

# Como calcular a amostra na pesquisa odontológica?



Mauro Henrique Nogueira Guimarães de Abreu

Universidade Federal de Minas Gerais  
2010



27ª Reunião Anual da SBPqO

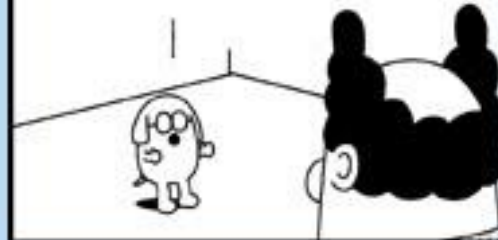
09 a 12 de setembro de 2010  
Águas de Lindóia - São Paulo - SP

# % DOGBERT, O ESTATÍSTICO %

Que culpa temos nós se nosso brilho natural ofusca a mente do chefão?



... E então, depois da análise exploratória do box-plot, passamos a etapa da Regressão Linear.



Depois geramos um Q-Q-Plot para checar a normalidade dos resíduos e daí testamos a homocedasticidade do modelo.



Falta de fala... Olhar confuso... Esse homem foi vítima de consultoria estatística.



# Referências

1. Babbie, E. Métodos de pesquisa de survey. Belo Horizonte: Editora UFMG, 2001. 519p.
2. Hulley, S.B. et al. Delineando a pesquisa clínica: uma abordagem epidemiológica. 3.ed. São Paulo: Artmed, 2008. 384p.
3. Luiz, R.R. et al. Planejamento amostral. In: Luiz, R.R.; Costa, A.J.L.; Nadanovsky, P. Epidemiologia e bioestatística na pesquisa odontológica. São Paulo: Atheneu, 2005. p.245-272.
4. Lwanga, S.K.; Lemeshow, S. Sample size determination in health studies. Geneva: WHO, 1991. 80p.
5. Pereira, M.G. Epidemiologia. Teoria e Prática. Rio de Janeiro: Guanabara Koogan, 1995. 583p.
6. Riffenburgh, R.H. Statistics in medicine. 2.ed. San Diego: Elsevier, 2006. 622p.
7. Silva, N.N. Amostragem probabilística. 2.ed. São Paulo: EDUSP, 2004. 235p.



# Conteúdo programático

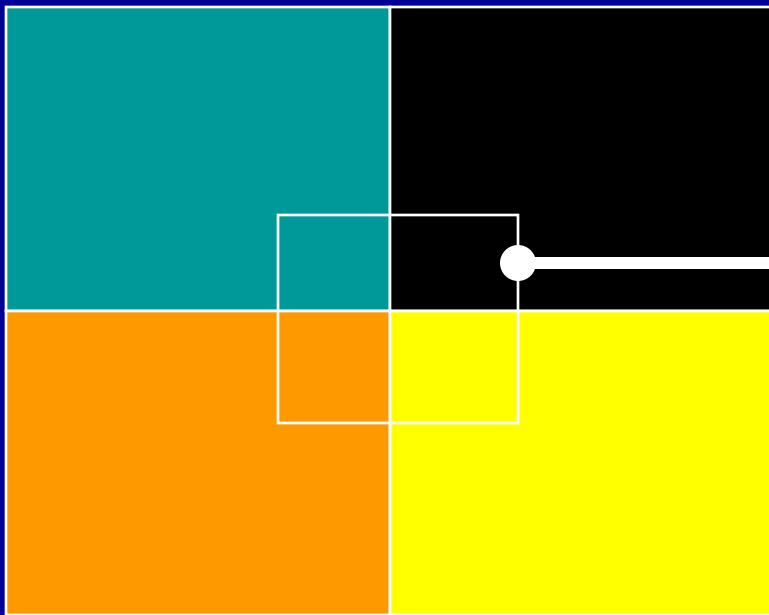
- Conceitos gerais
- Esquemas amostrais
- Tamanho da amostra

# Conceitos iniciais

- **População:** é a agregação de elementos da qual é extraída a amostra do meu estudo
- **Amostra representativa:** os objetos de estudos foram selecionados de forma a conseguir boa cobertura, os métodos de avaliação desses objetos são conhecidos e passíveis de reprodução/repetição (confiabilidade)

