Statistical Analysis

Point estimation

LU4 - "Step-by-Step"

Ana Cristina Costa

ccosta@novaims.unl.pt

Maria Henriques

mhenriques@novaims.unl.pt

NOVA IMS Information Management School

Master in Statistics and Information Management

2020

Topics

- **LU4 – Point estimation**

Notation and concepts

Populations can be characterized by **PARAMETERS**, which are fixed numbers.

$$\boldsymbol{\mu} \quad \boldsymbol{\sigma^2} \quad \boldsymbol{p}$$

$$\mu_1 - \mu_2 \qquad \frac{\sigma_1^2}{\sigma_1^2} \qquad P_1 - P_2$$

However, these parameters are generally non-observable, as they characterize large populations. That is why we need to **ESTIMATE** them.

Find an approximate value that we can consider close enough to the unknown parameter.

Notation and concepts

To estimate those parameters, we need to select **RANDOM SAMPLES**.

$$X_1 = (x_1, x_2, \dots, x_n)$$
$$X_2 = (x_1, x_2, \dots, x_n)$$
$$X_n = (x_1, x_2, \dots, x_n)$$

$X_1, X_2, \dots, X_n$
(iid random variables)

iid Independent and Identically Distributed

With the information collected from these Random Samples, we can use **ESTIMATORS** to estimate the parameters.

$$\bar{X} \rightarrow \mu$$
$$S^2 \rightarrow \sigma^2$$
$$\hat{p} \rightarrow p$$

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Estimators are **statistics** (functions of the r.v. X_i) that are used to produce **ESTIMATES**.

$$\bar{x} = 5$$
$$\bar{x} = 10$$

Notation and concepts

- **Issues addressed**
  - In case of more than one estimator of a parameter, how can we decide which one is better than another?
  - What are the desirable properties of an estimator?

- **Desirable properties of estimators**
  - Sufficiency
  - Unbiasedness
  - Efficiency
  - Consistency

Notation and concepts

- **Concepts**
  - **Sufficiency** → when the estimator takes all the relevant information about the population parameter from the sample.
  - **Unbiasedness** → in medium terms, the estimator reaches the actual value of the parameter.
  - **Efficiency** → the estimator is more efficient (i.e., the estimates are more *accurate*) the smaller the variability of its sampling distribution.
  - **Consistency** → for *large* samples, the estimator should be *approximately* equal to the parameter.

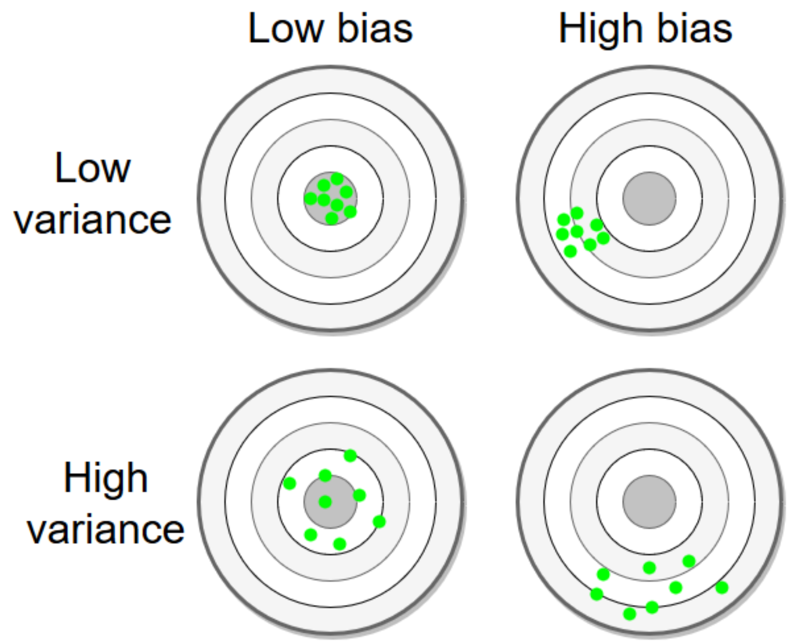
Notation and concepts

- **Unbiasedness**

Is there a bias?

- **Efficiency**

Is the variance low?



