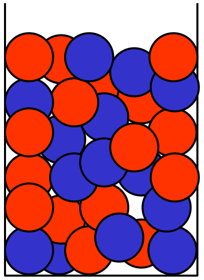
Estatística: Aplicação ao Sensoriamento Remoto

SER 204

Estimação Pontual

Camilo Daleles Rennó
camilo.renno@inpe.br
acesso do conteúdo do curso em Bibdigital do INPE ou GitHub

Inferência Estatística



Considere o experimento: retiram-se 3 bolas de uma urna (com reposição). Define-se uma v.a. X cujo valor representa o número total de bolas vermelhas dentre as 3 escolhidas. Qual a média e variância de X?

Quais os valores possíveis de X?

Valores inteiros (número de tentativas bem-sucedidas)

Mínimo 0 (nenhuma bola vermelha)

Máximo 3 (todas 3 são bolas vermelhas)

X: $\{0, 1, 2, 3\}$

Qual a distribuição de probabilidade de X?

X é discreto

A probabilidade de sucesso p é igual para todas tentativas (sorteio com reposição)

O número de sorteios é pré-definido ($n = 3$) e o número de sucessos é a v.a. X

Distribuição: Binomial

Quais os parâmetros que definem esta Binomial?

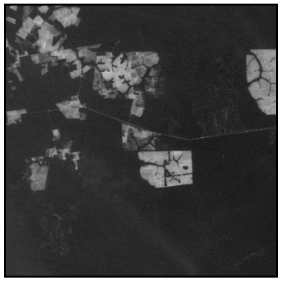
n e p

$n = 3$

$p = ?$ (precisaria conhecer toda a população)

DISTRIBUIÇÃO CONHECIDA
PARÂMETRO(S) DESCONHECIDO(S)

Inferência Estatística



Numa imagem, um *pixel* é selecionado ao acaso. Define-se uma v.a. X cujo valor representa seu valor digital. Qual a probabilidade deste *pixel* possuir valor entre 100 e 150?

Quais os valores possíveis de X?

Considerando uma imagem 8 *bits*...

Mínimo 0 (região escura)

Máximo 255 (região clara)

X: $\{0, 1, ..., 255\}$

Qual a distribuição de probabilidade de X?

X é discreto

Distribuição: Desconhecida (discreta)

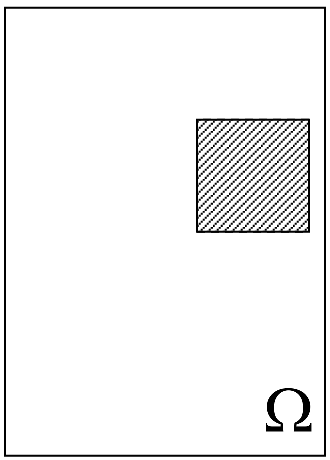
Que parâmetros são necessários para definir esta distribuição?

???????

DISTRIBUIÇÃO DESCONHECIDA

Inferência Estatística

inferir certas características da população



amostra

n elementos (ou objetos) da população
ex: sortear n *pixels* de uma imagem (com ou sem reposição)

n realizações da v.a.
ex: medir a reflectância de um objeto n vezes

$\Downarrow$

distribuição desconhecida
e/ou
parâmetros desconhecidos

a amostra constitui um conjunto de n v.a.
$X_1, X_2, \ldots, X_n$ com mesma distribuição (conhecida ou não)

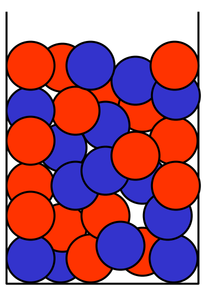
$\Downarrow$

$\boxed{\text{Amostra Aleatória}}$

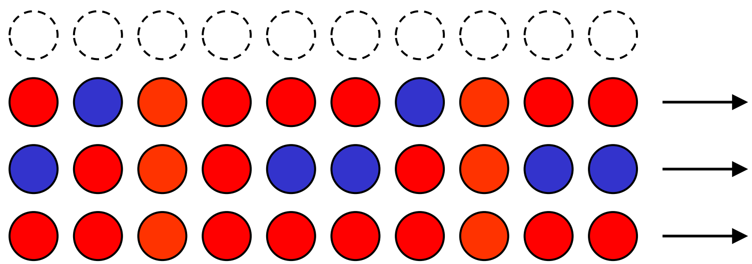
Amostra Aleatória

Como uma amostra aleatória é um conjunto de n v.a.:

cada amostragem resulta num conjunto distinto de valores e portanto pode levar a uma conclusão distinta



Sorteio de 10 bolas



Conclusão:

bolas vermelhas são maioria

mesma proporção

só há bolas vermelhas na urna!

Grandes questões:

- Quão representativa é a amostra disponível para a análise?
  - tamanho de amostra e métodos de obtenção das amostras
- Que características devem ser observadas para representar a população?
  - estimação de parâmetros
- Quão confiável é conclusão obtida pela pesquisa?
  - erros

Estimação de Parâmetros

População ⟶ Amostra

Distribuição de Probabilidade (ou FDP)

Distribuição Amostral (Frequências)

Parâmetros (valor fixo) ⟵ estimar ⟵ Estatísticas (variável aleatória)

Estimação
- pontual (estatísticas)
- por intervalo (intervalos de confiança)

OBS: estatística: é a v.a. que estima (pontualmente) um parâmetro (populacional) as vezes é chamada simplesmente de estimador

estimativa: é o valor do estimador obtido para uma amostra específica

Distribuição Amostral

É muito comum utilizar um conjunto de valores observados (amostra) para tentar "enxergar" a verdadeira distribuição da população.

Esta capacidade, é claro, depende do tamanho e representatividade da amostra.

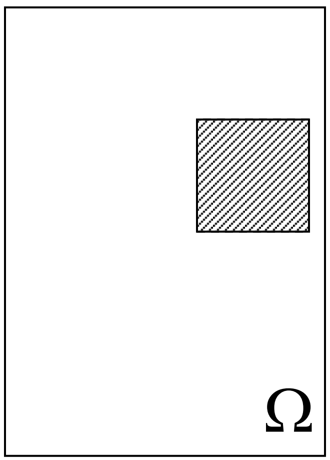
Obs1: Tipos de amostragem e tamanho ideal de uma amostra serão discutidos em "Teoria de Amostragem".

Obs2: Testes estatísticos formais que visam comprovar se uma população segue ou não uma distribuição específica serão discutidos durante o curso.

Existem pelos menos 3 representações gráficas que podem ser utilizadas para avaliar a distribuição amostral:

- histograma (gráfico de frequências)
- frequência acumulada
- boxplot

Estimação Pontual de um Parâmetro



amostra composta por n valores

parâmetro desconhecido θ

De que maneira os valores da amostra podem ser combinados a fim de se produzir uma "boa" estimativa desse parâmetro θ?

$\left\{\begin{array}{l}\text{método dos momentos}\\ \text{método da máxima verossimilhança}\end{array}\right.$ $\Rightarrow$ é preciso conhecer a distribuição!

Estimação Pontual de um Parâmetro

Considere que seja possível produzir m diferentes estimadores para θ, sendo que $\hat{\theta}_i$ representa o i-ésimo estimador de θ $(i = 1, \ldots, m)$

Como escolher qual estimador é melhor?

