

Amostragem

# Roteiro

- ◇ Introdução
  - ◇ População x amostra
  - ◇ Definições e composição da amostra
- ◇ Métodos de amostragem
  - ◇ Amostras probabilísticas e não-probabilísticas
  - ◇ Principais métodos de amostragem
- ◇ Cálculo do tamanho de amostra
  - ◇ Estimativa de média e proporção

# População x Amostra

## ◇ População

- ◇ Conjunto de elementos que compartilham alguma (ou mais) característica(s)
- ◇ Ex:
  - ◇ Cavalos
  - ◇ Cavalos de esporte
  - ◇ Cavalos de esporte que participam de olimpíadas
  - ◇ etc

## ◇ Amostra

- ◇ Subconjunto da população
- ◇ Conjunto de elementos selecionados para um estudo

# Introdução

**População**



**Amostra**

# Censo x Amostra

## ◇ Censo

- ◇ Quando **todos** os elementos de uma população são investigados
- ◇ Determina exatamente a distribuição de uma variável na população
- ◇ Geralmente são difíceis ou mesmo impossíveis de serem realizados
  - ◇ Logística, tempo, recursos financeiros

## ◇ Amostragem

- ◇ Quando **alguns** elementos de uma população são investigados
- ◇ Vantagens em relação ao censo
  - ◇ Maior velocidade na obtenção dos resultados
  - ◇ Menor custo

# População alvo e População de Estudo

- ◇ População alvo: População sobre a qual se quer uma resposta (objetivo)
- ◇ População de estudo: População da qual a amostra é obtida

População alvo = População de estudo (ideal)

População alvo  $\approx$  População de estudo (aceitável)

População alvo  $\neq$  População de estudo (ruim)

# Amostra

- ◇ Unidade amostral
  - ◇ Unidade observada da população de estudo, sobre a qual os dados são obtidos
    - ◇ Indivíduo (animal)
    - ◇ Agregado de indivíduos (rebanho, fazendas ou regiões administrativas)
    - ◇ País
- ◇ Quadro amostral
  - ◇ Lista que contém os membros da população de estudo

# Viés

- ◇ Objetivo da amostragem é providenciar uma estimativa sem viés (erro) da variável estudada em uma população
- ◇ Estimativas tendenciosas (com viés) acontecem quando
  - ◇ Lista de membros do quadro amostral é incompleta
  - ◇ Informação é obsoleta
  - ◇ Segmentos da população não são traçáveis
  - ◇ Recusa na participação ou produção de respostas erradas
  - ◇ Processo de seleção dos indivíduos não é aleatório
- ◇ **Vieses não são compensados** pelo aumento do tamanho de amostra

# Tipos de amostra

- ◇ Amostras não-probabilísticas
  - ◇ Critério de escolha é definido pelo investigador
  - ◇ Não são necessariamente representativas da população alvo
  - ◇ Não permitem a comprovação de hipóteses
    - ◇ Mas permitem a geração de hipóteses
- ◇ Amostras probabilísticas
  - ◇ Cada unidade amostral tem alguma probabilidade de ser selecionada (base da amostra aleatória)
  - ◇ São representativas da população alvo

# Tipos de amostra

- ◇ Amostras não-probabilísticas
  - ◇ Amostra de conveniência
  - ◇ Amostra intencional
  
- ◇ Amostras probabilísticas
  - ◇ Amostra aleatória simples
  - ◇ Amostra sistemática
  - ◇ Amostra estratificada
  - ◇ Amostra por aglomerados
  - ◇ Amostra multi-estágio

# Métodos de amostragem

- ◇ Amostragem por conveniência
  - ◇ Seleção de unidades amostrais mais acessíveis
  - ◇ Usualmente não representativa da população alvo
  - ◇ Ex.: estudo de câncer de mama em cães atendidos em um determinado hospital veterinário
  
- ◇ Amostragem intencional
  - ◇ Seleção intencional, baseada no conhecimento sobre o problema e a propósito do estudo
  - ◇ Ex.: Cães apenas de uma determinada raça

# Métodos de amostragem

## ◇ Amostragem aleatória simples

- ◇ Cada unidade amostral da população tem a mesma, e conhecida, probabilidade de ser selecionada a partir de um quadro amostral

1. Numeração de todos os indivíduos da população de estudo

2. Seleção do número necessário de indivíduos

3. Escolha dos indivíduos

- ◇ Sorteio
- ◇ Tabela de números aleatórios
- ◇ Geração de números aleatórios no computador