Source: http://www.machinelearningtutorial.net/2017/01/26/the-bias-variance-tradeoff/

Unbiasedness

- $\hat{\theta}$ **is an Unbiased estimator of** θ **if**

$$\mathrm{E}(\hat{\theta}) = \theta$$

Otherwise, the estimator is said to be **biased**, and its **bias** is given by

$$bias(\hat{\theta}) = E(\hat{\theta}) - \theta$$

- $\hat{\theta}$ **is an Asymptotically Unbiased estimator of** θ **if**

$$\lim_{n \to +\infty} E(\hat{\theta}) = \theta$$

REMEMBER

$$E(a) = a$$
$$E(aX) = aE(X)$$
$$E(X + Y) = E(X) + E(Y)$$

Where:

- a is a constant
- X and Y are random variables

Efficiency

- **The Efficiency of an estimator is its Mean Squared Error (MSE)**

$$MSE(\hat{\theta}) = E\left[\left(\hat{\theta} - \theta\right)^2\right] = V(\hat{\theta}) + \left[bias(\hat{\theta})\right]^2$$

An estimator will be more efficient (more accurate)

- The smaller the bias.
- The smaller the variance.

REMEMBER

$$V(a) = 0$$
$$\mathrm{V}(aX) = a^2 V(X)$$
$$V(X + Y) = V(X) + V(Y) \pm 2Cov(X, Y)$$

If X and Y are independent, then $Cov(X, Y) = 0$

Efficiency

- **The Relative Efficiency of two estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ of a parameter θ is given by**

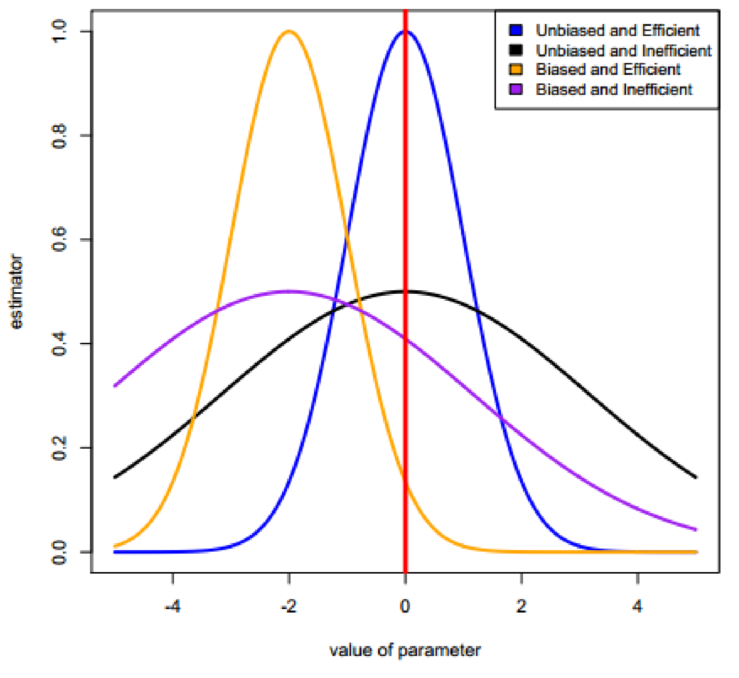
$$eff\left(\hat{\theta}_1, \hat{\theta}_2\right) = \frac{MSE\left(\hat{\theta}_1\right)}{MSE\left(\hat{\theta}_2\right)}$$

  - $\hat{\theta}_1$ is more efficient than $\hat{\theta}_2$ when $eff\left(\hat{\theta}_1, \hat{\theta}_2\right) < 1$

  - **Example**
    - If $eff\left(\hat{\theta}_1, \hat{\theta}_2\right) = 1.6$, than the variability associated to $\hat{\theta}_1$ is 1.6 higher than the variability associated to $\hat{\theta}_2$.

Efficiency

- **Efficiency and bias trade-off**
  - In practice, a biased estimator can be a better estimator than an unbiased one when its bias is small and its efficiency is higher



Source: http://www.zorro-trader.com/manual/en/Lecture%205.htm

Consistency

- **$\hat{\boldsymbol{\theta}}$ is a Consistent estimator of θ if and only if $\hat{\boldsymbol{\theta}}$ converges in probability to θ:**

$$\lim_{n\to+\infty} P\left(|\hat{\theta}-\theta| \leq \varepsilon\right) = 1 \quad \text{for all } \varepsilon>0$$

  - An estimator is consistent if increasing the sample size ($n \to +\infty$) implies an increase in the probability of the estimated value ($\hat{\theta}$) to be in a neighbourhood (ε) of the true value of the parameter (θ).

