Sampling Distributions (Distribusi Penarikan Contoh)

- **Sebaran (Distribusi) Peluang teoritis**
- Peubah Acak : Statistik ***Sample* , misal**

  Rata-rata dan proporsi *sample*
- Hasil semua kemungkinan *Sample* dg ukuran yg sama
- ***Sampling Distribution*: Distribusi peluang yang menyatakan peluang nilai-nilai yang mungking bagi suatu statistik contoh**

Ilustrasi *Sampling Distributions*

Misalkan ada suatu populasi……

Ukuran Populasi, *N* = 4

Peubah Acak, *X*, adalah Umur individu

Nilai-nilai *X*: 18, 20, 22, 24

diukur dalam tahun

Karakteristik Populasi

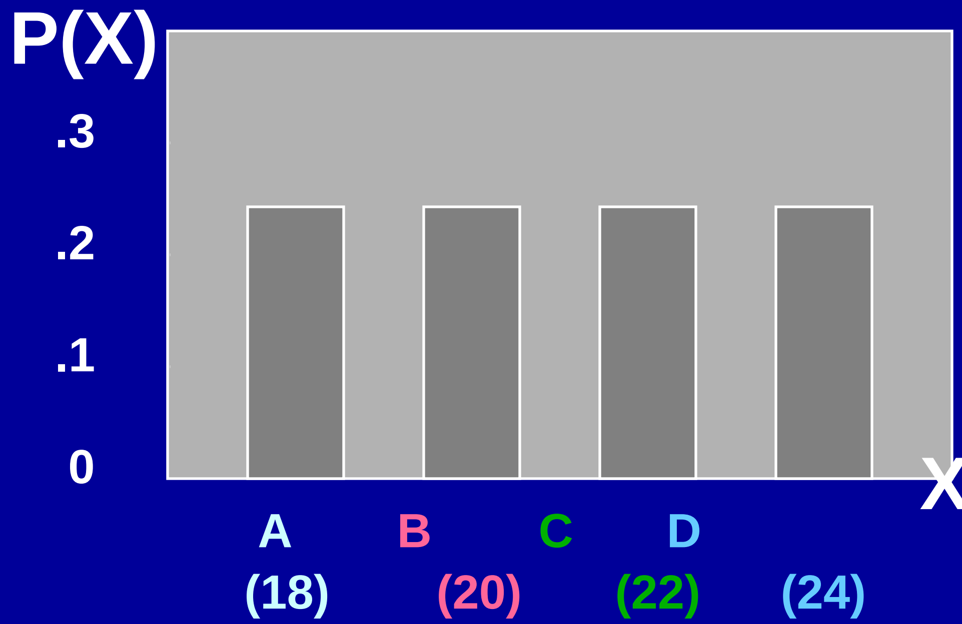
Ukuran Ringkas (Summary)

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N}$$

$$= \frac{18 + 20 + 22 + 24}{4} = 21$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} (X_i - \mu)^2}{N}} = 2.236$$

Distribusi Populasi



Distribusi Seragam

Semua kemungkinan *Sample* berukuran n = 2

1st	2nd Observation			
Obs	18	20	22	24
18	18,18	18,20	18,22	18,24
20	20,18	20,20	20,22	20,24
22	22,18	22,20	22,22	22,24
24	24,18	24,20	24,22	24,24

16 Kemungkinan *Sample*

Diambil dengan cara Pengembalian (*with replacement*)

16 Rataan *Sample*

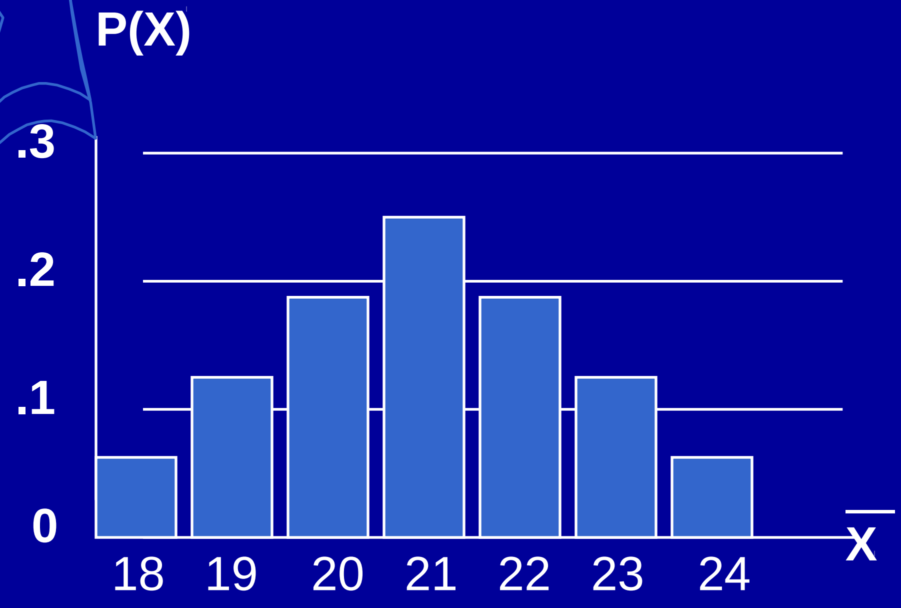
1st	2nd Observation			
Obs	18	20	22	24
18	18	19	20	21
20	19	20	21	22
22	20	21	22	23
24	21	22	23	24

Distribusi *Sampling* dari Semua kemungkinan rataan *Sample*

16 Rataan *Sample*

1st Obs	2nd Observation			
	18	20	22	24
18	18	19	20	21
20	19	20	21	22
22	20	21	22	23
24	21	22	23	24