Importante:

- lembre-se que todo estimador é uma v.a. e portanto seu valor (estimativa) varia de amostra para amostra
- dificilmente (ou é improvável) que uma amostra forneça uma estimativa igual ao parâmetro que se deseja estimar

Estimação Pontual de um Parâmetro

Considere que seja possível produzir m diferentes estimadores para θ, sendo que $\hat{\theta}_i$ representa o i-ésimo estimador de θ $(i = 1, \ldots, m)$

Para que $\hat{\theta}_k$ seja o melhor, então esse estimador deveria

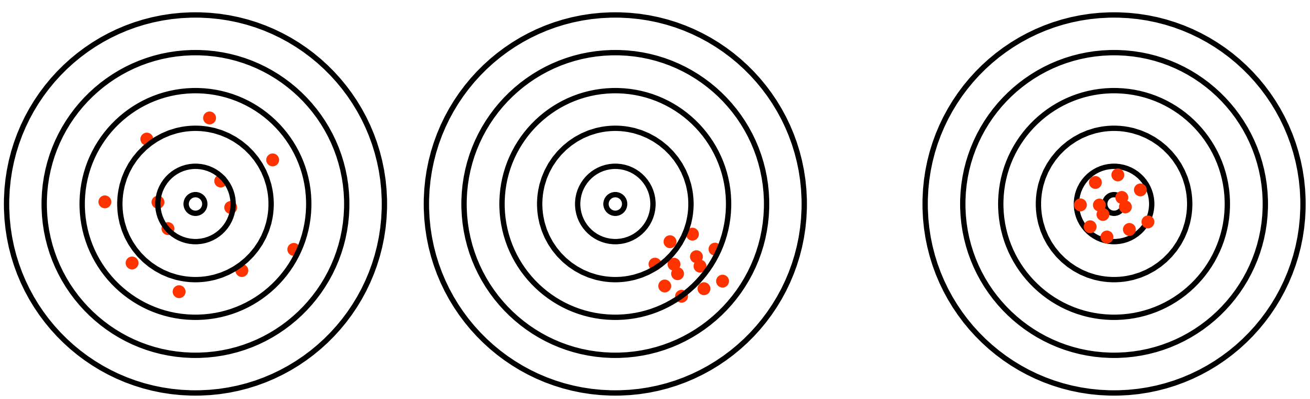
- ser não tendencioso (exatidão) $\Rightarrow$ A média das estimativas de todas as amostras de tamanho n possíveis de serem retiradas da população é igual ao verdadeiro valor do parâmetro

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

- ter variância mínima (precisão) $\Rightarrow$ O melhor estimador irá produzir estimativas mais próximas entre si (idealmente próximas ao verdadeiro valor do parâmetro)

  $$Var(\hat{\theta}_k) < Var(\hat{\theta}_j) \quad \forall k \neq j$$

Tiro ao alvo



Exato Impreciso | Inexato Preciso | Exato Preciso

Estimação Pontual de μ

Seja X uma v.a. com distribuição qualquer com média (μ) e variância (σ^2) também desconhecidas. Retira-se uma amostra de tamanho n com a finalidade de se estimar μ.

- média populacional μ

Ex. amostra com $n = 5$ $\{3,4;\ 4,5;\ 2,6;\ 3,8;\ 6,0\}$ $\bar{X} = \frac{3,4 + 4,5 + 2,6 + 3,8 + 6,0}{5} = 4,06$

De que maneira os valores da amostra podem ser combinados a fim de se produzir uma "boa" estimativa de μ?

Como não há nenhuma razão para acreditar que um valor da amostra é mais importante do que o outro:

$$\hat{\mu} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{X} = \underbrace{\sum_{j=1}^{N} x_j FR(X = x_j)}_{\text{dados agrupados (v.a. discreta)}}$$

média amostral

Estimação Pontual de μ

Seja X uma v.a. com distribuição qualquer com média (μ) e variância (σ^2) também desconhecidas. Retira-se uma amostra de tamanho n com a finalidade de se estimar μ.

- média populacional μ

Verificando a tendenciosidade de $\bar{X}$

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

$$= \frac{E\left(X_1\right) + E(X_2) + \cdots + E(X_n)}{n} = \frac{\cancel{n}\mu}{\cancel{n}} = \mu$$

estimador
não tendencioso

Interpretação (teórica): se calculássemos a média dos $\bar{X}$ de todas amostras (de tamanho n) possíveis de serem obtidas, o resultado seria μ

Estimação Pontual de μ

Seja X uma v.a. com distribuição qualquer com média (μ) e variância (σ^2) também desconhecidas. Retira-se uma amostra de tamanho n com a finalidade de se estimar μ.

- média populacional μ

Calculando a variância de $\bar{X}$ (avaliação de precisão)

$$Var(\bar{X}) = Var\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{Var\left(X_1 + X_2 + \cdots + X_n\right)}{n^2}$$

$$= \frac{Var(X_1) + Var(X_2) + \cdots + Var(X_n)}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

$\sigma^2 \quad + \quad \sigma^2 \quad + \ldots + \quad \sigma^2$

Se as amostras forem independentes, ou seja, se elas não guardarem nenhuma relação entre si.

Estimação Pontual de μ

Seja X uma v.a. com distribuição qualquer com média (μ) e variância (σ^2) também desconhecidas. Retira-se uma amostra de tamanho n com a finalidade de se estimar μ.

- média populacional μ

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad E(\bar{X}) = \mu \qquad \boxed{Var(\bar{X}) = \frac{\sigma^2}{n}}$$

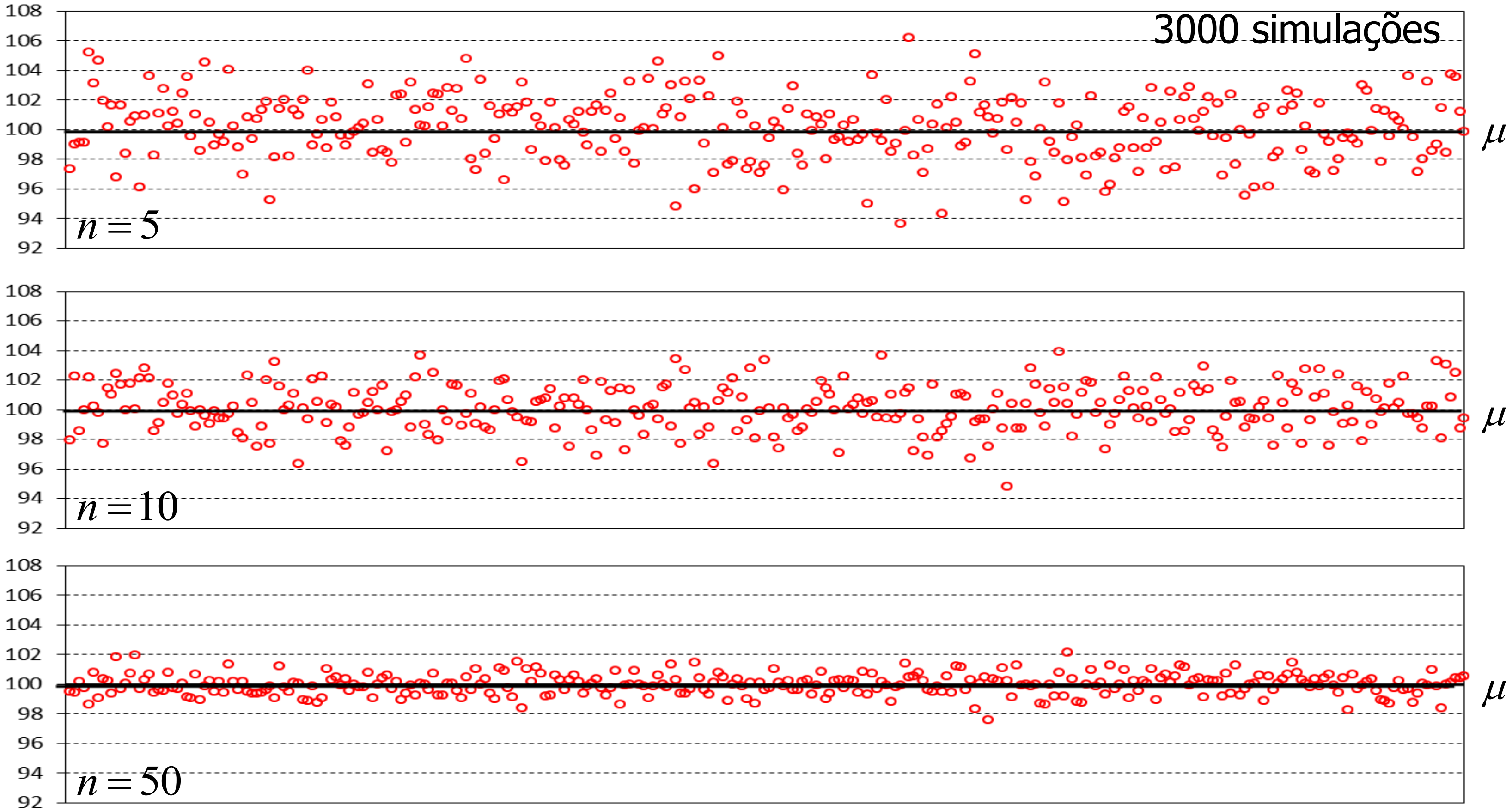
A precisão da média amostral depende da variação original dos dados (σ^2) e do tamanho da amostra (n)

