STA513 – Analisis Statistika untuk Bisnis, Ekonomi, dan Industri

Semester Ganjil 2020/2021

PERTEMUAN #2

Sebaran Penaraikan Contoh

disusun oleh:
Bagus Sartono
bagusco@apps.ipb.ac.id

Prodi Statistika dan Sains Data
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Pertanian Bogor

2020

IPB University
Bogor Indonesia

Outline

- Apa itu sebaran penarikan contoh (*sampling distribution*)
- Sebaran penarikan contoh untuk rata-rata contoh
- Sebaran penarikan contoh untuk proporsi contoh
- Sebaran penarikan contoh untuk ragam contoh

IPB University
Bogor Indonesia

Perhatikan Hasil Quick Count dari Beberapa Lembaga Survei Ini

Persentase suara untuk suatu pasangan berbeda-beda dari satu lembaga ke lembaga yang lain

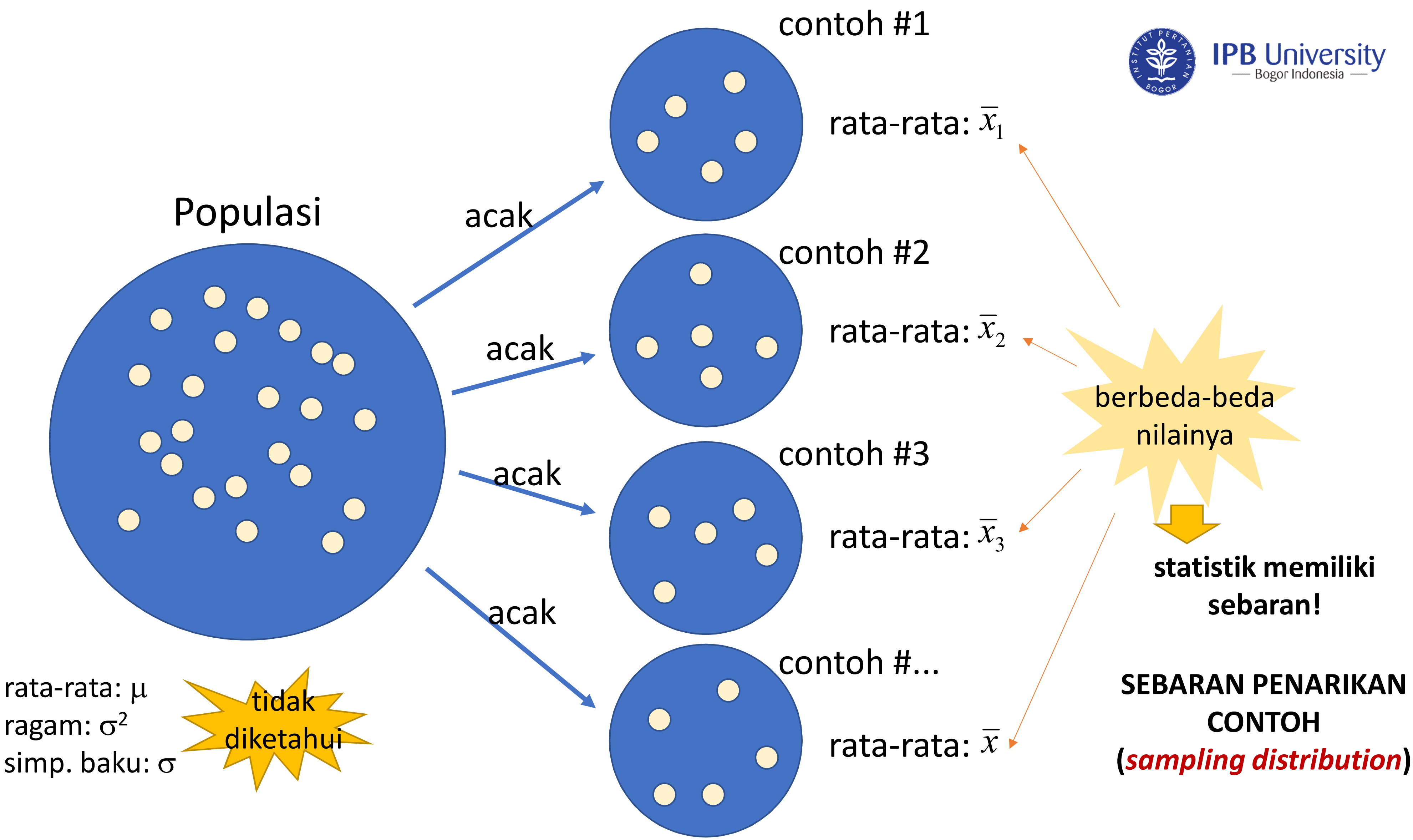
IPB University
Bogor Indonesia

Bandingkan Hasil Quick Count dengan Aktual-nya

Cagub-Cawagub	Putaran 2	
	Pemilih	%
Ahok – Djarot	2.350.366	42,04%
Anies – Sandi	**3.240.987**	**57,96%**
Jumlah suara sah	5.591.353	100,00%
Sumber	*pilkada2017.kpu.go.id*	

- Tidak ada yang sama dengan hasil KPU
- Ada yang lebih tinggi, ada yang lebih rendah