Distribusi Rataan *Sample*



in sample = 2,

in Sampling Distribution = 16

Ukuran Ringkas untuk Distribusi *Sampling*

$$\mu_{\bar{X}} = \frac{\sum_{i=1}^{N} \bar{X}_i}{N} = \frac{18+19+19+\cdots+24}{16} = 21$$

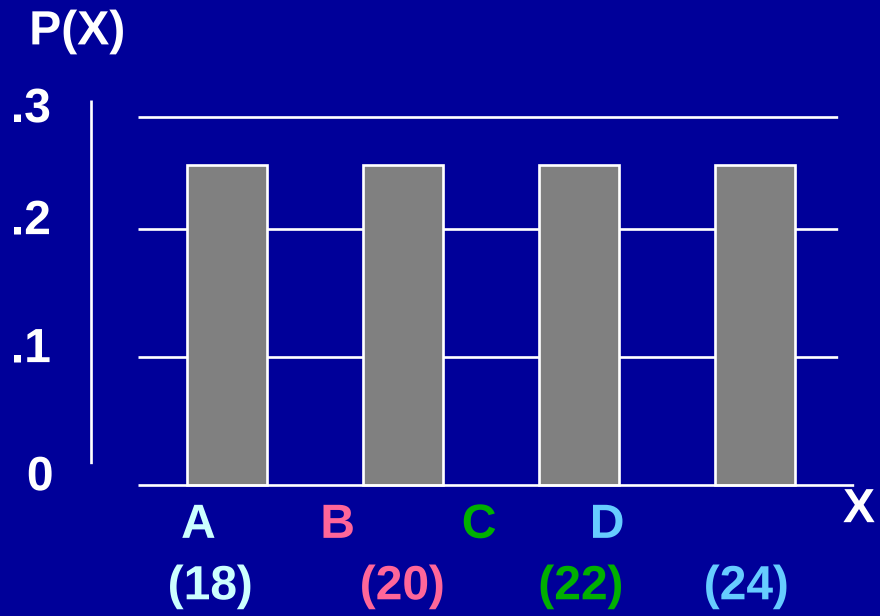
$$\sigma_{\bar{X}} = \sqrt{\frac{\sum_{i=1}^{N}\left(\bar{X}_i - \mu_{\bar{X}}\right)^2}{N}}$$

$$= \sqrt{\frac{(18-21)^2 + (19-21)^2 + \cdots + (24-21)^2}{16}} = 1.58$$

Membandingkan Populasi dgn Distribusi *Sampling*-nya

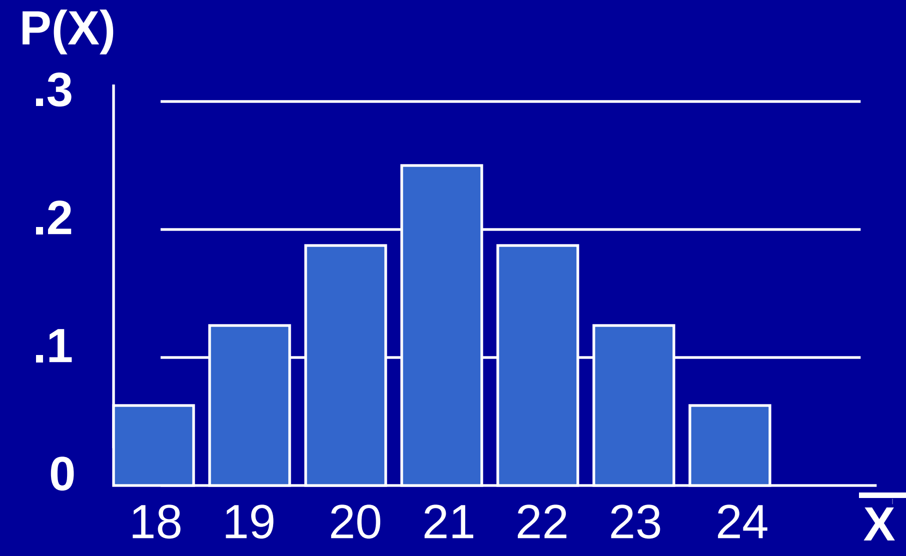
Populasi
N = 4

$$\mu = 21, \quad \sigma = 2.236$$



Distribusi Rataan *Sample*
n = 2

$$\mu_{\overline{x}} = 21 \quad \sigma_{\overline{x}} = 1.58$$



Sifat-Sifat dari Rataan Contoh (dugaan Rataan Populasi)

- **Nilai tengah Populasi sama dgn nilai harapan dugaanya** $\mu_{\bar{x}} = E(\bar{x}) = \mu$
- **Standar deviasi dugaan (dari distribusi *sampling*) kurang dari Standar Deviasi populasi**
- **Formula (*sampling with replacement*):**

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Jika n naik, maka $\sigma_{\bar{x}}$ turun

Sifat dari rataan contoh (Dugaan Rataan Populasi)

Unbiasedness (Tidak Bias)

- Nilai harapan (rata-rata dari semua kemungkinan) dugaan sama dgn nilai sebenarnya (rataan populasi)

