



# PENGANTAR BIOSTATISTIK

Prof. Bhisma Murti  
Bagian Ilmu Kesehatan Masyarakat  
Fakultas Kedokteran  
Universitas Sebelas Maret

# DEFINISI BIOSTATISTIK

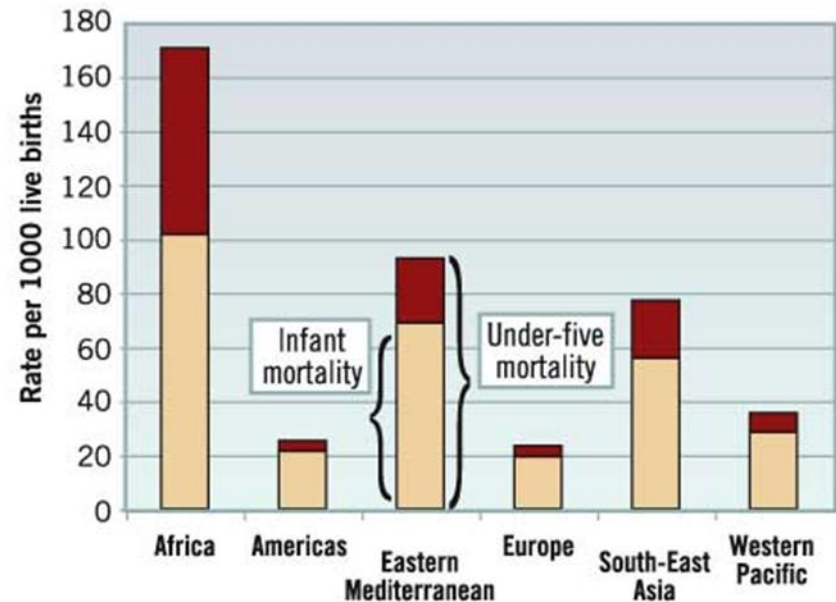
## ■ Biostatistik:

- Metode statistik yang diterapkan pada ilmu-ilmu terkait kesehatan, seperti kedokteran dan kesehatan masyarakat, untuk membantu memahami tentang karakteristik populasi, dan hubungan/ pengaruh variabel pada populasi

## ■ Statistik:

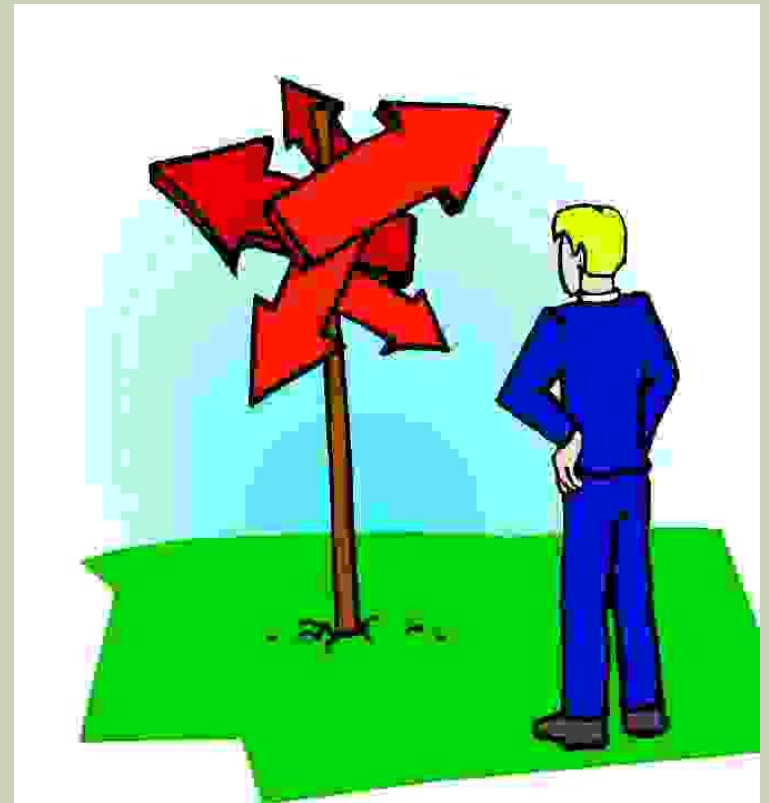
- Cabang matematika terapan yang berurusan dengan pengumpulan, analisis, dan interpretasi data, dan penggunaan teori probabilitas untuk menaksir parameter populasi

Under-five and infant mortality rates, by WHO Region, 2003



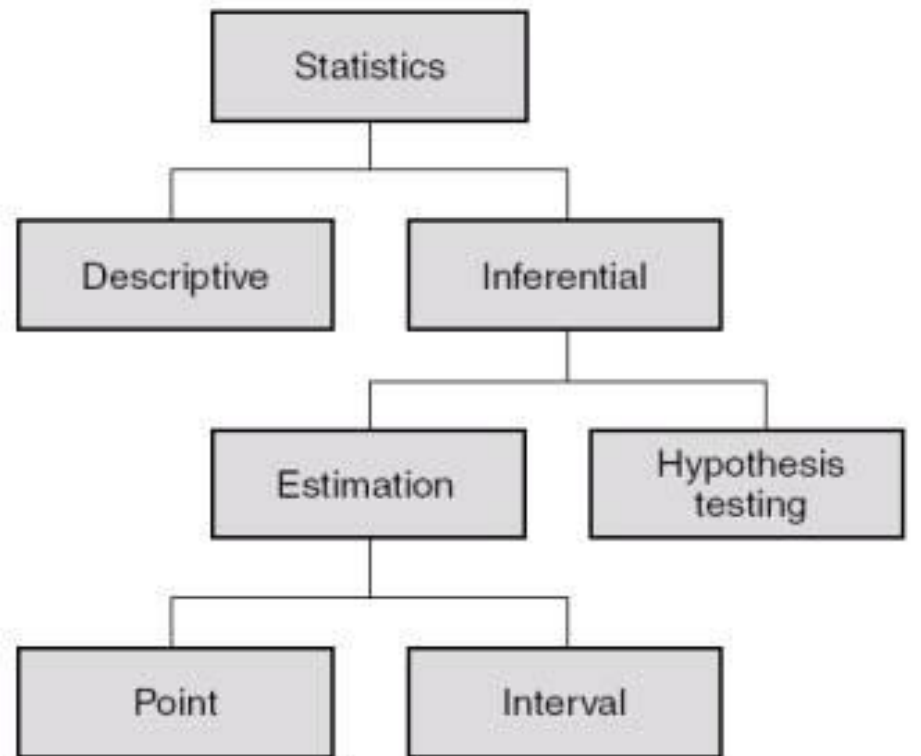
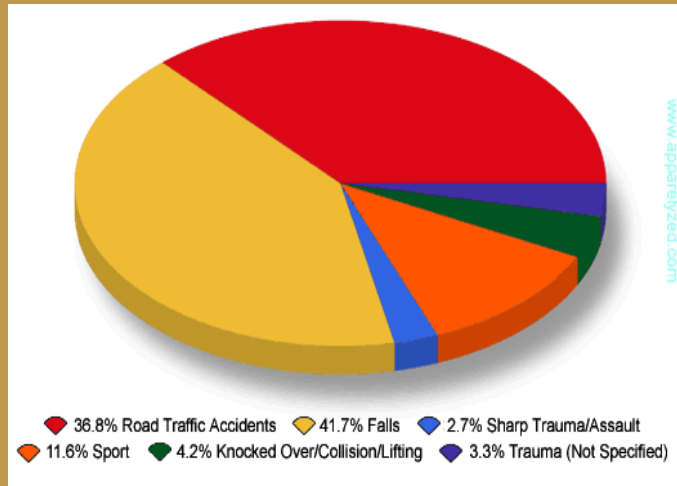
# KEGUNAAN BIOSTATISTIK