IPB University — Bogor Indonesia



Sebaran Penarikan Contoh untuk Rata-Rata Contoh

Perhatikan ilustrasi berikut

- **Andaikan terdapat suatu populasi** berukuran N=4
- Peubah yang dicatat adalah usia
- Nilai-nilai setiap individu adalah: 18, 20, 22, 24 (tahun)

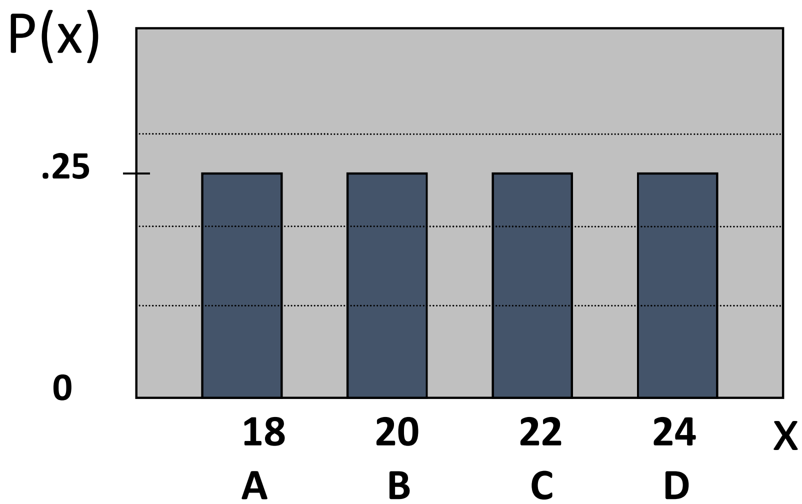
Sumber: Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc.

Ilustrasi (lanjutan)

Ringkasan parameter dari data populasi dan sebarannya adalah

$$\mu = \frac{\sum X_i}{N} = \frac{18+20+22+24}{4} = 21$$

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



Uniform Distribution

Ilustrasi (lanjutan)

Perhatikan semua kemungkinan contoh berukuran n = 2

Amatan #1	Amatan #2			
	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 contoh berukuran n = 2 (*sampling with replacement*)

16 kemungkinan nilai rata-rata contoh

Amatan #1	Amatan #2			
	18	20	22	24
18	18	19	20	21
20	19	20	21	22
22	20	21	22	23
24	21	22	23	24

Sumber: Statistics for Business and Economics, 6e © 2007 Pearson Education, Inc.

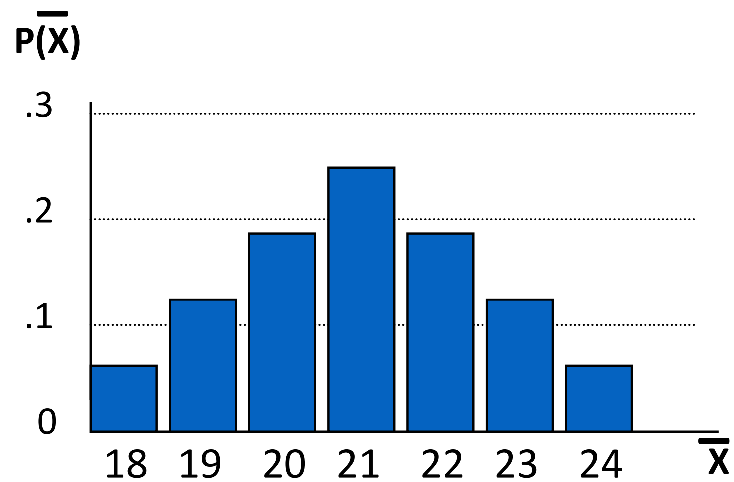
Ilustrasi (lanjutan)

