Unidade IV – Inferência estatística

4.1. Introdução e histórico
4.2. Conceitos fundamentais
4.3. Distribuições amostrais e Teorema central do limite
4.4. Estimação de parâmetros
4.5. Testes de hipóteses
4.6. Quebras das pressuposições no processo de inferência
4.7. Testes de qui-quadrado

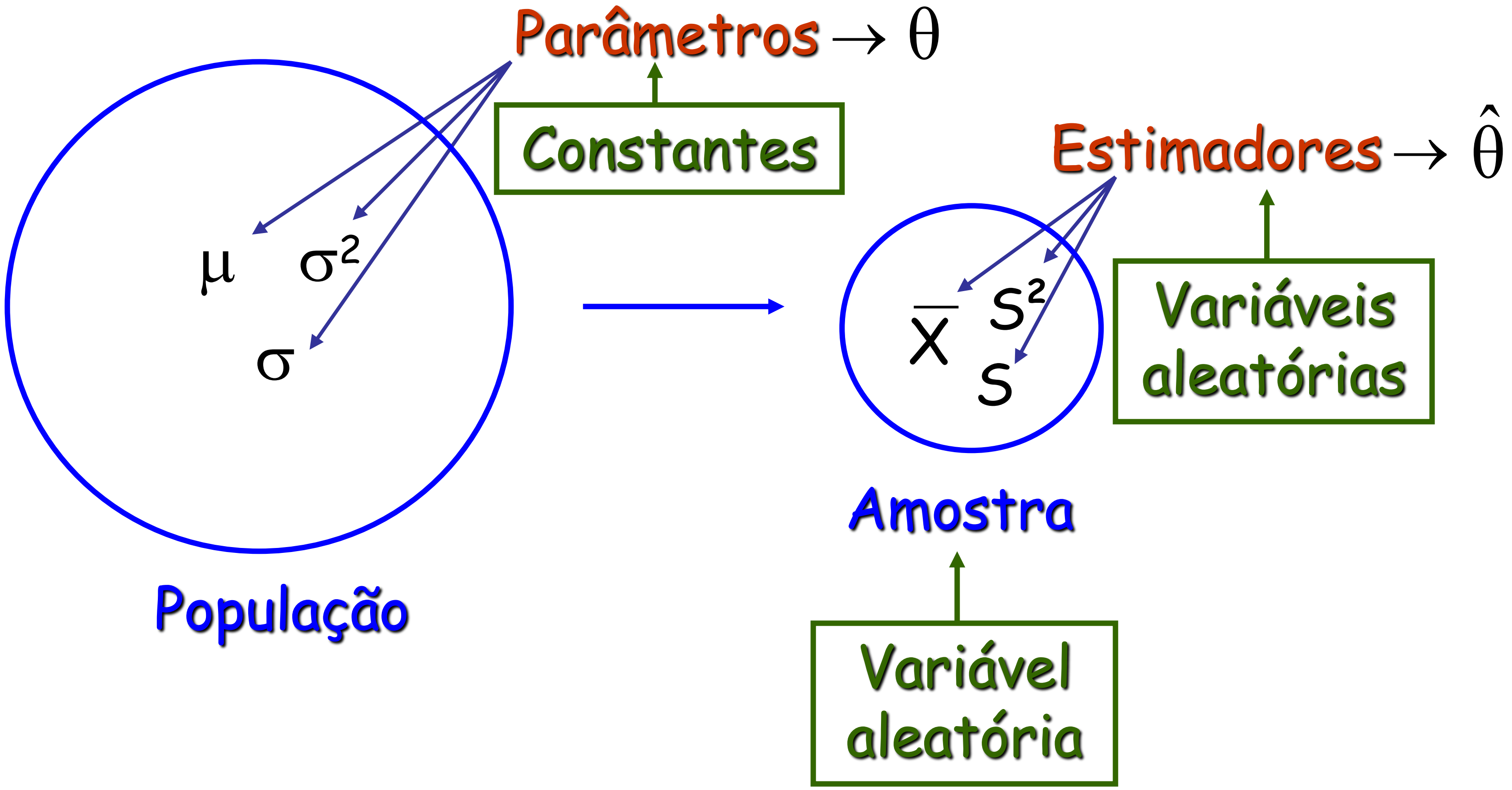
Profª. Clause Piana

Teoria da estimação de parâmetros

Profª. Clause Piana

4.2 Estimação de parâmetros

Conceitos fundamentais



Parâmetros:

- constantes que caracterizam uma população (distribuição)
- geralmente são valores desconhecidos
- são representados, genericamente, pela letra grega teta (θ)

Exemplos: média da população (μ)
variância da população (σ^2)

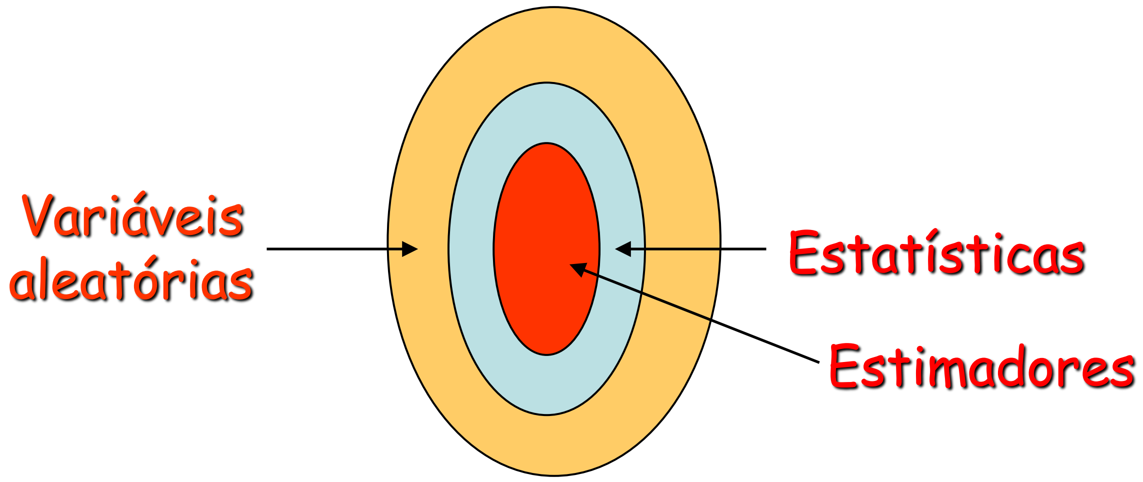
Estimadores:

- valores (medidas) calculados na amostra com objetivo de fornecer informação sobre os parâmetros
- são estatísticas, portanto, são variáveis aleatórias
- são representados, genericamente, pela letra teta com um acento circunflexo $(\hat{\theta})$, onde se lê teta chapéu