**Tabela. População de estudo de 15 cães. Amostrar aleatoriamente 5 cães.**

<b>Nome</b>	<b>Peso (Kg)</b>	<b>Idade (anos)</b>	<b>Sexo (M/F)</b>	<b>Espécie</b>	<b>Raça</b>
<b>Sansão</b>	<b>20</b>	<b>4</b>	<b>Macho</b>	<b>canina</b>	<b>SRD</b>
<b>Pepe</b>	<b>6</b>	<b>6</b>	<b>Macho</b>	<b>canina</b>	<b>Poodle</b>
<b>Pamela</b>	<b>29,1</b>	<b>8</b>	<b>Fêmea</b>	<b>canina</b>	<b>Afganhound</b>
<b>Nick</b>	<b>15,6</b>	<b>6</b>	<b>Macho</b>	<b>canina</b>	<b>Beagle</b>
<b>Lucas</b>	<b>14,5</b>	<b>7</b>	<b>Macho</b>	<b>canina</b>	<b>Beagle</b>
<b>July</b>	<b>8,9</b>	<b>3,75</b>	<b>Fêmea</b>	<b>canina</b>	<b>Bishon Frisé</b>
<b>Bree</b>	<b>20</b>	<b>3</b>	<b>Macho</b>	<b>canina</b>	<b>Boxer</b>
<b>Guga</b>	<b>26</b>	<b>0,8</b>	<b>Macho</b>	<b>canina</b>	<b>Boxer</b>
<b>Spike</b>	<b>39</b>	<b>7</b>	<b>Macho</b>	<b>canina</b>	<b>Boxer</b>
<b>Biba</b>	<b>32</b>	<b>6,7</b>	<b>Fêmea</b>	<b>canina</b>	<b>Boxer</b>
<b>Sky</b>	<b>21,8</b>	<b>3</b>	<b>Macho</b>	<b>canina</b>	<b>Chow chow</b>
<b>Duque</b>	<b>13</b>	<b>7</b>	<b>Macho</b>	<b>canina</b>	<b>Cocker</b>
<b>Votan</b>	<b>6,31</b>	<b>0,33</b>	<b>Macho</b>	<b>canina</b>	<b>Cocker</b>
<b>Dique</b>	<b>27,4</b>	<b>8</b>	<b>Macho</b>	<b>canina</b>	<b>Collie</b>
<b>Pituca</b>	<b>9</b>	<b>4</b>	<b>Fêmea</b>	<b>canina</b>	<b>Daschound</b>

# Métodos de amostragem

## ◇ Amostragem sistemática

◇ As unidades amostrais são selecionados em intervalos iguais, que podem ser definidos inicialmente, sendo a primeira selecionada ao acaso

◇  $I = N / n$

◇ I = intervalo

◇ N = tamanho da população alvo

◇ n = número de indivíduos a selecionar

◇ Útil em algumas situações de campo quando

◇ não se dispõe da numeração de todos os indivíduos

◇ é difícil individualizar o animal correspondente ao número selecionado



Image © 2008 DigitalGlobe

©2008 Google

<http://earth.google.com/>

# Métodos de amostragem

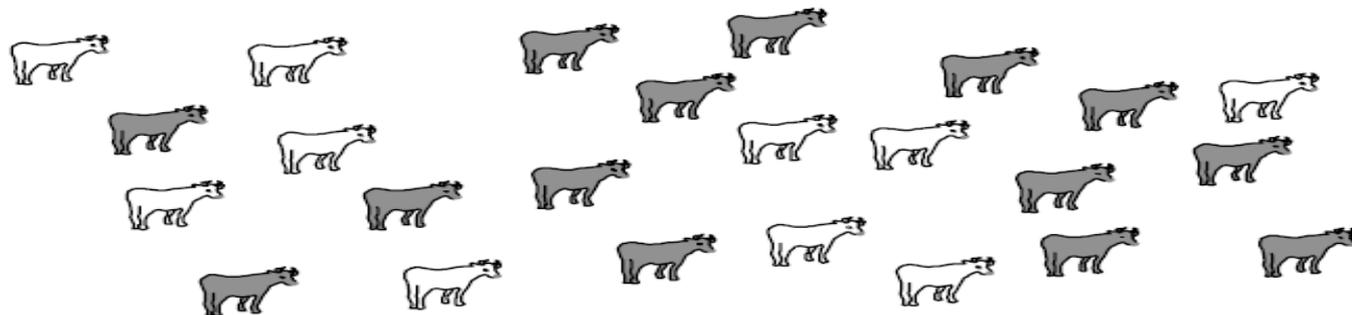
## ◇ Amostragem estratificada

- ◇ Permite uma representação adequada de todos os grupos da população alvo na composição final da amostra