Sebaran Penarikan Contoh dari Rata-Rarata

16 nilai rata-rata contoh

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

Sebaran Rata-Rata Contoh



tidak lagi UNIFORM (seragam)

cenderung berbentuk normal

Ilustrasi (lanjutan)

Ringkasan dari sebaran rata-rata contoh

$$E(\overline{X}) = \frac{\sum \overline{X}_i}{N} = \frac{18 + 19 + 21 + \cdots + 24}{16} = 21 = \mu$$

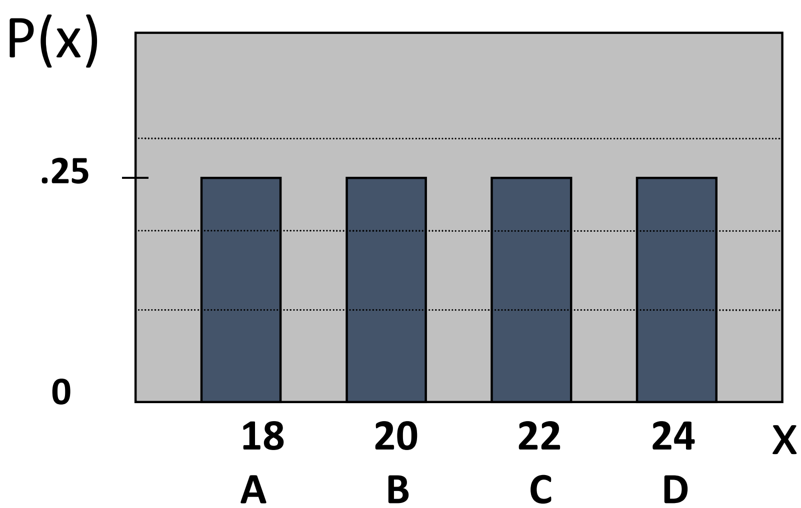
$$\sigma_{\overline{X}} = \sqrt{\frac{\sum (\overline{X}_i - \mu)^2}{N}}$$

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

Perbandingan Sebaran Populasi dan Sebaran Penarikan Contoh untuk Rata-Rata

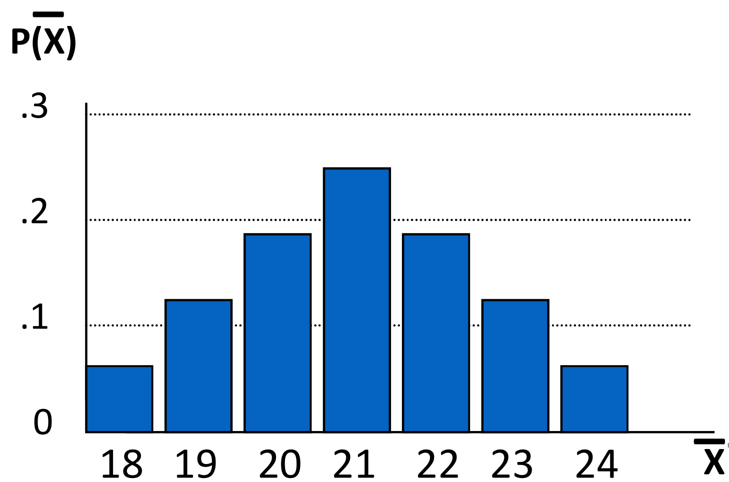
Populasi N = 4

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



Rata-Rata Contoh dengan n = 2

$\mu_{\bar{X}} = 21 \quad \sigma_{\bar{X}} = 1.58$



Dalil Limit Pusat *(central limit theorem)*

Jika dari suatu populasi dengan nilai harapan μ dan ragam σ^2 ditarik contoh secara acak berukuran n yang besar maka rata-rata contoh akan:

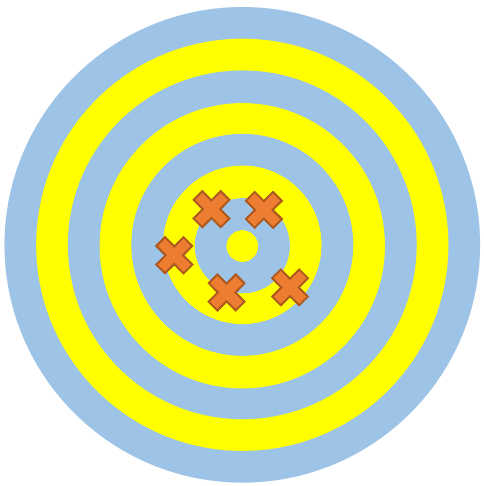
1. memiliki sebaran yang mendekati normal jika ukuran contoh (n) semakin besar
2. nilai harapan rata-rata contoh adalah μ
3. ragam dari rata-rata contoh adalah σ^2/n

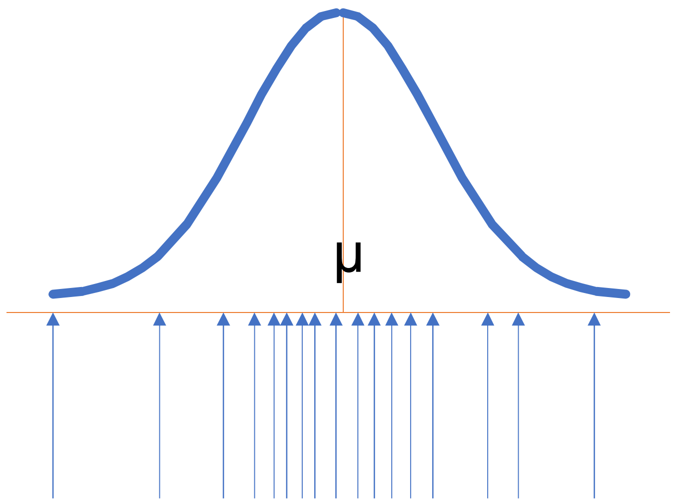
untuk $n \rightarrow \infty$

$$\bar{x} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

nilai harapan rata-rata contoh adalah μ

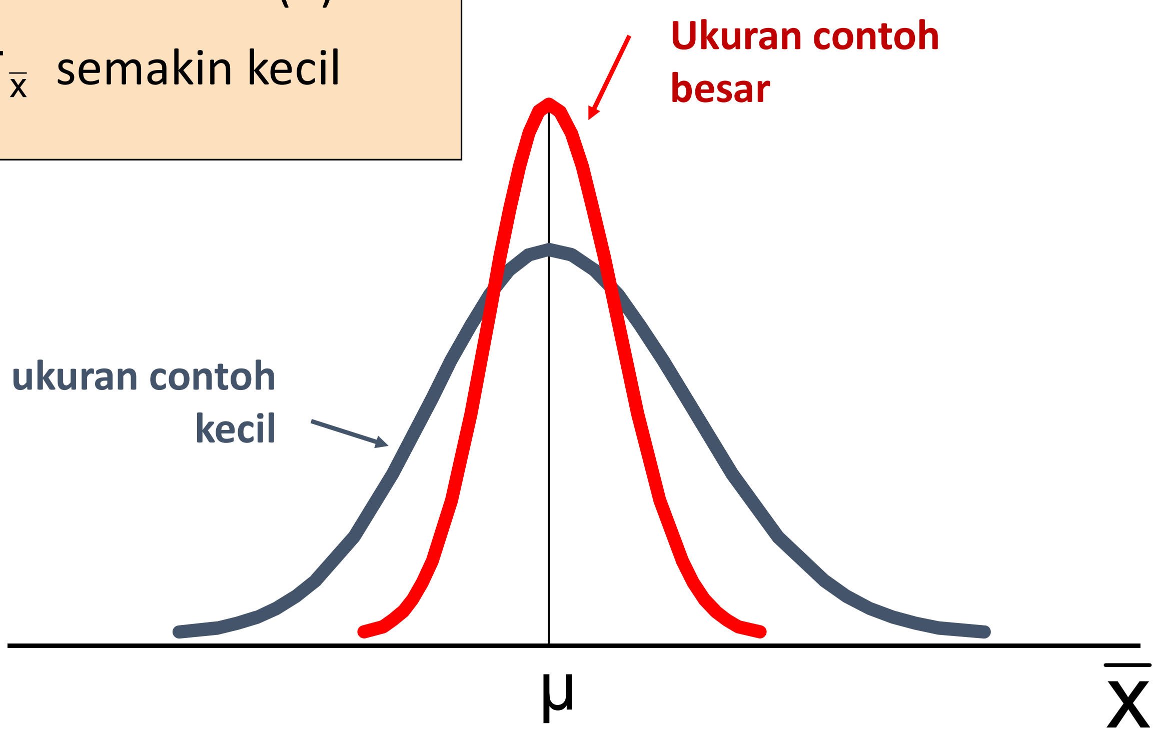
- nilai harapan rata-rata contoh sama dengan nilai harapan populasi
- rata-rata contoh adalah **penduga yang tak bias** bagi rata-rata populasi





