

Amostragem



-
- 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 para estimar média e proporção
-

- População

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

- Amostra

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

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

Amostra

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

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

População alvo x 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)

Viés

Objetivo da amostragem é providenciar uma estimativa sem viés

É um erro sistemático da variável estudada em uma população

Em inglês: bias

Viés

- Composição de amostras inadequadas levam à problemas de viés (erros de estimativas)
- Viés de seleção
- Viés de detecção
- Viés de exclusão
- Viés de sobrevivência
- ...
- Comparar população alvo ~ população de estudo

Viés

- 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

Vieses

não são compensados
pelo aumento do tamanho
de amostra

Não confundir viés de amostragem (ou viés estatístico, etc) com viés cognitivo (cognitive bias):

- Viés de confirmação
 - Viés da autoconveniência
 - Viés retrospectivo
 - Pareidolia
 - Memória
 - Etc
-

Tipos de amostra

- **Amostras não-probabilísticas**
 - Critério de escolha é definido pelo investigador
 - Não são 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
 - 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 conglomerados
 - 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

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



<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



Raça A

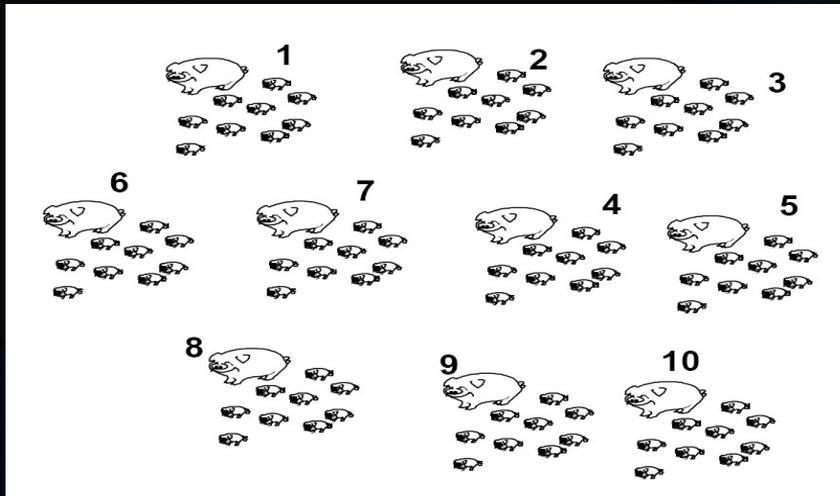
Raça B

Tomar uma amostra simples aleatória

Tomar uma amostra simples aleatória

Métodos de amostragem

- Amostragem por conglomerados
 - Aplicação em nível agregado (grupo) de unidades individuais
 - Indivíduo continua sendo a unidade de interesse
 - Selecionam-se grupos de indivíduos
 - Todos os indivíduos de um grupo são amostrados
 - Vantagem: logística
 - Desvantagem: grande variação inter-aglomerados (grupos muito diferentes um do outro) e pequena variação intra-aglomerados (indivíduos dentro de um mesmo grupo muito parecidos entre si) resultará em estimativas enviesadas (variância elevada).