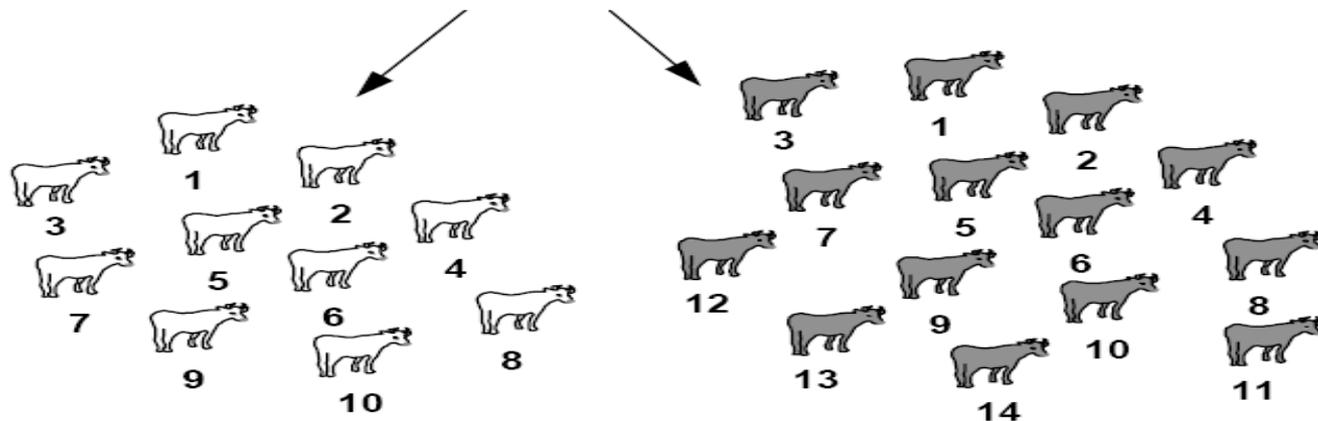
1. A população alvo é subdividida em grupos (**estratos**) com base em características importantes para o estudo

2. Os indivíduos em cada estrato são selecionados aleatoriamente (várias amostras simples aleatórias)

- ◇ Diminui os efeitos (sub ou superestimação) de uma amostra simples aleatória
- ◇ Deve-se conhecer antecipadamente a composição da população em relação a essas características



**Estratificação por raça**



**Raça A**

**Raça B**

**Tomar uma amostra  
simples aleatória**

**Tomar uma amostra  
simples aleatória**

**Tabela. Amostragem estratificada de vacas em diferentes regiões da Grã-Bretanha, baseada em amostragem de 5% do total das vacas. (Dados extraídos de Wilson *et al.* 1983)**

Região	Número de vacas	Número amostrado
Devon e Cornwall	302.647	$302.647 \times 0,05 = 15.132$
Sul da Inglaterra	469.486	$469.486 \times 0,05 = 23.474$
Gales	342.346	$342.346 \times 0,05 = 17.117$
Norte da Inglaterra	273.838	$273.838 \times 0,05 = 13.692$
Total	1.388.317	69.415

Adaptado de: THRUSFIELD, M. *Epidemiologia Veterinária*. p. 227. 2004.

# Métodos de amostragem

- ◇ Amostragem por aglomerados
  - ◇ Aplicação em nível agregado (grupo) de unidades individuais
    - ◇ Indivíduo continua sendo a unidade de interesse
    - ◇ Unidade amostral transforma-se em grupo de indivíduos
  - ◇ Amostragem de grupo em um estágio
    - ◇ Seleção de alguns grupos
    - ◇ Amostragem de todos os indivíduos nos grupos selecionados
  - ◇ Amostragem de grupo em dois estágios
    - ◇ Seleção de alguns grupos
    - ◇ Amostragem de alguns indivíduos nos grupos selecionados