Quanto maior o tamanho da amostra (n), mais precisa será a estimativa de μ

Estimação Pontual de μ

- média populacional μ $\quad E(\bar{X}) = \mu \quad Var(\bar{X}) = \dfrac{\sigma^2}{n}$

Simulando-se $\bar{X}$ a partir de amostras de uma v.a. $X \sim N(\mu = 100, \sigma^2 = 25)$



Estimação Pontual de σ^2

Seja X uma v.a. com distribuição qualquer com média (μ) e variância (σ^2) também desconhecidas. Retira-se uma amostra de tamanho n com a finalidade de se estimar σ^2.

- variância populacional σ^2

De que maneira os valores da amostra podem ser combinados a fim de se produzir uma "boa" estimativa de σ^2?

Como não há nenhuma razão para acreditar que um valor da amostra é mais importante do que o outro e μ é desconhecido:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2}{n}$$

Mas será um estimador tendencioso?

Estimação Pontual de σ^2

Seja X uma v.a. com distribuição qualquer com média (μ) e variância (σ^2) também desconhecidas. Retira-se uma amostra de tamanho n com a finalidade de se estimar σ^2.

- variância populacional σ^2

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2}{n}$$

$$\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = \sum_{i=1}^{n}\left(X_i^2 - 2\bar{X}X_i + \bar{X}^2\right)$$

$$= \sum_{i=1}^{n} X_i^2 - 2\bar{X}\sum_{i=1}^{n} X_i + n\bar{X}^2$$

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

$$= \sum_{i=1}^{n} X_i^2 - 2n\bar{X}^2 + n\bar{X}^2$$

$$= \sum_{i=1}^{n} X_i^2 - n\bar{X}^2$$

Estimação Pontual de σ^2

Seja X uma v.a. com distribuição qualquer com média (μ) e variância (σ^2) também desconhecidas. Retira-se uma amostra de tamanho n com a finalidade de se estimar σ^2.

- variância populacional σ^2

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2}{n} \qquad E\left(\hat{\sigma}^2\right) = E\left(\frac{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}{n}\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i^2\right) - E\left(\bar{X}^2\right)$$

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

$$Var(X_i) = \sigma^2 = E\left(X_i^2\right) - \left(E\left(X_i\right)\right)^2 = E\left(X_i^2\right) - \mu^2 \Rightarrow E\left(X_i^2\right) = \sigma^2 + \mu^2$$

Estimação Pontual de σ^2

Seja X uma v.a. com distribuição qualquer com média (μ) e variância (σ^2) também desconhecidas. Retira-se uma amostra de tamanho n com a finalidade de se estimar σ^2.

- variância populacional σ^2

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2}{n}$$

$$E\left(\hat{\sigma}^2\right) = E\left(\frac{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}{n}\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} X_i^2\right) - E\left(\bar{X}^2\right)$$

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

$$Var(\bar{X}) = \frac{\sigma^2}{n} = E\left(\bar{X}^2\right) - \left(E\left(\bar{X}\right)\right)^2 = E\left(\bar{X}^2\right) - \mu^2 \Rightarrow E\left(\bar{X}^2\right) = \frac{\sigma^2}{n} + \mu^2$$

Estimação Pontual de σ^2

Seja X uma v.a. com distribuição qualquer com média (μ) e variância (σ^2) também desconhecidas. Retira-se uma amostra de tamanho n com a finalidade de se estimar σ^2.

- variância populacional σ^2

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2}{n}$$

$$E\left(\hat{\sigma}^2\right) = E\left(\frac{\sum_{i=1}^{n} X_i^2 - n\bar{X}^2}{n}\right) = \frac{1}{n}E\left(\sum_{i=1}^{n} X_i^2\right) - E\left(\bar{X}^2\right)$$

$$E\left(X_i^2\right) = \sigma^2 + \mu^2$$

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

$$E\left(\bar{X}^2\right) = \frac{\sigma^2}{n} + \mu^2$$

$$= \sigma^2 + \cancel{\mu^2} - \frac{\sigma^2}{n} - \cancel{\mu^2} = \frac{n\sigma^2 - \sigma^2}{n} = \boxed{\frac{n-1}{n}\sigma^2}$$

estimador tendencioso!

Estimação Pontual de σ^2

Seja X uma v.a. com distribuição qualquer com média (μ) e variância (σ^2) também desconhecidas. Retira-se uma amostra de tamanho n com a finalidade de se estimar σ^2.

- variância populacional σ^2

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2}{\cancel{n}} \frac{\cancel{n}}{n-1} \qquad s^2 = \frac{\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2}{n-1} \qquad E\left(s^2\right) = \sigma^2$$

estimador
não tendencioso

variância amostral

Interpretação (teórica): se calculássemos a média dos s^2 de todas amostras (de tamanho n) possíveis de serem obtidas, o resultado seria σ^2