# Conceitos iniciais

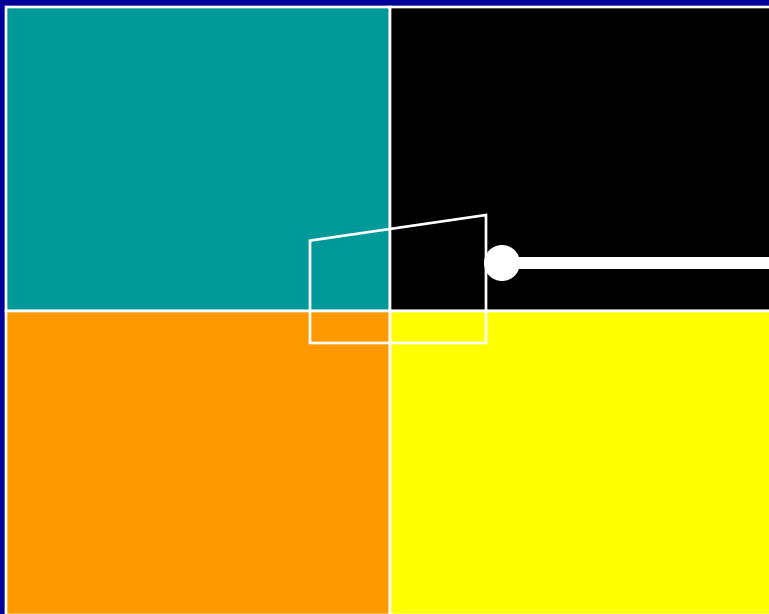
- Amostras probabilísticas: são amostras obtidas por aleatorização.



A idéia de aleatorização é conseguir uma amostra com características bem próximas da população

# Conceitos iniciais

- Amostras probabilísticas: são amostras obtidas por aleatorização.



Essa amostra não pode ser considerada representativa da população

•  
•  
•  
Quando utilizar amostras?

Quando trabalhar com população?

- Depende do meu objeto de estudo:
  - Avaliar a prevalência de cárie entre os 45 escolares com necessidade especial de uma determinada escola do meu município?

**Amostra**

**OU**

**População**

•  
•  
•  
Quando utilizar amostras?

Quando trabalhar com população?

- Depende do meu objeto de estudo:
  - Avaliar a prevalência de cárie entre os brasileiros em 2003?

**Amostra**

**OU**

**População**

# Utilidade da amostragem

- População muita extensa
- Não há tempo e/ou recursos suficientes para o estudo de toda a população
- O estudo objetiva avaliar vários detalhes de cada unidade do estudo
- A característica a ser estudada é muito variável, sujeita a alterações bruscas em curto período de tempo

- 
- 
- 

## Para determinar a população-alvo:

- Determinar critérios de exclusão e inclusão específicos
  - Inclusão: aspectos que serão importantes para o estudo: características demográficas, clínicas, geográficas, temporais.
  - Exclusão: questões que podem dificultar a coleta de dados por falta de adesão, incapacidade de fornecer dados, alto risco de efeitos adversos.

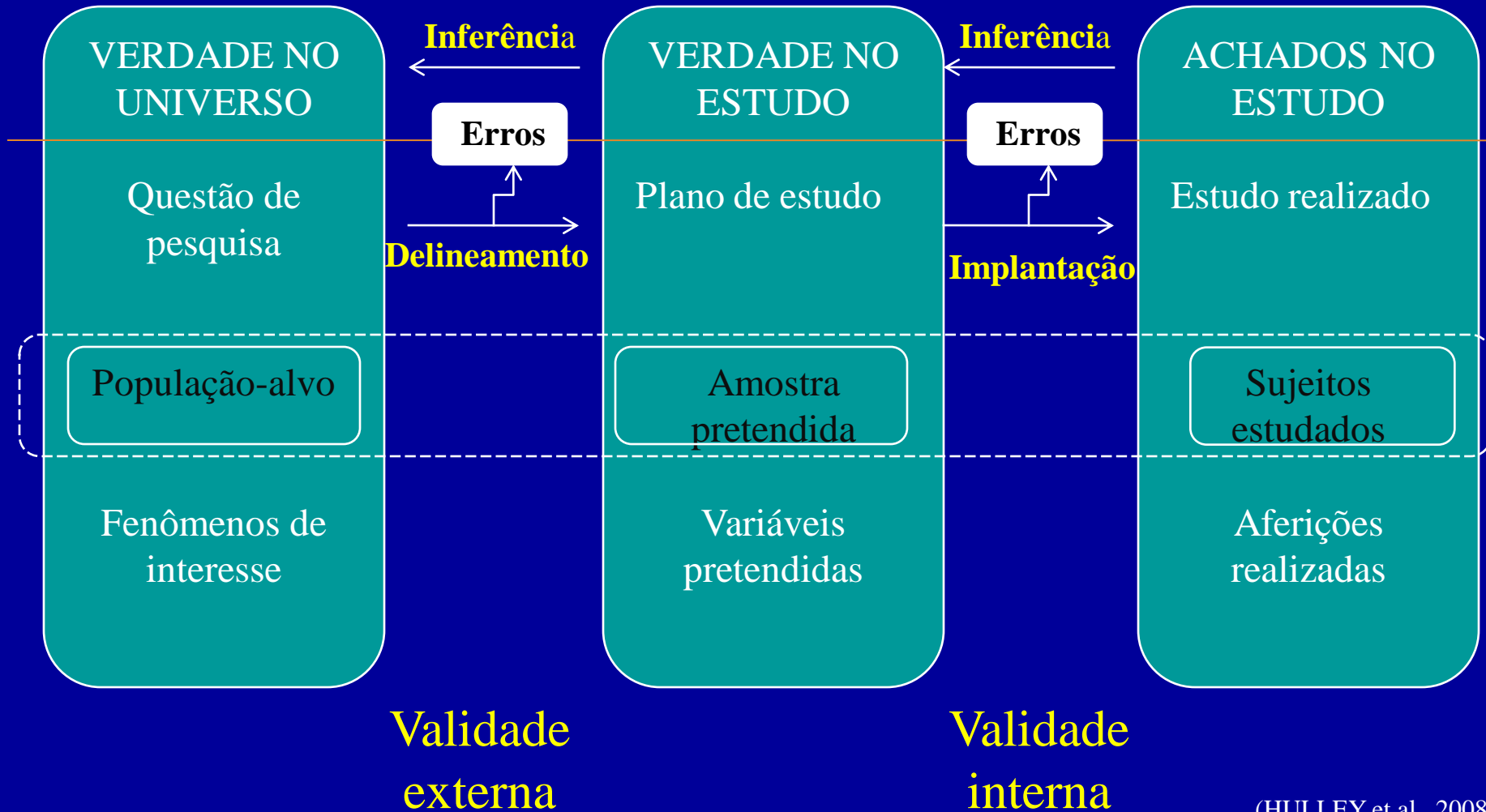
# Para determinar a população-alvo:

- Estudo sobre a adesão de pais/responsáveis a medidas de controle mecânico-químico de placa em pacientes com deficiências físicas (Abreu *et al.*, 2002)
- Critérios de inclusão: diagnóstico médico da referida deficiência, idade entre 7 e 21 anos, ambos os sexos, pais autorizaram
- Critérios de exclusão: outros diagnósticos médicos (Ex. Síndrome de Down)

# Unidade de observação

- Depende do objeto do estudo:
  - Cárie dentária: dente, grupo de dentes, indivíduo, comunidade
  - Biofilme/Placa dentária: superfície, dente, grupo de dentes, indivíduo, comunidade
  - Oclusopatias: indivíduo, comunidade
  - Opinião/percepção: indivíduo, comunidade

# Escolhendo sujeitos para o estudo que representem a população de interesse



# Esquemas amostrais

- Com a determinação da unidade amostral e da população-alvo, passa-se à definição do esquema probabilístico mais adequado para a seleção da amostra.
- Mais uma vez, o esquema a ser escolhido depende do meu objetivo.

# Esquemas amostrais

- Amostragem aleatória simples
  - Abreu *et al.* (2002):
    - População-alvo: escolares 7-21 anos com deficiência física, matriculados na Escola João Moreira Salles.
    - Há uma listagem atualizada e confiável desses alunos.
    - Numero todos os indivíduos e faço sorteio ou uso “Tabela de Números Aleatórios” (dando a mesma chance para todos) para selecionar a amostra.

# Esquemas amostrais

- Amostragem aleatória simples
  - Tabela de números aleatórios



8	0	7	7	8	7	0	2	7	6
5	8	6	8	4	0	0	7	0	8
9	7	3	3	1	1	4	6	1	8
2	0	8	5	2	5	4	0	5	6
4	8	0	6	6	3	8	6	9	6
7	5	4	9	9	4	3	6	1	5

# Esquemas amostrais

- Amostragem aleatória simples
  - Tabela de números aleatórios
    - Numerei os escolares de 1 a 176
    - Vou escolher uma forma de ler a tabela de números aleatórios: por exemplo, sempre da esquerda para a direita e de cima para baixo.
    - Assim teremos os seguintes selecionados:

# Esquemas amostrais

- Tabela de números aleatórios

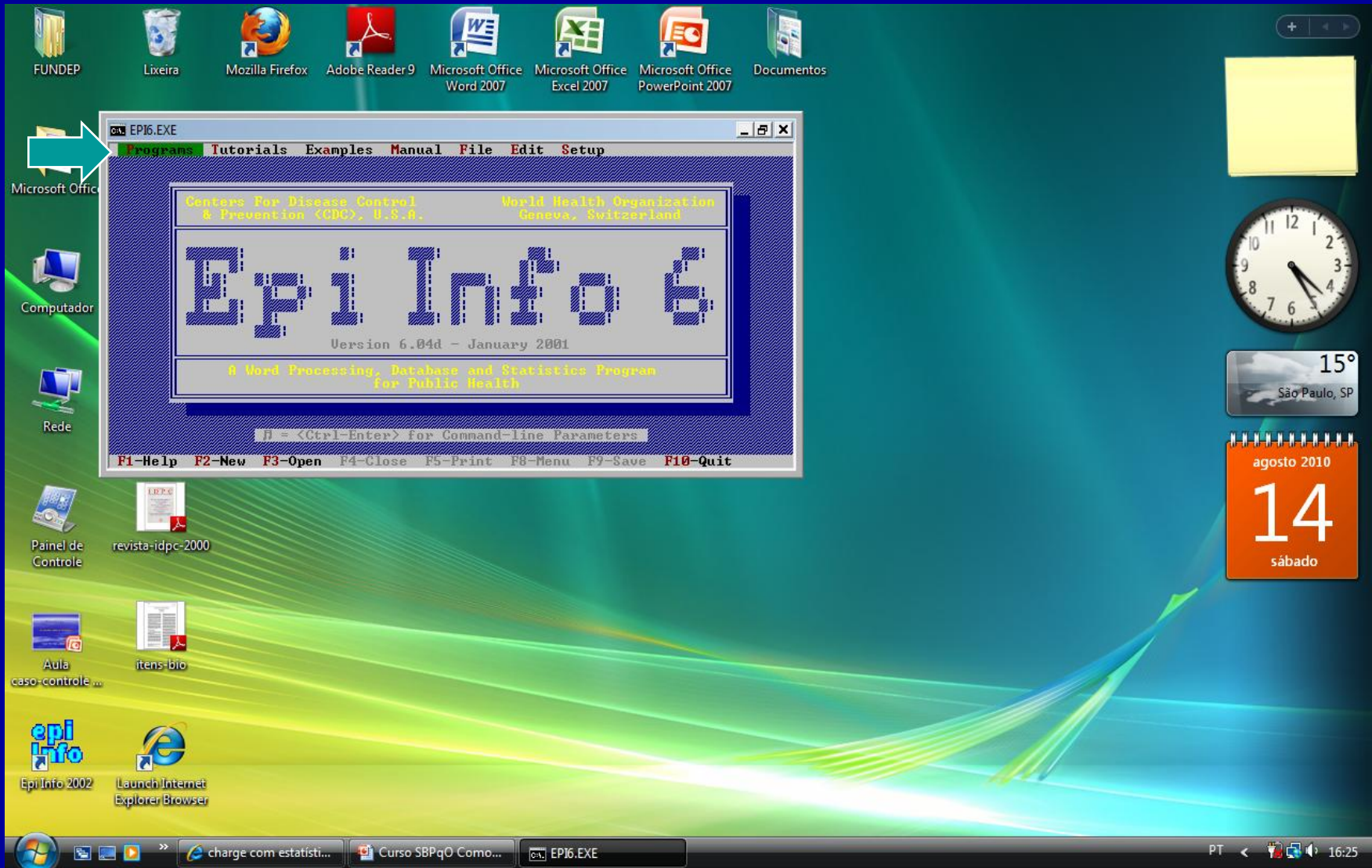


8	0	7	7	8	7	0	2	7
5	8	6	8	4	0	0	7	0
9	7	3	3	1	1	4	6	1
1	0	8	5	2	5	4	0	5
4	8	0	6	6	3	8	6	9
7	5	4	9	9	4	3	6	1



- 
- 
- 

# Como sortear no Epi Info?



FUNDEP

Lixeira

Mozilla Firefox

Adobe Reader 9

Microsoft Office Word 2007

Microsoft Office Excel 2007

Microsoft Office PowerPoint 2007

Documentos

EPI6.EXE

Programs Tutorials Examples Manual File Edit Setup

Centers For Disease Control & Prevention (CDC), U.S.A. World Health Organization Geneva, Switzerland

Epi Info

Version 6.04d - January 2001

A Word Processing, Database and Statistics Program For Public Health

H = <Ctrl-Enter> for Command-line Parameters

F1-Help F2-New F3-Open F4-Close F5-Print F8-Menu F9-Save F10-Quit

Microsoft Office

Computador

Rede

Painel de Controle

revista-idpc-2000

Aula

itens-bio

caso-control...

Epi Info

EpiInfo 2002

Launch Internet Explorer Browser

charge com estatisti...

Curso SBPqO Como...

EPI6.EXE

PT < 16:25

15°  
São Paulo, SP

agosto 2010  
14  
sábado

Windows XP desktop environment with a green and blue abstract background. The taskbar at the bottom shows the Start button, taskbar icons for Mail.com Message, Curso SBPqO Como..., and EPI6.EXE. The system tray on the right shows the time as 16:26 and the language as PT.

The desktop contains several icons: FUNDEP, Lixeira, Mozilla Firefox, Adobe Reader 9, Microsoft Office Word 2007, Microsoft Office Excel 2007, Microsoft Office PowerPoint 2007, Documentos, Microsoft Office, Computador, Rede, Painel de Controle, revista-idpc-2000, Aula caso-controle..., itens-bio, Epi-Info 2002, and Launch Internet Explorer Browser.

A yellow sticky note is attached to the top right of the desktop. A clock shows the time is approximately 3:25. A weather widget displays 15° in São Paulo, SP. A calendar shows the date is August 14, 2010, a Saturday.