Exemplos: média da amostra $\left(\overline{X} \text{ ou } \hat{\mu}\right)$
variância da amostra $\left(S^2 \text{ ou } \hat{\sigma}^2\right)$

Estimativa é um valor particular que o estimador assume

O estimador é uma estatística e todas as estatísticas são variáveis aleatórias



População (X)

μ, σ^2, σ — Parâmetros (constantes)

Amostra aleatória: $[X_1, X_2, ..., X_n]$ ← Variável aleatória

Estimadores → $\bar{X}$ S^2 ← Variáveis aleatórias

Estimativas: $\bar{x}_1$, $\bar{x}_2$, . . . , $\bar{x}_k$

Estimativas: s_1^2, s_2^2, . . . , s_k^2

População (X)

μ

σ^2

σ

Parâmetros

Amostra aleatória

$[X_1, X_2, ..., X_n]$

Estimadores → $\bar{X}$ S^2

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

Todos os estimadores são variáveis aleatórias e têm uma distribuição de probabilidade.

Parâmetro: constante populacional desconhecida

μ (média populacional)

Estimador: expressão matemática da medida amostral

$$\bar{X} = \frac{\sum X_i}{n}$$ (média amostral)

Estimativa: valor numérico que o estimador assume numa determinada amostra

$$\bar{x}_1 = \frac{1+1}{2} = 1$$ (média amostral calculada)

Exemplo:

Uma população é constituída por quatro valores ($N = 4$):

$X = x$	1	2	3	4
$P(X=x)$	0,2	0,3	0,3	0,2

$\mu=2,5$

$\sigma^2=1,05$

Desta população, retiramos uma amostra aleatória de tamanho dois ($n=2$) $\rightarrow$ $[X_1, X_2]$

Quantas e quais são as amostras de tamanho dois que podemos extrair desta população de tamanho quatro?

$$k = N^n = 4^2 = 16 \text{ possíveis amostras}$$

Parâmetro		$\mu = 2{,}5$	$\sigma^2 = 1{,}05$
Estimador		$\bar{X} = \frac{\sum X_i}{n}$	$S^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$
Estimativas	Amostra 1: (1, 1)	$\bar{x}_1 = \frac{1+1}{2} = 1$	$s_1^2 = \frac{(1-1)^2 + (1-1)^2}{2-1} = 0$
	Amostra 2: (1, 2)	$\bar{x}_2 = \frac{1+2}{2} = 1{,}5$	$s_2^2 = \frac{(1-1{,}5)^2 + (2-1{,}5)^2}{2-1} = 0{,}5$
	Amostra 3: (1, 3)	$\bar{x}_3 = 2$	$s_3^2 = 2$
	Amostra 4: (1, 4)	$\bar{x}_4 = 2{,}5$	$s_4^2 = 4{,}5$
	Amostra 5: (2, 1)	$\bar{x}_5 = 1{,}5$	$s_5^2 = 0{,}5$
	Amostra 6: (2, 2)	$\bar{x}_6 = 2$	$s_6^2 = 0$
	Amostra 7: (2, 3)	$\bar{x}_7 = 2{,}5$	$s_7^2 = 0{,}5$
	Amostra 8: (2, 4)	$\bar{x}_8 = 3$	$s_8^2 = 2$
	Amostra 9: (3, 1)	$\bar{x}_9 = 2$	$s_9^2 = 2$
	Amostra 10: (3, 2)	$\bar{x}_{10} = 2{,}5$	$s_{10}^2 = 0{,}5$
	Amostra 11: (3, 3)	$\bar{x}_{11} = 3$	$s_{11}^2 = 0$
	Amostra 12: (3, 4)	$\bar{x}_{12} = 3{,}5$	$s_{12}^2 = 0{,}5$
	Amostra 13: (4, 1)	$\bar{x}_{13} = 2{,}5$	$s_{13}^2 = 4{,}5$
	Amostra 14: (4, 2)	$\bar{x}_{14} = 3$	$s_{14}^2 = 2$
	Amostra 15: (4, 3)	$\bar{x}_{15} = 3{,}5$	$s_{15}^2 = 0{,}5$
	Amostra 16: (4, 4)	$\bar{x}_{16} = 4$	$s_{16}^2 = 0$

Podem existir vários estimadores para um mesmo parâmetro.

Exemplos:

$$\bar{X} = \frac{\sum X_i}{n}$$

$$\bar{X}_p = \frac{\sum X_i p_i}{\sum p_i}$$

$$X_1$$

estimadores de μ

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

$$S_n^2 = \frac{\sum (X_i - \bar{X})^2}{n}$$

estimadores de σ^2

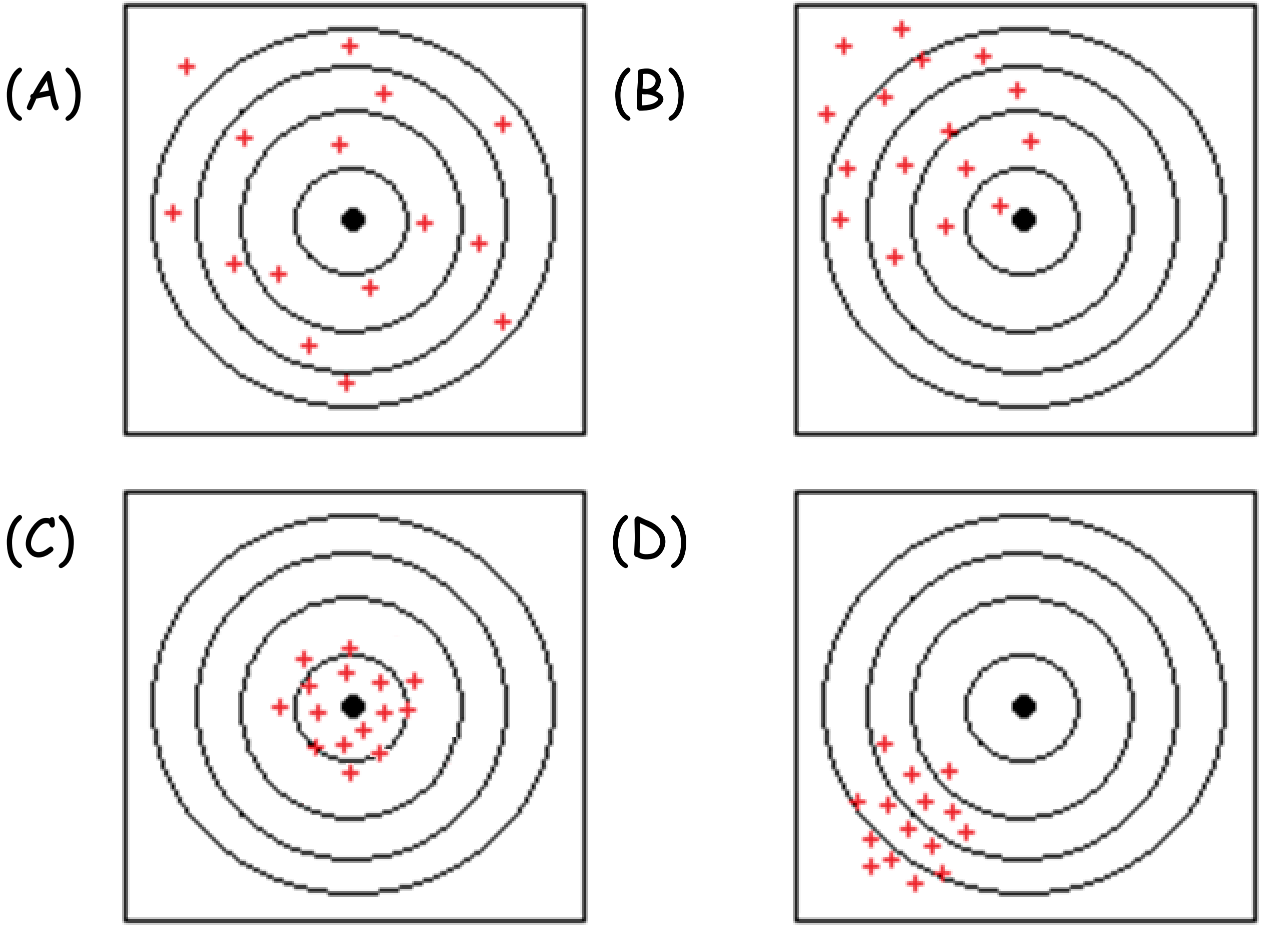
Propriedades dos estimadores

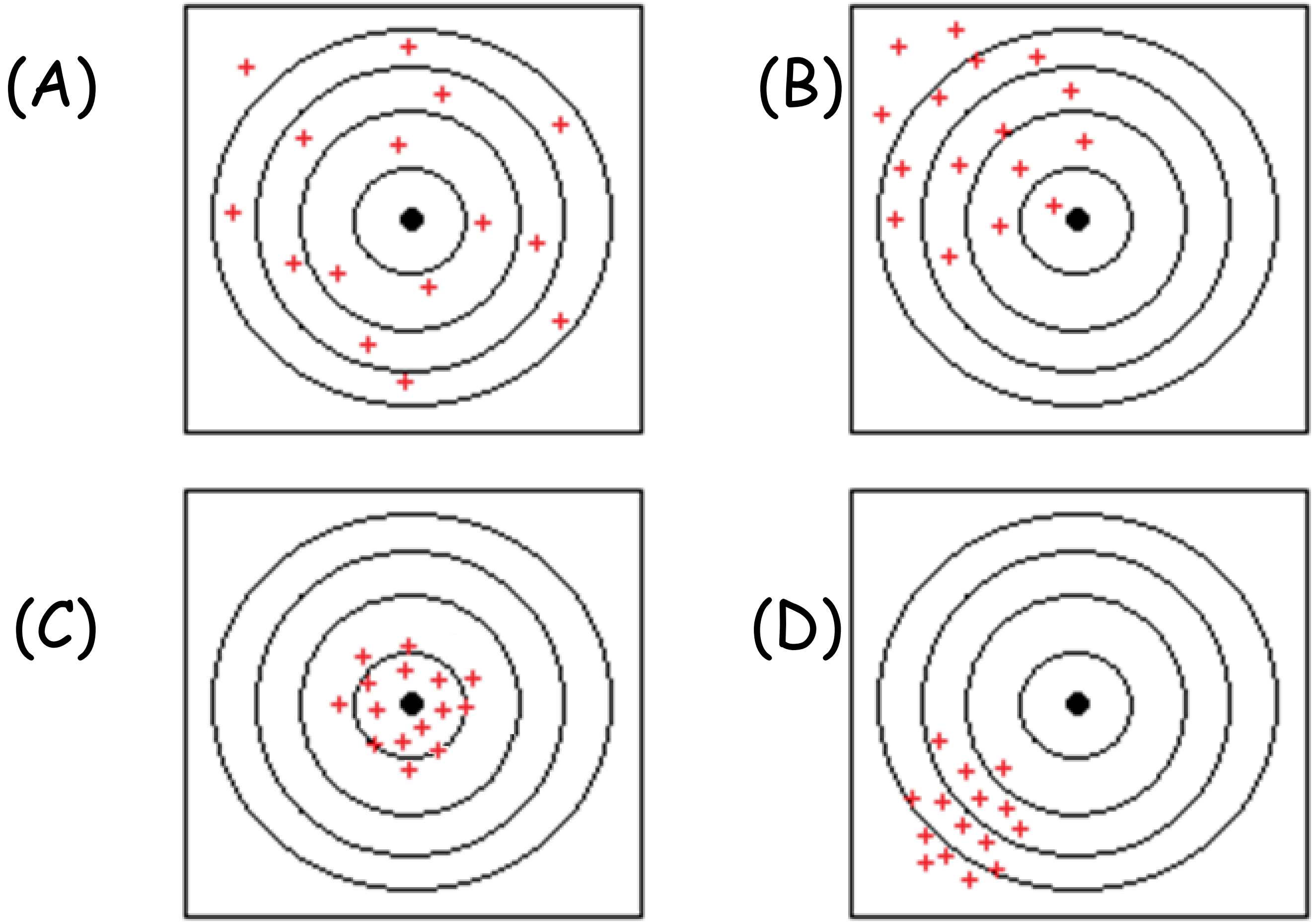
- Imparcialidade ou não tendenciosidade
- Eficiência ou variância mínima
- Consistência