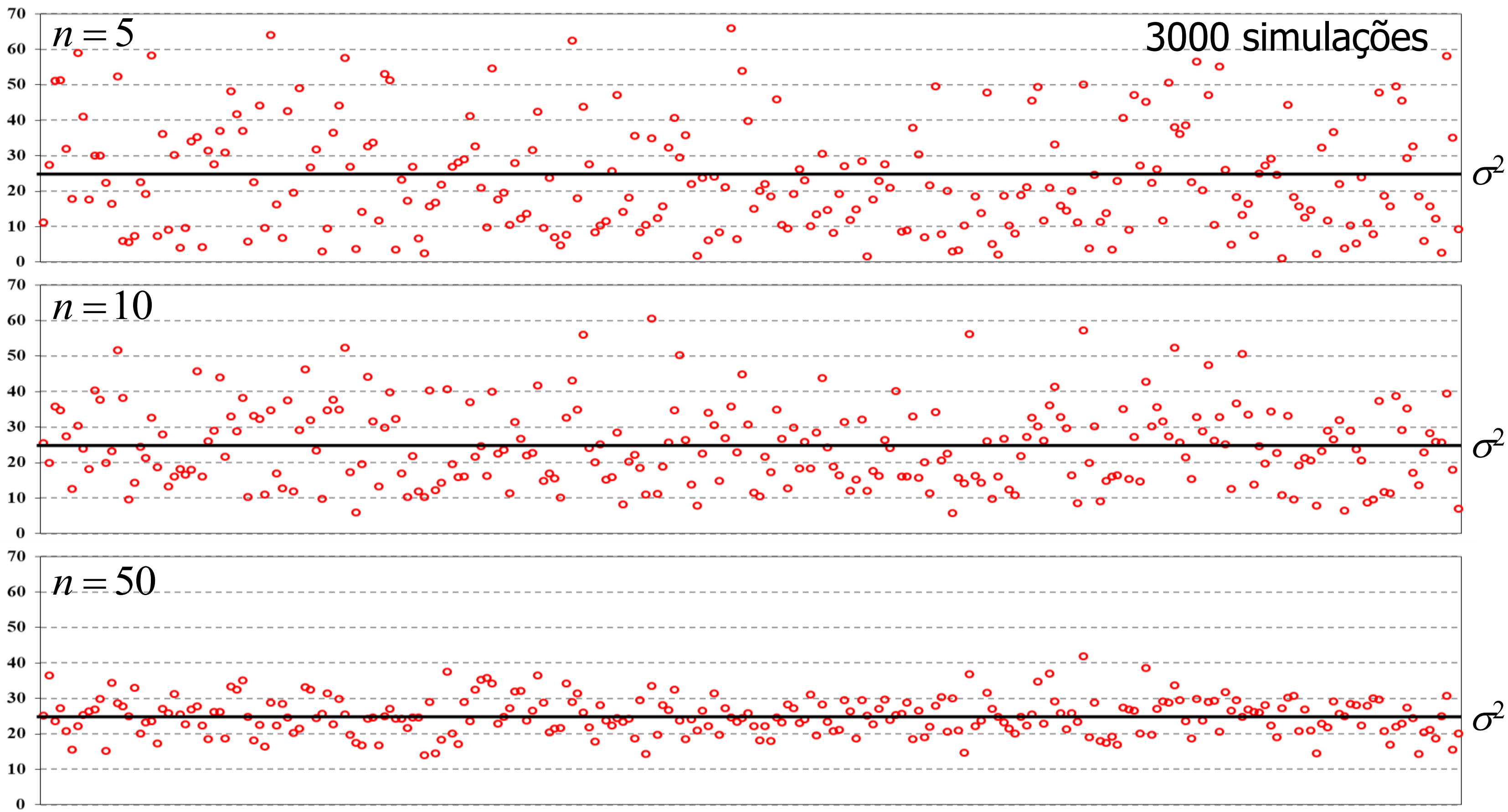
Curiosidade: $Var\left(s^2\right) = \dfrac{2\sigma^4}{n-1}$ (precisão aumenta com o tamanho da amostra!)

(ver Estimadores.xlsx)

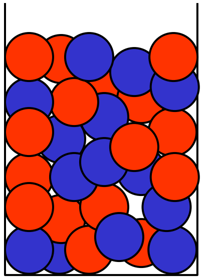
Estimação Pontual de σ^2

- variância populacional σ^2 $E(s^2) = \sigma^2$

Simulando-se s^2 a partir de amostras de uma v.a. $X \sim N(\mu = 100, \sigma^2 = 25)$



Estimação Pontual de p



Numa urna, há N bolas, sendo K vermelhas e $N - K$ azuis. Assim, pode-se dizer que K/N representa a proporção p de bolas vermelhas na urna (que por sua vez, representa a probabilidade de se selecionar uma bola vermelha desta urna).

Mas se N e K são desconhecidos, como estimar p?

Considere que n bolas são escolhidas ao acaso (com reposição), definindo-se Y como o número de bolas vermelhas entre as n selecionadas, qual a distribuição de Y?

$Y \sim$ Binomial(n, p) $\qquad Y = \sum_{i=1}^{n} X_i \qquad X_i \sim$ Bernoulli $\qquad p = P(X_i = 1) \Leftrightarrow P$(sucesso)

Qual é o melhor estimador pontual de p?

$$\frac{Y}{n} = \hat{p}$$

Proporção Amostral

$$E(\hat{p}) = E\left(\frac{Y}{n}\right) = \frac{E(Y)}{n} = \frac{np}{n} = p$$

estimador
não tendencioso

$$Var(\hat{p}) = Var\left(\frac{Y}{n}\right) = \frac{Var(Y)}{n^2} = \frac{npq}{n^2} = \frac{pq}{n}$$

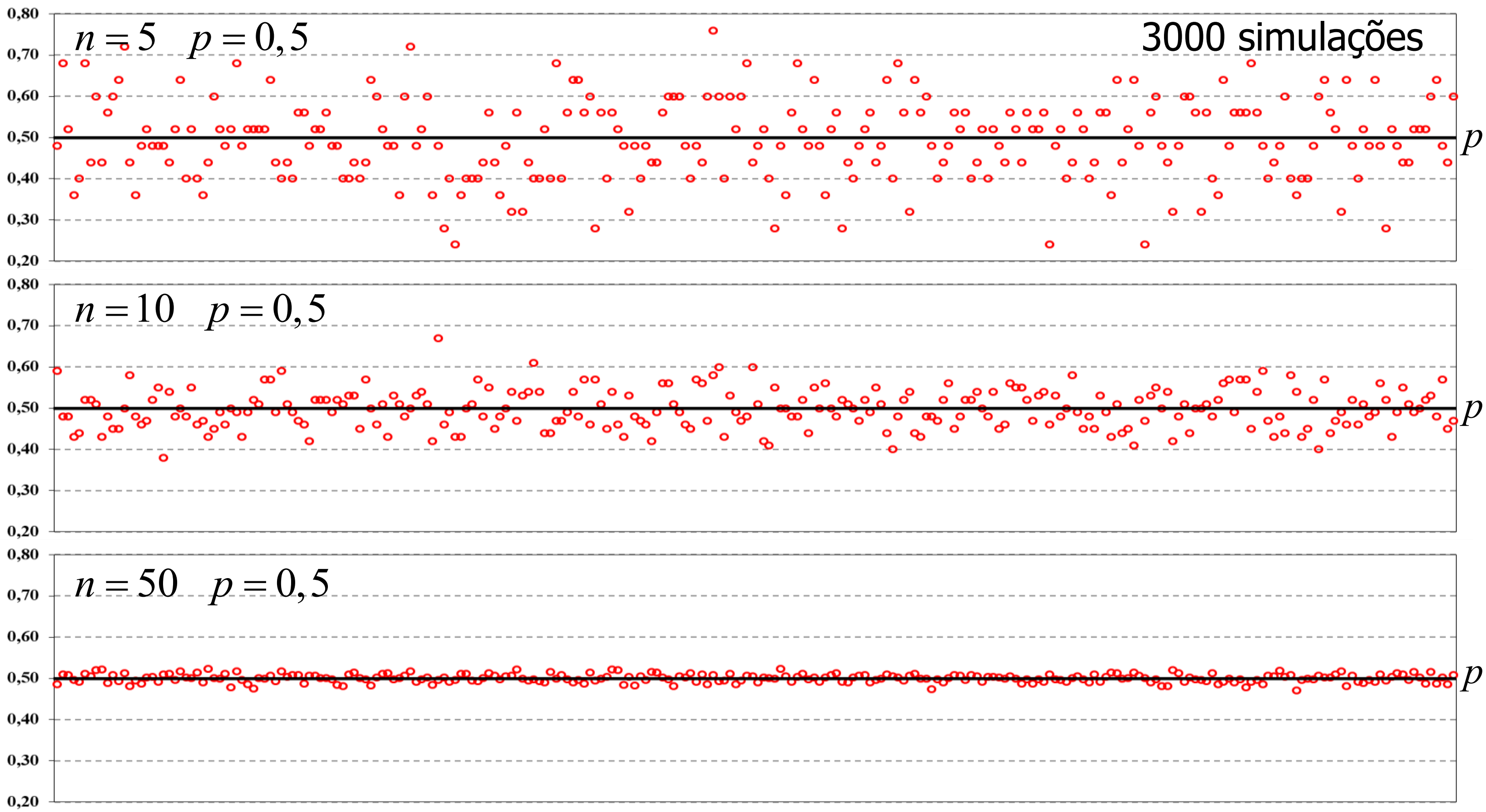
Quanto maior o tamanho da amostra (n), mais precisa será a estimativa de p

