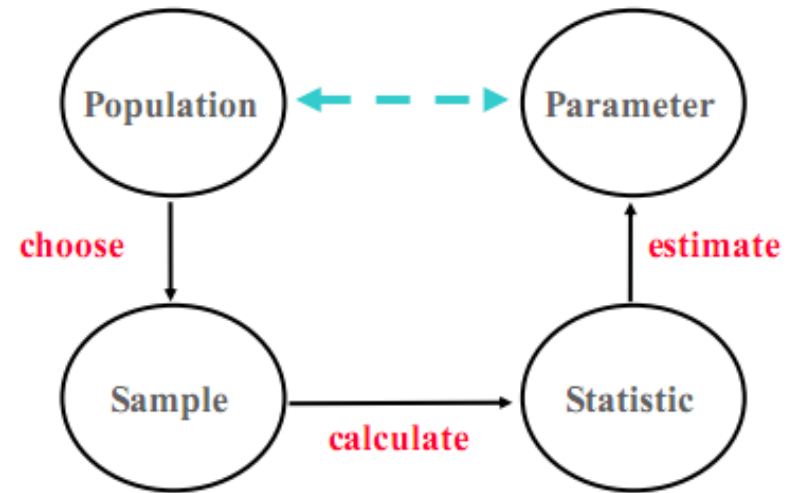
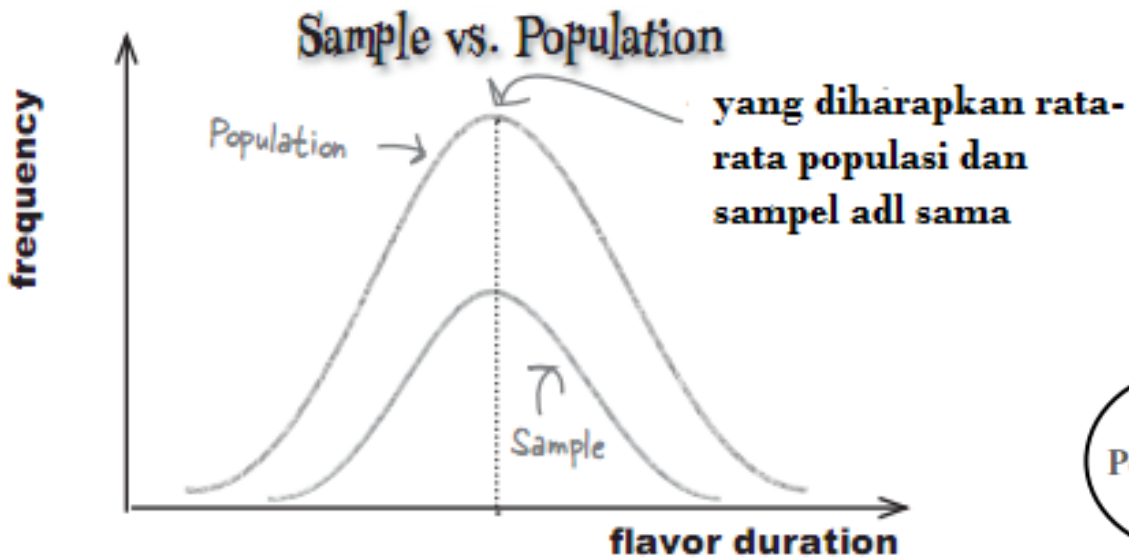




