:  $\bar{x} = \sum x / N$
- Deviasi baku sampel:  $S = \sqrt{[\sum (x - \bar{x})^2 / (N-1)]}$
- Deviasi baku rata-rata (kesalahan baku):  $S / \sqrt{N}$
- Taraf signifikansi (tingkat kecocokan), peluang bahwa kesimpulan yang kita ambil salah.
- Interval keyakinan berdasarkan asumsi distribusi (n < 30 distribusi t, sebaliknya menggunakan distribusi normal)

# MATEMATIK VS STOKASTIK

- Secara matematik  $81 \neq 83$
- Dalam konteks kehidupan sehari-hari 81 tidak berbeda signifikan jika dikaitkan dengan perolehan skor ujian yang akhirnya sama-sama menjadi A
- Dalam konteks statistik **signifikan** tidaknya suatu **ukuran pemusatan** sangat bergantung pada **ukuran penyebarannya**

# STOKASTIK

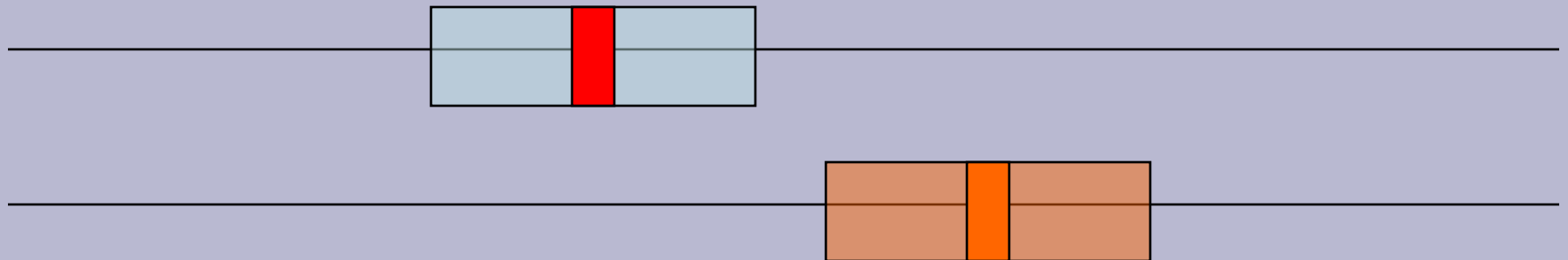
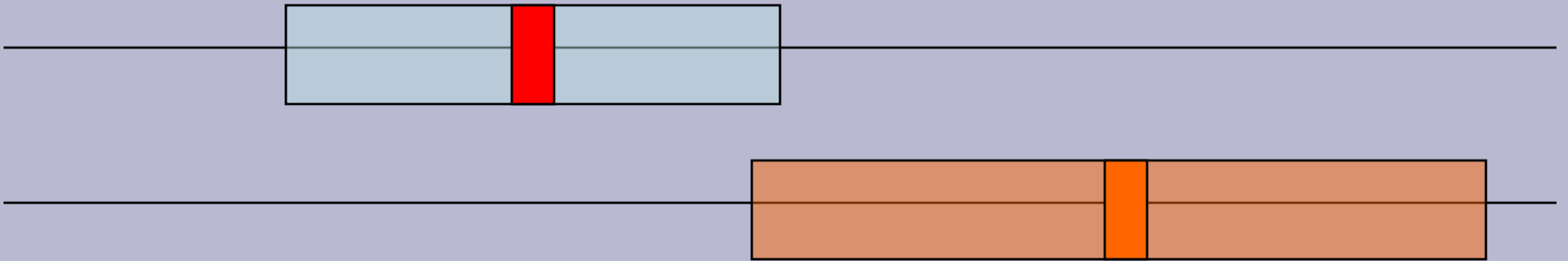
Parameter



**XT tidak beda signifikan dengan semua titik antara  $X_0$  dan  $X_1$   
(Hipotesis bahwa titik hijau merupakan parameter populasi adalah benar)**

# STOKASTIK

Kedua kelompok tidak beda signifikan



Kedua kelompok beda signifikan



# UJI T

- Untuk uji taksiran parameter populasi berdasarkan statistik sampel
- Untuk uji beda rata-rata dengan banyak kelompok dua
- Kedua kelompok mungkin **saling bebas** (beda subjek, misalnya laki, perempuan) mungkin **tidak saling bebas** (subjek sama beda objek, misalnya *pre-post treatment*)

# SYARAT UJI BEDA MEAN

## Syarat Penggunaan Uji Beda Mean

- Data diambil secara **acak**
- Ada variabel yang berupa **kuantitatif (skala interval)**, misalnya tinggi badan, berat badan, IQ, nilai ujian
- Ada variabel yang bersifat **faktor (kelompok) dengan dua kategori** (laki-perempuan, desa-kota, eksperimen-kontrol)