- An <u>unbiased</u> estimator $\hat{\boldsymbol{\theta}}$ is **Consistent** if $\lim_{n\to+\infty} V\left(\hat{\theta}\right) = 0$
- An estimator $\hat{\boldsymbol{\theta}}$ is **Consistent in Mean Square Error** if $\lim_{n\to+\infty} MSE\left(\hat{\theta}\right) = 0$

Consistency

- **Properties**
  - If $\widehat{\Theta}$ and $\widehat{\Theta}'$ are consistent estimators of θ and θ', respectively, then
    - $\widehat{\Theta} + \widehat{\Theta}'$ is a consistent estimator of $\theta+\theta'$
    - $\widehat{\Theta} \times \widehat{\Theta}'$ is a consistent estimator of $\theta\times\theta'$
    - $\widehat{\Theta}/\widehat{\Theta}'$ is a consistent estimator of θ/θ', with $\theta'\neq0$
    - If $g(\cdot)$ is a real continuous function in θ, then $g\big(\widehat{\Theta}\big)$ is a consistent estimator of $g(\theta)$

Summary

- **If you are asked to**
  - Study the <u>bias</u> of an estimator or
  - Evaluate if an estimator is <u>biased/unbiased</u>

1. Find the Expected Value of that Estimator: $E(\hat{\theta})$
2. Calculate its bias: $bias(\hat{\theta}) = E(\hat{\theta}) - \theta$

  - Study the <u>consistency</u> of an estimator

1. Evaluate the estimator's bias.
2. If it is **unbiased**, then evaluate if $\lim_{n\to+\infty} V(\hat{\theta}) = 0$
3. If it is **biased**, then evaluate if $\lim_{n\to+\infty} MSE(\hat{\theta}) = 0$
4. An estimator is consistent if it is unbiased and $\lim_{n\to+\infty} V(\hat{\theta}) = 0$ or if $\lim_{n\to+\infty} MSE(\hat{\theta}) = 0$

  - Evaluate the <u>Relative Efficiency</u> between two estimators or
  - Compare the <u>efficiency</u> between two estimators

1. Calculate their bias.
2. Calculate their variance.
3. Calculate their MSE.
4. Calculate the Relative Efficiency:

$$eff(\hat{\theta}, \tilde{\theta}) = \frac{MSE(\hat{\theta})}{MSE(\tilde{\theta})}$$

Example 1

a) **Population with mean μ and variance σ^2. Show that $\bar{X}$ is an unbiased estimator of μ.**

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} * n * \mu = \mu$$

$\therefore \bar{X}$ is an unbiased estimator of μ

Exercise 1

- **Population with mean μ and variance σ^2. Show that $\bar{X}$ is a consistent estimator of μ.**
  - Step 1: Evaluate the estimator's bias.
    - Step 1.1. Find the Expected Value of $\bar{X}$.

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} Xi\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} n\mu = \mu$$

    - Step 1.2. Calculate the bias.

$$bias(\bar{X}) = E(\bar{X}) - \mu = \mu - \mu = 0 \Rightarrow \bar{X}\ is\ unbiased$$

  - Step 2: As $\bar{X}$ is unbiased, then evaluate if $\lim_{n\to+\infty} V(\bar{X}) = 0.$
    - Step 2.1. Find the Variance of $\bar{X}$.

$$\mathrm{V}(\bar{X}) = V\left(\frac{1}{n}\sum_{i=1}^{n} Xi\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

    - Step 2.2. Evaluate if $\lim_{n\to+\infty} V(\hat{\theta}) = 0$

When n increases and approaches +∞, $\frac{\sigma^2}{n}$ gets close to 0.

$\therefore$ $\bar{X}$ is a consistent estimator of μ as it is an unbiased estimator and its variance approaches zero as n increases.

Exercise 1

- **Population with mean μ and variance σ^2. Consider two alternative estimators of μ, both unbiased, and their variances:**

$$\hat{\mu}_1 = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \hat{\mu}_2 = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i \qquad V(\hat{\mu}_1) = \frac{1}{n}\sigma^2 \qquad V(\hat{\mu}_2) = \frac{1}{n-1}\sigma^2$$

Show that $\hat{\mu}_1$ is more efficient than $\hat{\mu}_2$.

- Step 1: Calculate their bias. In this case, both are unbiased:

$$bias(\hat{\mu}_1) = bias(\hat{\mu}_2) = 0$$

- Step 2: Calculate their variances.

$$V(\hat{\mu}_1) = \frac{1}{n}\sigma^2; \; V(\hat{\mu}_2) = \frac{1}{n-1}\sigma^2$$

- Step 3: Calculate their MSE.