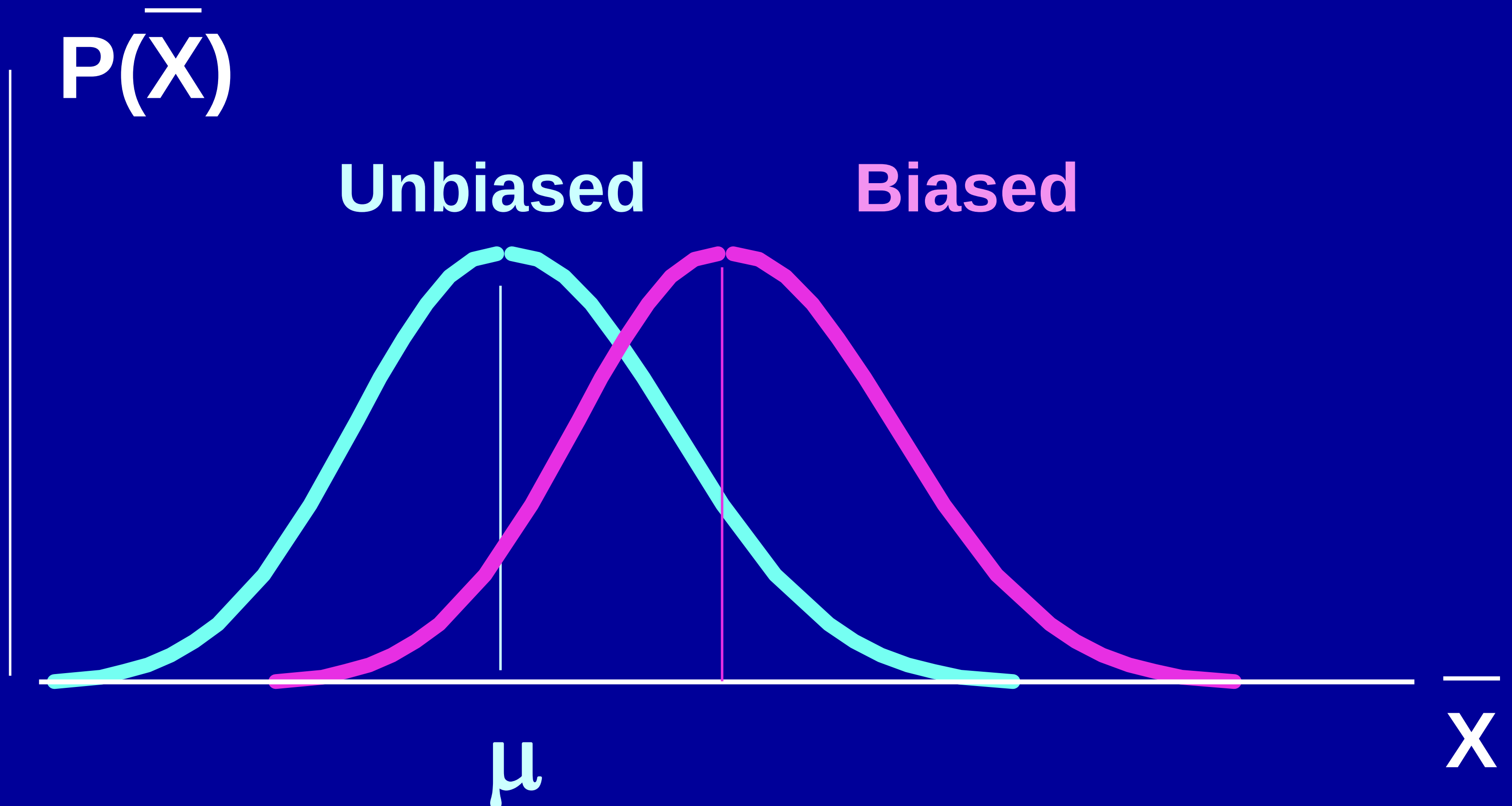
Efficiency (efisien)

- Rataan contoh variasinya lebih kecil dari penduga tak-bias lainnya

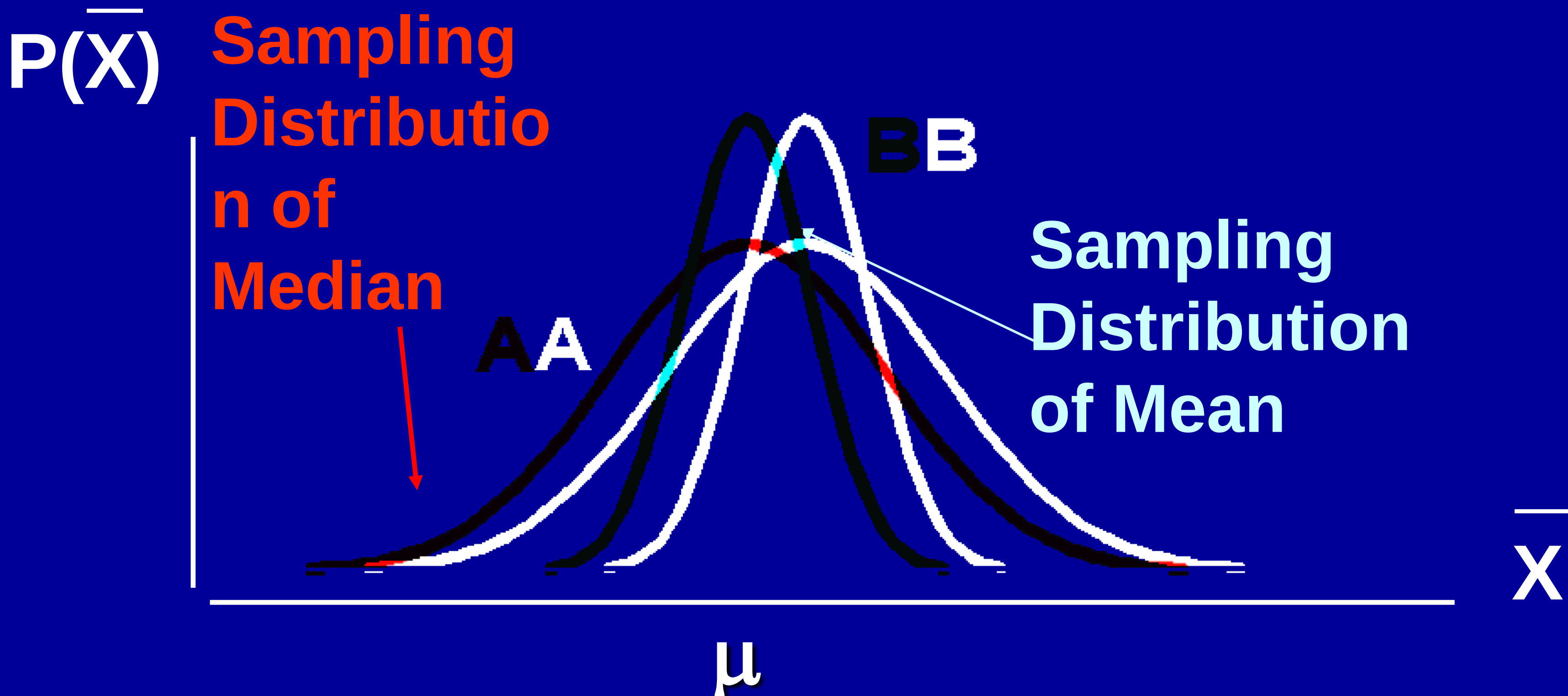
Consistency (Konsisten)