Qual a prevalência de *E. coli* em

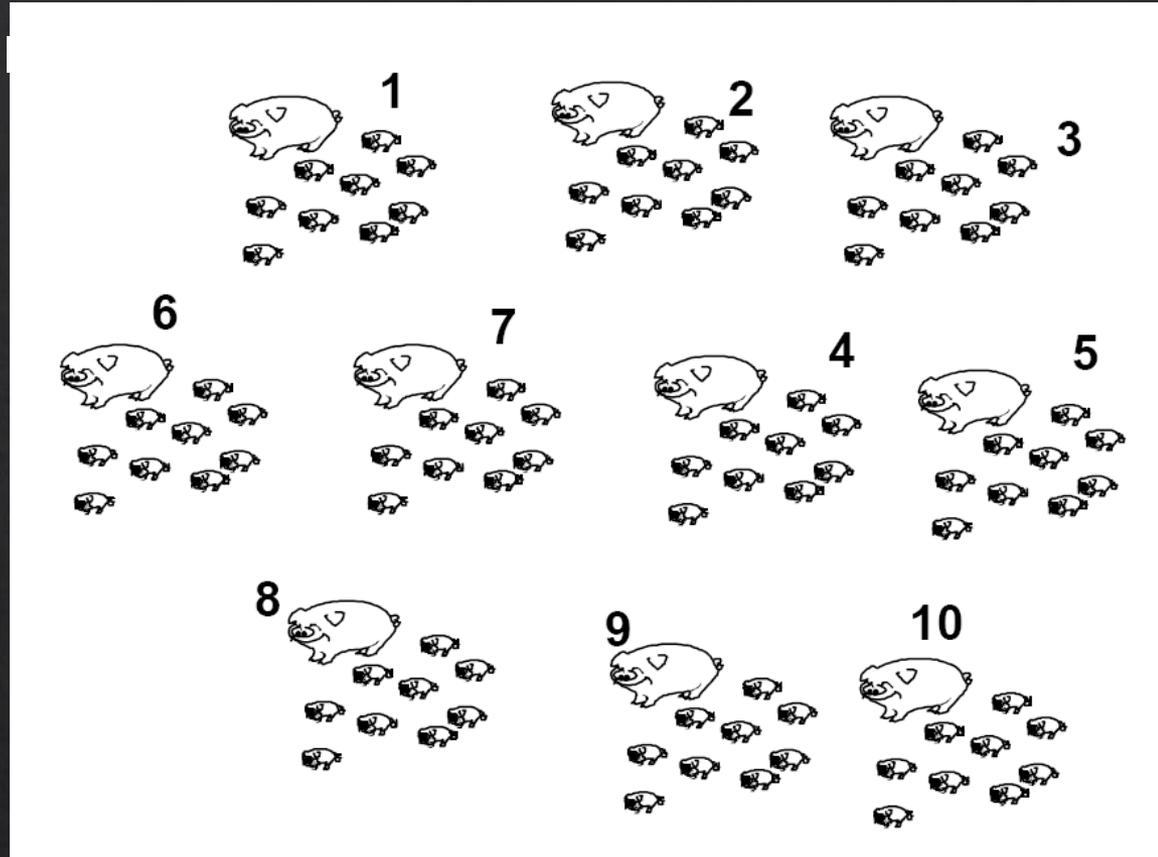
Amostra de 30 leitões



9 leitões por porca



Amostrar 4 porcas



**Porca = aglomerado de leitões**

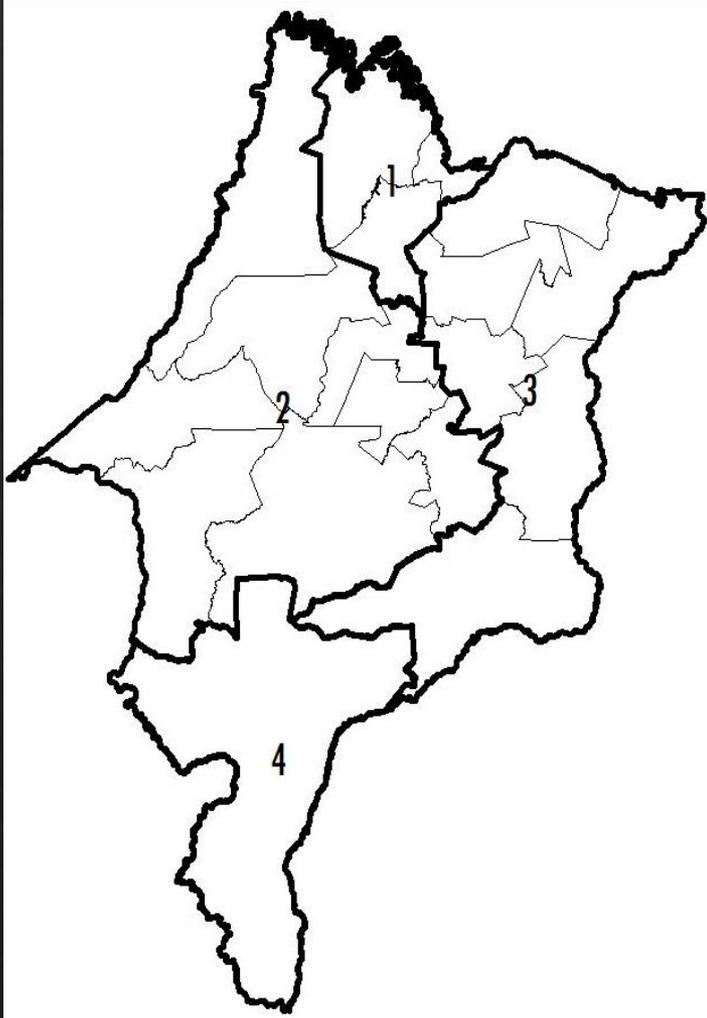
Adaptado de: PFEIFFER, D. Veterinary Epidemiology - An Introduction. 2002.

**Desvantagem: grande variação inter-aglomerados e pequena variação intra-aglomerados resultará em estimativas enviesadas (variância elevada).**

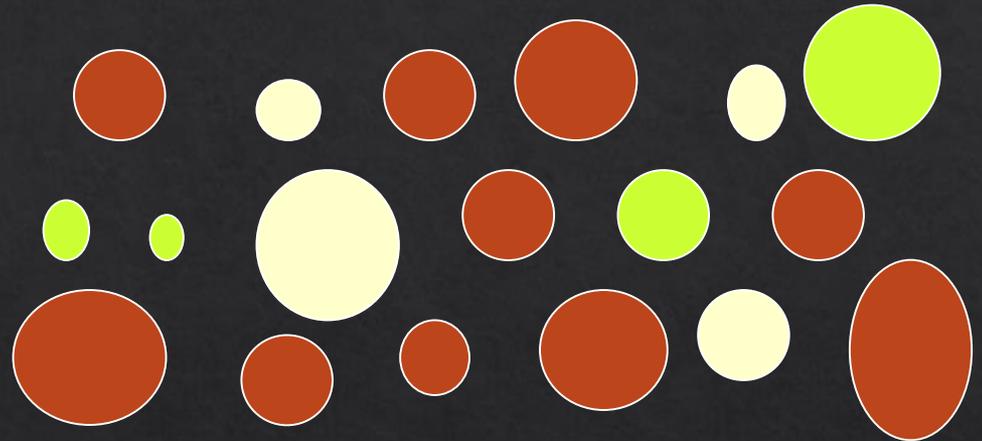