The main window is titled "EPI6.EXE" and contains a menu-driven interface for an epidemiological calculator. The menu items are:

- EPED word processor
- ENTERX (uses high men)
- ENTER data
- ANALYSIS of data
- CHECK customize entry
- IMPORT files
- EXPORT files
- MERGE files
- STATCALC calculator
- CSAMPLE analyze surveys
- EPI6TABLE calculator** (highlighted)
- EPINUT anthropometry
- VALIDATE duplicate entry
- QUIT

The main window content displays the following text:

```

World Health Organization
Geneva, Switzerland

Info 6

on 6.04d - January 2001

Database and Statistics Program
For Public Health
  
```

At the bottom of the window, it says: "F1 = <Ctrl-Enter> for Command-line Parameters" and "F1-Help F10-Quit | « Calculator for various epidemiologic statistics".



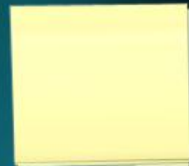
Microsoft Office EPI6.EXE

Describe Compare Study **Sample** Probability Setup

- Sample size
- Lower calculation
- Random number table
- Random number list

Epitable

F1-Help F3-Move F4-Zoom F5-Print F6-Next F9-0 F10-Quit Men:186216



15°  
São Paulo, SP

agosto 2010  
**14**  
sábado





Windows XP desktop environment with a green and blue abstract background. A taskbar at the bottom shows the Start button and several open applications: Mail.com Message..., Curso SBPqO Como..., and EPI6.EXE. The system tray on the right shows the date and time as 16:30 on August 14, 2010.

The EPI6.EXE application window is open, displaying a menu with options: Describe, Compare, Study, Sample, Probability, and Setu. A sub-window titled "Random number List generator" is active, showing a form with the following fields and values:

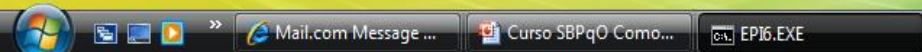
Field	Value
How many random numbers	384
Minimum range of numbers	1
Maximum range of numbers	1200

Below the form, there is a checkbox for "Drawing with replacement" which is currently unchecked. At the bottom of the sub-window are three buttons: "Calculate", "Reset", and "Quit". A large green arrow points to the "Calculate" button. The status bar at the bottom of the EPI6.EXE window shows keyboard shortcuts: F1-Help, F2-Move, F4-Zoom, F5-Print, F6-Next, F7-0, F10-Quit, and Mem:185032.



Weather widget showing 15° and São Paulo, SP.

Calendar widget showing agosto 2010, 14, sábado.



FUNDEP Lixeira Mozilla Firefox Adobe Reader 9 Microsoft Office Word 2007 Microsoft Office Excel 2007 Microsoft Office PowerPoint 2007 Documentos

Microsoft Office Saúde Coletiva UFMG EPI6 - Ata

Computador resultad\_final03

Rede Dissertacao

Painel de Controle revista-idpc-2000

Aula caso-control... itens-bio

Epi Info 2002 Launch Internet Explorer Browser

EPI6.EXE Describe Compare Study Sample Probability Setup

**Random number List generator**

How many random numbers 384  
 Minimum range of numbers 1

Files	Edit	Search				
1	8	11	13	18	24	
25	29	31	32	35	37	
38	41	47	49	52	53	
54	60	68	70	71	75	
78	80	81	85	87	90	
91	93	95	96	97	103	
104	106	109	114	117	119	
130	132	133	134	136	137	
138	140	143	145	147	155	
161	162	163	167	169	170	
*-----*-----*						
65:2						