# LANGKAH UJI BEDA

## MANUAL

- Rumuskan Hipotesis dan tingkat signifikansi (batas peluang kesalahan kesimpulan)
- Hitung Statistik

$$\bar{x} = \sum x / N; \bar{y} = \sum y / N_y$$

$$sx = \sqrt{\sum [(x - \bar{x})^2 / (Nx - 1)]}; sy = \sqrt{\sum [(y - \bar{y})^2 / (Ny - 1)]}$$

$$sebm = \sqrt{\frac{sx^2 (Nx - 1) + sy^2 (Ny - 1)}{Nx + Ny - 2}}$$

$$z_0 = \frac{\bar{x} - \bar{y}}{sebm(\sqrt{1/nx + 1/ny})};$$

# LANGKAH UJI BEDA

3. Tentukan titik kritis ( $t_k$ , atau  $z_k$ ) sesuai uji 1 atau 2 arah, misalnya untuk sampel besar  $z$  satu dua arah = 1,96 sedangkan satu arah = 1,65
4. Ambil kesimpulan
  1. jika  $z_0$  atau  $t_0 < t_k$ , maka  $H_0$  diterima
  2. Jika  $z_0$  atau  $t_0 \geq t_k$ , maka  $H_0$  ditolak
5. Untuk sampel relatif besar (dua arah)
  1. jika  $z_0 < 1,96$ , maka  $H_0$  diterima
  2. Jika  $z_0 \geq 1,96$ , maka  $H_0$  ditolak

# LANGKAH UJI BEDA

UNTUK UJI SATU KELOMPOK (MENAKSIR  
BESAR PARAMETER)