Algumas ideias relacionadas a essas propriedades

- Acurácia
- Precisão
- Viés

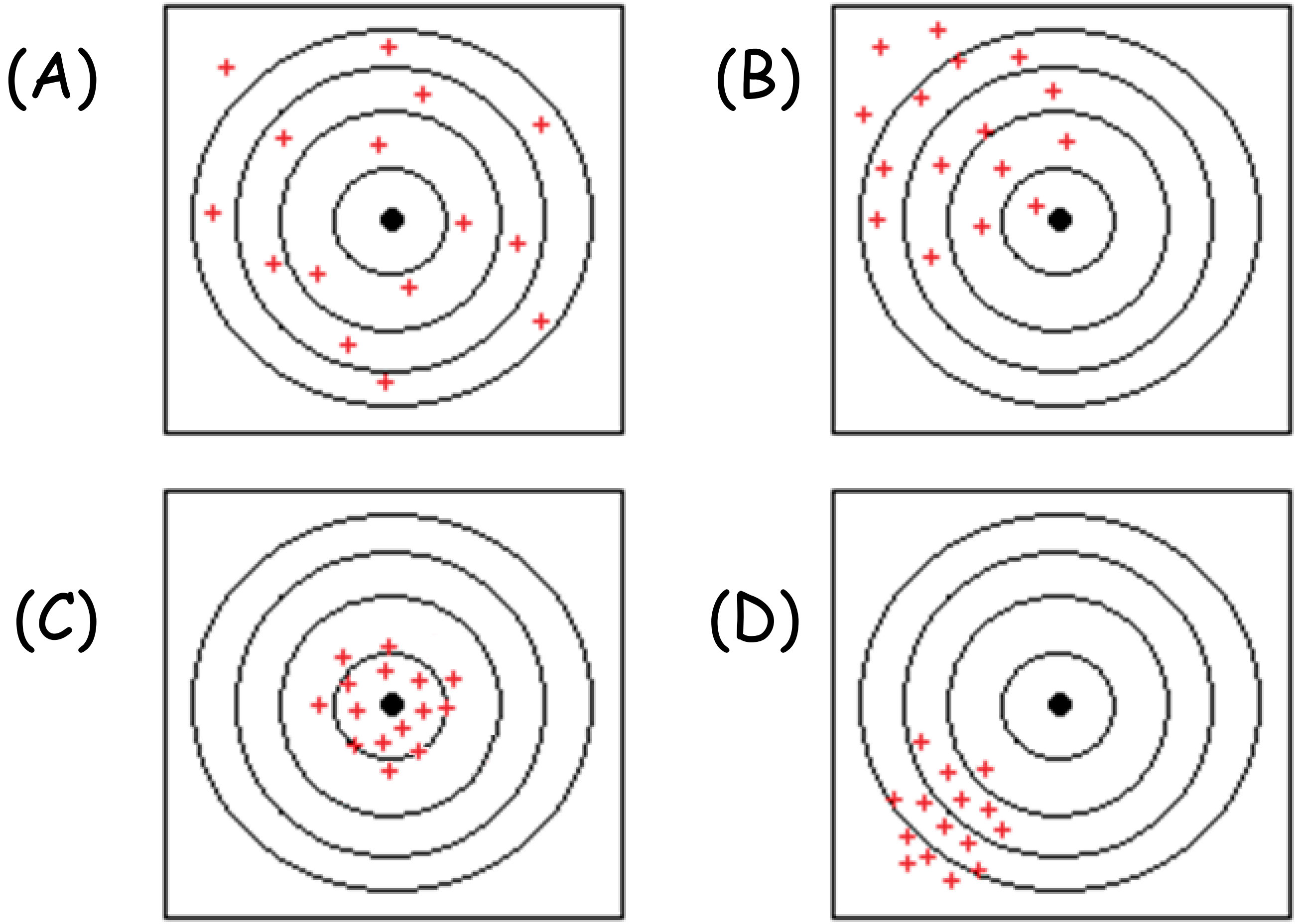
Exemplo: Desejamos comprar um rifle e, após algumas seleções, restaram quatro alternativas, que chamaremos de rifles A, B, C e D. Foi feito um teste com cada rifle que consistiu em fixá-lo num cavalete, mirar o centro de um alvo e disparar 15 tiros. Os resultados estão ilustrados na figura abaixo.