Quanto mais p se aproxima de 0,5 (50%), menos precisa será sua estimativa

Estimação Pontual de p

- proporção populacional p $\qquad E(\hat{p}) = p \qquad Var(\hat{p}) = \dfrac{pq}{n}$

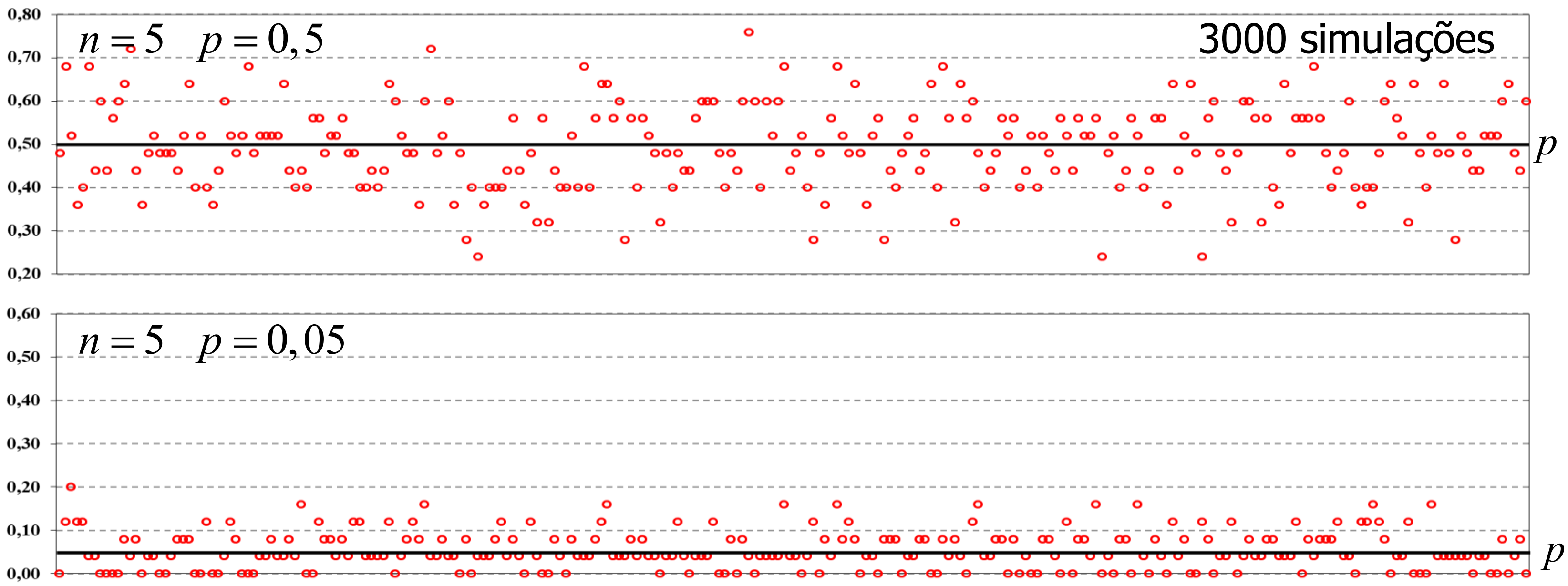
Simulando-se $\hat{p}$ a partir de amostras de uma v.a. $\quad X \sim Binomial(n, p = 0,5)$



Estimação Pontual de p

- proporção populacional p $\qquad E(\hat{p}) = p \qquad Var(\hat{p}) = \dfrac{pq}{n}$

Simulando-se $\hat{p}$ a partir de amostras de uma v.a. $\qquad X \sim Binomial(n=5, p)$



Estimação Pontual de μ e σ^2

Exemplo: uma amostra ($n = 12$) é retirada de uma população e os seguintes valores são observados: 0, 2, 3, 5, 2, 1, 2, 1, 3, 3, 4, 2. Calcule a média e variância amostrais.

distribuição amostral

Valor	Freq. Absoluta	Freq. Relativa
0	1	1/12
1	2	1/6
2	4	1/3
3	3	1/4
4	1	1/12
5	1	1/12
Total	12	1

- média amostral $\bar{X}$

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$ (dados brutos)

$$\bar{X} = \frac{0+2+3+\ldots+2}{12} = \frac{7}{3}$$

$$\bar{X} = \frac{\sum_{j=1}^{N} x_j FA(X = x_j)}{n} = \sum_{j=1}^{N} x_j FR(X = x_j)$$ (dados agrupados)

$$\bar{X} = \frac{0*1+1*2+2*4+3*3+4*1+5*1}{12} = \frac{7}{3}$$ (usando FA)

$$\bar{X} = 0*\frac{1}{12}+1*\frac{1}{6}+2*\frac{1}{3}+3*\frac{1}{4}+4*\frac{1}{12}+5*\frac{1}{12} = \frac{7}{3}$$ (usando FR)

Estimação Pontual de μ e σ^2

Exemplo: uma amostra ($n = 12$) é retirada de uma população e os seguintes valores são observados: 0, 2, 3, 5, 2, 1, 2, 1, 3, 3, 4, 2. Calcule a média e variância amostrais.

- variância amostral s^2

$$\bar{X} = \frac{7}{3}$$

distribuição amostral

Valor	Freq. Absoluta	~~Freq. Relativa~~
0	1	~~1/12~~
1	2	~~1/6~~
2	4	~~1/3~~
3	3	~~1/4~~
4	1	~~1/12~~
5	1	~~1/12~~
Total	12	~~1~~

$$s^2 = \frac{\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2}{n-1} = \frac{\sum_{i=1}^{n} x_i^2 - n\bar{X}^2}{n-1}$$ (dados brutos)

$$s^2 = \frac{(0-\frac{7}{3})^2 + (2-\frac{7}{3})^2 + \ldots + (2-\frac{7}{3})^2}{11} = \frac{(0^2 + 2^2 + \ldots + 2^2) - 12 * \left(\frac{7}{3}\right)^2}{11} = 1,88$$

$$s^2 = \frac{\sum_{j=1}^{N}\left(x_j - \bar{X}\right)^2 FA(X = x_j)}{n-1} = \frac{\sum_{j=1}^{N} x_j^2 FA(X = x_j) - n\bar{X}^2}{n-1}$$ (dados agrupados)

$$s^2 = \frac{(0-\frac{7}{3})^2 * 1 + (1-\frac{7}{3})^2 * 2 + \ldots + (5-\frac{7}{3})^2 * 1}{11} = \frac{(0^2 * 1 + 1^2 * 2 + \ldots + 5^2 * 1) - 12 * \left(\frac{7}{3}\right)^2}{11} = 1,88$$

Estimação Pontual de μ, σ^2 e p

Observações:

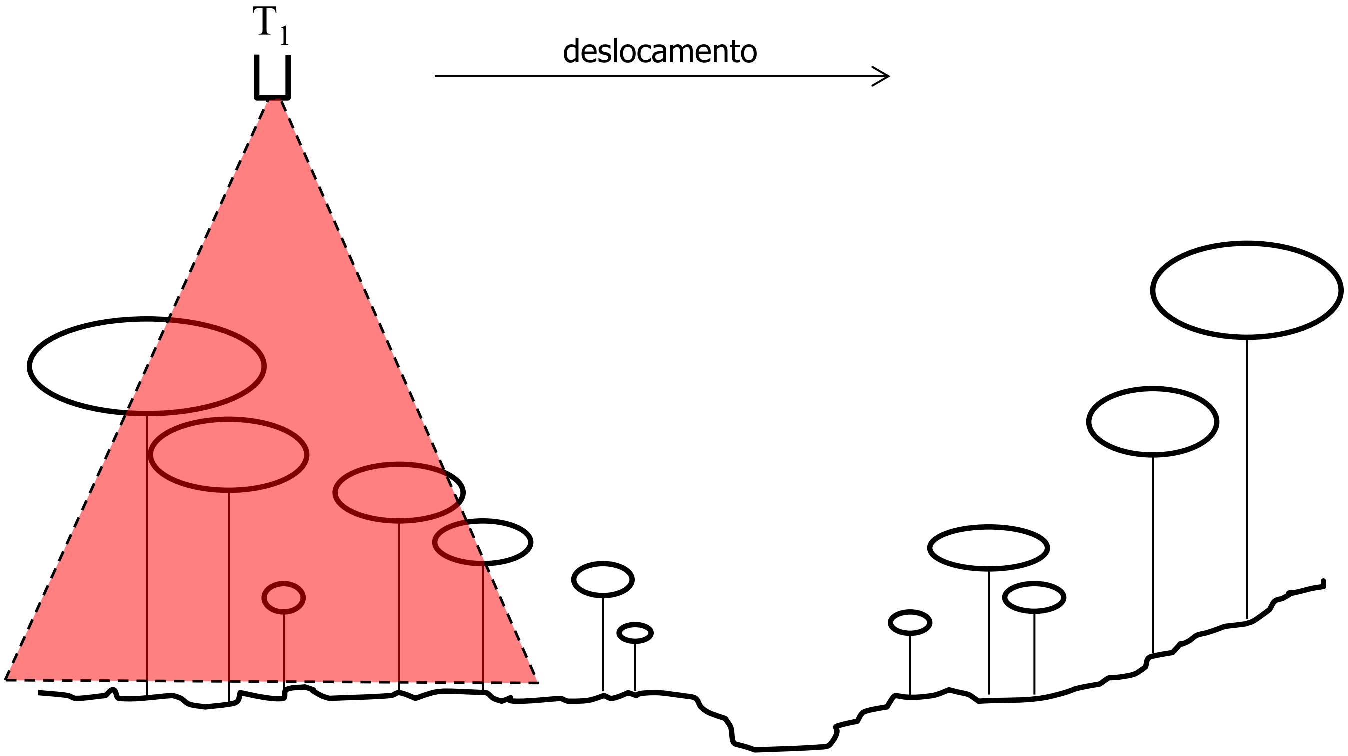
- μ, σ^2 e p são parâmetros que representam a população e portanto são valores fixos sendo, em geral, desconhecidos

- $\bar{X}$, s^2 e $\hat{p}$ são estatísticas calculadas a partir da amostra e representam variáveis aleatórias (cada conjunto de amostras pode apresentar um valor diferente)

- Não confunda variância amostral (s^2) com variância da média amostral ($\mathrm{Var}(\bar{X})$)

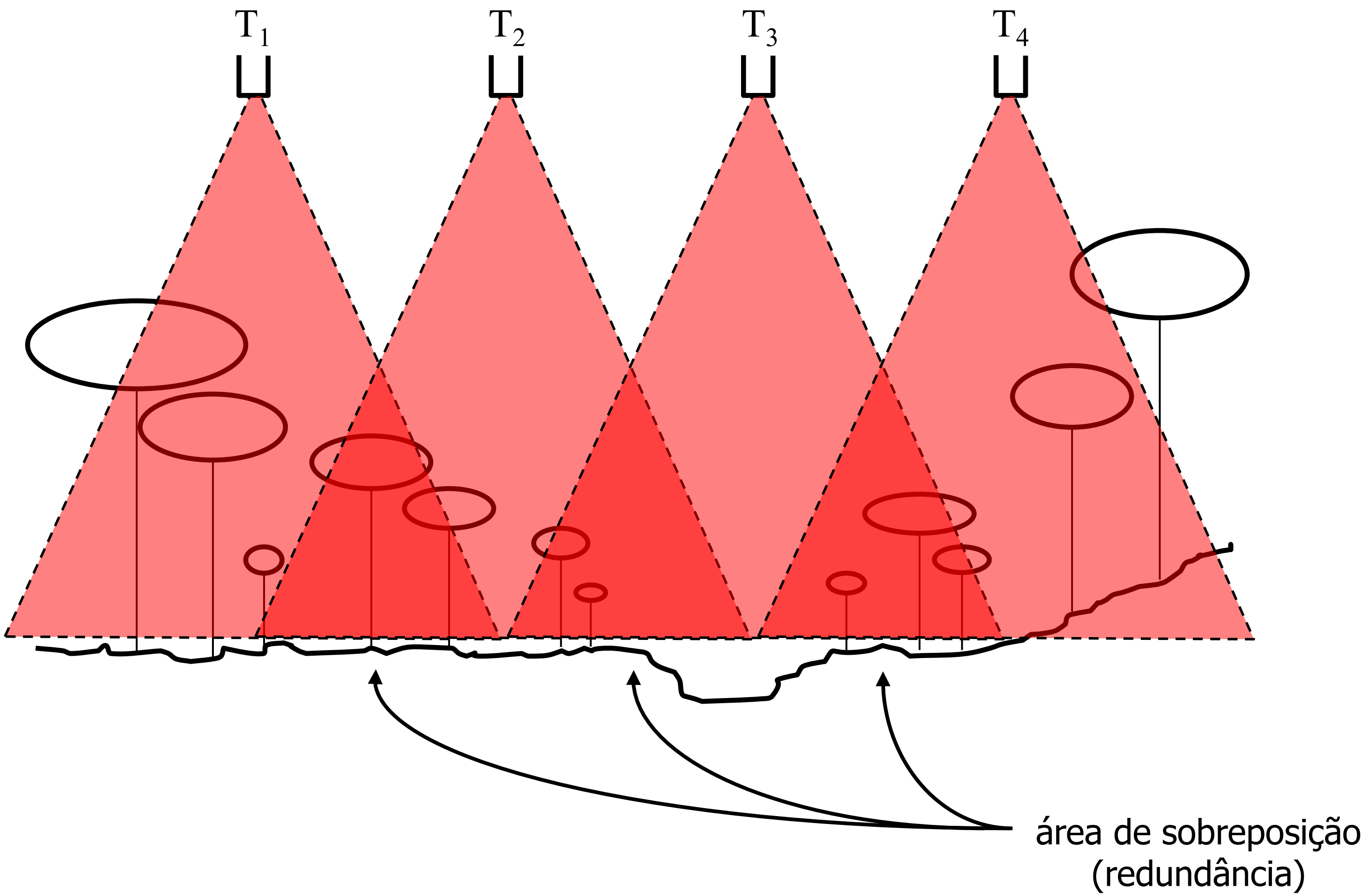
- De modo geral, as amostras devem ser obtidas de modo independente uma das outras, ou seja, o valor de uma amostra não deve ter relação com o(s) valor(es) das outras amostras (exceção em estudos de séries temporais ou dados espaciais, onde estuda-se exatamente esta relação)