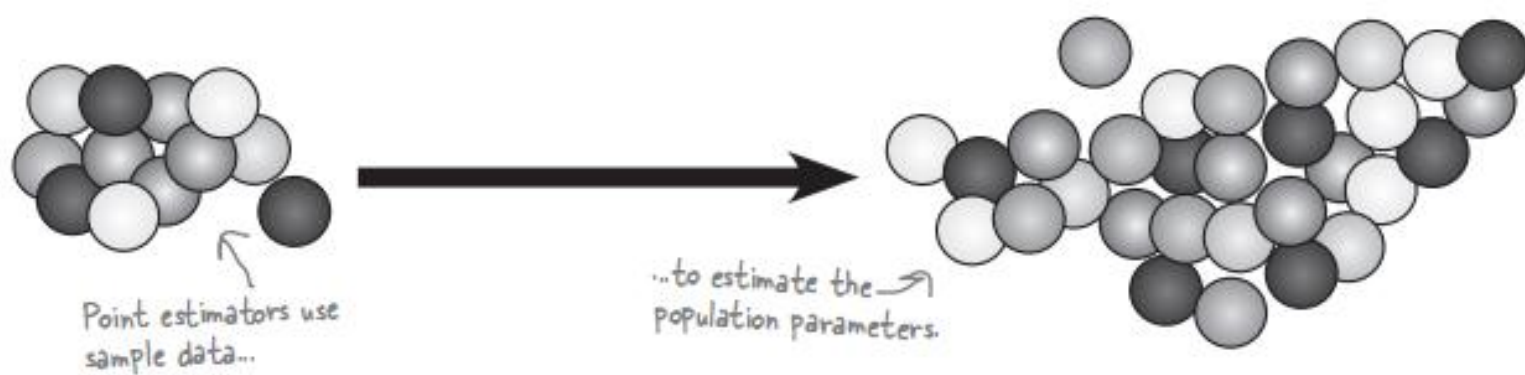
BAB 3

ESTIMASI

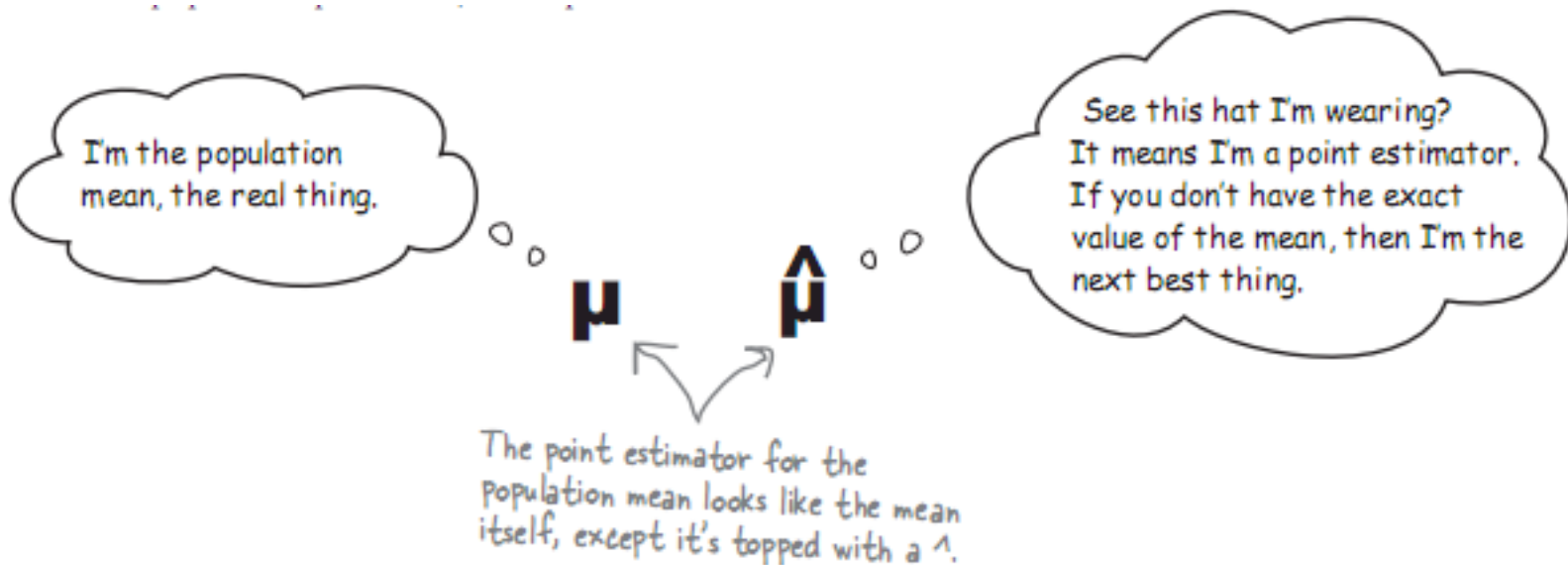
Ilustrasi



- Mungkin rata-ratanya tidaklah sama tapi estimasi terbaik dapat ditentukan
- Rata-rata sampel → estimator titik untuk rata-rata populasi
→ Artinya jika dihitung sampel data akan mengestimasi rata-rata populasi



- Estimator titik dapat menghasilkan nilai pendekatan suatu parameter populasi



A **point estimate** of some population parameter θ is a single numerical value $\hat{\theta}$ of a statistic $\hat{\Theta}$. The statistic $\hat{\Theta}$ is called the **point estimator**.

Estimasi titik untuk rata-rata

$\bar{x}$ is the mean of the sample. $\rightarrow \bar{\mathbf{x}} = \frac{\Sigma \mathbf{x}}{\mathbf{n}}$ $\leftarrow$ Add together the numbers in the sample, and divide by how many there are.

We estimate the mean of the population... $\rightarrow \hat{\mu} = \bar{\mathbf{x}} \leftarrow$...using the mean of the sample.

Misalkan variabel random X berdistribusi Normal dengan rata-rata μ tidak diketahui. Misalkan diambil $x_1=25$, $x_2=30$, $x_3=29$ dan $x_4=31$, maka dapat diestimasi rata-ratanya adalah