# Métodos de amostragem

- ◇ Amostragem multi-estágios
  - ◇ Deve ser efetuada em dois ou mais estágios
  - ◇ Em cada estágio pode ser aplicado um dos métodos vistos anteriormente
  - ◇ Método aplicado em estudos amostrais em grande escala, com populações heterogêneas
    1. Amostrar determinadas regiões administrativas
    2. Amostrar fazendas leiteiras em cada região
    3. Amostrar vacas em cada fazenda

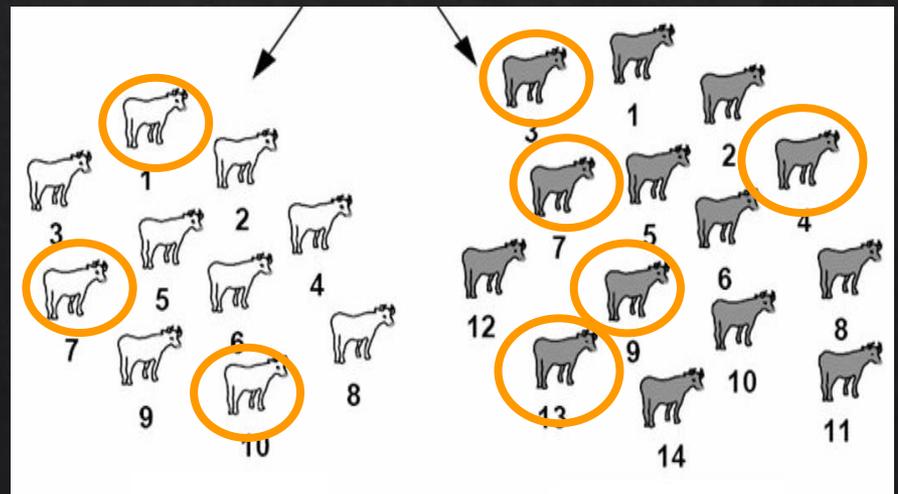
◇ Estimativa de animais e propriedades positivas para a brucelose bovina