1º critério: em média acertar o alvo

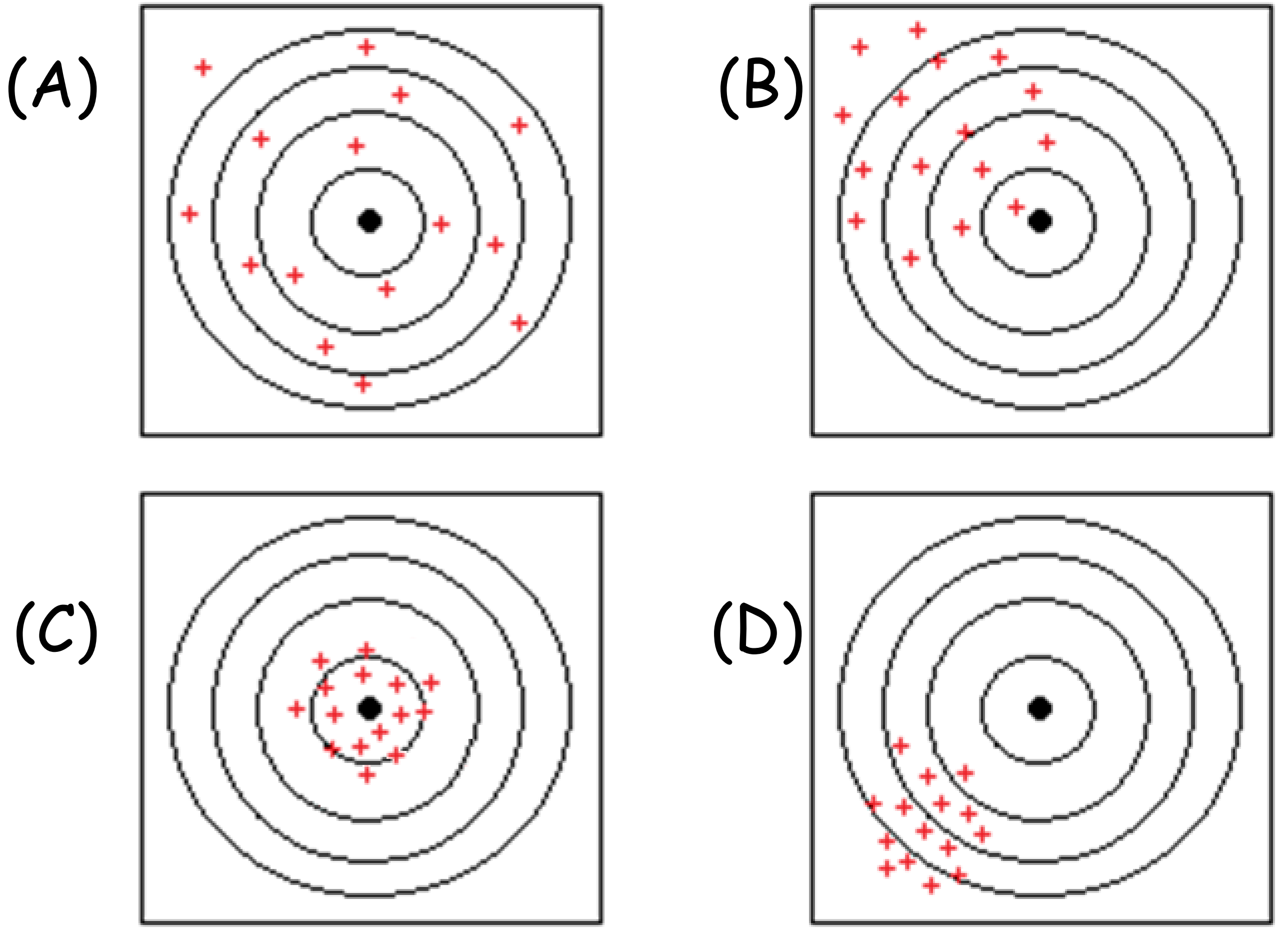
2º critério: não ser muito dispersivo



Acurácia: proximidade de cada observação em relação ao valor real

Precisão: proximidade de cada observação em relação à média das observações

Viés: diferença entre a média das observações e o valor real



Rifle A: não viesada, pouco acurada e tem baixa precisão

Rifle B: viesada, pouco acurada e tem média precisão

Rifle C: não viesada, muito acurada e tem alta precisão

Rifle D: viesada, pouco acurada e tem alta precisão

Propriedades dos estimadores

1. Imparcialidade ou não tendenciosidade

Um estimador $\hat{\theta}$ é um estimador imparcial do parâmetro θ se o valor esperado de $\hat{\theta}$ for igual a θ.

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

Exemplos:

$\bar{X}$ é um estimador imparcial de μ, pois $E(\bar{X}) = \mu$

$\bar{X}_p$ é um estimador imparcial de μ, pois $E(\bar{X}_p) = \mu$

X_1 é um estimador imparcial de μ, pois $E(X_1) = \mu$

S^2 é um estimador imparcial de σ^2, pois $E(S^2) = \sigma^2$

S_n^2 não é um estimador imparcial σ^2, pois $E(S_n^2) = \frac{n-1}{n}\sigma^2$

Distribuição de probabilidade da população

$X = x$	40	45	50	Σ
$P(X = x)$	0,2	0,3	0,5	1

$E(X) = \mu = 46,5$

$V(X) = \sigma^2 = 15,25$

$n = 2 \rightarrow [X_1, X_2]$

Amostra	$[X_1, X_2]$	$P[X_1, X_2]$	$\bar{X}$	X_+	S^2	S_n^2		
1	(40, 40)	$0,2 \times 0,2 = 0,04$	40	80	0	0		
2	(40, 45)	$0,2 \times 0,3 = 0,06$	42,5	85	12,5	6,25		
3	(40, 50)	$0,2 \times 0,5 = 0,10$	45	90	50	25		
4	(45, 40)	$0,3 \times 0,2 = 0,06$	42,5	85	12,5	6,25		
5	(45, 45)	$0,3 \times 0,3 = 0,09$	45	90	0	0		
6	(45, 50)	$0,3 \times 0,5 = 0,15$	47,5	95	12,5	6,25		
7	(50, 40)	$0,5 \times 0,2 = 0,10$	45	90	50	25		
8	(50, 45)	$0,5 \times 0,3 = 0,15$	47,5	95	12,5	6,25		
9	(50, 50)	$0,5 \times 0,5 = 0,25$	50	100	0	0		

$$V(X) = \sigma^2 = 15,25$$

Distribuição amostral de S^2 das amostras de tamanho 2

$S^2 = s^2$	0	12,5	50	Σ
$P(S^2 = s^2)$	0,38	0,42	0,2	1

$$E(S^2) = \mu_{S^2} = \sum s^2 p(s^2) = 0 \times 0,38 + 12,5 \times 0,42 + 50 \times 0,2 = 15,25$$

$$E(S^2) = \sigma^2 = 15,25$$

Distribuição amostral de S_n^2 das amostras de tamanho 2

$S_n^2 = s_n^2$	0	6,25	25	Σ
$P(S_n^2 = s_n^2)$	0,38	0,42	0,2	1

$$E(S_n^2) = \mu_{S_n^2} = \sum s_n^2 p(s_n^2) = 0 \times 0,38 + 6,25 \times 0,42 + 25 \times 0,2 = 7,625$$

$$E(S_n^2) = \frac{n-1}{n} \sigma^2$$

Propriedades dos estimadores

1. Imparcialidade ou não tendenciosidade

Um estimador $\hat{\theta}$ é um estimador imparcial do parâmetro θ se o valor esperado de $\hat{\theta}$ for igual a θ.

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

Exemplos:

$\bar{X}$ é um estimador imparcial de μ, pois de $E(\bar{X}) = \mu$

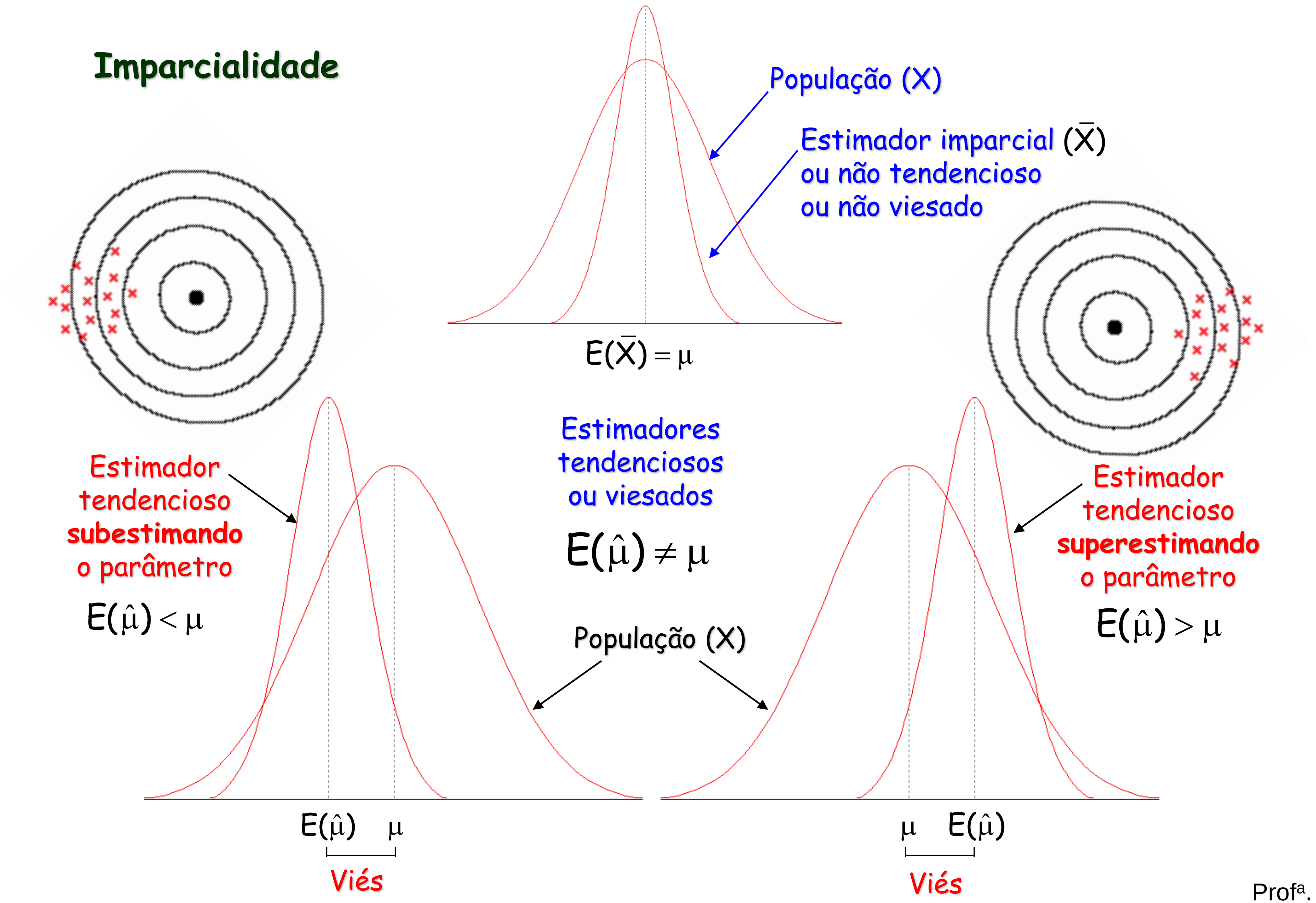
$\bar{X}_p$ é um estimador imparcial de μ, pois de $E(\bar{X}_p) = \mu$