As they are unbiased, their MSE=Variance

- Step 4: Calculate the efficiency:

$$eff(\hat{\mu}_1, \hat{\mu}_2) = \frac{MSE(\hat{\mu}_1)}{MSE(\hat{\mu}_2)} = \frac{\frac{1}{n}\sigma^2}{\frac{1}{n-1}\sigma^2} = \frac{n-1}{n} < 1$$

$\therefore \hat{\mu}_1$ is more efficient than $\hat{\mu}_2$

Exercise 2

- **Population with mean μ and variance σ^2**

a) **Find the value of a that makes $\hat{\mu}$ an unbiased estimator.**

$$\boldsymbol{\hat{\mu} = 0,4X_1 + aX_3 + 0,3X_5 + 0,2X_n}$$

- Step 1: Find the Expected Value of $\hat{\mu}$: $E(\hat{\mu})$

$$E(\hat{\mu}) = E(0{,}4X_1 + aX_3 + 0{,}3X_5 + 0{,}2X_n)$$
$$= 0{,}4E(X_1) + aE(X_3) + 0{,}3E(X_5) + 0{,}2E(X_n)$$
$$= 0{,}4\mu + a\mu + 0{,}3\mu + 0{,}2\mu = (0{,}9 + a)\mu$$

- Step 2: Find a so that $\hat{\mu}$ is unbiased: $bias(\hat{\mu}) = 0$.

$$bias(\hat{\mu}) = 0$$
$$\Leftrightarrow E(\hat{\mu}) - \mu = 0$$
$$\Leftrightarrow (0{,}9 + a)\mu - \mu = 0$$
$$\Leftrightarrow 0{,}9 + a = 1$$
$$\Leftrightarrow a = 0{,}1$$

Exercise 2

b) **Study $\hat{\mu}$ as to its consistency in Mean Square Error (MSE).**

$$\hat{\boldsymbol{\mu}} = \mathbf{0,4}\boldsymbol{X_1} + \mathbf{0,1}\boldsymbol{X_3} + \mathbf{0,3}\boldsymbol{X_5} + \mathbf{0,2}\boldsymbol{X_n}$$

- Step 1: Evaluate if $\hat{\mu}$ is biased/unbiased

We know $\hat{\mu}$ is unbiased from the previous exercise.

- Step 2: Find the Variance of $\hat{\mu}$ - $V(\hat{\mu})$

$$\begin{aligned} V(\hat{\mu}) &= V(0{,}4X_1 + aX_3 + 0{,}3X_5 + 0{,}2X_n) \\ &= 0{,}4^2V(X_1) + 0{,}1^2V(X_3) + 0{,}3^2V(X_5) + 0{,}2^2V(X_n) \\ &= 0{,}16\sigma^2 + 0{,}01\sigma^2 + 0{,}09\sigma^2 + 0{,}04\sigma^2 = 0{,}30\sigma^2 \end{aligned}$$

- Step 3: Find the MSE of $\hat{\mu}$ - $\mathrm{MSE}(\hat{\mu})$

$$MSE(\hat{\mu}) = V(\hat{\mu}) + bias(\hat{\mu}) = 0{,}30\sigma^2$$

$\therefore$ $\hat{\mu}$ is not consistent in MSE because $MSE(\hat{\mu})$ does not depend on n and, therefore, $\lim_{n \to +\infty} MSE(\hat{\theta}) \neq 0$

Exercise 2

c) **Compare $\hat{\mu}$ with the following alternative estimator as to the Relative Efficiency.**

$$\tilde{\boldsymbol{\mu}} = \frac{\mathbf{1}}{\mathbf{8}}(\mathbf{2}\boldsymbol{X}_{\mathbf{1}} + \mathbf{5}\boldsymbol{X}_{\mathbf{2}} - \mathbf{3}\boldsymbol{X}_{\mathbf{3}} + \mathbf{4}\boldsymbol{X}_{\boldsymbol{n}})$$

- Step 1: Evaluate if $\tilde{\mu}$ is biased/unbiased

$$E(\tilde{\mu}) = E\left(\frac{1}{8}(2X_1 + 5X_2 - 3X_3 + 4X_n)\right)$$

$$= \frac{1}{8}[2E(X_1) + 5E(X_2) - 3E(X_3) + 4E(X_n)] = \frac{1}{8}(2\mu + 5\mu - 3\mu + 4\mu) = \frac{1}{8}(8\mu) = \mu$$

$$bias(\tilde{\boldsymbol{\mu}}) = E(\tilde{\boldsymbol{\mu}}) - \mu = \mu - \mu = 0$$