- Jika ukuran *sample* naik, variasi rataan *sample* dari rataan populasi turun

Unbiasedness

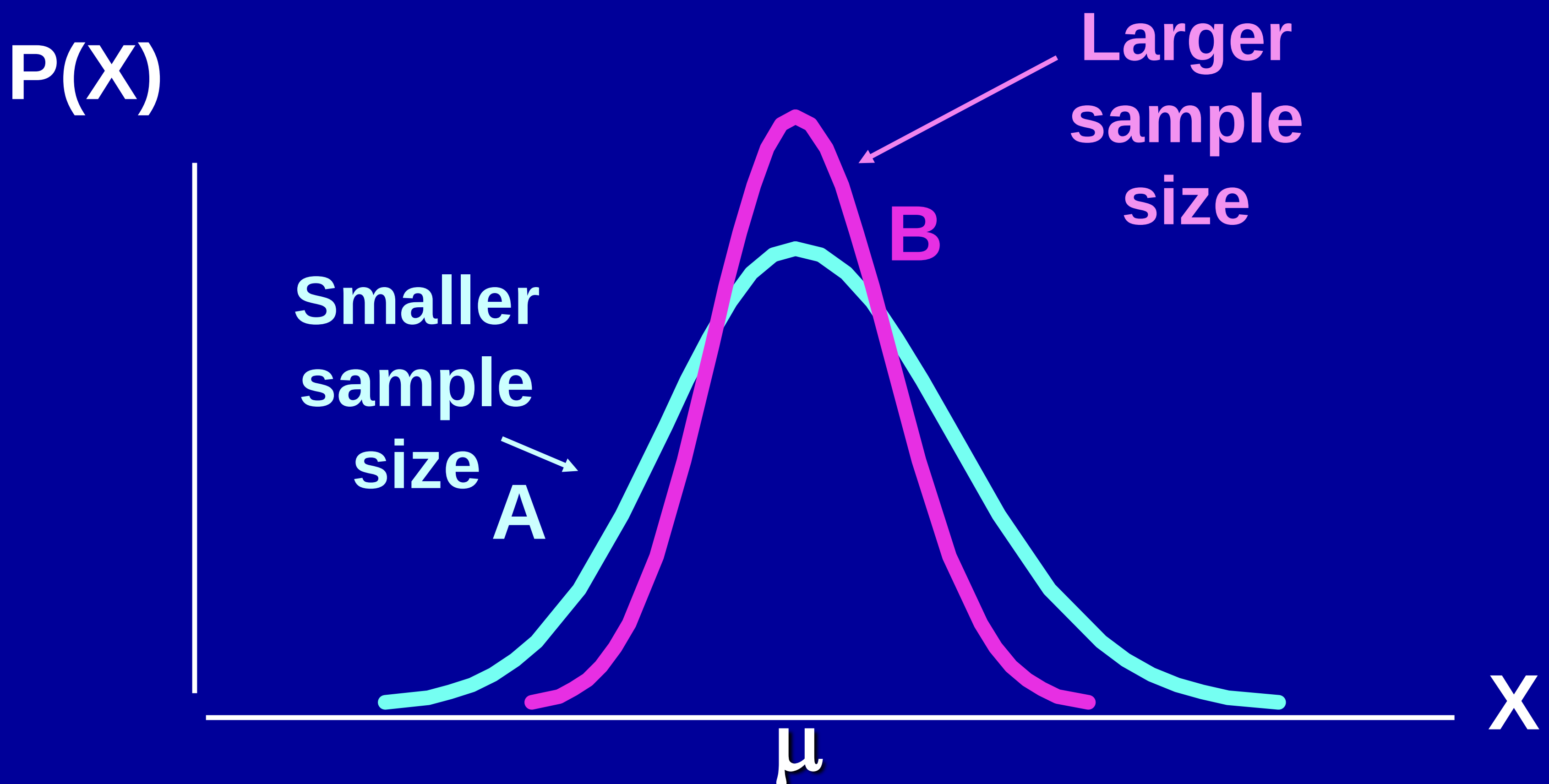


Efficiency



P($\bar{X}$)
Sampling Distributio n of Median
A
B
Sampling Distribution of Mean
$\bar{X}$
μ