X_1 é um estimador imparcial de μ, pois de $E(X_1) = \mu$

S^2 é um estimador imparcial de σ^2, pois de $E(S^2) = \sigma^2$

S_n^2 não é um estimador imparcial σ^2, pois de $E(S_n^2) = \frac{n-1}{n}\sigma^2$

Imparcialidade



Propriedades dos estimadores

1. Imparcialidade ou não tendenciosidade

Um estimador $\hat{\theta}$ é um estimador imparcial do parâmetro θ se o valor esperado de $\hat{\theta}$ for igual a θ.

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

Exemplos:

$\bar{X}$ é um estimador imparcial de μ, pois de $E(\bar{X}) = \mu$

$\bar{X}_p$ é um estimador imparcial de μ, pois de $E(\bar{X}_p) = \mu$

X_1 é um estimador imparcial de μ, pois de $E(X_1) = \mu$

S^2 é um estimador imparcial de σ^2, pois de $E(S^2) = \sigma^2$

S_n^2 não é um estimador imparcial σ^2, pois de $E(S_n^2) = \frac{n-1}{n}\sigma^2$

2. Eficiência ou variância mínima

Se dois ou mais estimadores de um mesmo parâmetro são imparciais, é mais eficiente aquele que possui a menor variância.

Exemplo:

Dentre todos os estimadores imparciais de μ ($\bar{X}$, $\bar{X}_p$ e X_1), a média simples ($\bar{X}$) é o mais eficiente porque tem a menor variância.

Demonstração:

Consideremos uma amostra aleatória de tamanho três (n = 3)

$$[X_1, X_2, X_3]$$

Média simples: $\bar{X} = \frac{\sum X_i}{n} \longrightarrow \bar{X} = \frac{X_1 + X_2 + X_3}{3} = \frac{1}{3}(X_1 + X_2 + X_3)$

Média ponderada: $\bar{X}_p = \frac{\sum X_i p_i}{\sum p_i} \longrightarrow \bar{X}_p = \frac{1X_1 + 2X_2 + 1X_3}{4} = \frac{1}{4}(X_1 + 2X_2 + X_3)$

Média simples

$$\overline{X} = \frac{1}{3}(X_1 + X_2 + X_3)$$

$$V(\overline{X}) = V\left[\frac{1}{3}(X_1 + X_2 + X_3)\right]$$

$$V(\overline{X}) = \frac{1}{9}[V(X_1 + X_2 + X_3)]$$

$$V(\overline{X}) = \frac{1}{9}[V(X_1) + V(X_2) + V(X_3)]$$

$$V(\overline{X}) = \frac{1}{9}(\sigma^2 + \sigma^2 + \sigma^2)$$

$$V(\overline{X}) = \frac{3\sigma^2}{9} = 0{,}33\sigma^2$$

Média ponderada

$$\overline{X}_p = \frac{1}{4}(X_1 + 2X_2 + X_3)$$

$$V(\overline{X}_p) = V\left[\frac{1}{4}(X_1 + 2X_2 + X_3)\right]$$

$$V(\overline{X}_p) = \frac{1}{16}[V(X_1) + V(2X_2) + V(X_3)]$$

$$V(\overline{X}_p) = \frac{1}{16}[V(X_1) + 4V(X_2) + V(X_3)]$$

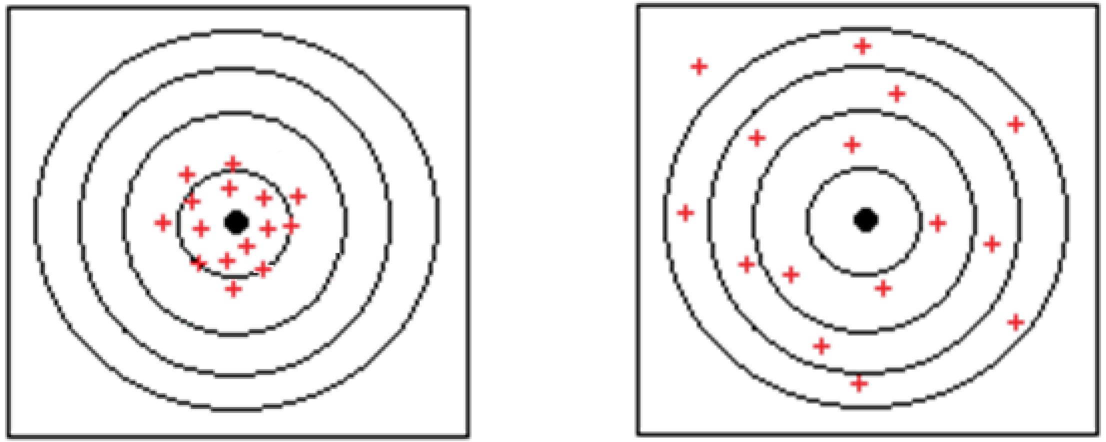
$$V(\overline{X}_p) = \frac{1}{16}(\sigma^2 + 4\sigma^2 + \sigma^2)$$