- Step 2: Find the Variance of $\tilde{\mu}$ - $V(\tilde{\mu})$

$$V(\tilde{\boldsymbol{\mu}}) = V\left(\frac{\mathbf{1}}{\mathbf{8}}(\mathbf{2}\boldsymbol{X}_{\mathbf{1}} + \mathbf{5}\boldsymbol{X}_{\mathbf{2}} - \mathbf{3}\boldsymbol{X}_{\mathbf{3}} + \mathbf{4}\boldsymbol{X}_{\boldsymbol{n}})\right) = \frac{1}{64}[4V(X_1) + 25V(X_2) + 9V(X_3) + 16V(X_n) = \frac{1}{64}(54\sigma^2) = 0{,}84\sigma^2$$

- Step 3: Find the MSE of $\hat{\mu}$ - $\mathrm{MSE}(\hat{\mu})$

$$MSE(\tilde{\mu}) = V(\tilde{\mu}) + bias(\tilde{\mu}) = 0{,}84\sigma^2$$

- Step 4: Calculate the Relative Efficiency

$$eff(\hat{\mu}, \tilde{\mu}) = \frac{MSE(\hat{\mu})}{MSE(\tilde{\mu})} = \frac{0{,}30\sigma^2}{0{,}84\sigma^2} = 0{,}36 < 1$$

$\therefore$ $\hat{\mu}$ is more efficient than $\tilde{\mu}$

Example 1

b) Population N(μ, σ). Show that S^2 is an unbiased estimator of σ^2.

From the form, is known that "If $X_1, X_2, \dots, X_n$ are iid $N(\mu, \sigma)$ then $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$" and that the $E(\chi^2_n) = n$. Therefore:

$$E\left(\frac{(n-1)S^2}{\sigma^2}\right) = E\left(\chi^2_{(n-1)}\right) \Leftrightarrow \left(\frac{n-1}{\sigma^2}\right) E(S^2) = n-1 \Leftrightarrow E(S^2) = \frac{(n-1)*\sigma^2}{n-1} = \sigma^2$$

$\therefore\ S^2$ is an unbiased estimator of σ^2

Example 1

c) Population N(μ, σ). Show that $M_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ is an asymptotically unbiased estimator of σ^2 and derive its bias.

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

Therefore, $M_2 = \frac{n-1}{n} S^2$ and

$$E(M_2) = E\left(\frac{n-1}{n}S^2\right) = \left(\frac{n-1}{n}\right)E(S^2) = \frac{n-1}{n}\sigma^2 \xrightarrow[n\to+\infty]{} \sigma^2$$

M_2 is an asymptotically unbiased estimator of σ^2. As n increases, the expected value of M_2 approaches the parameter's true value.

$$\text{bias}(M_2) = E(M_2) - \sigma^2 = \frac{n-1}{n}\sigma^2 - \sigma^2 = \left(\frac{n-1-n}{n}\right)\sigma^2 = \frac{-\sigma^2}{n}$$

$$\lim_{n\to+\infty} bias(M_2) = \lim_{n\to+\infty} \frac{-\sigma^2}{n} = 0$$

∴ For large n, M_2 is approximately unbiased.

Example 2

Population N(μ, σ). Consider the two estimators of σ^2:

$$S^2 \text{ and } M_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

a) Show that $V(S^2) = \frac{2\sigma^4}{n-1}$ **and** $V(M_2) = \frac{n-1}{n^2}2\sigma^4$

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} \qquad V(\chi^2_{(n)}) = 2n$$

$$V\left(\frac{(n-1)S^2}{\sigma^2}\right) = V\left(\chi^2_{(n-1)}\right) \Leftrightarrow \left(\frac{n-1}{\sigma^2}\right)^2 V(S^2) = 2(n-1) \Leftrightarrow V(S^2) = \frac{2(n-1) * \sigma^4}{(n-1)^2} = \frac{2\sigma^4}{n-1}$$

$$M_2 = \frac{n-1}{n}S^2$$

$$V(M_2) = V\left(\frac{(n-1)}{n}S^2\right) = \frac{(n-1)^2}{n^2}V(S^2) = \frac{(n-1)^2}{n^2} * \frac{2\sigma^4}{n-1} = \frac{n-1}{n^2}2\sigma^4$$