Consistency



Jika Populasi Menyebar Normal

Central Tendency

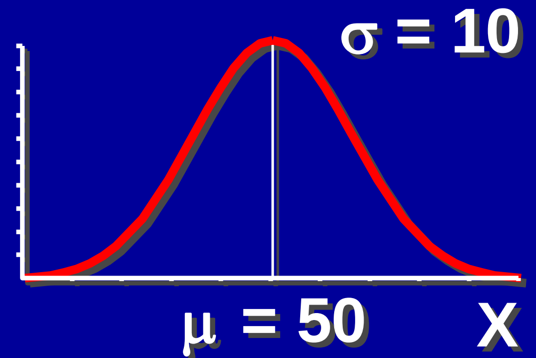
$$\mu_{\bar{X}} = \mu$$

Variation

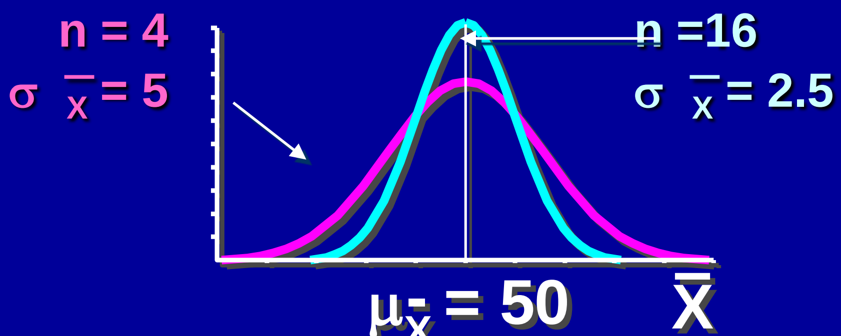
$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

***Sampling* dgn pengembalian**

Population Distribution

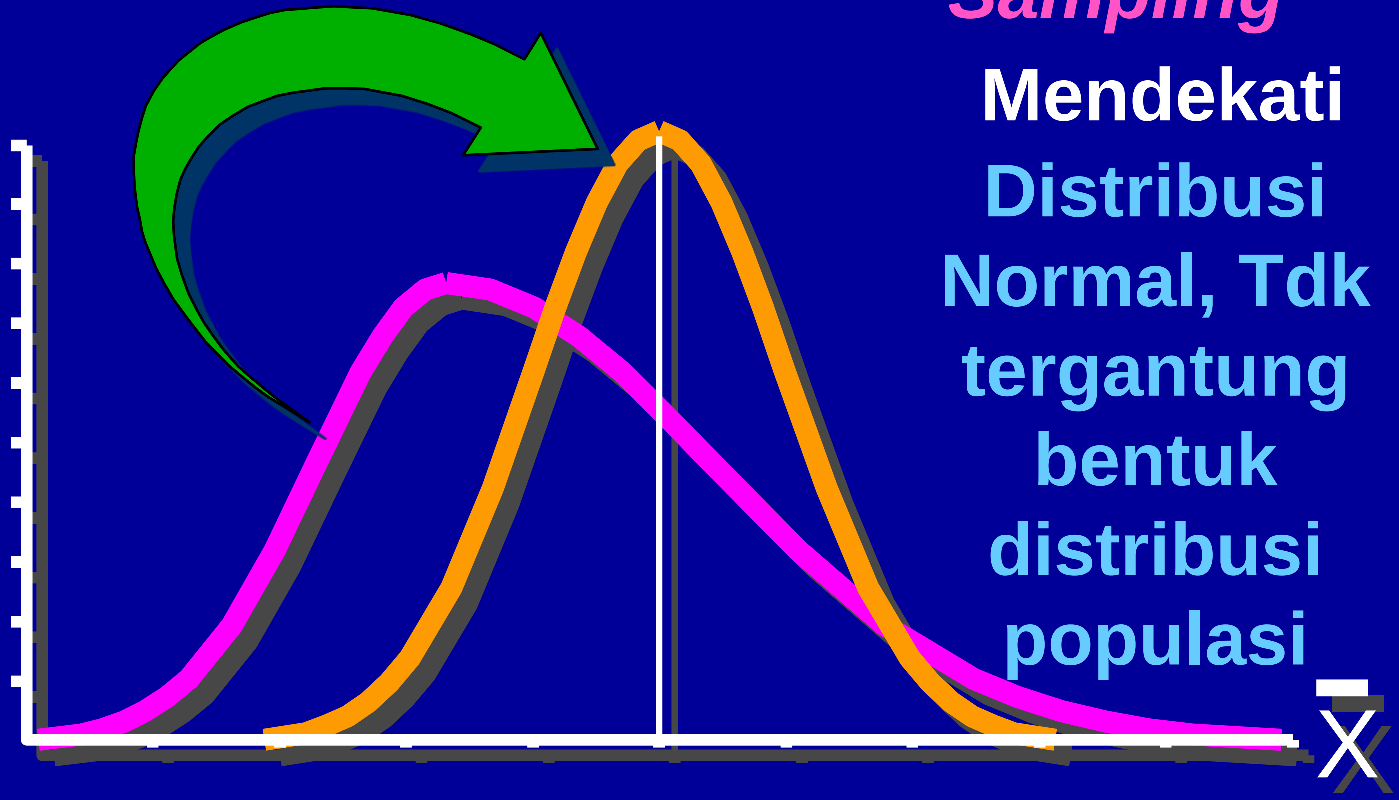


Sampling Distributions



Central Limit Theorem
(Dalil Limit Pusat)

Jika
Sample
Size
Cukup
Besar



Distribusi
Sampling
Mendekati
Distribusi
Normal, Tdk
tergantung
bentuk
distribusi
populasi