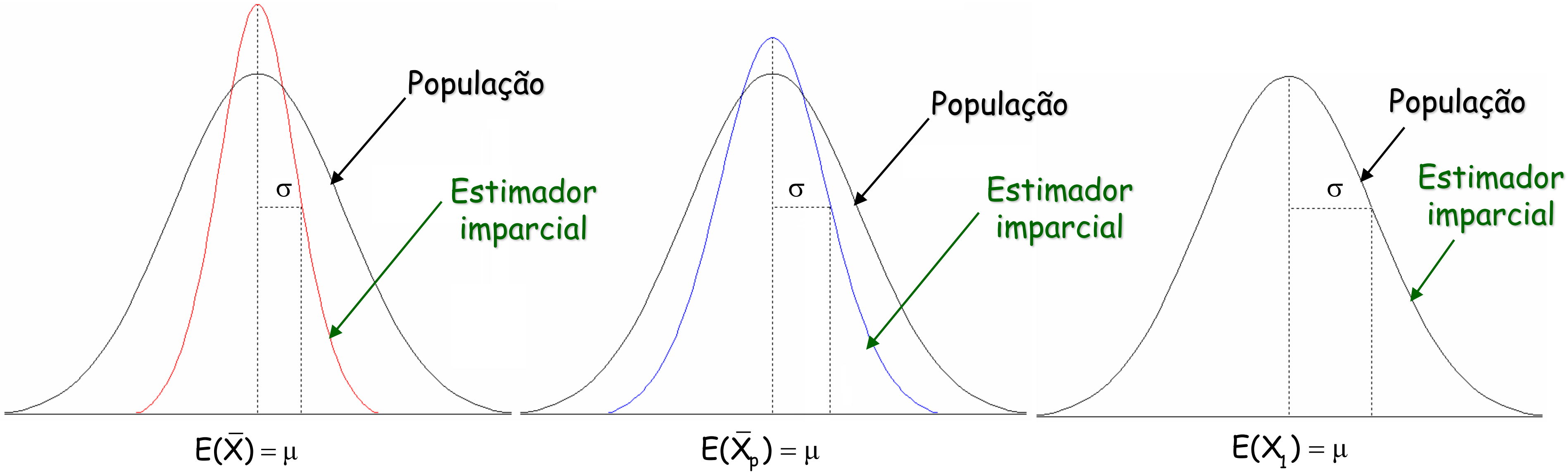
$$V(\overline{X}_p) = \frac{6\sigma^2}{16} = 0{,}38\sigma^2$$

Mais eficiente

$$V(\overline{X}) = 0{,}33\sigma^2 < V(\overline{X}_p) = 0{,}38\sigma^2 < V(X_1) = \sigma^2$$

Eficiência ou variância mínima

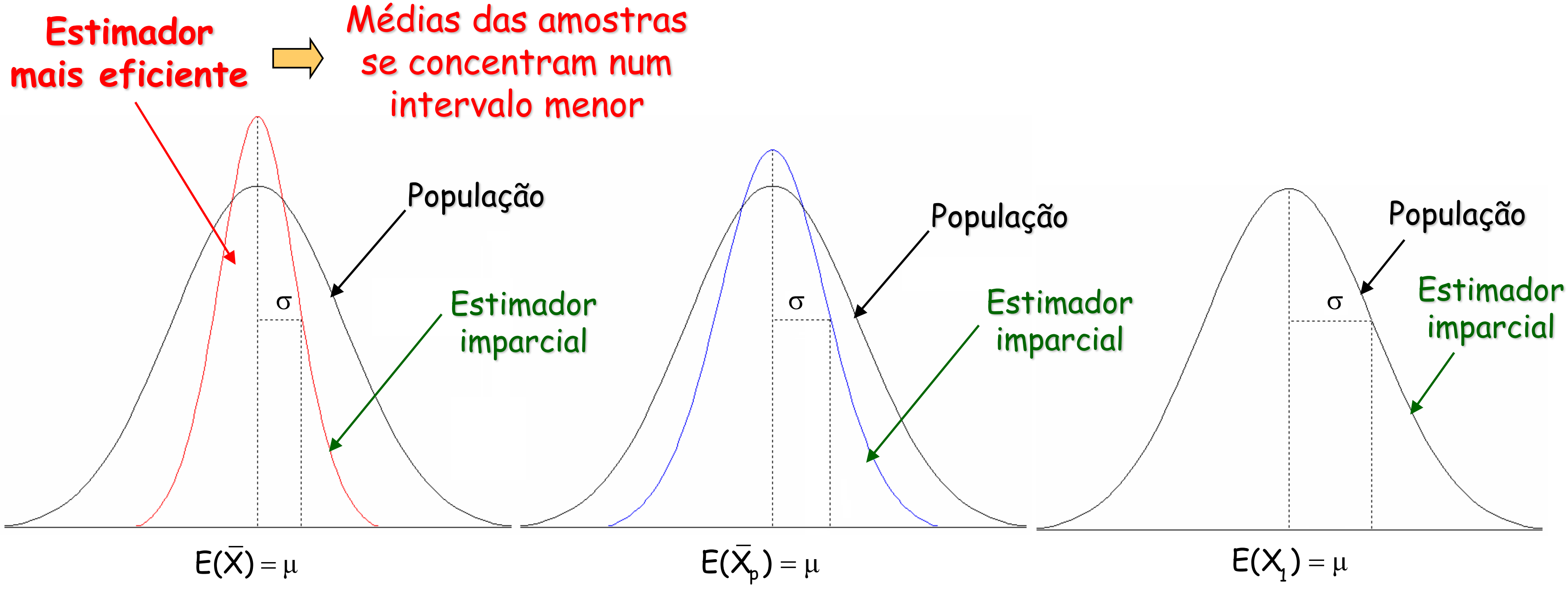




$$V(\bar{X}) = 0{,}33\sigma^2 < V(\bar{X}_p) = 0{,}38\sigma^2 < V(X_1) = \sigma^2$$

Eficiência ou variância mínima

Estimadores imparciais de μ



$$V(\bar{X}) = 0,33\sigma^2 < V(\bar{X}_p) = 0,38\sigma^2 < V(X_1) = \sigma^2$$

3. Consistência

Um estimador é consistente se à medida que o tamanho da amostra aumenta o valor do estimador se aproxima do parâmetro.

$$\hat{\theta} \xrightarrow{n \to N} \theta$$

Exemplo:

S_n^2 é um estimador consistente de σ^2

Com base nessa propriedade, podemos concluir que:

→ Se a amostra for pequena, devemos utilizar S^2 para estimar σ^2

→ Se a amostra for grande, podemos utilizar S_n^2 ou S^2 para estimar σ^2

Propriedades dos estimadores

- Imparcialidade ou não tendenciosidade

Um estimador $\hat{\theta}$ é um estimador imparcial do parâmetro θ se o valor esperado de $\hat{\theta}$ for igual a θ.