- Biostatistik berguna untuk memberikan informasi tentang:
  1. Karakteristik populasi
    - Berapa persen dari populasi yang menderita TB paru?
    - Berapa rata-rata tekanan darah sistolik populasi obes ( $BMI > 30$ )?
  2. Hubungan/ pengaruh variabel pada populasi
    - Apakah merokok berhubungan dengan peningkatan risiko penyakit jantung koroner (PJK)?
    - Apakah pemberian metilprednisolon dapat mengurangi mortalitas pasien dengan tetanus?
- Informasi itu berguna untuk membantu mengambil keputusan, membuat perencanaan, atau memecahkan masalah



# PEMBAGIAN BESAR STATISTIK

Kausa Cedera Trauma Tulang Belakang



# VARIABEL, DATA, DAN INFORMASI

## ■ Variabel

Karakteristik subjek yang memiliki variasi atribut atau nilai

- Jenis kelamin: laki-laki, perempuan
- Kadar gula darah: 110, 111, 112, .....400, 401, 402 mg/dl

## ■ Data

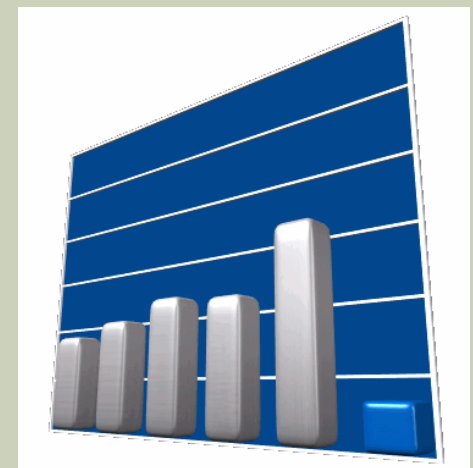
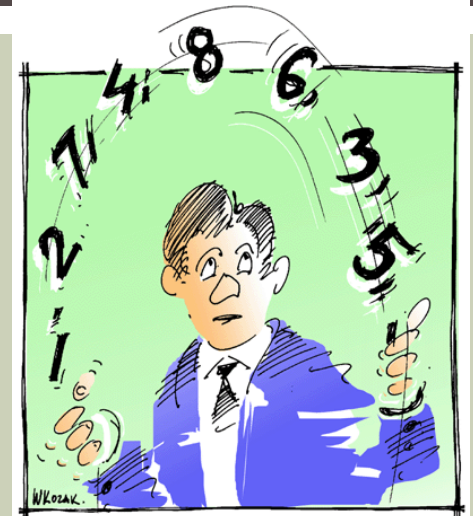
Kumpulan angka kuantitatif hasil pengukuran variabel, biasanya belum memiliki makna

- Subjek A: Seks= laki-laki, TB= 170cm, BB=90kg, TDS= 150mmHg
- Subjek B: Seks= perempuan, TB=165cm; BB=55kg, TDS=120mmHg

## ■ Informasi

Hasil analisis data sehingga memiliki makna yang berguna untuk membuat keputusan

- Rata-rata TDS populasi obes= 150mmHg, populasi overweight= 140mmHg, populasi normal= 120mmHg
- Rasio L:P di Indonesia=0.95, di India= 1.05



# JENIS VARIABEL

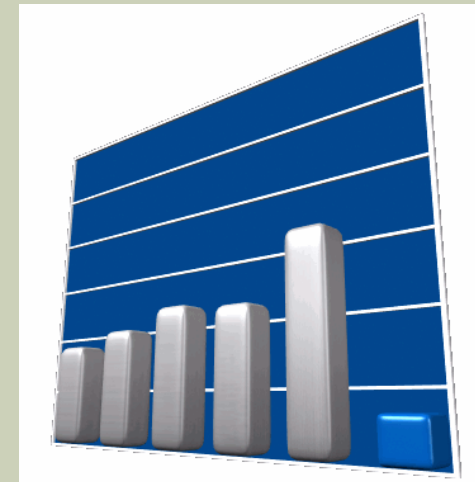
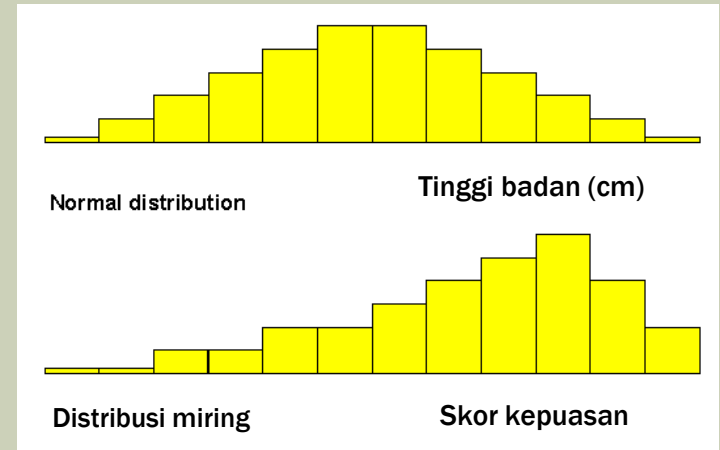
## ■ Jenis variabel:

### 1. Kontinu

- Terdapat kontinuitas antar nilai-nilai variabel
- TB (cm), BB (kg), luas infark (%), tekanan darah sistolik (mmHg), gula darah (mg/dl), Hb (g/dl), skor kepuasan
- Deskripsi: n, Mean, SD, Min, Maks

### 2. Kategorikal

- Tidak terdapat kontinuitas antar atribut variabel
- Jenis kelamin (Laki-laki, Perempuan); Stadium kanker (I, II, III, IV); Area geografis (Rural, Urban)
- Deskripsi, n, persen