$$\bar{x} = \frac{25 + 30 + 29 + 31}{4} = 28.75$$

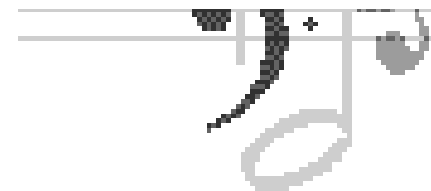


Estimasi titik untuk variansi

Population variance $\rightarrow \sigma^2 = \frac{\sum (x - \mu)^2}{n}$ $\leftarrow$ Population mean
 $\leftarrow$ Size of the population

Point estimator
for the
population
variance $\rightarrow \hat{\sigma}^2 = s^2$

Estimator for the
population variance $\rightarrow \hat{\sigma}^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$ $\leftarrow$ Take each item in the sample, subtract the sample mean, square the result, then add the lot together.
 $\leftarrow$ Divide by the number in the sample minus 1.



latihan

Tentukan estimasi titik untuk rata-rata dan variansi untuk data :

61.9 62.6 63.3 64.8 65.1 66.4 67.1 67.2 68.7 69.9



Estimasi untuk Proporsi

- Jika X menggambarkan jumlah sukses dalam suatu populasi maka X mengikuti distribusi $\text{Bin}(n, p)$
- Misal akan diestimasi rata-rata populasi ???
- Estimasi rata-rata populasi $\rightarrow$ rata-rata sampel
- Proporsi sukses dalam populasi $\rightarrow$ proporsi sukses dalam sampel