**1º. Amostragem aleatória de propriedades/rebanhos**



**2º. Amostragem aleatória de fêmeas com idade  $\geq 2$  anos**



## Conveniência



## Intencional



## Aleatória



## Sistemática



# Tamanho da amostra

- ◇ Quantos indivíduos devem ser selecionados para o estudo?  
Depende dos objetivos e das circunstâncias do estudo.
- ◇ Tamanho da amostra é determinado por:
  - ◇ Fatores não-estatísticos:
    - ◇ Infra-estrutura
    - ◇ Existência de quadros amostrais
    - ◇ Disponibilidade de recursos financeiros e logísticos
  - ◇ Fatores estatísticos:
    - ◇ Erro desejado
    - ◇ Desvio-padrão

# Estimativa de uma média

## Amostra aleatória simples

- ◇ Distribuição normal

$$n = 1,96^2 \frac{\sigma^2}{\varepsilon^2}$$

- ◇  $n$  = tamanho da amostra
- ◇  $\sigma$  = desvio padrão
- ◇  $\varepsilon$  = erro máximo desejado

# Estimativa de proporção

## Amostra aleatória simples

- ◇ Aproximação distribuição normal da binomial

$$n = \frac{1,96^2 P_{esp} (1 - P_{esp})}{d^2}$$

- ◇  $n$  = tamanho da amostra
- ◇  $P_{esp}$  = proporção esperada
- ◇  $d$  = erro desejado

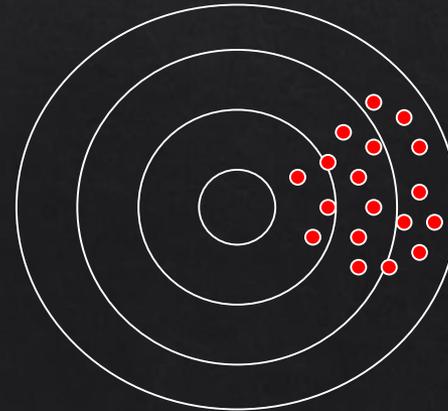
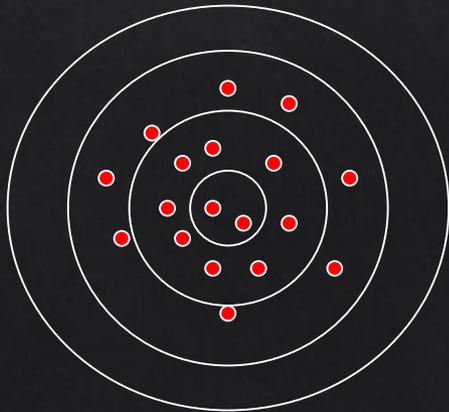
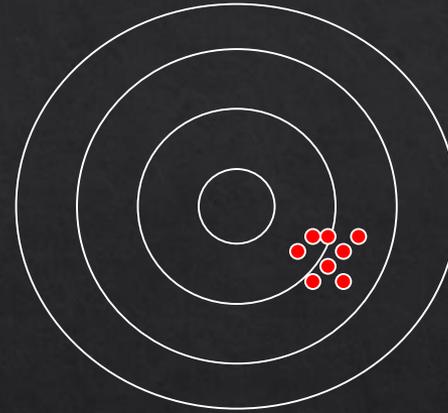
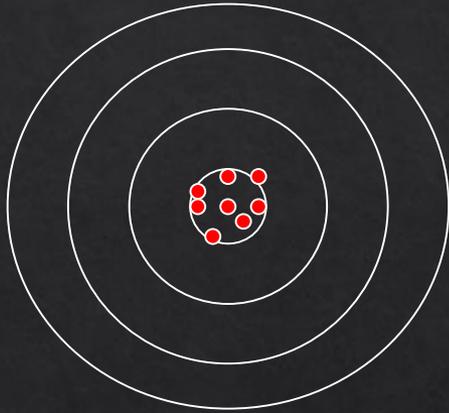
# Populações finitas