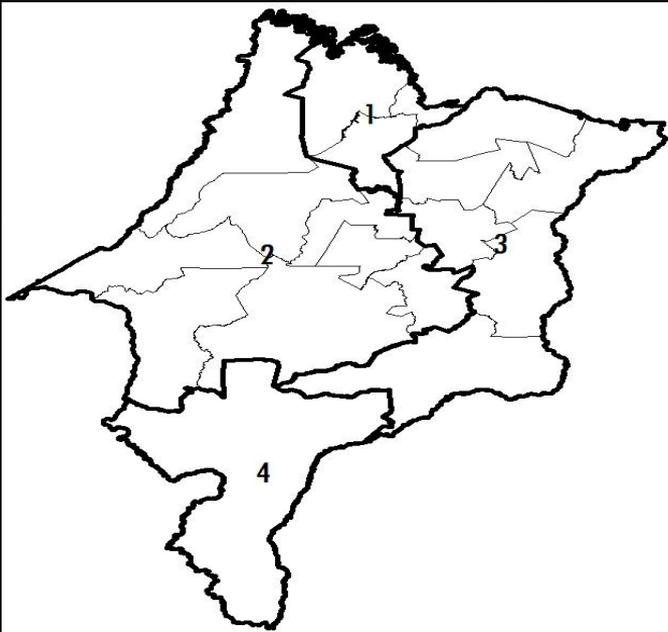
Porca = conglomerado de leitões

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

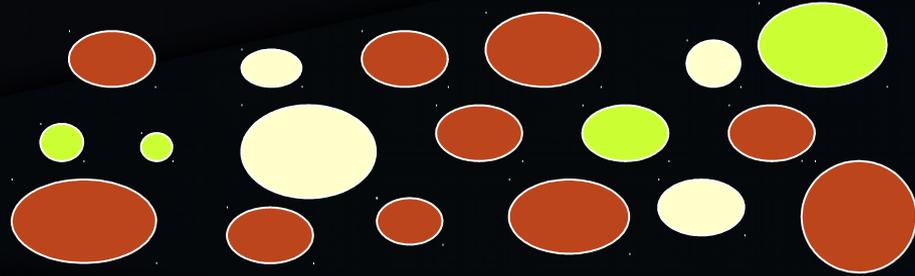
Métodos de amostragem

- 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

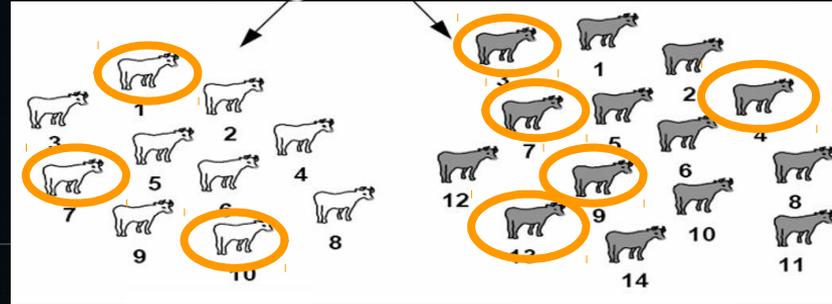
Estimativa de animais e propriedades positivas para a brucelose bovina



1º. Amostragem aleatória de propriedades/rebanhos



2º. Amostragem aleatória de fêmeas



Resumo

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

- n = tamanho da amostra
- σ = desvio padrão
- ε = erro máximo desejado

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

Estimativa de proporção

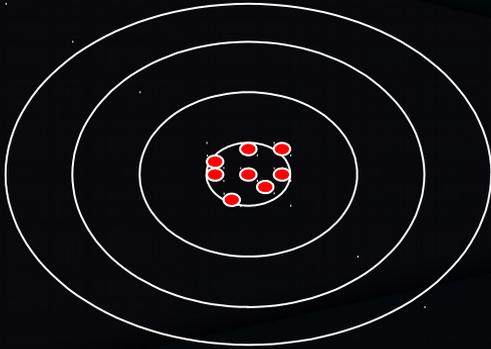
Amostra aleatória simples

- Aproximação distribuição normal da binomial
- n = tamanho da amostra
- P_{esp} = proporção esperada
 - (não sabe, chuta 50%)
- d = erro desejado

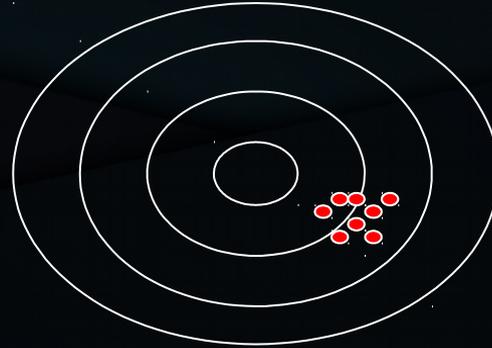
$$n = \frac{1,96^2 P_{esp} (1 - P_{esp})}{d^2}$$

Precisão x Acurácia

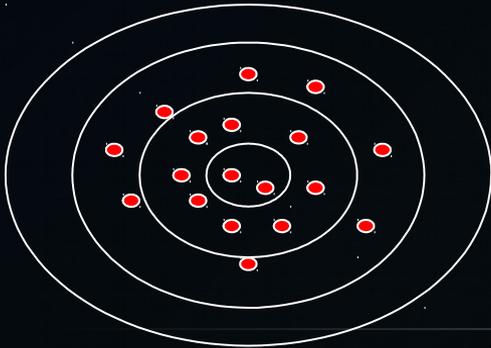
**Preciso e
acurado**



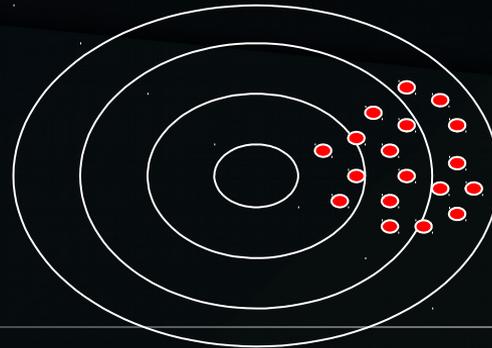
**Preciso e
não-acurado**



**Impreciso
e acurado**



**Impreciso e
não-acurado**



Obrigado