Point estimator for the proportion of successes in the population $\rightarrow \hat{p} = p_s \leftarrow$ Proportion of successes in the sample

dengan

$$p_s = \frac{\text{number of successes}}{\text{number in sample}}$$



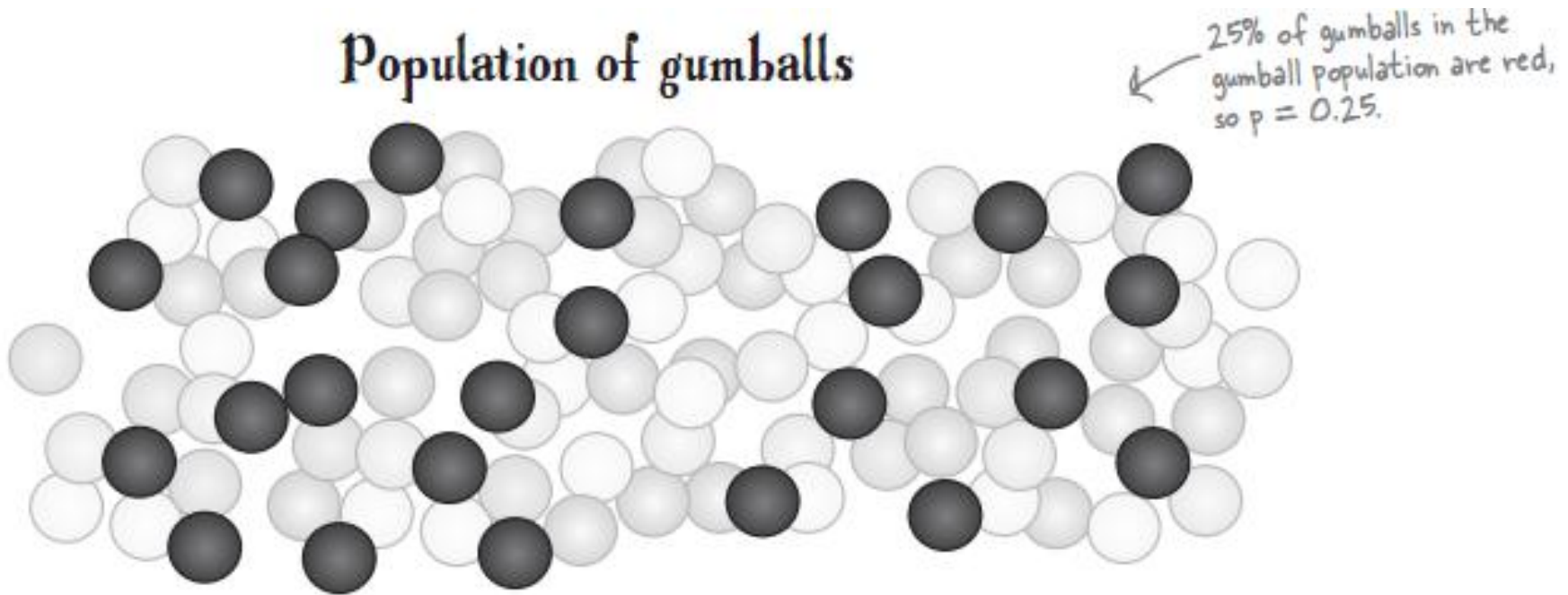
contoh

- Misal diteliti hobi anak-anak terhadap game online di suatu daerah XZ. Diambil sampel sebanyak 40 anak, dan ternyata 32 anak menyukai game online
→ $P_s=0.8$; estimasi titik untuk proporsi sukses (anak menyukai game online) dalam populasi adalah 0.8



Distribusi sampling dari proporsi

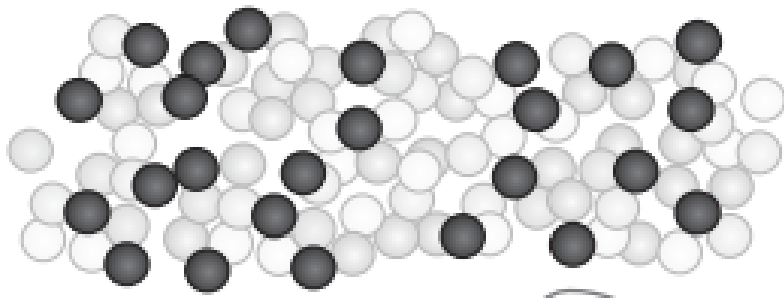
- Bagaimana menentukan distribusi dari sampel proporsi?
- Contoh



- Jika X menggambarkan jumlah **red gumball** dalam sampel maka $X \sim \text{Bin}(n, p)$. Misal $n=100$, $p=0.25$
- Proporsi red gumball dalam sampel tergantung dari X yaitu jumlah red gumball dalam sampel



Sample



$X \sim B(n, p)$

We don't know the exact number of red gumballs in the sample, but we know its distribution.