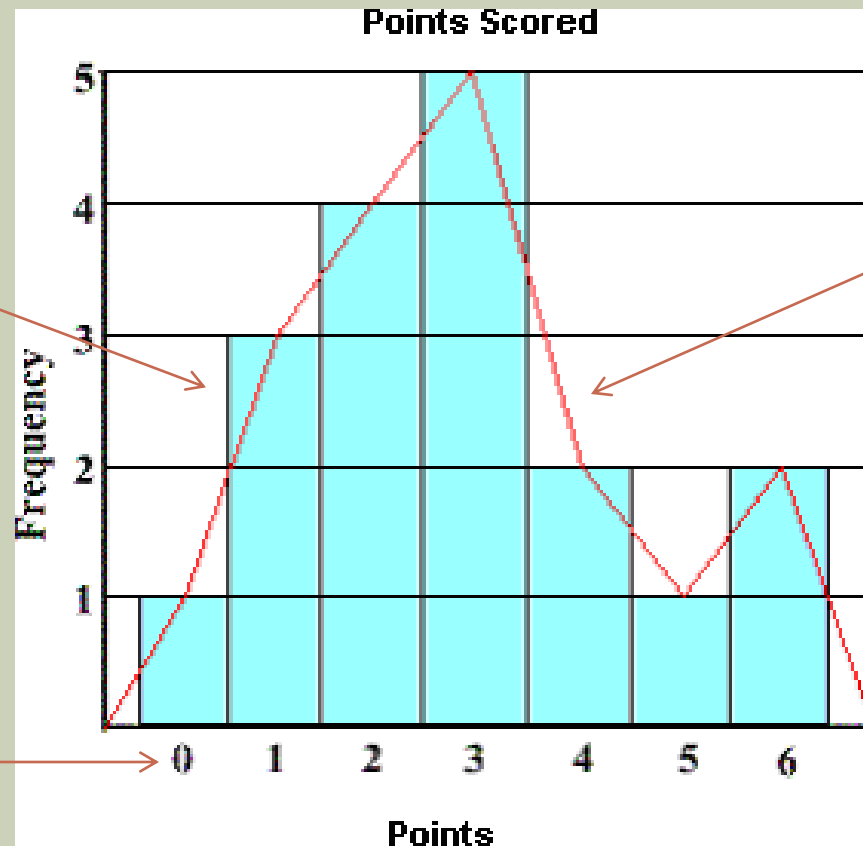


# DATA KONTINU, HISTOGRAM, DAN POLIGON

Histogram

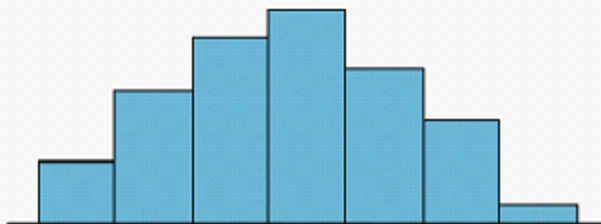
Poligon

Data kontinu

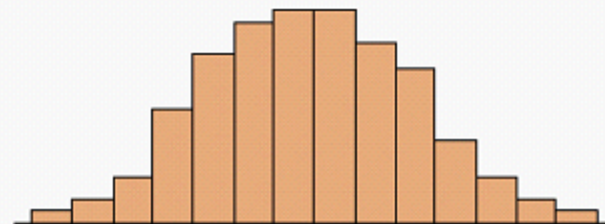


# DATA KONTINU, HISTOGRAM, DAN DISTRIBUSI NORMAL

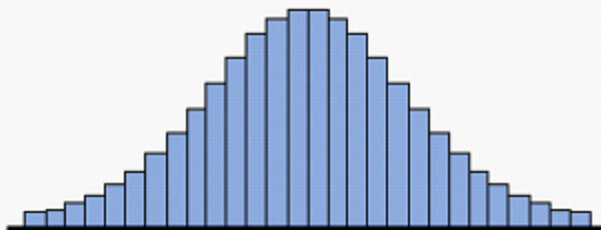
## Histograms for the distribution of heights



(a) Random sample of 100 women



(b) Sample size increased and class width decreased



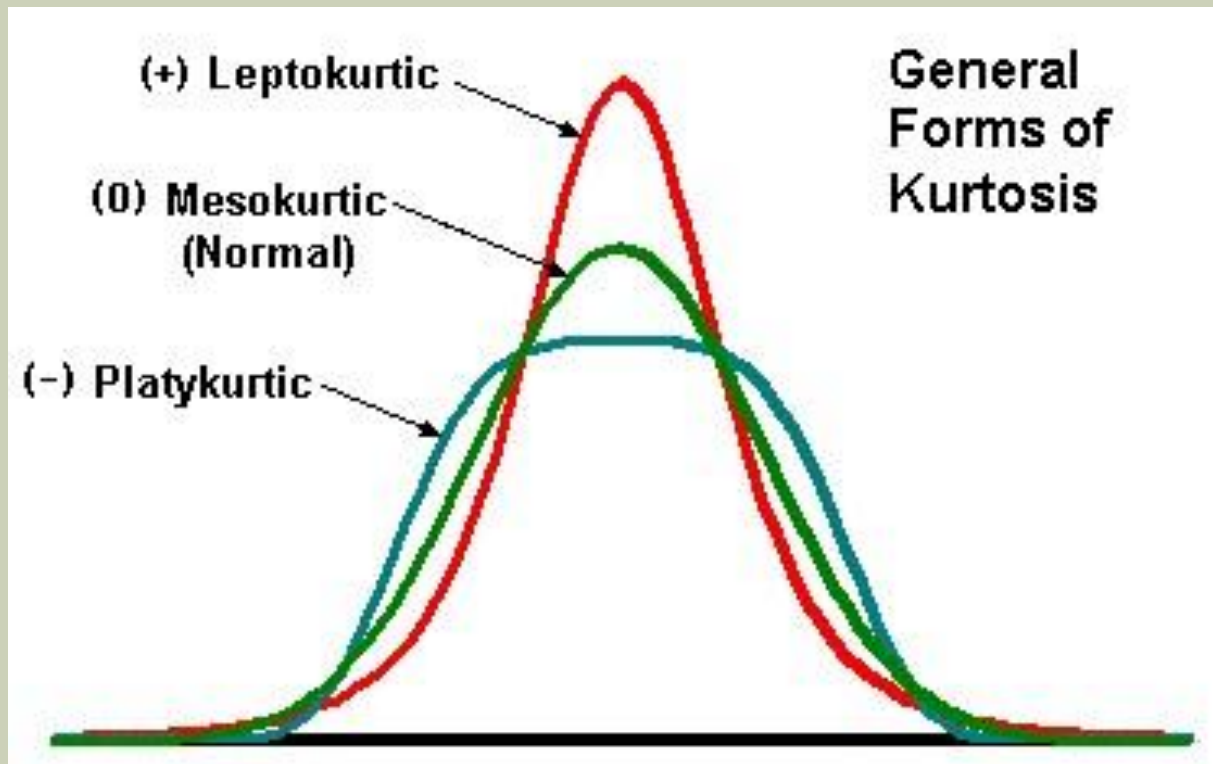
(c) Sample size increased and class width decreased further