Utilização de amostras não independentes

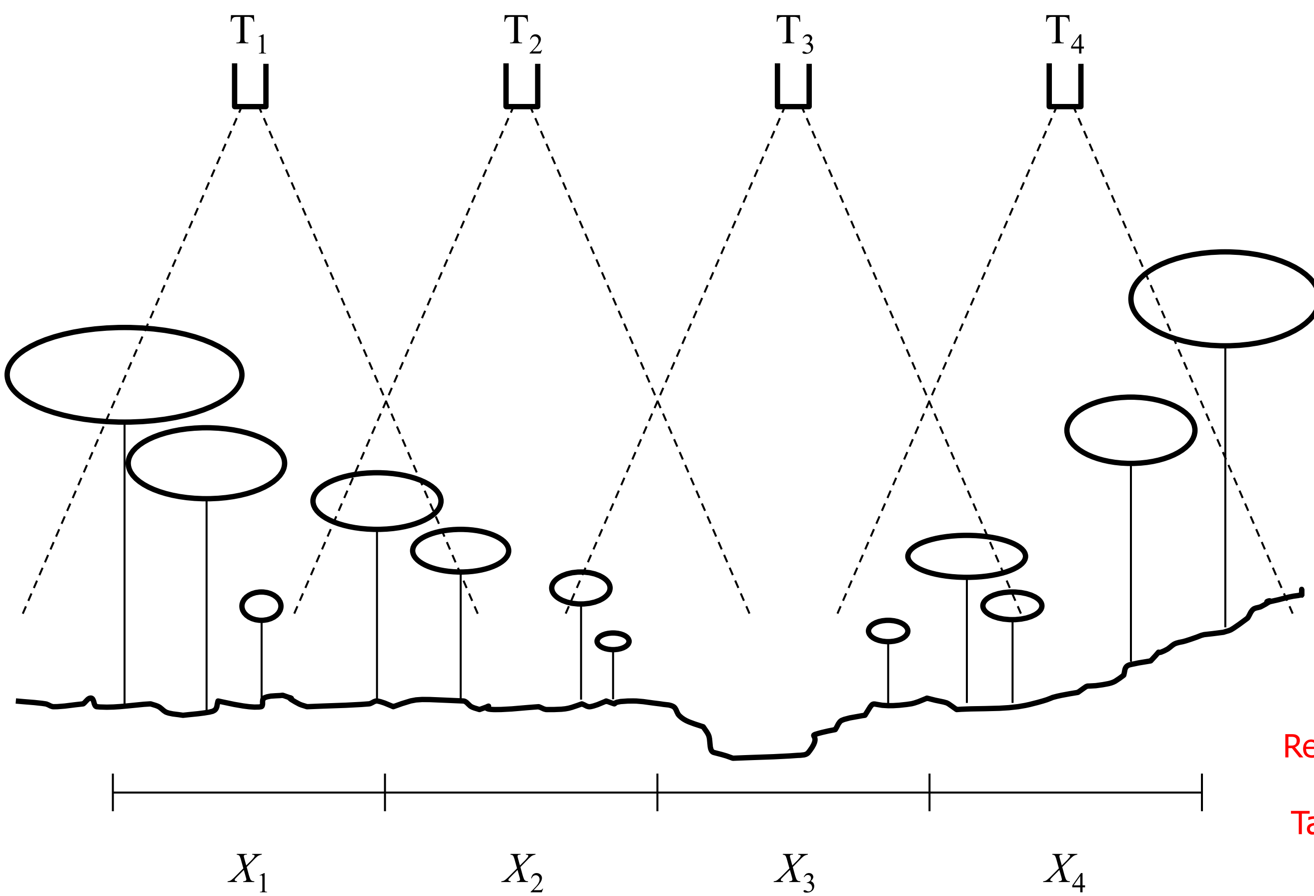


A resposta do sensor representa a integração das respostas de todos objetos que estão no campo de visada

Utilização de amostras não independentes



Utilização de amostras não independentes



Resolução Espacial
≠
Tamanho do Pixel

X_i: valor que representa a resposta do sensor no tempo T_i (≡ elemento de resolução)

Note que estes valores não são independentes (devido a sobreposição)

Utilização de amostras não independentes

Suponha que X representa um conjunto de amostras independentes de uma v.a. qualquer obtidas numa determinada sequência (série temporal por exemplo)

X	X'	X''	X'''
9			
6			
2			
3	3,2		
6			
2			
6			
10			
6			
5			
7			
1			
7			
8			
5			

X', X'' e X''' resultam do cálculo de médias móveis (tamanho 3) aplicado sobre X

$$X'_i = 0,1X_{i-1} + 0,8X_i + 0,1X_{i+1}$$

Utilização de amostras não independentes

Suponha que X representa um conjunto de amostras independentes de uma v.a. qualquer obtidas numa determinada sequência (série temporal por exemplo)

X	X'	X''	X'''
9			
6	5,9		
2	2,5		
3	3,2		
6	5,3		
2	2,8		
6	6		
10	9,2		
6	6,3		
5	5,3		
7	6,2		
1	2,2		
7	6,5		
8	7,6		
5			

X', X'' e X''' resultam do cálculo de médias móveis (tamanho 3) aplicado sobre X

$$X'_i = 0,1X_{i-1} + 0,8X_i + 0,1X_{i+1}$$

$$X''_i = 0,2X_{i-1} + 0,6X_i + 0,2X_{i+1}$$

$$X'''_i = \left(X_{i-1} + X_i + X_{i+1}\right)/3$$

Utilização de amostras não independentes

Suponha que X representa um conjunto de amostras independentes de uma v.a. qualquer obtidas numa determinada sequência (série temporal por exemplo)

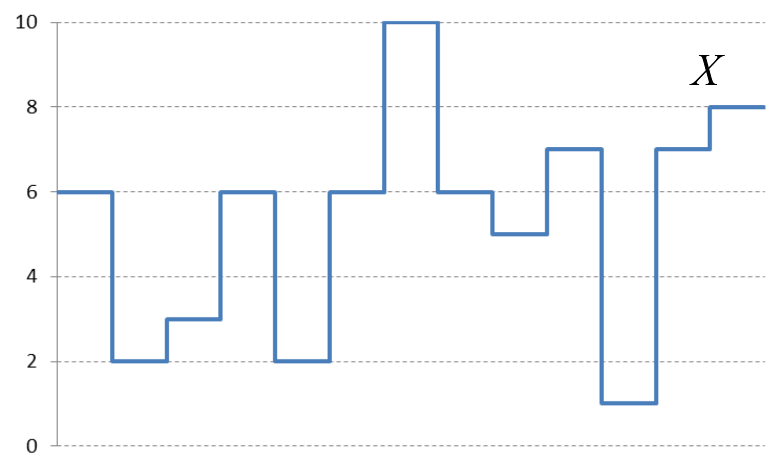
X	X'	X''	X'''
~~9~~			
6	5,9	5,8	5,67
2	2,5	3	3,67
3	3,2	3,4	3,67
6	5,3	4,6	3,67
2	2,8	3,6	4,67
6	6	6	6
10	9,2	8,4	7,33
6	6,3	6,6	7
5	5,3	5,6	6
7	6,2	5,4	4,33
1	2,2	3,4	5
7	6,5	6	5,33
8	7,6	7,2	6,67
~~5~~			

X', X'' e X''' resultam do cálculo de médias móveis (tamanho 3) aplicado sobre X

$$X_i' = 0,1X_{i-1} + 0,8X_i + 0,1X_{i+1}$$

$$X_i'' = 0,2X_{i-1} + 0,6X_i + 0,2X_{i+1}$$

$$X_i''' = \left(X_{i-1} + X_i + X_{i+1}\right)/3$$



Utilização de amostras não independentes

Suponha que X representa um conjunto de amostras independentes de uma v.a. qualquer obtidas numa determinada sequência (série temporal por exemplo)