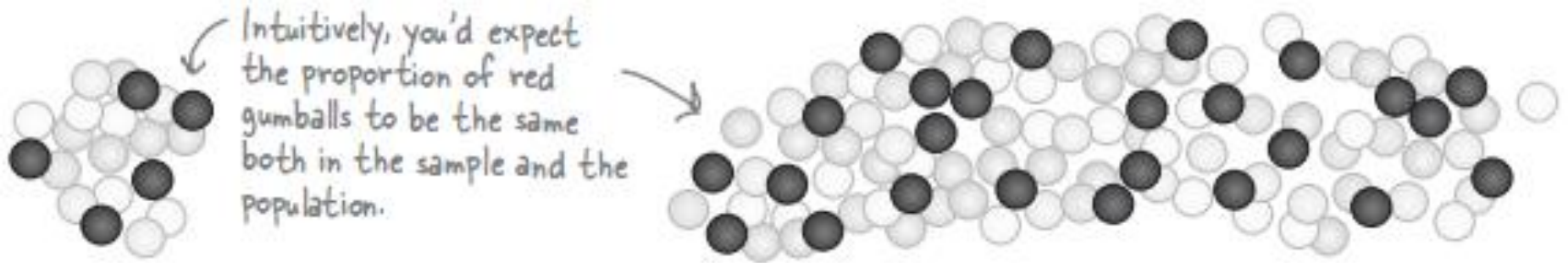
$$P_s = \frac{X}{n}$$

P_s represents the proportion of successes in the sample.

- Distribusi dari sampel proporsi dapat ditentukan dengan menggunakan **distribusi sampling proporsi** atau distribusi dari P_s



Ekspektasi & Variansi dari P_s ?



$$E(P_s) = E\left(\frac{X}{n}\right)$$

$$= \frac{E(X)}{n}$$

$$= \frac{np}{n} \leftarrow E(X) = np$$

$$= p$$

$$\text{Var}(P_s) = \text{Var}\left(\frac{X}{n}\right)$$

$$= \frac{\text{Var}(X)}{n^2}$$

$$= \frac{npq}{n^2} \leftarrow \text{Var}(X) = npq$$

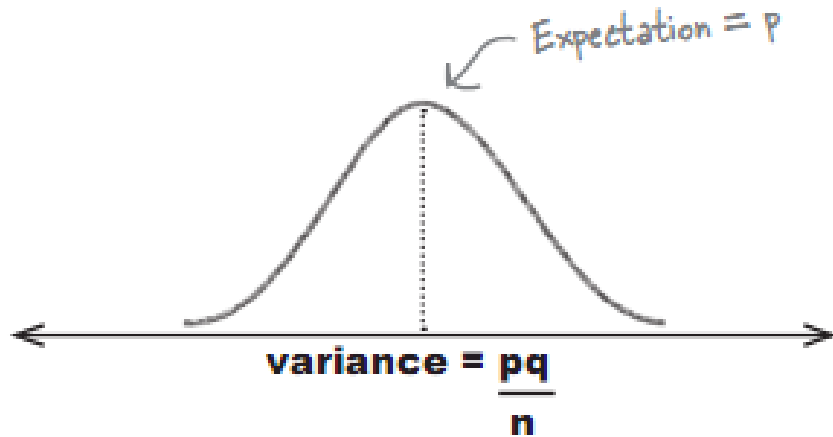
$$= \frac{pq}{n}$$

This comes from $\text{Var}(aX) = a^2 \text{Var}(X)$.
In this case, $a = 1/n$.