- ◇ Correção do tamanho amostral para populações finitas

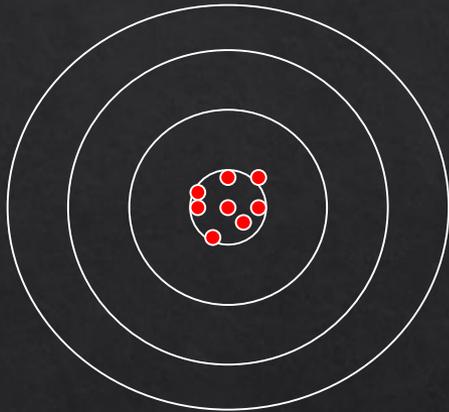
$$n_{ajus} = \frac{N * n}{N + n}$$

- $n_{ajus}$  = tamanho da amostra ajustado
- $N$  = tamanho da população total
- $n$  = tamanho inicial da amostra

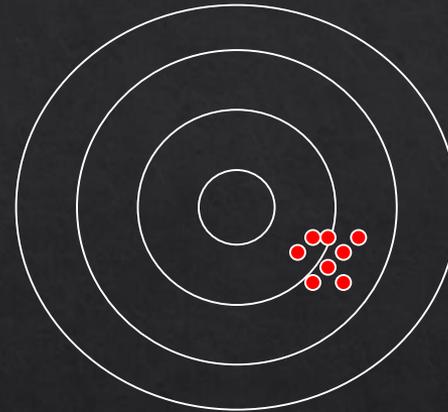
# Precisão x Acurácia



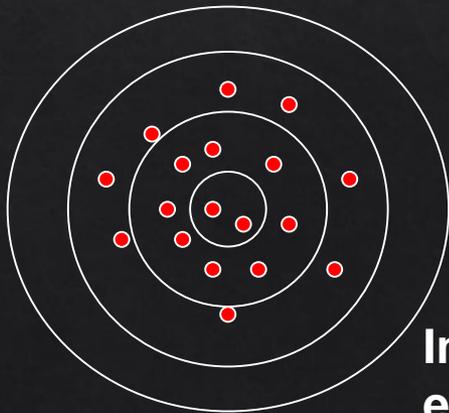
# Precisão x Acurácia



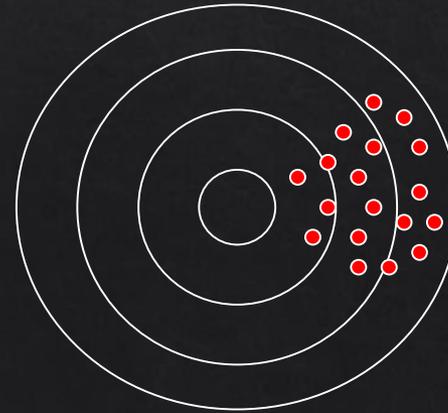
**Preciso  
e  
acurado**



**Preciso e  
não-  
acurado**



**Impreciso  
e acurado**



**Impreciso  
e não-  
acurado**

# Referências

1. COCHRAN, W.G. **Sampling techniques**. 3ª ed, John Wiley & Sons, 1967.
2. PEREIRA, M.G. **Epidemiologia - Teoria e Prática**. Guanabara Koogan, p. 337-357, 2005.
3. STEVENSON, M. **An introduction to veterinary epidemiology**. EpiCentre, IVABS, p. 62-71, 2007.  
[http://epicentre.massey.ac.nz/Portals/0/EpiCentre/Downloads/Education/227-407/Stevenson\\_intro\\_epidemiology\\_2007.pdf](http://epicentre.massey.ac.nz/Portals/0/EpiCentre/Downloads/Education/227-407/Stevenson_intro_epidemiology_2007.pdf)
4. THRUSFIELD, M. **Epidemiologia Veterinaria**. 2ª ed, Roca, p.223-239, 2004.
5. THRUSFIELD, M. **Veterinary epidemiology**. 3ª ed, Blackwell Publishing, p.228-246, 2005.