(d) Normal distribution for the population



# KURVA NORMAL DAN KURTOSIS



# UKURAN TENDENSI SENTRAL, KURVA NORMAL, KURVA MIRING

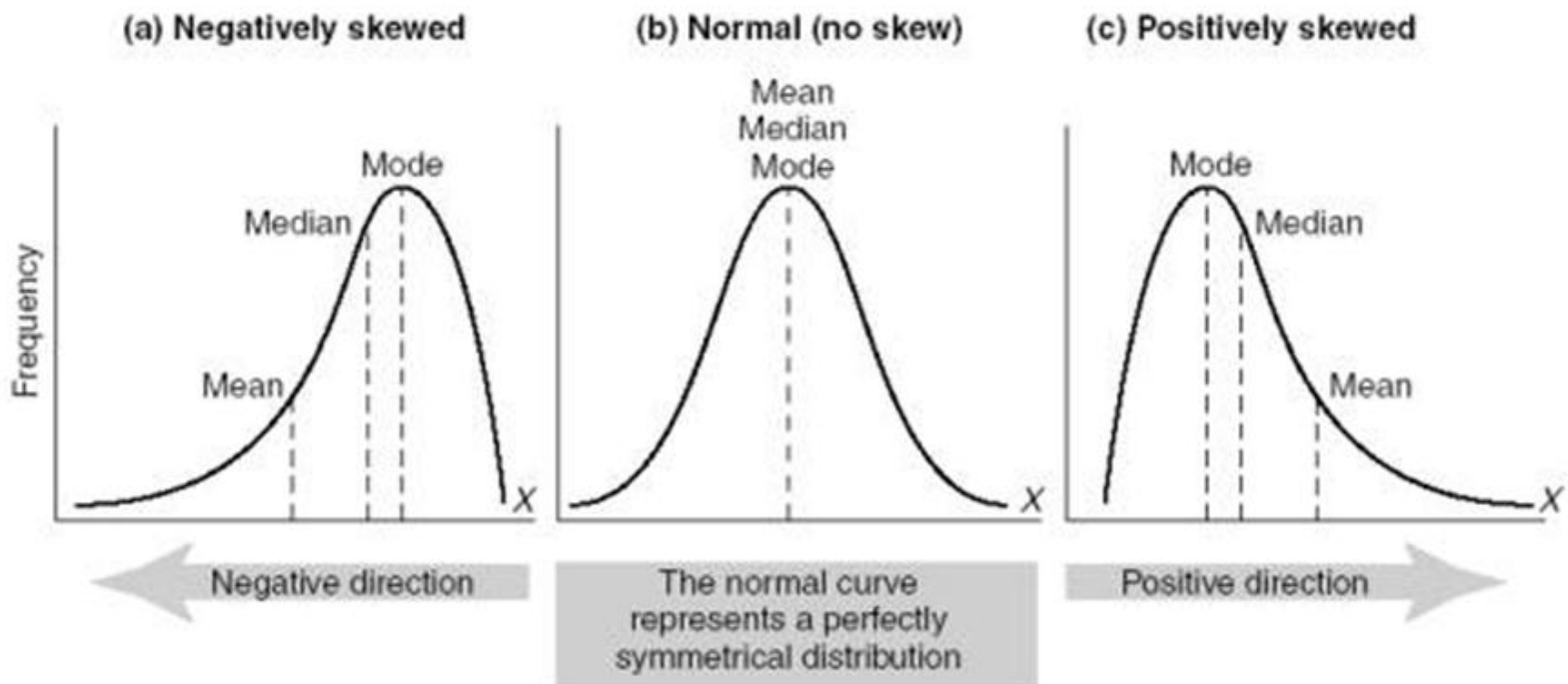
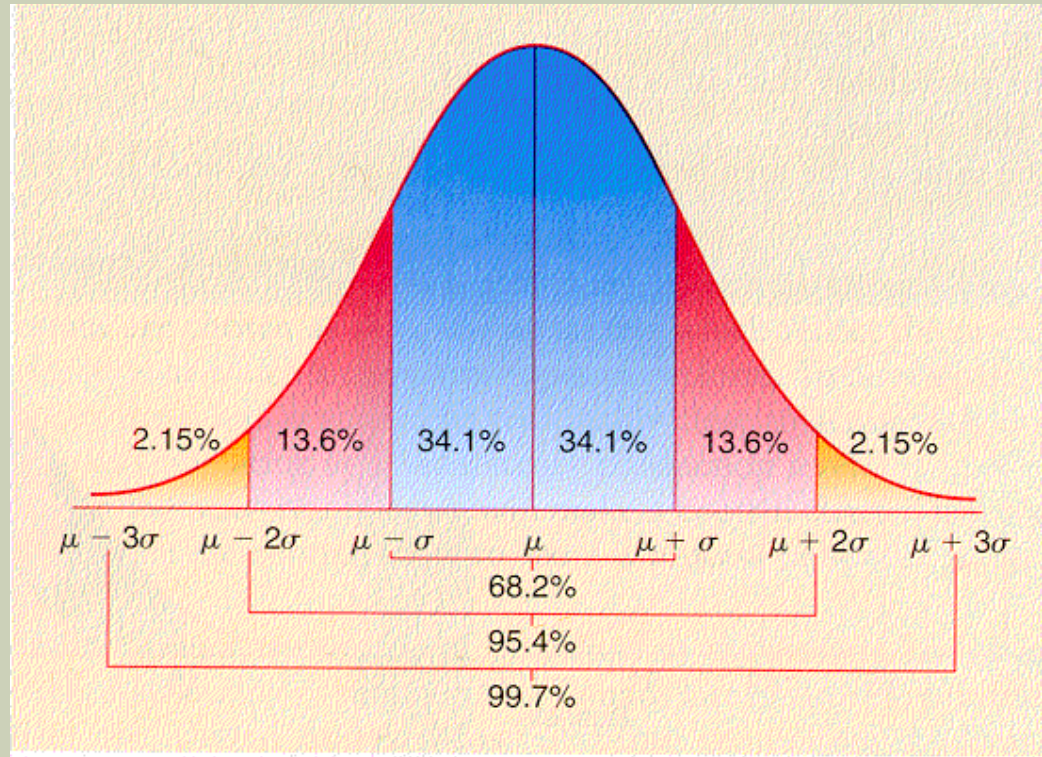