So...

- Distribusi P_s tergantung dari n ...



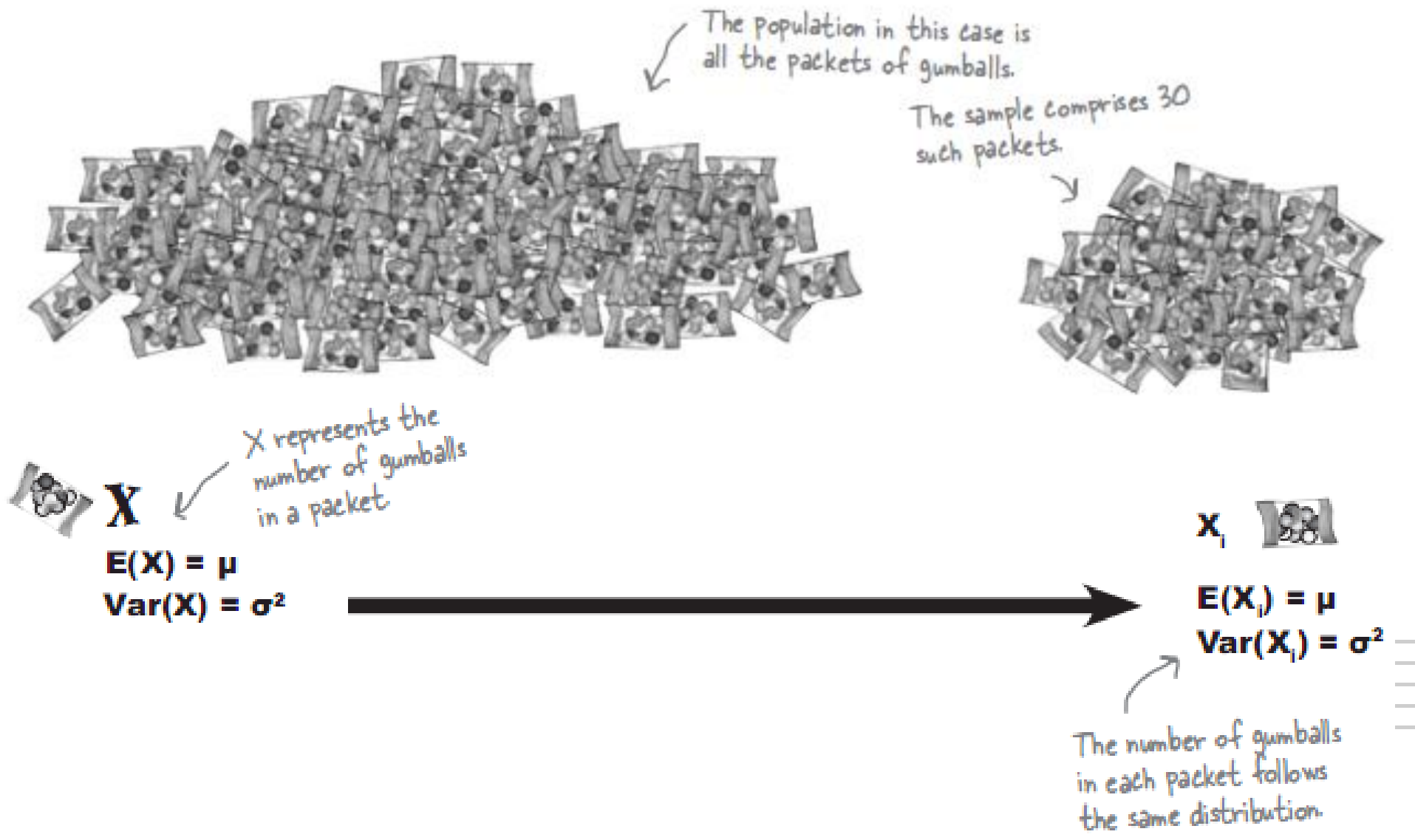
- Jika n bertambah besar (ukuran besar > 30) maka distribusi P_s akan mengikuti distribusi Normal, atau

