

Métodos de amostragem
Cálculo do tamanho de amostra

Prof. Grisi
e Prof. Mauro Borba (UFRGS)

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 prevalência
 - Detecção de doença

Fontes de variabilidade

- Variabilidade biológica (genética, ambiente, gênero (M/F), idade)
ex. dois grupos de animais submetidos a diferentes dietas.
Pergunta: Qual dos dois grupos apresenta a melhor conversão alimentar?



Fontes de variabilidade

- **Erros técnicos ou erros de medição**
Fatores que geram erros de medição:
 - **humanos** (variações entre observadores, fadiga, cuidado com observação, estimativa do último dígito)
 - **instrumentais**

Fontes de variabilidade

- Valor verdadeiro:

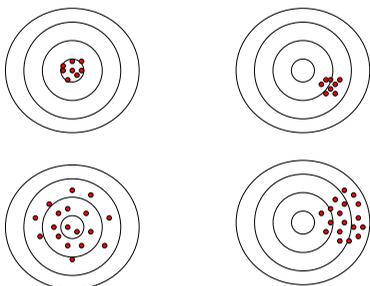
$$\varepsilon = X - X_0$$

erro = valor medido – valor verdadeiro

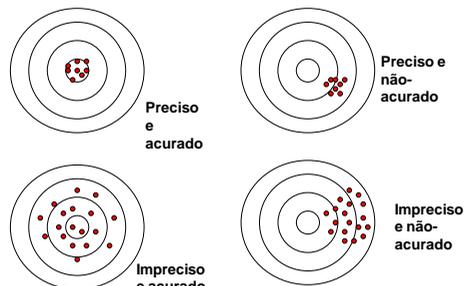
Nem todas as grandezas possuem valor verdadeiro.

- **Erros estatísticos:** gerados por variações incontroláveis e aleatórias no processo de medição.
Ex. variações na contagem do número de células por unidade de área em uma lâmina observada em microscópio.
- **Erros sistemáticos:** gerados por variações de caráter não-aleatório.
Ex. balança que mede sempre 1 kg a mais.
- **Erros grosseiros:** erros de digitação, falhas na anotação de um número.

Precisão x Acurácia



Precisão x Acurácia



Introdução

Desafio do profissional da saúde em um estudo epidemiológico: reunir, adequadamente, um número suficiente de indivíduos para a composição do(s) grupo(s) a ser(em) investigado(s) (Pereira, 2005).

- Todos os indivíduos da população devem ser incluídos no estudo ou apenas uma amostra?
- Quantos indivíduos são necessários para formar uma amostra?
- Como obter corretamente uma amostra?

Objetivos e circunstâncias do estudo.

Introdução



Introdução

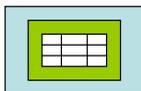
- **População**
 - Conjunto de elementos que compartilham ou apresentam, em comum, uma ou mais características
 - Classificada em dois tipos
 - Finita (ex.: número de clínicas veterinárias no Município de São Paulo)
 - Infinita (ex.: número de zooplâncton marinho presente nos oceanos)
- **Amostra**
 - Subconjunto da população
 - Conjunto de elementos selecionados para um estudo

Introdução

- **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
 - Pode-se obter estimativas acuradas de uma variável na população
 - Vantagens em relação ao censo
 - Estudo operacionalmente mais simples e mais detalhado
 - Maior velocidade na obtenção dos resultados
 - Menor custo

Introdução

- **População alvo**
 - População total sobre a qual a informação é requerida
- **População de estudo**
 - População da qual a amostra é obtida
 - Deveria ser a população alvo, mas nem sempre é possível
- **Unidade elementar (amostral)**
 - Divisão menor da população de estudo (em geral é o animal)
- **Estrato**
 - Conjunto de unidades amostrais agrupadas de acordo com características comuns
- **Quadro amostral**
 - Lista que contém os membros da população de estudo



Introdução

- **Natureza das unidades amostrais**
 - Indivíduo (animal)
 - Agregado de indivíduos (rebanho, fazendas ou regiões administrativas)
- **Unidade amostral x unidade epidemiológica**
 - Estudos de doenças infecciosas
- **Composição da amostra**
 - Critérios de inclusão
 - Critérios de exclusão

Introdução

- **Objetivo da amostragem é providenciar uma estimativa sem vícios da variável estudada em uma população**
- **Estimativas tendenciosas (vícios) 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
- **Esses vícios não são compensados pelo aumento do tamanho de amostra**