Jika Populasi Tidak Normal

Ukuran pemusatan

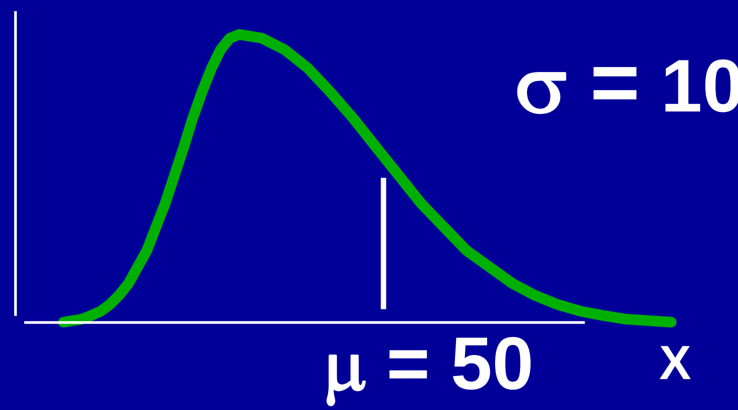
$$\mu_{\bar{x}} = \mu$$

Variasi

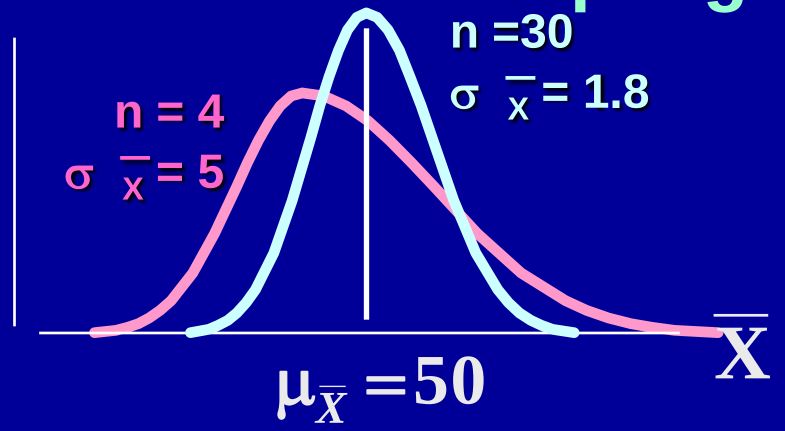
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Sampling with Replacement

Distribusi Populasi



Distribusi Sampling



Dalil Limit Pusat. Jika dari suatu populasi yang besar atau takhingga, yang mempunyai nilai tengah μ dan ragam σ^2, diambil contoh berukuran n maka rata-rata contoh $\bar{X}$ akan menyebar normal dengan nilai tengah $\mu_{\bar{x}} = \mu$ dan ragam $\sigma_{\bar{x}}^2 = \sigma^2/n$. Oleh karenanya,

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

merupakan suatu nilai bagi peubah acak normal baku Z.

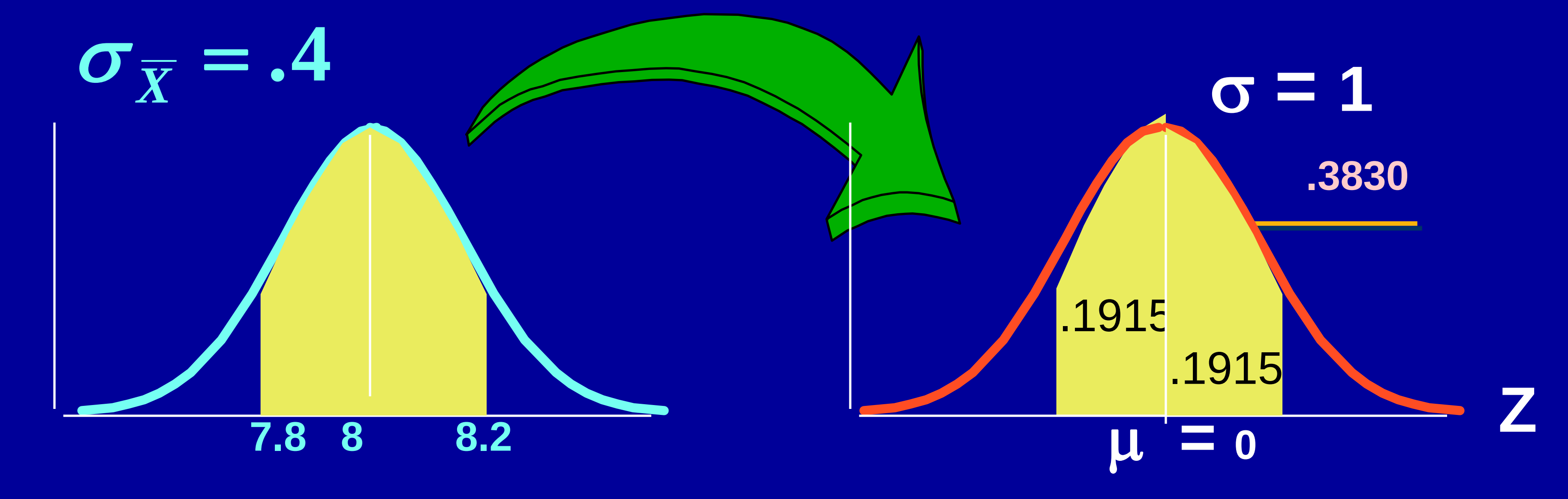
Teladan 6.1. Sebuah perusahaan memproduksi bohlam. Jika umur bohlam tersebut menyebar normal dengan nilai tengah 800 jam dan simpangan baku 40 jam, hitunglah **peluang** bahwa suatu contoh acak 16 bohlam akan mempunyai umur rata-rata kurang dari 775 jam.

Jawab. Sebaran penarikan contoh bagi $\bar{X}$ adalah normal dengan $\mu_{\bar{x}} = 800$ dan $\sigma_{\bar{x}} = 40/\sqrt{16} = 10$. Karena $z = (775-800)/10 = -2.5$, maka $P(\bar{X} < 775) = P(Z < -2.5) = 0.0062$

Teladan: Distribusi Sampling

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{7.8 - 8}{2 / \sqrt{25}} = -.50$$

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{8.2 - 8}{2 / \sqrt{25}} = .50$$



6.3 Sebaran t

Dalam praktek, karena kendala biaya dan waktu, kita seringkali hanya mengamati contoh untuk menyimpulkan karakteristik populasi. Karena tidak mengamati populasi, jarang sekali kita mengetahui ragam populasi, darimana contoh kita berasal. Untuk contoh berukuran $n \geq 30$, nilai ragam contoh s^2 dapat digunakan sebagai dugaan yang baik bagi σ^2, sehingga nilai-nilai $(\bar{x} - \mu)/(s/\sqrt{n})$ masih menyebar mendekati sebaran normal baku Z sehingga dalil limit pusat masih tetap berlaku.