$$P_s \sim N\left(p, \frac{pq}{n}\right)$$

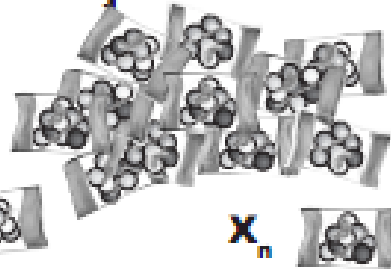


Distribusi sampling untuk rata-rata

- contoh



Sample of X



Each X_i is an independent observation of X , so each packet has the same expectation and variance for the number of gumballs.

This is the mean of the sample, the mean number of gumballs in the packets.

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

$$E(X_1) = \mu$$

$$\text{Var}(X_1) = \sigma^2$$

$$E(X_n) = \mu$$

$$\text{Var}(X_n) = \sigma^2$$

Ekspektasi $\bar{X}$

$$E(\bar{X}) = E\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)$$

These two expressions are the same, just written in a different way.

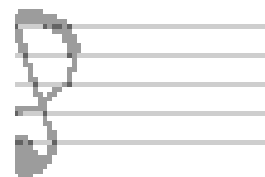
$$= E\left(\frac{1}{n}X_1 + \frac{1}{n}X_2 + \dots + \frac{1}{n}X_n\right)$$

$$= E\left(\frac{1}{n}X_1\right) + E\left(\frac{1}{n}X_2\right) + \dots + E\left(\frac{1}{n}X_n\right)$$

We can split this into n separate expectations because $E(X + Y) = E(X) + E(Y)$.

Every expectation includes $1/n$, so we can take it out of the expression. This comes from $E(aX) = aE(X)$.

$$\rightarrow = \frac{1}{n} (E(X_1) + E(X_2) + \dots + E(X_n))$$



Variansi $\bar{X}$?

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{X_1 + X_2 + \dots + X_n}{n}\right)$$

$$= \text{Var}\left(\frac{1}{n}X_1 + \frac{1}{n}X_2 + \dots + \frac{1}{n}X_n\right)$$

$$= \text{Var}\left(\frac{1}{n}X_1\right) + \text{Var}\left(\frac{1}{n}X_2\right) + \dots + \text{Var}\left(\frac{1}{n}X_n\right)$$

$$= \left(\frac{1}{n}\right)^2 (\text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_n))$$