Sifat keragaman rata-rata contoh

semakin besar
ukuran contoh (n)
$\sigma_{\bar{x}}$ semakin kecil



$\sigma_{\bar{x}}$ disebut standard error (galat baku)

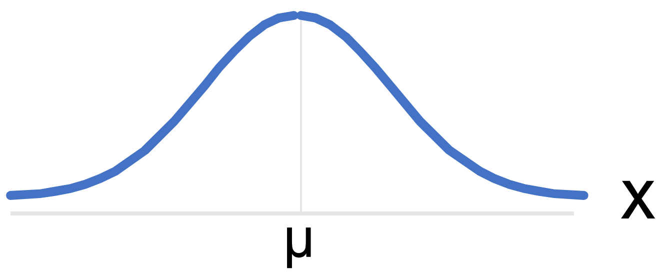
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

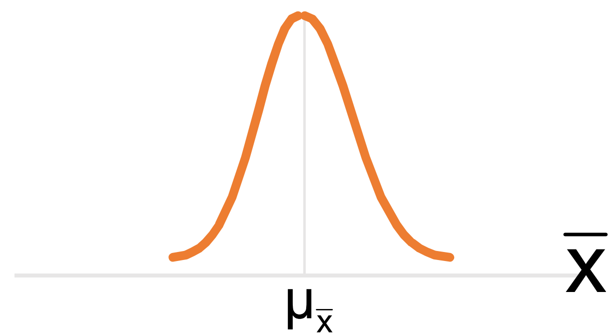
semakin besar n,
simpangan rata-rata contoh terhadap μ cenderung lebih kecil

Bentuk Sampling Distribution dari Rata-Rata Contoh

Jika contoh berasal dari populasi yang menyebar normal

... maka rata-rata contoh akan menyebar normal, berapapun *n*.

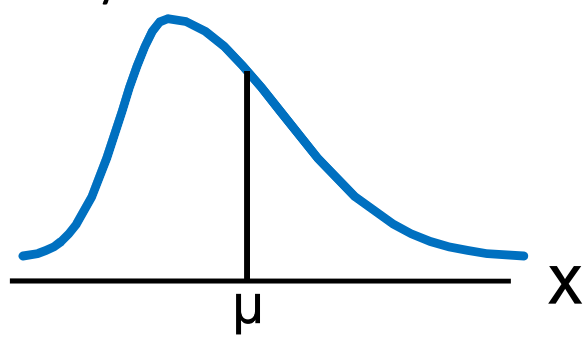


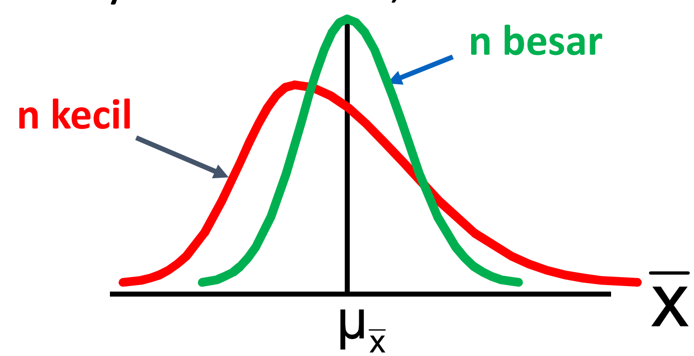


dalil limit pusat

Jika contoh berasal dari populasi yang menyebar tidak normal

... maka rata-rata contoh akan menyebar normal, asalkan *n* besar.