F1-Help F2-Move F3-Zoom F5-Print F6-Next F9-0 F10-Quit Men:183888

15° São Paulo, SP

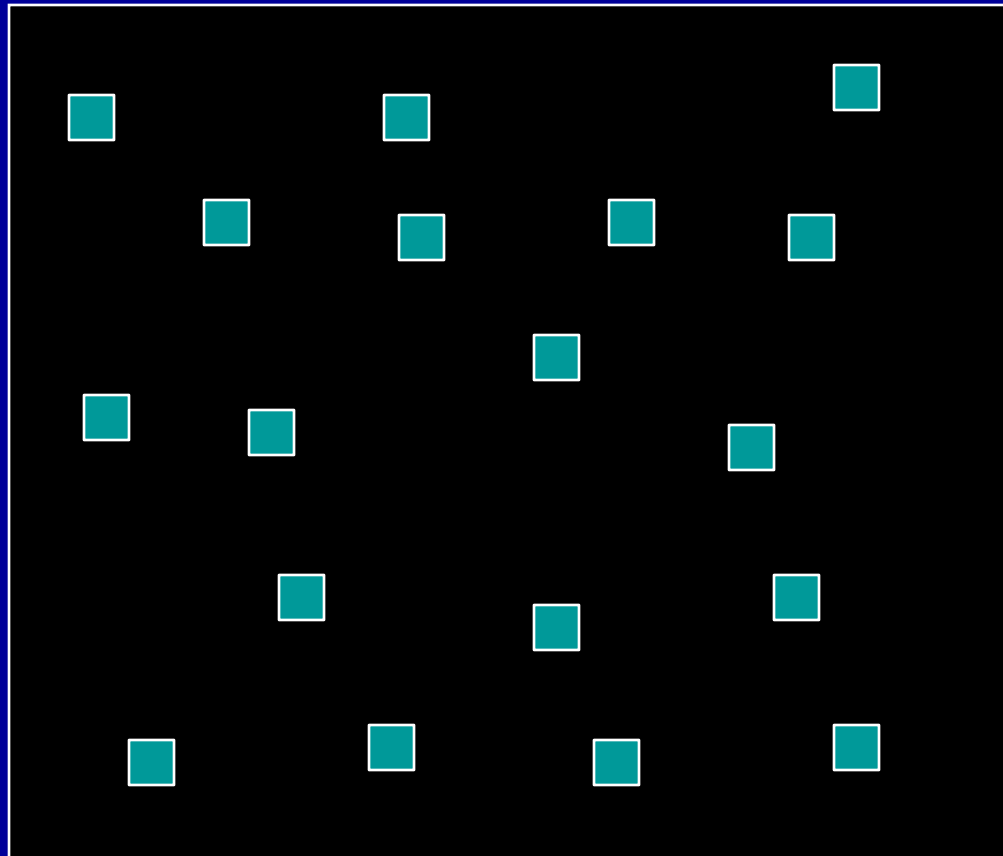
agosto 2010  
**14**  
 sábado

Mail.com Message ... Curso SBPqO Como... EPI6.EXE

PT 16:31

- 
- 
- 

# AMOSTRAGEM ALEATÓRIA SIMPLES

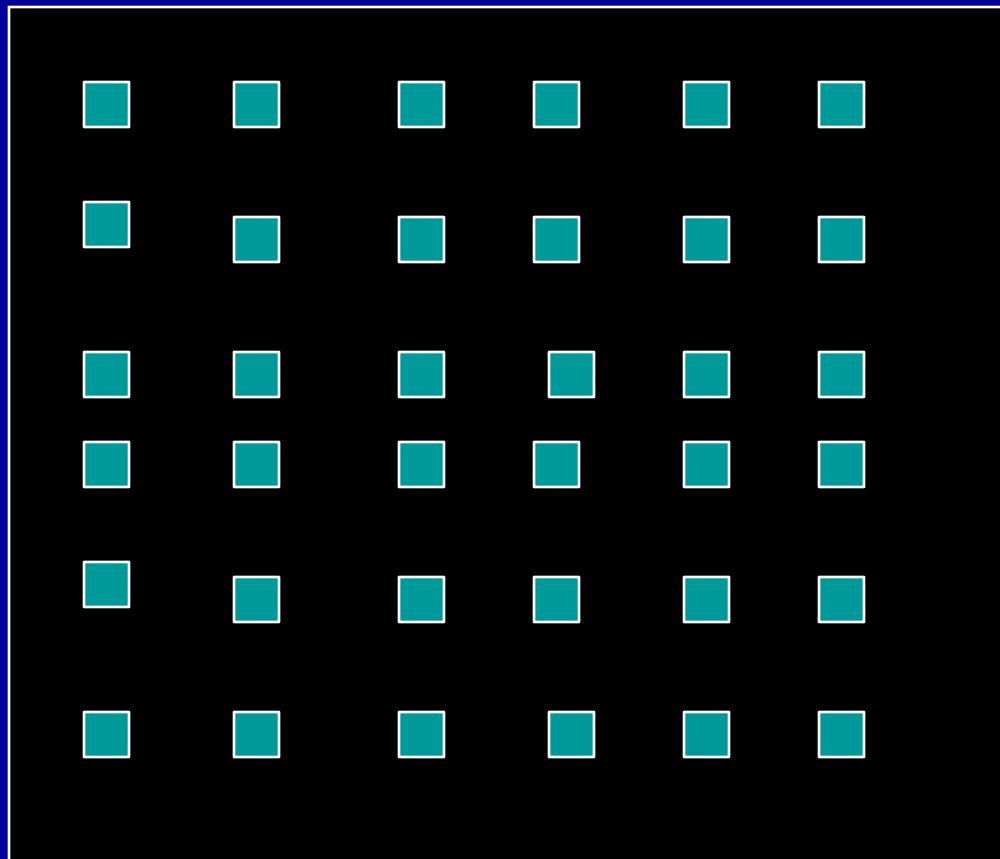


# Esquemas amostrais

- Amostragem sistemática
  - Você sorteia o primeiro indivíduo e segue um intervalo de amostragem
  - Exemplo: selecionamos aleatoriamente o indivíduo número 15 e, a partir de então, selecionamos o no. 25, 35, 45, 55, etc... (com um intervalo de amostragem igual a 10).