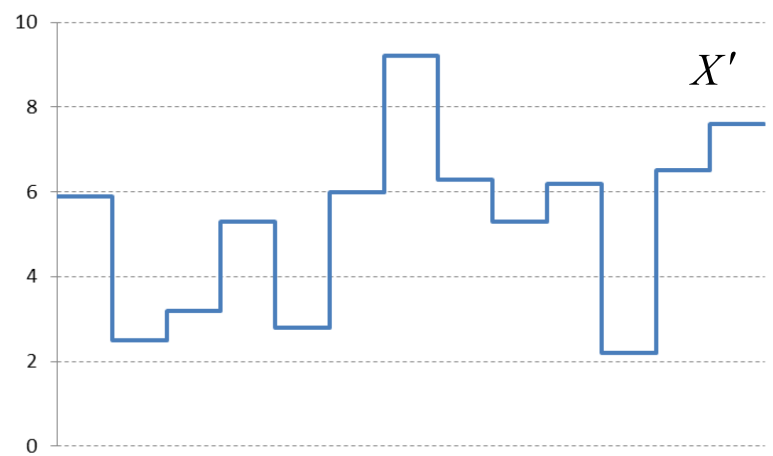
X	X'	X''	X'''
~~9~~			
6	5,9	5,8	5,67
2	2,5	3	3,67
3	3,2	3,4	3,67
6	5,3	4,6	3,67
2	2,8	3,6	4,67
6	6	6	6
10	9,2	8,4	7,33
6	6,3	6,6	7
5	5,3	5,6	6
7	6,2	5,4	4,33
1	2,2	3,4	5
7	6,5	6	5,33
8	7,6	7,2	6,67
~~5~~			

X', X'' e X''' resultam do cálculo de médias móveis (tamanho 3) aplicado sobre X

$$X_i' = 0,1X_{i-1} + 0,8X_i + 0,1X_{i+1}$$

$$X_i'' = 0,2X_{i-1} + 0,6X_i + 0,2X_{i+1}$$

$$X_i''' = \left(X_{i-1} + X_i + X_{i+1}\right)/3$$



Utilização de amostras não independentes

Suponha que X representa um conjunto de amostras independentes de uma v.a. qualquer obtidas numa determinada sequência (série temporal por exemplo)

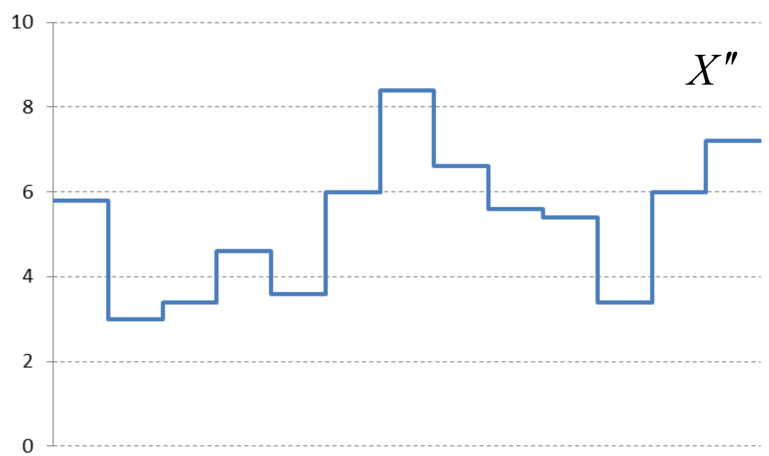
X	X'	X''	X'''
~~9~~			
6	5,9	5,8	5,67
2	2,5	3	3,67
3	3,2	3,4	3,67
6	5,3	4,6	3,67
2	2,8	3,6	4,67
6	6	6	6
10	9,2	8,4	7,33
6	6,3	6,6	7
5	5,3	5,6	6
7	6,2	5,4	4,33
1	2,2	3,4	5
7	6,5	6	5,33
8	7,6	7,2	6,67
~~5~~			

X', X'' e X''' resultam do cálculo de médias móveis (tamanho 3) aplicado sobre X

$$X_i' = 0,1X_{i-1} + 0,8X_i + 0,1X_{i+1}$$

$$X_i'' = 0,2X_{i-1} + 0,6X_i + 0,2X_{i+1}$$

$$X_i''' = \left(X_{i-1} + X_i + X_{i+1}\right)/3$$



Utilização de amostras não independentes

Suponha que X representa um conjunto de amostras independentes de uma v.a. qualquer obtidas numa determinada sequência (série temporal por exemplo)

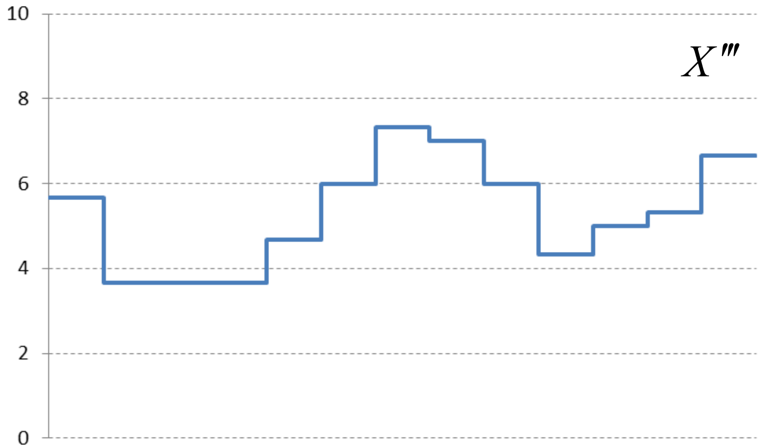
X	X'	X''	X'''
~~9~~			
6	5,9	5,8	5,67
2	2,5	3	3,67
3	3,2	3,4	3,67
6	5,3	4,6	3,67
2	2,8	3,6	4,67
6	6	6	6
10	9,2	8,4	7,33
6	6,3	6,6	7
5	5,3	5,6	6
7	6,2	5,4	4,33
1	2,2	3,4	5
7	6,5	6	5,33
8	7,6	7,2	6,67
~~5~~			

X', X'' e X''' resultam do cálculo de médias móveis (tamanho 3) aplicado sobre X

$$X_i' = 0,1X_{i-1} + 0,8X_i + 0,1X_{i+1}$$

$$X_i'' = 0,2X_{i-1} + 0,6X_i + 0,2X_{i+1}$$

$$X_i''' = \left(X_{i-1} + X_i + X_{i+1}\right)/3$$



Utilização de amostras não independentes

Suponha que X representa um conjunto de amostras independentes de uma v.a. qualquer obtidas numa determinada sequência (série temporal por exemplo)

X	X'	X''	X'''
~~9~~			
6	5,9	5,8	5,67
2	2,5	3	3,67
3	3,2	3,4	3,67
6	5,3	4,6	3,67
2	2,8	3,6	4,67
6	6	6	6
10	9,2	8,4	7,33
6	6,3	6,6	7
5	5,3	5,6	6
7	6,2	5,4	4,33
1	2,2	3,4	5
7	6,5	6	5,33
8	7,6	7,2	6,67
~~5~~			

X', X'' e X''' resultam do cálculo de médias móveis (tamanho 3) aplicado sobre X

$$X_i' = 0,1X_{i-1} + 0,8X_i + 0,1X_{i+1}$$

$$X_i'' = 0,2X_{i-1} + 0,6X_i + 0,2X_{i+1}$$

$$X_i''' = \left(X_{i-1} + X_i + X_{i+1}\right)/3$$

	$\bar{X}$	s^2
X	5,31	6,90
X'	5,31	4,39
X''	5,31	2,68
X'''	5,31	1,62

Conclusão: a utilização de amostras não independentes (autocorrelacionadas) afetam mais a estimação da variância do que a estimação da média