Introdução

Aplicações de amostragem em Medicina Veterinária

- Caracterização da população de cães e gatos domiciliados em São Paulo
- Detecção dos fatores de risco de lesões musculares em cavalos de corrida
- Estimativa de animais e propriedades positivas para a brucelose bovina
- Detecção da prevalência de *Salmonella* spp. em granjas de suínos utilizando alimentação seca ou líquida
- Estimativa da densidade de aves criadas em sistema fundo de quintal
- Detecção sorológica de agentes infecciosos em animais selvagens
- Quantificação da densidade populacional de camarões em viveiros

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 probabilidade igual 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
 - Obtenção de informações de forma mais simples e rápida
 - Cuidado ao extrapolar resultados para a população alvo
 - Ex.: estudo de câncer de mama em cães atendidos em um determinado hospital veterinário
- **Amostragem intencional**
 - Seleção intencional e subjetiva de unidades amostrais
 - Cuidado ao extrapolar resultados para a população alvo
 - Ex.: estudo de produção de leite em bovinos participantes de uma feira agropecuária

Tabela. População de estudo de 15 cães. Amostragem aleatoriamente 5 cães.

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

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

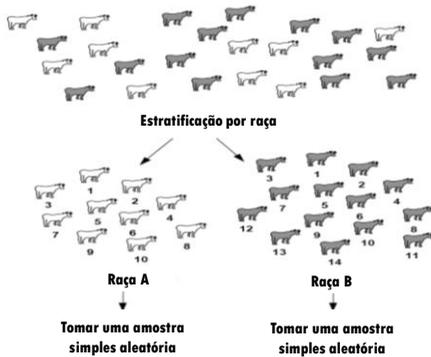
- **Amostragem sistemática**
 - As unidades amostrais são selecionadas em intervalos iguais, que podem ser definidos inicialmente, sendo a primeira selecionada ao acaso
 - $I = N / n$
 - I = intervalo
 - N = tamanho da população alvo
 - n = número de indivíduos a selecionar
 - Útil em algumas situações 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



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

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

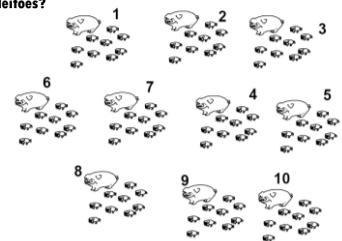
- **Amostragem por aglomerados**
 - Aplicação em nível agregado (grupo) de unidades individuais
 - Indivíduo continua sendo a unidade de interesse
 - Unidade amostral transforma-se em grupo de indivíduos
 - Amostragem de grupo em um estágio
 - Seleção de alguns grupos
 - Amostragem de todos os indivíduos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

Amostra de 30 leitões

9 leitões por porca

Amostrar 4 porcas



Porca = aglomerado de leitões

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

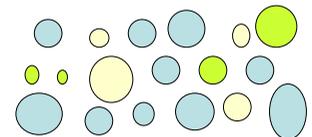
Desvantagem: grande variação inter-aglomerados e pequena variação intra-aglomerados resultará em estimativas enviesadas (variância elevada).

Métodos de amostragem

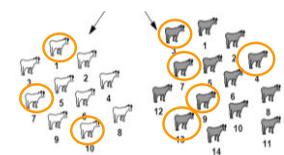
- **Amostragem multi-estágios**
 - Deve ser efetuada em dois ou mais estágios
 - Em cada estágio pode ser aplicado um dos métodos vistos anteriormente
 - Método aplicado em estudos amostrais em grande escala, com populações heterogêneas
 1. Amostrar determinadas regiões administrativas
 2. Amostrar fazendas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

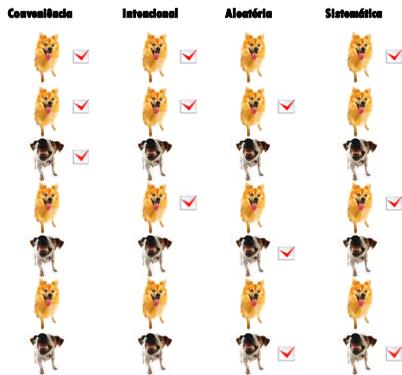


1°. Amostragem aleatória de propriedades/rebanhos



2°. Amostragem aleatória de fêmeas com idade ≥ 2 anos





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"
 - Prevalência esperada

Estimativa de média

Amostra aleatória simples

- Distribuição normal
- População infinita

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

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

Correção para populações pequenas

- Distribuição normal
- População finita

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

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

A) Qual o tamanho de amostra necessária para se estimar o peso médio de cervos em uma dada população (considerada infinita) de estudo. Sabe-se que o desvio padrão de peso para animais dessa idade é de 30 kg. Utilize um erro de 10 kg e confiança de 95%. B) Se essa população fosse composta por 200 cervos, qual o número de animais deveria ser amostrado?