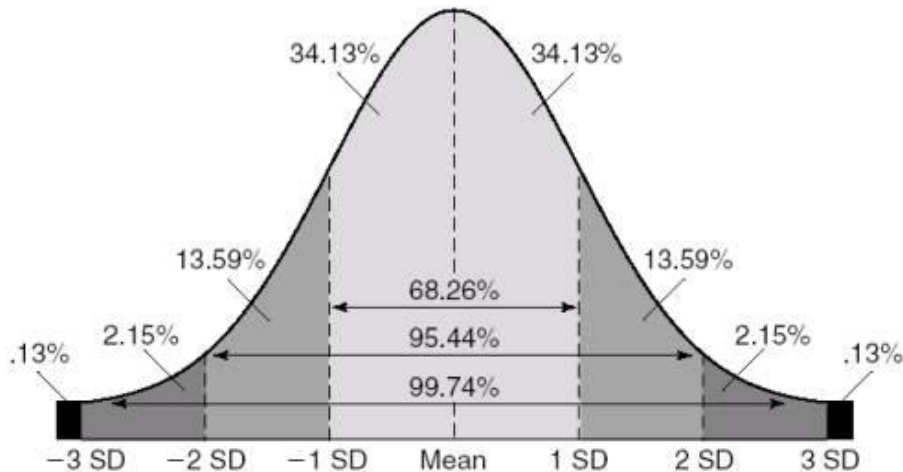
FIGURE 15.6 Examples of normal and skewed distributions

# KURVA NORMAL

- 95% nilai-nilai terletak antara  $\mu - 2\sigma$  dan  $\mu + 2\sigma$
- 5% nilai-nilai terletak di luar antara  $\mu - 2\sigma$  dan  $\mu + 2\sigma$

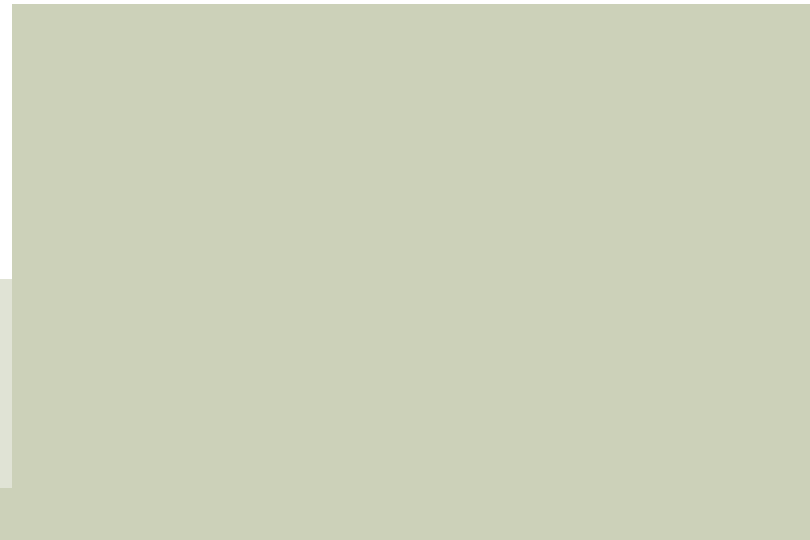
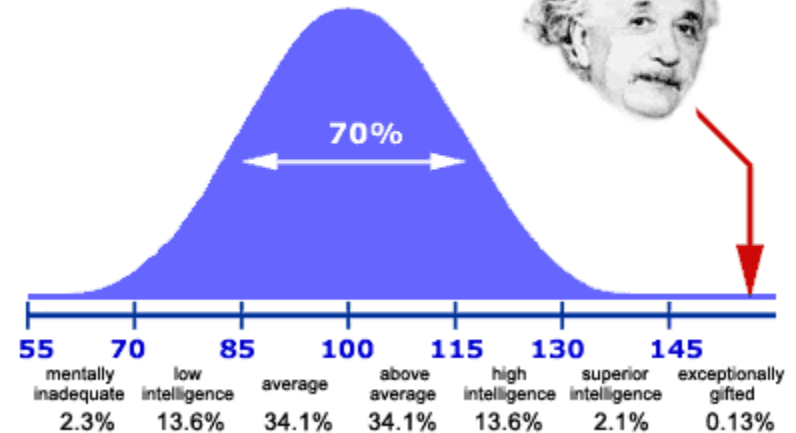
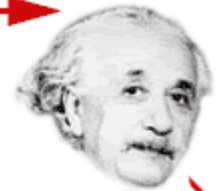


# PERSENTIL, SKOR STANDAR, DAN KURVA NORMAL



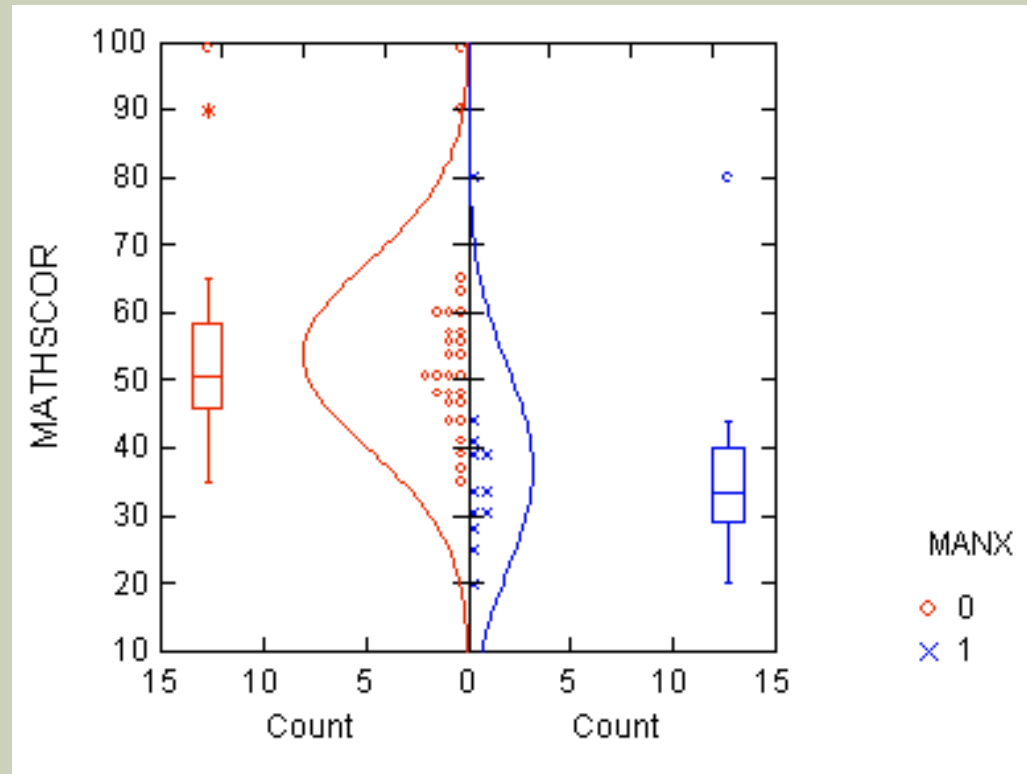
z-scores	-3	-2	-1	0	1	2	3
Percentile ranks	0.1	2	16	50	84	98	99.9
IQ	55	70	85	100	115	130	145
SAT	200	300	400	500	600	700	800