Jika ukuran contoh kecil ($n<30$), nilai-nilai s^2 dari suatu contoh ke contoh lainnya berfluktuasi cukup besar, sehingga nilai-nilai $(\bar{x} - \mu)/(s/\sqrt{n})$ tidak lagi menyebar mendekati sebaran normal baku Z. untuk ukuran contoh kecil ini, kita dapat menggunakan sebaran **t-Student**, yang nilai-nilainya adalah: $t = \dfrac{\bar{x} - \mu}{s / \sqrt{n}}$

Proporsi Populasi

- **Peubah kategori (misalnya, Jenis kelamin)**
- **% populasi yg punya karakteristik tertentu**
- **Jika 2 kejadian, distribusi binomial**

  - Punya atau tdk punya karakteristik tertentu
- **Proporsi *Sample* (p_s)**

$$P_s = \frac{X}{n} = \frac{\text{jumlah sukses}}{\text{ukuran sample}}$$

Distribusi *Sampling* Proporsi

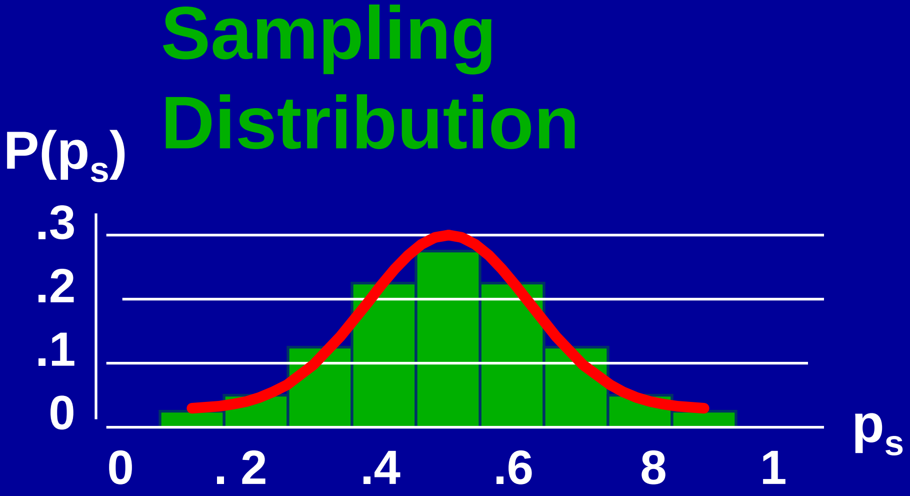
Didekati dgn distribusi normal

- $n \cdot p \geq 5$
- $n \cdot (1 - p) \geq 5$

Mean $\mu_P = p$

Standard error

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



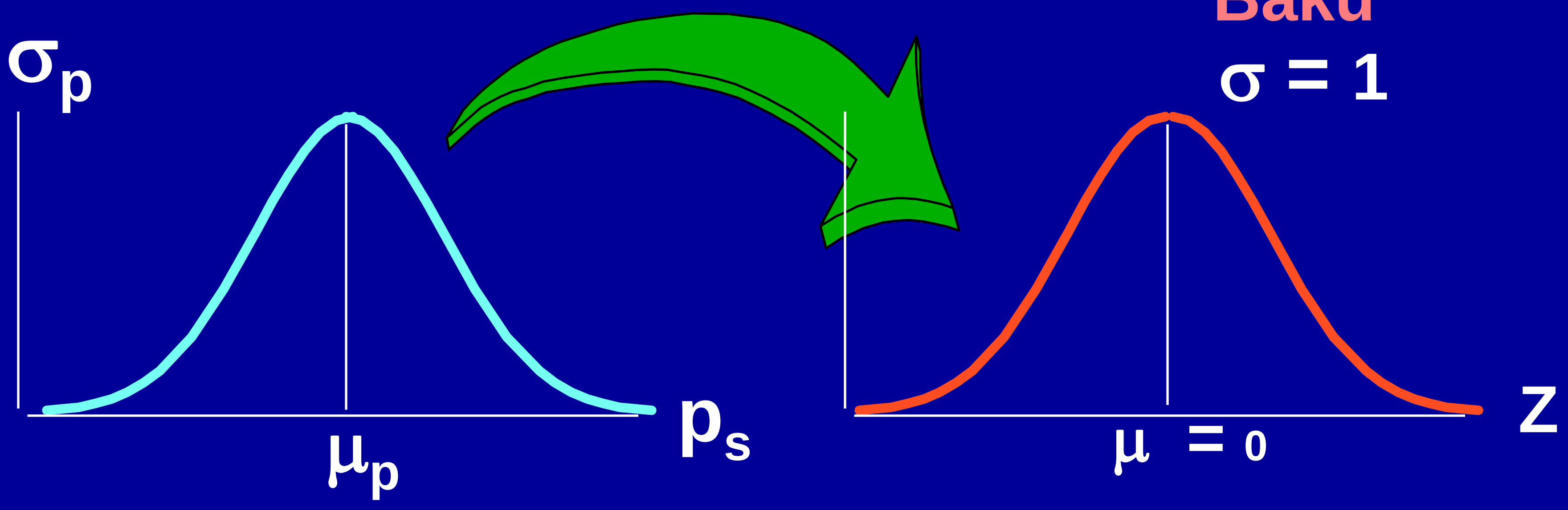
p = *proporsi populasi*

Standardisasi
Distribusi *Sampling* Proporsi

$$Z \cong \frac{p_s - \mu_p}{\sigma_p} = \frac{p_s - p}{\sqrt{\frac{p(1-p)}{n}}}$$