$\sigma = 30$
 $\varepsilon = 10$

$N = 200$

A)

$$n = 1,96^2 * \frac{30^2}{10^2}$$

$$n = 3,84 * 9 = 34,57$$

$$n = 35$$

B)

$$n = \frac{200 * 35}{200 + 35}$$

$$n = \frac{7000}{235} = 29,78$$

$$n = 30$$

Estimativa de prevalência

Amostra aleatória simples

- Aproximação distribuição normal da binomial
- População infinita

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

- n = tamanho da amostra
- P_{esp} = prevalência esperada
- d = erro

Correção para populações pequenas

- Aproximação distribuição normal da binomial
- População finita

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

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

Uma determinada doença em bovinos apresenta uma prevalência esperada de 30% (dados de literatura). A) Quantos animais devem ser amostrados para se poder estimar a prevalência desta doença em uma dada população (considerada infinita) de estudo. Utilize um erro de 5% e confiança de 95%. B) Se essa população fosse composta por 900 bovinos, qual o número de animais deveria ser amostrado?

$P_{esp} = 30\%$
 $d = 5\%$

$N = 900$

A)

$$n = \frac{1,96^2 * 0,30 * (1 - 0,30)}{0,05^2}$$

$$n = \frac{3,84 * 0,21}{0,0025} = \frac{0,8}{0,0025} = 322,69$$

$$n = 323$$

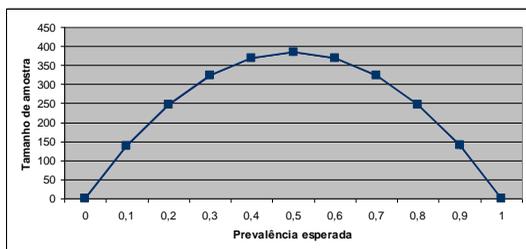
B)

$$n = \frac{900 * 323}{900 + 323}$$

$$n = \frac{290700}{1223} = 237,69$$

$$n = 238$$

Tamanho da amostra



Detecção de doença

Detecção de doença sem preocupação com a prevalência

$$n = [1 - (1 - p_1)^{\frac{1}{d}}] * [N - \frac{d}{2}] + 1$$

- n = tamanho da amostra
- p_1 = probabilidade de se encontrar, ao menos, um animal infectado
- d = número de animais infectados
- N = tamanho da população

Um proprietário de um tanque de criação de peixes deseja saber se os seus peixes estão livres de um vírus. Qual o número de peixes que devem ser examinados de modo a detectar pelo menos 1 animal positivo com confiança de 95%, sabendo-se que se o vírus estiver presente, deve estar infectando 20% da população de 5000 peixes do tanque.

$p_1 = 95\%$
 $d = 20\% = 1000$
 $N = 5000$

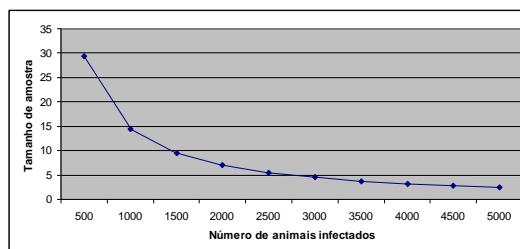
$$n = [1 - (1 - 0,95)^{\frac{1}{1000}}] * [5000 - \frac{1000}{2}] + 1$$

$$n = [1 - 0,9970] * [5000 - 500] + 1$$

$$n = [0,003] * [4500] + 1 = 14,46$$

$$n = 15$$

Tamanho da amostra



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
4. THRUSFIELD, M. **Epidemiologia Veterinária**. 2ª ed, Roca, p.223-239, 2004.
5. THRUSFIELD, M. **Veterinary epidemiology**. 3ª ed, Blackwell Publishing, p.228-246, 2005.