**Einstein's IQ = 160+**  
**What about yours ?**



# BOX-PLOT DAN DISTRIBUSI FREKUENSI

- Box-plot menunjukkan persentil 25, 50, 75,  $\pm 2SD$ , outlier ( $>3SD$ ,  $<-3SD$ )
- Box-plot mendeskripsikan nilai tendensi sentral dan normalitas distribusi



# STATISTIK DESKRIPTIF DAN STATISTIK INFERENSIAL

## ■ Statistik deskriptif

Metode statistik untuk mendeskripsikan statistik sampel atau parameter populasi

- Prevalensi TB paru pada populasi 1%
- Rata-rata tekanan darah sistolik populasi obes 150 mmHg

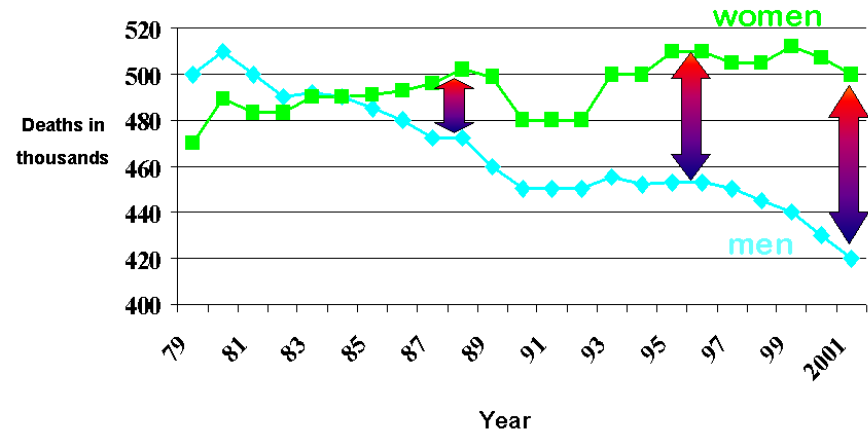
## ■ Statistik inferensial

Metode menarik kesimpulan tentang parameter populasi dengan menggunakan statistik sampel

- Merokok meningkatkan risiko sebesar 7 kali untuk mengalami PJK
- Pemberian metilprednisolon menurunkan kematian pasien dengan tetanus sebesar 20%

## Cardiovascular Disease Mortality Trends

*Women's rates are not declining in line with men's*



# STATISTIK INFERENSIAL

## ■ Statistik inferensial

Metode statistik untuk menarik kesimpulan tentang parameter populasi menggunakan statistik sampel:

1. Uji hipotesis – menguji signifikansi statistik tentang beda/ hubungan/ pengaruh variabel
2. Estimasi (penaksiran) – menaksir besarnya beda/ kekuatan hubungan/ pengaruh variabel



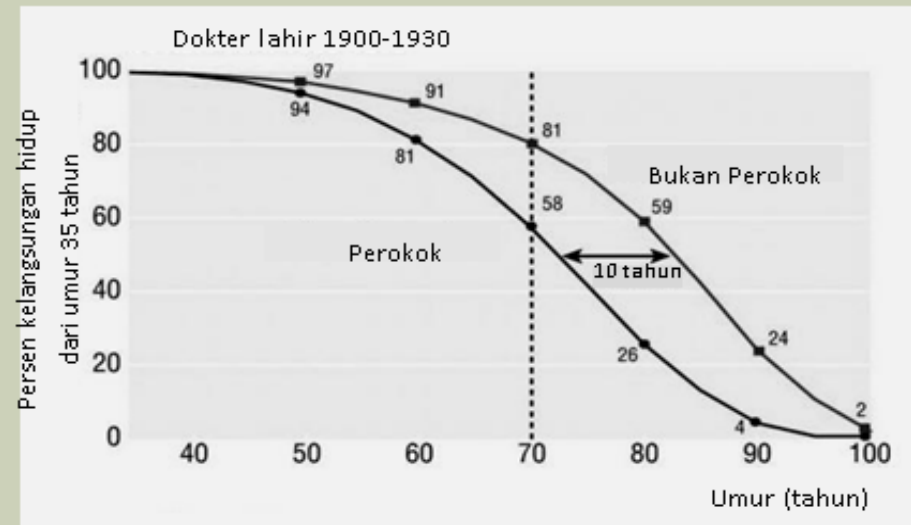
# UJI HIPOTESIS DAN ESTIMASI

## ■ Uji hipotesis

- Apakah pemberian probiotik mempercepat episode diare pada balita dengan signifikan secara statistik?
- Apakah penggunaan telepon seluler > 10 tahun berhubungan dengan peningkatan risiko neuroma otak dengan signifikan secara statistik?

## ■ Estimasi

- Berapa hari lebih pendek episode diare pada balita yang diberi probiotik?
- Berapa besar peningkatan risiko neuroma otak pada pengguna telepon seluler >10 tahun?
- Pada usia 35 tahun, berapa probabilitas perokok untuk bisa melangsungkan hidup sampai usia 70 tahun?



Gambar The British Doctors Study: Kelangsungan hidup sejak usia 35 tahun pada dokter perokok dan non-perokok.

Sumber: Doll et al. (2004)



# UJI STATISTIK

## Uji statistik:

Prosedur formal untuk menguji secara kuantitatif kesesuaian antara hasil pengamatan (data) dan hipotesis

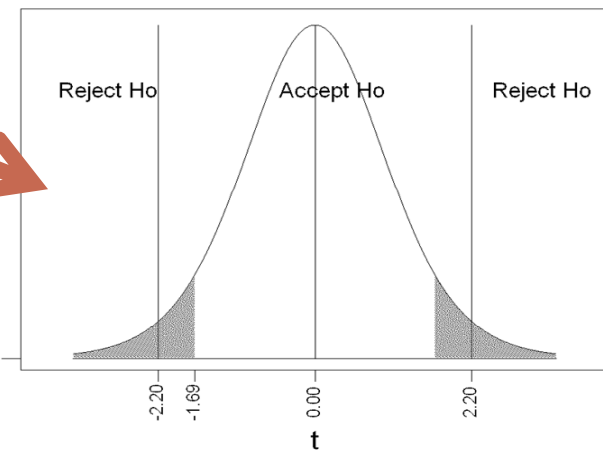
- Menolak hipotesis nol
- Mendukung hipotesis nol