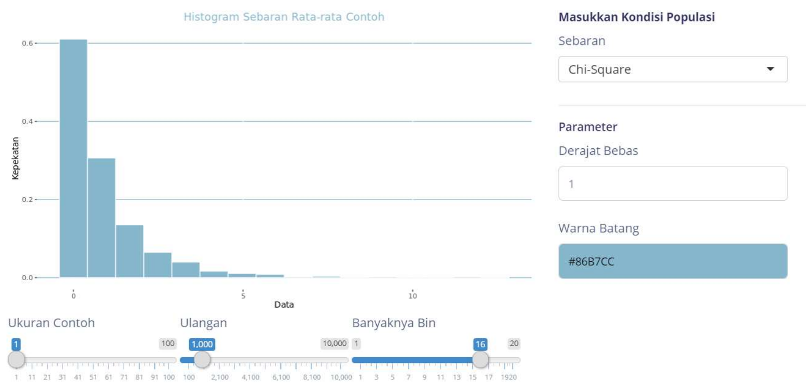
Seberapa besar n? Agar sebaran rata-rata contoh cukup dekat dengan sebaran normal…

- pada umumnya, untuk berbagai bentuk sebaran data populasi kita dapat mencapai itu ketika $n > 25$ (beberapa buku menyebut $n \geq 30$)

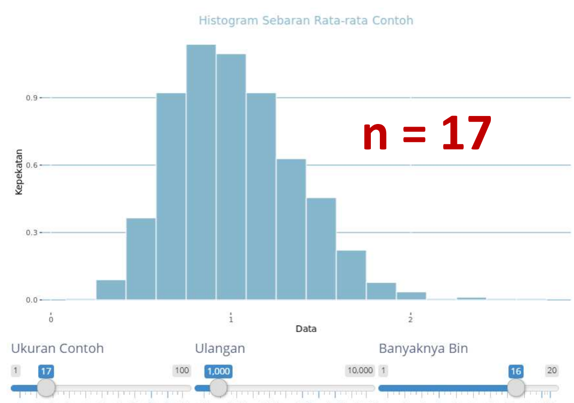
- pada sebaran data populasi yang sangat tidak simetris, diperlukan n yang lebih besar lagi

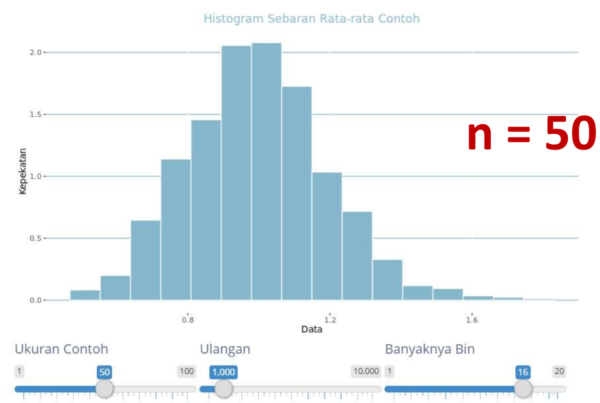
- gunakan aplikasi http://shiny.stat.ipb.ac.id/bagusco/stk211/ pada menu SIMULASI SEBARAN CONTOH untuk melakukan simulasi tentang hal ini

Seberapa besar n? Agar sebaran rata-rata contoh cukup dekat dengan sebaran normal…



URL aplikasi:
http://shiny.stat.ipb.ac.id/bagusco/stk211/





Ilustrasi

- Andaikan kita memiliki suatu populasi data banyaknya buku yang dipinjam siswa dalam satu semester dengan $\mu = 8$ dan simpangan baku $\sigma = 3$.
- Sebuah contoh acak $n = 36$ diambil dari populasi tersebut.
- Berapa peluang rata-rata contoh yang diperoleh berada pada rentang 7.8 and 8.2?
- Meskipun tidak ada keterangan bahwa sebaran populasi normal, berdasarkan dalil limit pusat kita dapat menyebutka bahwa sebaran rata-rata contoh mendekati normal dengan rata-rata $\mu = 8$ dan simpangan baku $3/\sqrt{36} = 3/6 = 0.5$

sehingga $P(7.8 < \bar{x} < 8.2) = 0.3108$

=NORM.DIST(8.2, 8, 0.5, 1) - NORM.DIST(7.8, 8, 0.5, 1)

mean
(μ)

stdev

Ilustrasi

- Andaikan kita memiliki suatu populasi data banyaknya buku yang dipinjam siswa dalam satu semester dengan $\mu = 8$ dan simpangan baku $\sigma = 3$.
- Sebuah contoh acak $n = 36$ diambil dari populasi tersebut.
- Berapa peluang rata-rata contoh yang diperoleh berada pada rentang 7.8 and 8.2?
- Meskipun tidak ada keterangan bahwa sebaran populasi normal, berdasarkan dalil limit pusat kita dapat menyebutka bahwa sebaran rata-rata contoh mendekati normal dengan rata-rata $\mu = 8$ dan simpangan baku $3/\sqrt{36} = 3/6 = 0.5$

sehingga
$$P(7.8 < \bar{x} < 8.2) = P\left(\frac{7.8-8}{0.5} < Z < \frac{8.2-8}{0.5}\right)$$
$$= P(-0.4 < Z < 0.4)$$
$$= P(Z < 0.4) - P(Z < -0.4)$$
$$= 0.6554 - 0.3446 = 0.3108$$