- 
- 
- 

# AMOSTRAGEM SISTEMÁTICA

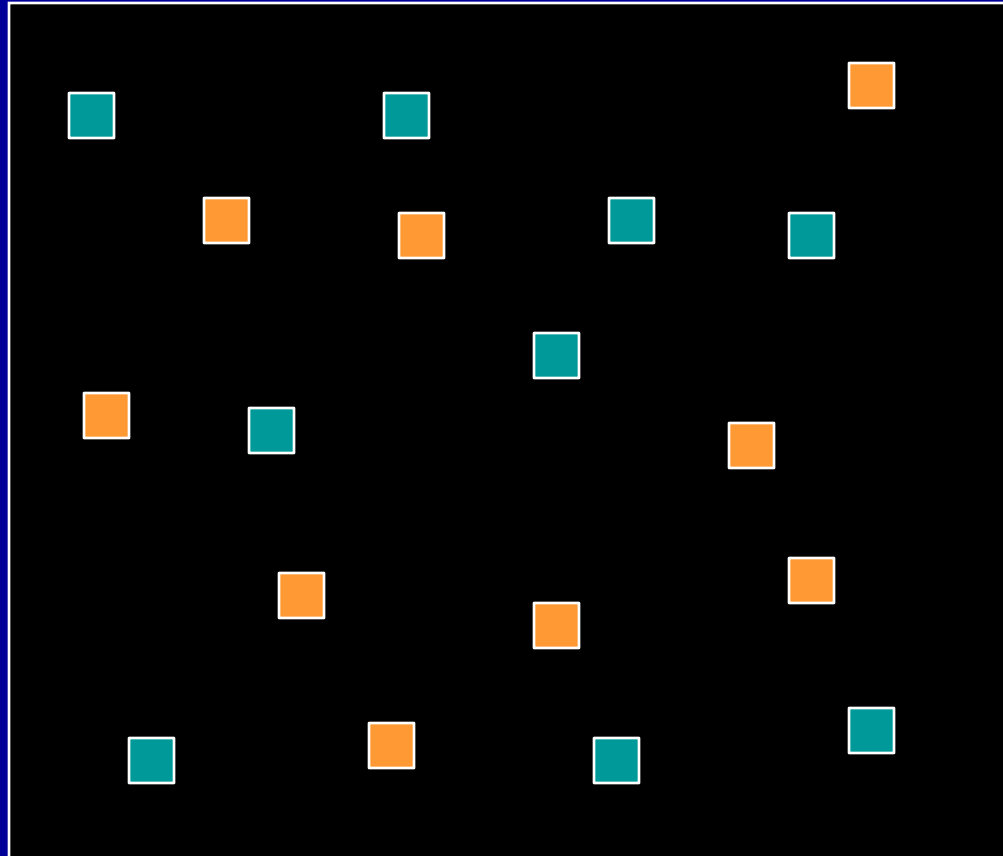


# Esquemas amostrais

- Amostragem aleatória estratificada
  - Vamos supor que o objetivo da minha pesquisa é avaliar a opinião de CD com até 10 anos de formados e com mais de 10 anos de formados, na cidade de Itajubá sobre o atendimento odontológico ao paciente com aids.
  - O objetivo da estratificação é diminuir a possibilidade de erro. Desta forma, realiza-se o processo de AAS ou sistemática em cada grupo separadamente.