## Jenis uji statistik:

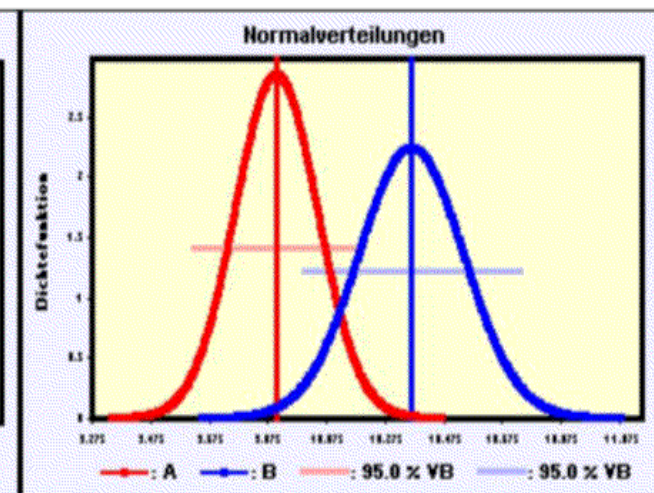
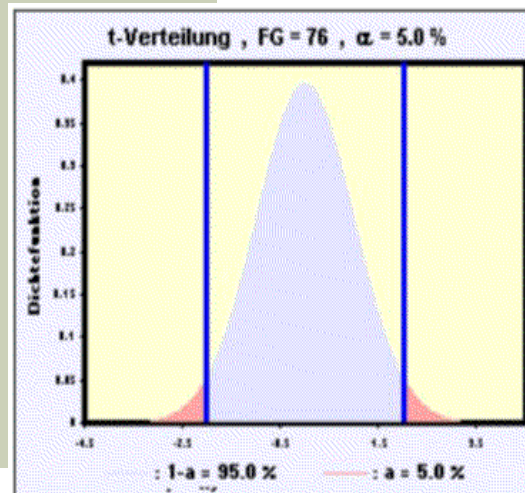
- Statistik parametrik
- Statistik non-parametrik

Statistik Parametrik

pdf



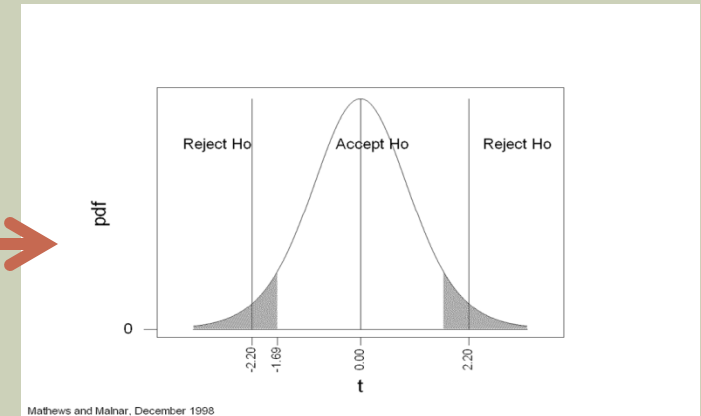
Mathews and Malnar, December 1998



# STATISTIK PARAMETRIK DAN STATISTIK NON PARAMETRIK

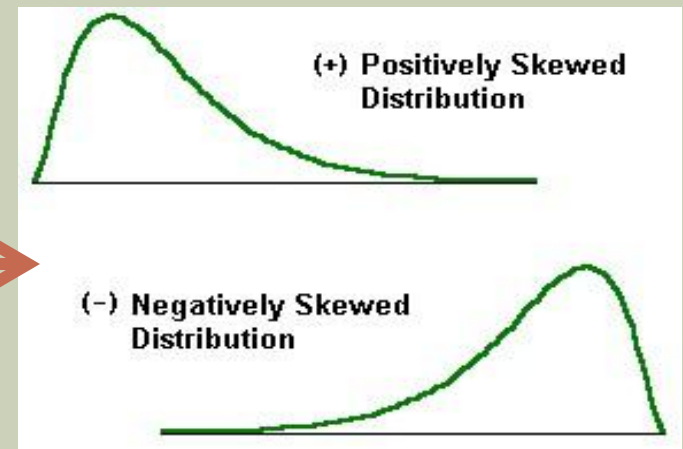
## ■ Statistik parametrik

- Prosedur pengujian hipotesis dan estimasi dengan menggunakan parameter mean dan asumsi normalitas distribusi frekuensi
- Uji t, F (Anova)

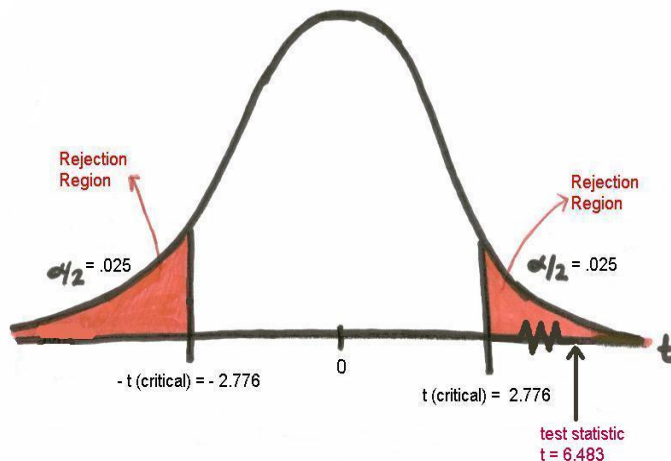
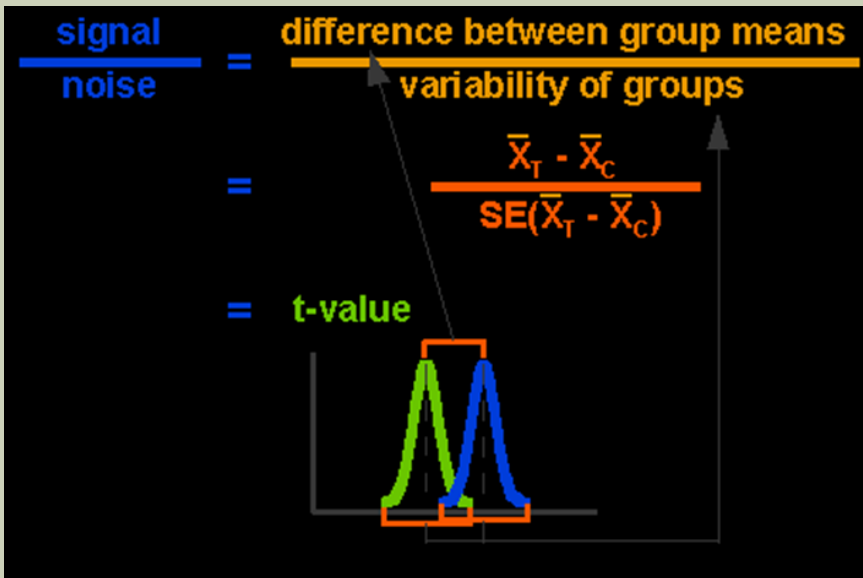


## ■ Statistika non-parametrik

- Prosedur pengujian hipotesis dan estimasi tanpa menggunakan parameter mean maupun asumsi normalitas distribusi frekuensi (distribution-free statistics)
- Uji Chi Kuadrat, Mann-Whitney, Wilcoxon, Kruskal-Wallis