=NORM.S.DIST(0.4,1) - NORM.S.DIST(-0.4,1)

Sebaran Penarikan Contoh untuk Proporsi Contoh

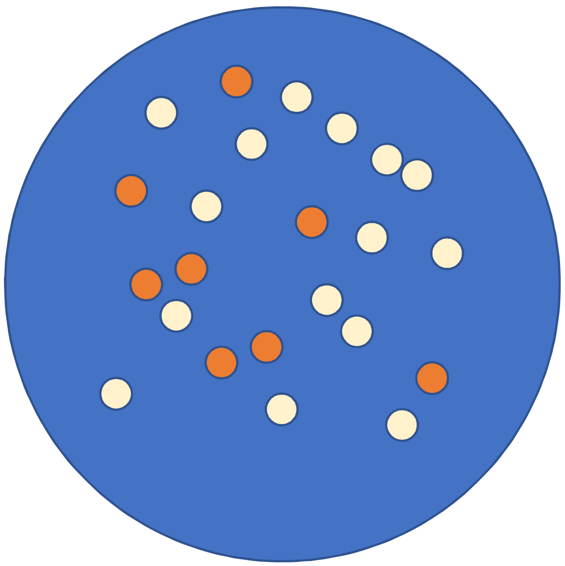
IPB University
Bogor Indonesia

Sebaran Penarikan Contoh dari Proporsi

- proporsi adalah fraksi antara banyaknya kejadian tertentu dibagi dengan banyaknya amatan
- misal: dari 200 orang terdapat 120 perempuan, maka proporsi perempuan adalah 120/200 atau 0.6 atau 60%
- andaikan nilai pada data kita tuliskan 1 untuk perempuan dan 0 untuk laki-laki…. maka proporsi 0.6 juga dapat dituliskan sebagai jumlah dari 200 buah data bernilai 1 dan 0 dibagi 200
- formula proporsi menyerupai rata-rata
- karenanya… sifat sebaran penarikan contoh dari proporsi juga sama dengan rata-rata

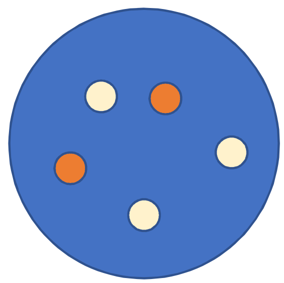
Sebaran Penarikan Contoh dari Proporsi

Populasi

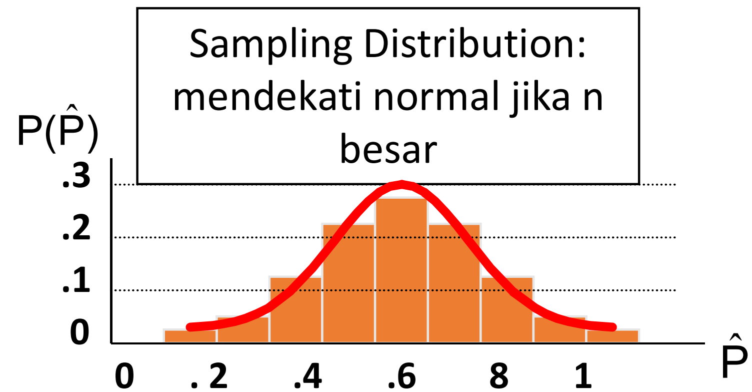


proporsi pada populasi = p

contoh



proporsi pada contoh = $\hat{p}$
(baca: p hat)



$$E(\hat{p}) = p$$

$$\sigma_{\hat{p}}^2 = \frac{p(1-p)}{n}$$

Ilustrasi

- Jika populasi yang sesungguhnya dari pendukung pasangan A adalah p = 0.4, dan dari populasi dilakukan survei dengan mengambil contoh 200 orang, berapa peluang survei menghasilkan proporsi antara antara 0.5 dan 0.6?