Example 2

b) **Considering that *n* is large, determine the relative efficiency of S^2 and M_2.**

$$V(S^2) = \frac{2\sigma^4}{n-1} \qquad V(M_2) = \frac{n-1}{n^2}2\sigma^4$$

$$eff(M_2, S^2) = \frac{V(M_2)}{V(S^2)} = \frac{\frac{n-1}{n^2}2\sigma^4}{\frac{2\sigma^4}{n-1}} = \frac{n-1}{n^2}2\sigma^4 * \frac{n-1}{2\sigma^4} = \frac{(n-1)^2}{n^2} < 1$$

$\therefore$ For large samples, M_2 is more efficient than S^2.

Exercise 4

Population N(μ, σ).

a) **Show that $E\left(\tilde{\sigma}^2\right) = \sigma^2$.**

$$\tilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$

From the first page of the form - 1st Theorem: $X_i \sim N(\mu, \sigma)\ then\ \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi_n^2$

From the last page of the form – Mean ans Variance of χ_n^2 : $E(\chi_n^2) = n \quad V(\chi_n^2) = 2n$

$$E\left[\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2\right] = E(\chi_n^2) = n \Leftrightarrow E\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2\right] = n \Leftrightarrow \frac{1}{\sigma^2}E\left[\sum_{i=1}^{n}(X_i-\mu)^2\right] = n$$

$$\Leftrightarrow E\left[\sum_{i=1}^{n}(X_i-\mu)^2\right] = n\sigma^2 \Leftrightarrow \frac{1}{n}E\left[\sum_{i=1}^{n}(X_i-\mu)^2\right] = \sigma^2 \Leftrightarrow E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right] = \sigma^2$$

$$\Leftrightarrow \boldsymbol{E\left(\tilde{\sigma}^2\right) = \sigma^2}$$

$\therefore\ \tilde{\sigma}^2$ is an unbiased estimator of σ^2.

Exercise 4

b) **Compare the efficiency of $\tilde{\boldsymbol{\sigma}}^2$ and $\hat{\boldsymbol{\sigma}}^2 = \boldsymbol{S}^2 = \frac{\mathbf{1}}{\boldsymbol{n}-\mathbf{1}}\sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{n}}(\boldsymbol{X}_{\boldsymbol{i}} - \overline{\boldsymbol{X}})^{\mathbf{2}}$.**

Since both $\tilde{\sigma}^2$ and S^2 are unbiased, we simply need to compare their variances to determine which one is more efficient. Determine $\mathrm{V}(\tilde{\sigma}^2) = \mathrm{V}\left[\frac{1}{\mathrm{n}}\sum_{i=1}^{n}(X_i - \mu)^2\right]$.

$$V\left[\sum_{i=1}^{n}\left(\frac{X_i - \mu}{\sigma}\right)^2\right] = V(\chi_n^2) = 2n \Leftrightarrow V\left[\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \mu)^2\right] = 2n \Leftrightarrow \frac{1}{\sigma^4}V\left[\sum_{i=1}^{n}(X_i - \mu)^2\right] = 2n$$

$$\Leftrightarrow V\left[\sum_{i=1}^{n}(X_i - \mu)^2\right] = 2n\sigma^4 \Leftrightarrow \frac{1}{n^2}V\left[\sum_{i=1}^{n}(X_i - \mu)^2\right] = \frac{2n\sigma^4}{n^2} \Leftrightarrow V\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2\right] = \frac{2\sigma^4}{n}$$

$$\Leftrightarrow \boldsymbol{V}\left(\tilde{\boldsymbol{\sigma}}^2\right) = \frac{\mathbf{2}\boldsymbol{\sigma}^4}{\boldsymbol{n}}$$

We have showned that $V(S^2) = \frac{2\sigma^4}{n-1}$ in example 2.

$$eff(\tilde{\sigma}^2, S^2) = \frac{V(\tilde{\sigma}^2)}{V(S^2)} = \frac{\frac{2\sigma^4}{n}}{\frac{2\sigma^4}{n-1}} = \frac{2\sigma^4}{n} * \frac{n-1}{2\sigma^4} = \frac{n-1}{n} < 1$$

$\therefore V(\tilde{\sigma}^2) < V(S^2)$, therefore $\tilde{\sigma}^2$ is more efficient than S^2.