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

- Eficiência ou variância mínima

Se dois ou mais estimadores de um mesmo parâmetro são imparciais, é mais eficiente aquele que possui a menor variância.

- Consistência

Um estimador é consistente se à medida que o tamanho da amostra aumenta o valor do estimador se aproxima do parâmetro.

$$\hat{\theta} \xrightarrow{n \to N} \theta$$

Melhores estimadores

Melhor estimador → $$\boxed{\bar{X} = \frac{\sum X_i}{n}}$$

imparcial e mais eficiente

$$\bar{X}_p = \frac{\sum X_i p_i}{\sum p_i}$$

$$X_1$$

estimadores de μ

Melhor estimador → $$\boxed{S^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}}$$

imparcial

$$S_n^2 = \frac{\sum (X_i - \bar{X})^2}{n}$$

estimadores de σ^2

Processos de estimação

Existem duas maneiras de obter a estimativa de um parâmetro:

- **Estimação por ponto (ou pontual)**
- **Estimação por intervalo**

Processos de estimação

Estimação por ponto (ou pontual)

É o processo através do qual obtemos um único ponto, ou seja, **um único valor** para estimar o parâmetro.

Exemplo: Amostra (1, 3, 2)

$$\bar{x} = \frac{\sum x_i}{n} = \frac{1+3+2}{3} = 2$$

estimativa pontual de μ

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} = \frac{(1-2)^2 + (3-2)^2 + (2-2)^2}{3-1} = 1$$

estimativa pontual de σ^2

Exemplo:

Uma população é constituída por quatro valores ($N = 4$):

$X = x$	1	2	3	4
$P(X=x)$	0,2	0,3	0,3	0,2

$\mu=2,5$

$\sigma^2=1,05$

Desta população, retiramos uma amostra aleatória de tamanho dois ($n=2$) $\rightarrow$ $[X_1, X_2]$

Quantas e quais são as amostras de tamanho dois que podemos extrair desta população de tamanho quatro?

$k = N^n = 4^2 = 16$ possíveis amostras

Parâmetro		$\mu = 2{,}5$	$\sigma^2 = 1{,}05$
Estimador		$\bar{X} = \frac{\sum X_i}{n}$	$S^2 = \frac{\sum (X_i - \bar{X})^2}{n-1}$
Estimativas pontuais	Amostra 1: (1, 1)	$\bar{x}_1 = \frac{1+1}{2} = 1$	$s_1^2 = \frac{(1-1)^2 + (1-1)^2}{2-1} = 0$
	Amostra 2: (1, 2)	$\bar{x}_2 = \frac{1+2}{2} = 1{,}5$	$s_2^2 = \frac{(1-1{,}5)^2 + (2-1{,}5)^2}{2-1} = 0{,}5$
	Amostra 3: (1, 3)	$\bar{x}_3 = 2$	$s_3^2 = 2$
	Amostra 4: (1, 4)	$\bar{x}_4 = 2{,}5$	$s_4^2 = 4{,}5$
	Amostra 5: (2, 1)	$\bar{x}_5 = 1{,}5$	$s_5^2 = 0{,}5$
	Amostra 6: (2, 2)	$\bar{x}_6 = 2$	$s_6^2 = 0$
	Amostra 7: (2, 3)	$\bar{x}_7 = 2{,}5$	$s_7^2 = 0{,}5$
	Amostra 8: (2, 4)	$\bar{x}_8 = 3$	$s_8^2 = 2$
	Amostra 9: (3, 1)	$\bar{x}_9 = 2$	$s_9^2 = 2$
	Amostra 10: (3, 2)	$\bar{x}_{10} = 2{,}5$	$s_{10}^2 = 0{,}5$
	Amostra 11: (3, 3)	$\bar{x}_{11} = 3$	$s_{11}^2 = 0$
	Amostra 12: (3, 4)	$\bar{x}_{12} = 3{,}5$	$s_{12}^2 = 0{,}5$
	Amostra 13: (4, 1)	$\bar{x}_{13} = 2{,}5$	$s_{13}^2 = 4{,}5$
	Amostra 14: (4, 2)	$\bar{x}_{14} = 3$	$s_{14}^2 = 2$
	Amostra 15: (4, 3)	$\bar{x}_{15} = 3{,}5$	$s_{15}^2 = 0{,}5$
	Amostra 16: (4, 4)	$\bar{x}_{16} = 4$	$s_{16}^2 = 0$

$X = x$	1	2	3	4
$P(X=x)$	0,2	0,3	0,3	0,2

μ=2,5

σ^2=1,05

Distribuição amostral da média das amostras de tamanho 2

$\bar{X} = \bar{x}$	1	1,5	2	2,5	3	3,5	4	Σ
$P(\bar{X} = \bar{x})$	0,04	0,12	0,21	0,26	0,21	0,12	0,04	1

0,68

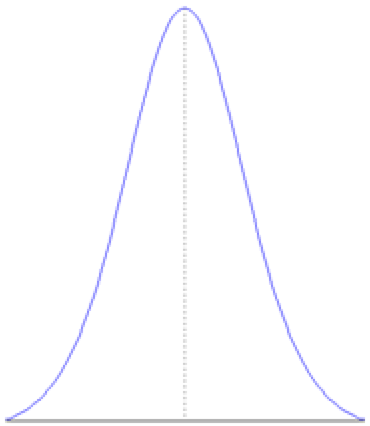
Distribuição amostral da média das amostras de tamanho 3

$\bar{X} = \bar{x}$	1	1,33	1,67	2	2,33	2,67	3	3,33	3,67	4	Σ
$P(\bar{X} = \bar{x})$	0,008	0,044	0,09	0,159	0,207	0,207	0,159	0,09	0,036	0,008	1

0,72

Se **n** é grande,

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



$E(\bar{X}) = \mu$

Estimação por intervalo

É um processo que permite obter os limites de um intervalo onde, com uma determinada probabilidade (nível de confiança), podemos esperar encontrar o verdadeiro valor do parâmetro.

dois valores

Limite inferior → $\underbrace{Li < \theta < Ls}_{\text{Intervalo de confiança}}$ ← Limite superior

As estimativas por intervalo são preferíveis porque indicam a precisão, estabelecendo limites que, com uma determinada probabilidade, devem compreender o parâmetro.