Jika p = 0.4 dan n = 200, berapa P(0.5 < phat < 0.6)?

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{.4(1-.4)}{200}} = .03464$$

$$P(0.5 < \hat{p} < 0.6) = 0.00195$$

=NORM.DIST(0.6, 0.4, 0.03464, 1) - NORM.DIST(0.5, 0.4, 0.03464, 1)

Sebaran Penarikan Contoh untuk Ragam Contoh

IPB University
Bogor Indonesia

Ragam

- Andaikan $x_1, x_2, . . . , x_n$ adalah contoh acak dari suatu populasi. Ragam contoh adalah

$$\boxed{s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

- Akar kuadrat dari ragam disebut simpangan baku

Sebaran dari Ragam Contoh

- Sebaran penarikan contoh dari s^2 memiliki nilai harapan dan ragam sebagai berikut:

$$\boxed{E(s^2) = \sigma^2} \qquad \boxed{Var(s^2) = \frac{2\sigma^4}{n-1}}$$

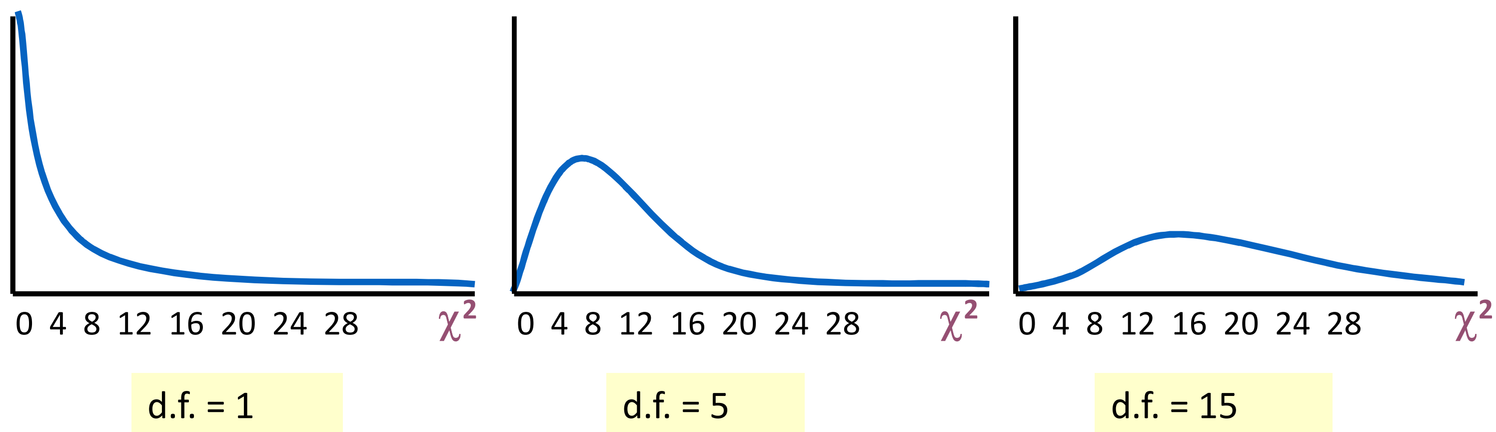
- Jika populasi memiliki sebaran normal, maka

$$\boxed{\frac{(n-1)s^2}{\sigma^2}}$$

memiliki sebaran χ^2 (chi-square, khi-kuadrat) dengan derajat bebas n – 1

Bentuk Sebaran Chi-Square

- Bentuk sebarannya tergantung pada nilai derajat bebas (degree of freedom, df)



Tentang Derajat Bebas

ide: banyaknya amatan yang nilainya bebas setelah rata-rata contoh diketahui

Ilustrasi: Rata-rata dari 3 bilangan adalah 8

X1 = 7
X2 = 8
X3 = ?

Karena rata-ratanya adalah 8 maka X3 haruslah bernilai 9 (nilai X3 tidak bebas)

Jadi untuk n = 3, derajat bebasnya adalah df = 3 – 1 = 2

(2 nilai bisa sembarang, tapi yang ketiga tidak jika rata-ratanya sudah diketahui)

INSTITUT PERTANIAN BOGOR

IPB University

— Bogor Indonesia —

Inspiring Innovation with Integrity

in Agriculture, Ocean and Biosciences for a Sustainable World