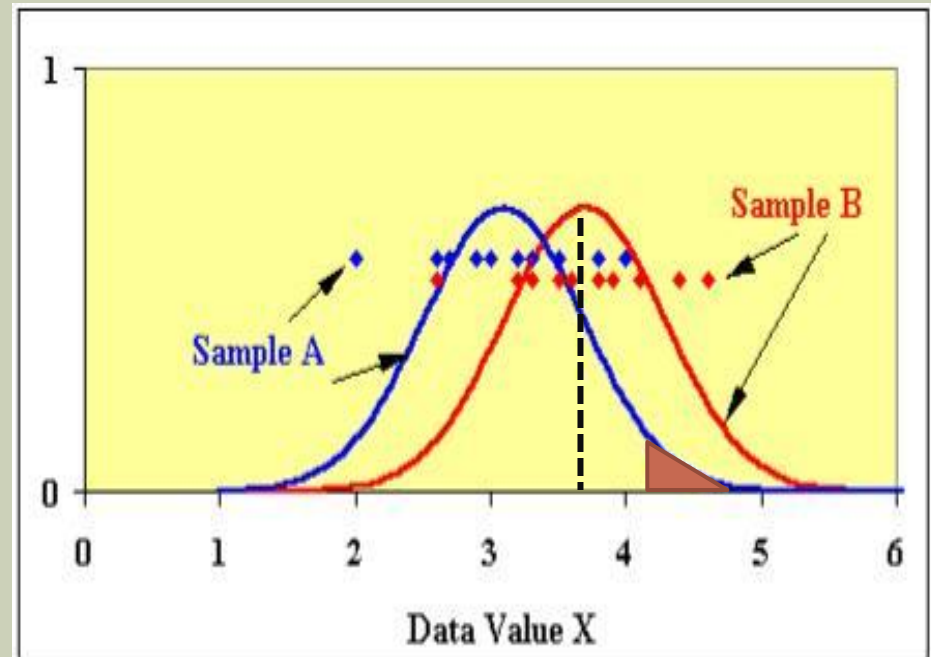
# PRINSIP UJI STATISTIK PARAMETRIK



- Populasi biru (kanan) memiliki mean  $>$  mean populasi hijau (kiri)
- Seberapa besar beda mean disebabkan karena faktor peluang (kebetulan)? Lakukan uji statistik dan lihat nilai p.
- Contoh: Uji statistik yang tepat untuk menguji beda mean dua populasi dengan distribusi frekuensi normal adalah uji t

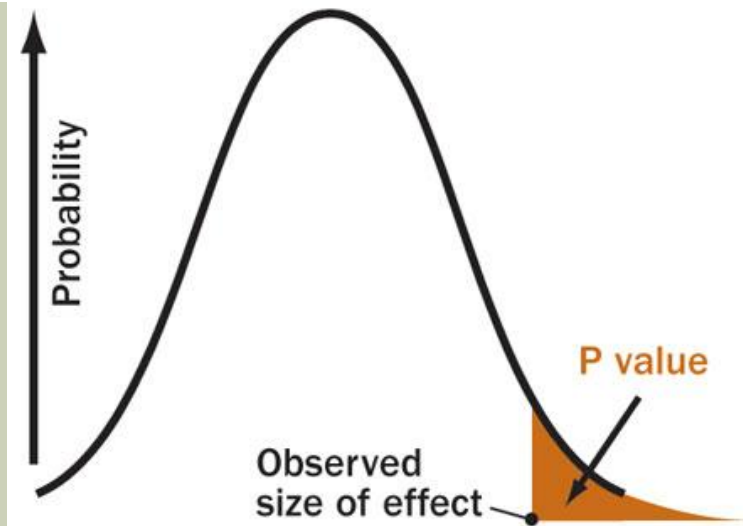
# UJI STATISTIK T

- Jika mean dari populasi merah (B) terletak di area penerimaan  $H_0$  populasi biru (A), maka walaupun ada beda tetapi perbedaan mean itu dikatakan secara statistik tidak signifikan ( $p \geq \alpha$ )
- Jika mean dari populasi merah (B) terletak di area penolakan  $H_0$  populasi biru (A), maka perbedaan mean itu dikatakan secara statistik signifikan ( $p < \alpha$ )



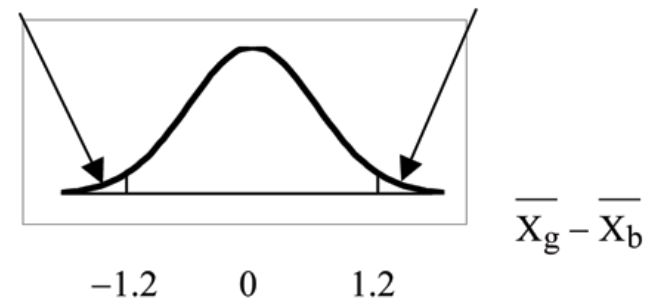
# NILAI P

- Nilai p (p value) adalah probabilitas untuk menarik kesimpulan SALAH bahwa terdapat beda/ hubungan/ pengaruh sebesar atau lebih besar daripada yang teramati, ketika  $H_0$  benar (tidak ada beda/ hubungan/ pengaruh)
- Nilai p menunjukkan besarnya peran peluang (kebetulan. Makin kecil nilai p, makin kecil beda/ hubungan/ pengaruh yang teramati terjadi karena kebetulan
- Jika nilai  $p \geq \alpha$ , maka beda itu secara statistik tidak signifikan, peran peluang besar
- Jika nilai  $p < \alpha$ , maka beda itu secara statistik signifikan, peran peluang kecil



$$\frac{1}{2} (\text{p-value}) = 0.0028$$

$$\frac{1}{2} (\text{p-value}) = 0.0028$$



From  $H_0$ ,  $\mu_g - \mu_b = 0$

# VERY IMPORTANT! FINAL WORDS

- Hasil uji statistik ditunjukkan oleh nilai p. Jadi lihatlah selalu nilai p dari uji statistik apapun
- Nilai p menunjukkan signifikansi statistik.
- Signifikansi statistik mengandung makna konsistensi temuan ketika temuan itu diulangi berkali-kali.
- Contoh: Probiotik bisa memperpendek lama diare sebesar 3 hari dengan  $p=0.002$ . Artinya, jika penelitian ini diulangi 1000 kali, maka anda akan menemukan 998 kali di antaranya memberikan kesimpulan yang sama dengan kesimpulan anda. Temuan itu baik karena menunjukkan konsistensi.
- Penting
  1. Jangan sekali kali mengebiri  $p=0.002$  menjadi  $p<0.05$  yang artinya secara statistik signifikan pada  $\alpha=0.05$ . Laporkan p apa adanya, 3 angka di belakang koma, misalnya  $p=0.002$ . Ingat  $p=0.02$  dan  $p=0.002$  sama-sama  $p<0.05$ , tetapi,  $p=0.002$  lebih konsisten daripada  $p=0.02$ , informasi itu hilang jika anda tuliskan  $p<0.05$
  2. Nilai p tidak menunjukkan validitas (kebenaran) hasil penelitian!