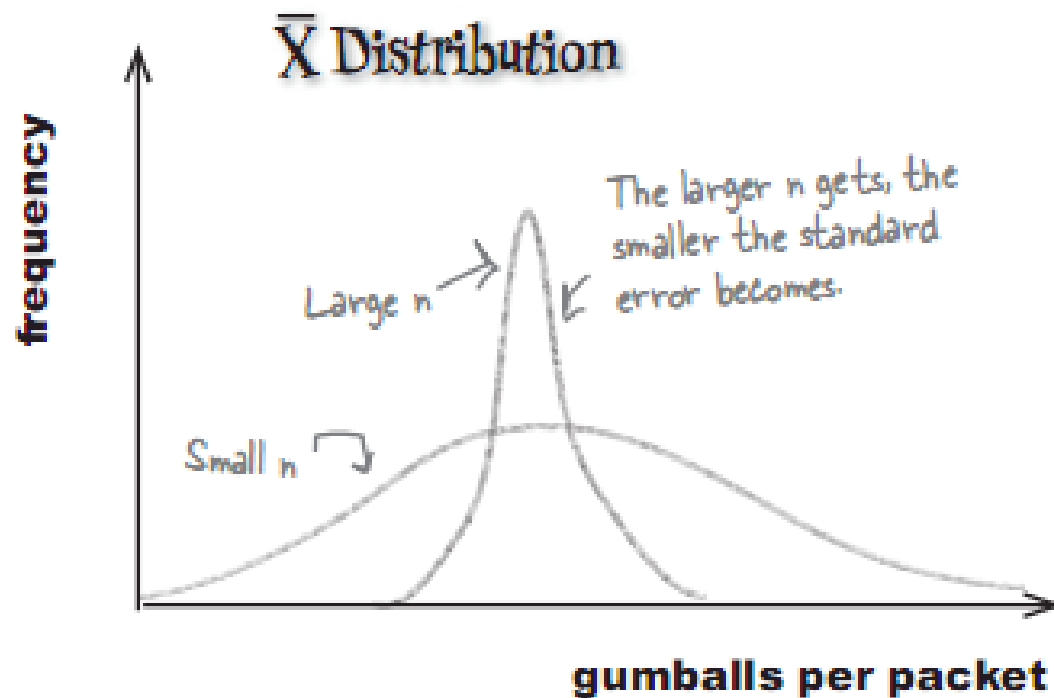
$$= \frac{1}{n^2} (\sigma^2 + \sigma^2 + \dots + \sigma^2)$$

$$= n \times \frac{1}{n^2} \sigma^2$$

$$= \frac{\sigma^2}{n}$$



So...



If $X \sim N(\mu, \sigma^2)$, then $\bar{X} \sim N(\mu, \sigma^2/n)$

These are the mean and variance of $\bar{X}$ that we found earlier.

Teorema Limit Pusat (CLT)

Jika sampel diambil dari suatu populasi dg distribusi non normal dan ukuran sampel besar maka rata-rata $\bar{X}$ diaproksimasikan berdistribusi Normal dengan rata-rata dan variansi populasi adalah μ, σ^2

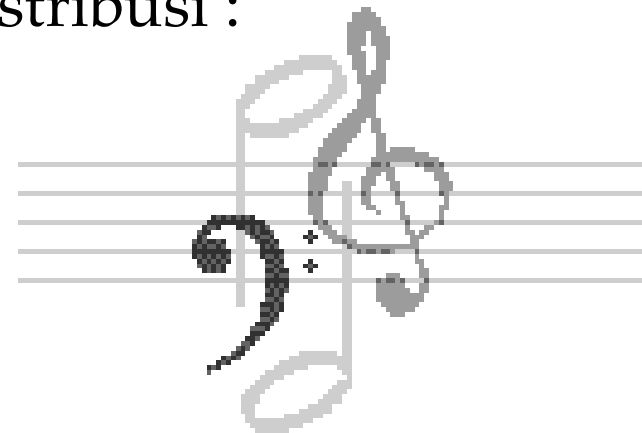
$$\bar{X} \sim N(\mu, \sigma^2/n)$$

This is the mean and variance of $\bar{X}$.

Latihan

Cari distribusi rata-rata dari $\bar{X}$ yang berdistribusi :

- $X \sim \text{POI}(\lambda)$
- $X \sim \text{Bin}(n, p)$



Latihan

7-12. Data on oxide thickness of semiconductors are as follows: 425, 431, 416, 419, 421, 436, 418, 410, 431, 433, 423, 426, 410, 435, 436, 428, 411, 426, 409, 437, 422, 428, 413, 416.

- (a) Calculate a point estimate of the mean oxide thickness for all wafers in the population.
- (b) Calculate a point estimate of the standard deviation of oxide thickness for all wafers in the population.
- (c) Calculate the standard error of the point estimate from part (a).
- (d) Calculate a point estimate of the median oxide thickness for all wafers in the population.
- (e) Calculate a point estimate of the proportion of wafers in the population that have oxide thickness greater than 430 angstrom.