1. Rata-rata  $y$  diganti dengan angka taksiran
2.  $Sey=0$  dan  $Ny=1$

$$\bar{x} = \sum x / N;$$

$$sx = \sqrt{\sum \left[ (x - \bar{x})^2 / (Nx - 1) \right]};$$

$$z_0 = \frac{\bar{x} - \mu_0}{sx / \sqrt{n}};$$

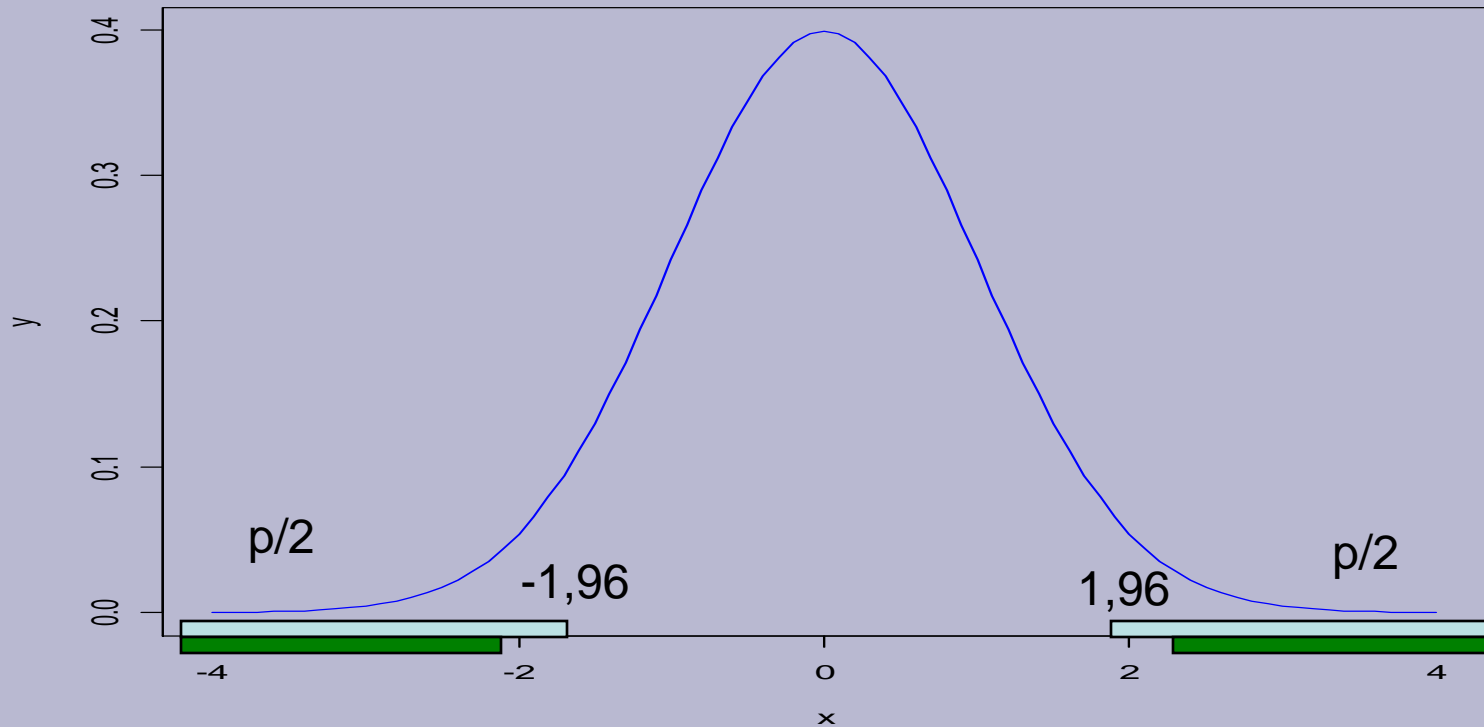
# LANGKAH UJI BEDA

## KOMPUTER

1. Rumuskan hipotesis (satu arah dua arah)
2. Tentukan jenis beda (satu kelompok, dua kelompok saling bebas atau tidak saling bebas)
3. Lakukan uji
4. **Lihat nilai p**
5. **Jika  $p < 5\%$  maka  $H_0$  ditolak dan  $H_a$  diterima**

# NILAI p

Peluang Normal



**Ha: dua arah, signifikan jika  $p < 5\%$   
Ekuivalen dengan  $|z| > 1,96$**

# UJI BEDA BERPASANGAN

## MANUAL

1. Hitung selisih pasangan
2. Lakukan uji t satu kelompok pada selisih pasangan dengan  $\mu_0=0$
3. Selanjutnya sama



# UJI BEDA BERPASANGAN

KOMPUTER (R, SPSS, MINITAB)

1. Pilih menu yang sesuai
2. Selanjutnya sama

# UJI PROPORSI

## Tujuan

Mengetahui ada tidaknya beda proporsi keberadaan antara dua kelompok

1. 60% mahasiswa S2 adalah laki-laki
2. Proporsi laki-laki dari jember dan dari luar jember berbeda

## Hipotesis

Ho: Proporsi kelompok A adalah  $p_0 = x\%$

Ha: Proporsi kelompok A tidak sama dengan  $p_0 = x\%$

# UJI PROPORSI

Langkah

1. Hitung  $p$  dan  $se$
2. Hitung statistik  $t$
3. Hitung p-value

$$\hat{p} = x / n$$

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

$$t = \frac{\hat{p} - p_0}{se} \sim N(0,1)$$

# UJI BEDA PROPORSI

Hipotesis

$H_0$ : proporsi kelompok X di wilayah A sama dengan di wilayah B

$H_a$ : proporsi kelompok X di wilayah A tidak sama sama dengan di wilayah B

(Misalnya proporsi laki-laki di desa dan di kota, proporsi buta huruf di desa dan di kota)

# UJI BEDA PROPORSI

Langkah

1. Hitung  $p$  dan  $se$
2. Hitung statistik  $t$
3. Hitung  $p$ -value

$$\hat{p}_1 = x/n_1; \hat{p}_2 = y/n_2$$

$$se = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

$$t = \frac{\hat{p}_1 - \hat{p}_2}{se} \sim N(0,1)$$

# ANALISIS VARIANSI

## Tujuan

- Untuk memeriksa adanya beda mean dari 3 kelompok atau lebih (misalnya kelompok ekonomi kuat, menengah dan lemah; Jember, Jawa Timur dan luar Jawa Timur)

# ANALISIS VARIANSI

## Syarat

- Ada variabel faktor dengan 3 kelompok atau lebih (misalnya kelompok ekonomi kuat, menengah dan lemah; Jember, Jawa Timur dan luar Jawa Timur)
- Ada variabel kuantitatif dengan skala interval (prestasi belajar, tinggi badan, berat badan)

# ANALISIS VARIANSI

## Syarat

- Data diambil secara acak
- Variansi kelompok bersifat homogen



# ANALISIS VARIANSI

Kenapa analisis variansi?

- Karena pada dasarnya variansi kelompok bersifat homogen
- Jika terjadi perbedaan signifikan antara variansi (rata-rata kuadrat sampel) dalam kelompok dengan antara kelompok maka itu pasti disebabkan oleh adanya beda mean (lihat ilustrasi)

# ANALISIS VARIANSI

## Hipotesis

- $H_0$  semua kelompok memiliki mean yang sama
- $H_a$  paling tidak ada dua kelompok yang memiliki mean yang tidak sama