Distribusi *Sampling*

Distribusi Normal Baku



Teladan: Distribusi Sampling Proporsi

✓ $np \geq 5$

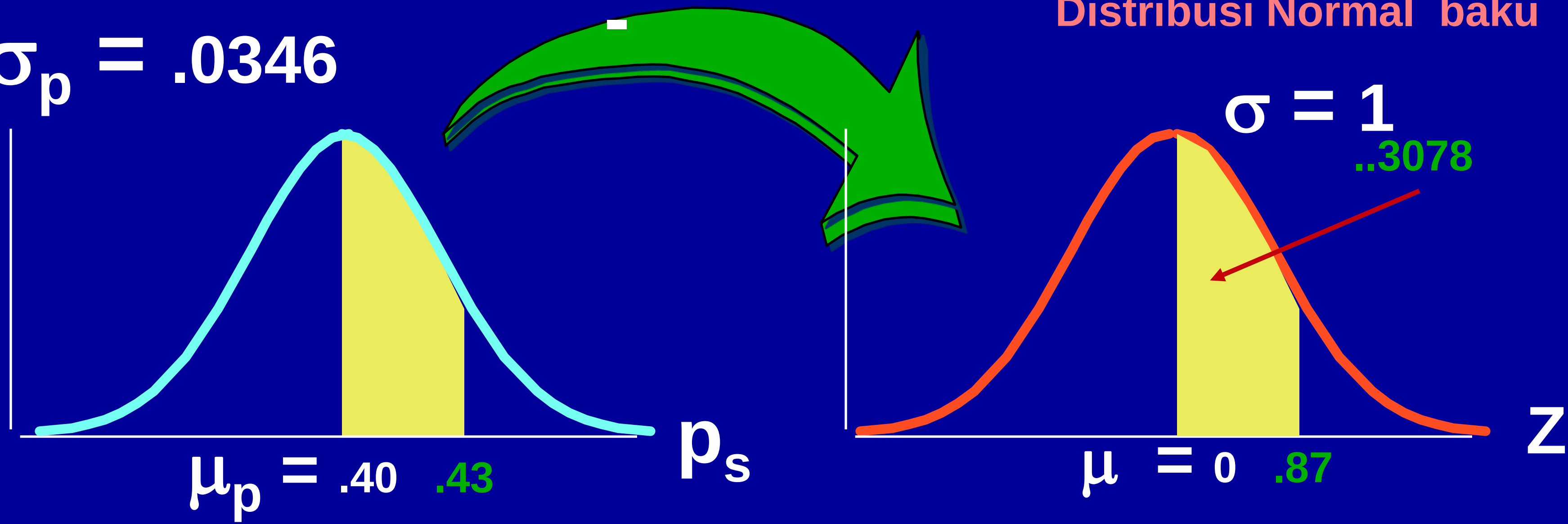
$n(1-p) \geq 5$

$$Z \cong \frac{p_s - p}{\sqrt{\dfrac{p(1-p)}{n}}} = \frac{.43 - .40}{\sqrt{\dfrac{.40 \times (1 - .40)}{200}}} = .87$$

Distribusi *Sampling*

Distribusi Normal baku

$\sigma_p = .0346$



Sampling dari Populasi Terbatas

- Modifikasi *Standard Error* jika ukuran *Sample* (*n*) besar Relatif terhadap ukuran Populasi (*N*)

$$n > .05 \cdot N \quad (\text{atau } n/N > .05)$$

- Gunakan Faktor Koreksi Populasi Terhingga
- *Standard errors* jika $n/N > .05$:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}} \qquad \sigma_P = \sqrt{\frac{p \cdot (1-p)}{n}} \cdot \sqrt{\frac{(N-n)}{(N-1)}}$$

Pengujian Hipotesis

Hipotesis: kesimpulan sementara dari penelitian, yang akan dibuktikan dengan data empiris

Utk diuji secara statistik → **hipotesis statistik** (Ho vs H1) **:** *pernyataan* (*dugaan*) *mengenai satu atau lebih parameter populasi*.

Dapat berbentuk suatu model atau nilai parameter tertentu.

Uji statistik pada hakekatnya membandingkan *apa yang diharapkan berdasarkan hipotesis* dengan *apa yang sesungguhnya diungkapkan dalam data empiris*.

Hipotesis Statistik

Ada 2 kemungkinan H_0 benar ataukah H_1 benar, tapi tidak tahu mana yg benar jika hanya mengamati data contoh.

Kemudian berdasarkan data contoh kita harus memutuskan apakah harus *terima H_0* (tolak H_1) atau *tolak H_0* (terima H_1). Dari tabel tersebut ada 4 kemungkinan kombinasi keputusan dan keadaan yang sebenarnya, yaitu **mengambil keputusan**:

Hipotesis Statistik

- 1) ***Terima H_0*** (tolak H_1) dan populasi sebenarnya memang ***H_0*** benar
- = P (terima $\mathbf{H_0}$ **/** pop $\mathbf{H_0}$)
- 2) ***Terima H_0*** (tolak H_1) padahal populasi sebenarnya ***H_1*** = P (terima $\mathbf{H_0}$ **/** pop $\mathbf{H_1}$) = β
- 3) ***Terima H_1*** (tolak H_0) dan populasi sebenarnya memang ***H_1*** benar
- = P (terima $\mathbf{H_1}$ **/** pop $\mathbf{H_1}$)
- 4) ***Terima H_1*** (tolak H_0) padahal populasi sebenarnya ***H_0*** = P (terima $\mathbf{H_1}$ **/** pop $\mathbf{H_0}$) = α