- 
- 
- 

# AMOSTRAGEM ALEATÓRIA ESTRATIFICADA

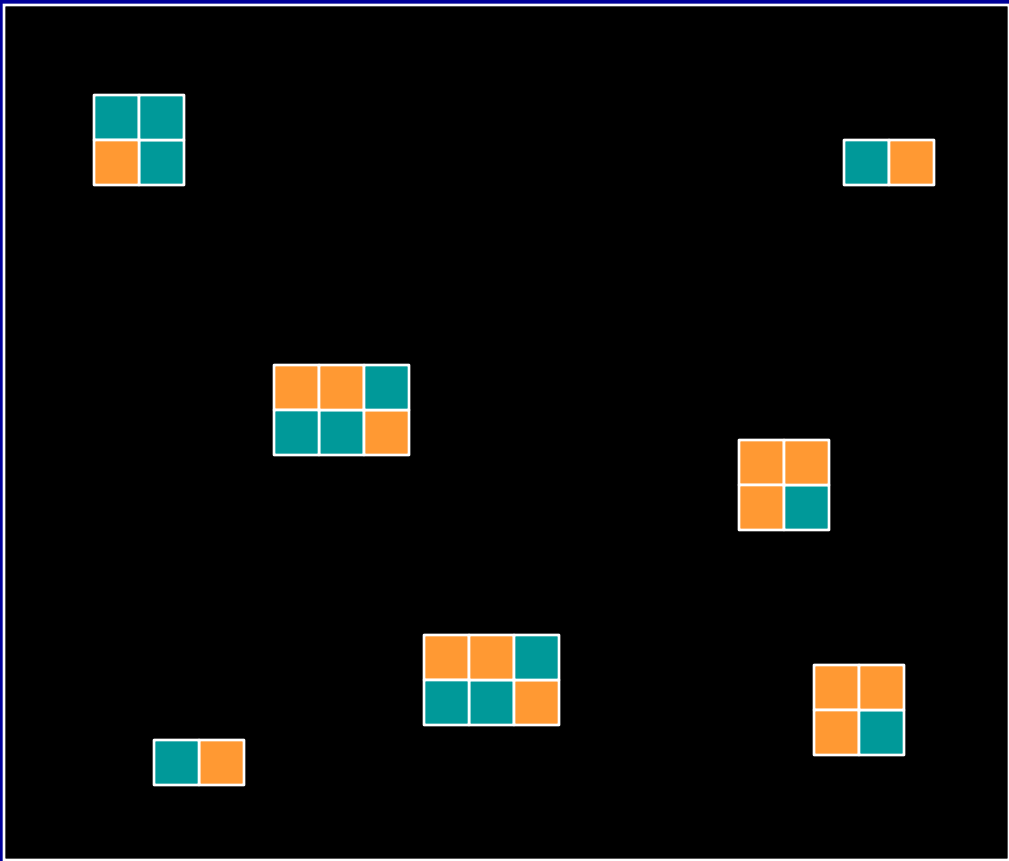


# Esquemas amostrais

- Amostragem por conglomerados
  - Vamos supor que o objetivo da minha pesquisa avaliar a experiência de cárie dentária entre adultos de uma cidade.
  - **Não há lista atualizada desses adultos.**
  - Deve-se sortear setores censitários ou quadras e trabalhar com amostra domiciliar.
  - Efeito do Desenho – Design of **EFF**ect (DEFF)

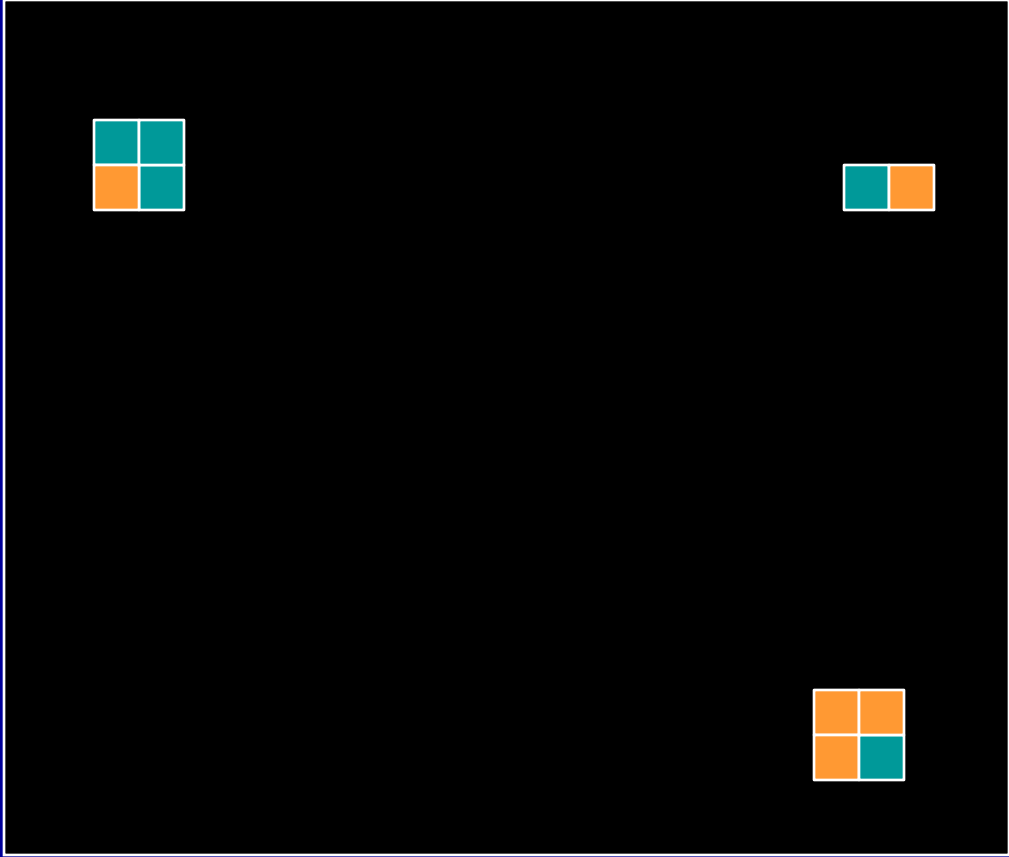
- 
- 
- 

# AMOSTRAGEM POR CONGLOMERADOS



- 
- 
- 

# AMOSTRAGEM POR CONGLOMERADOS



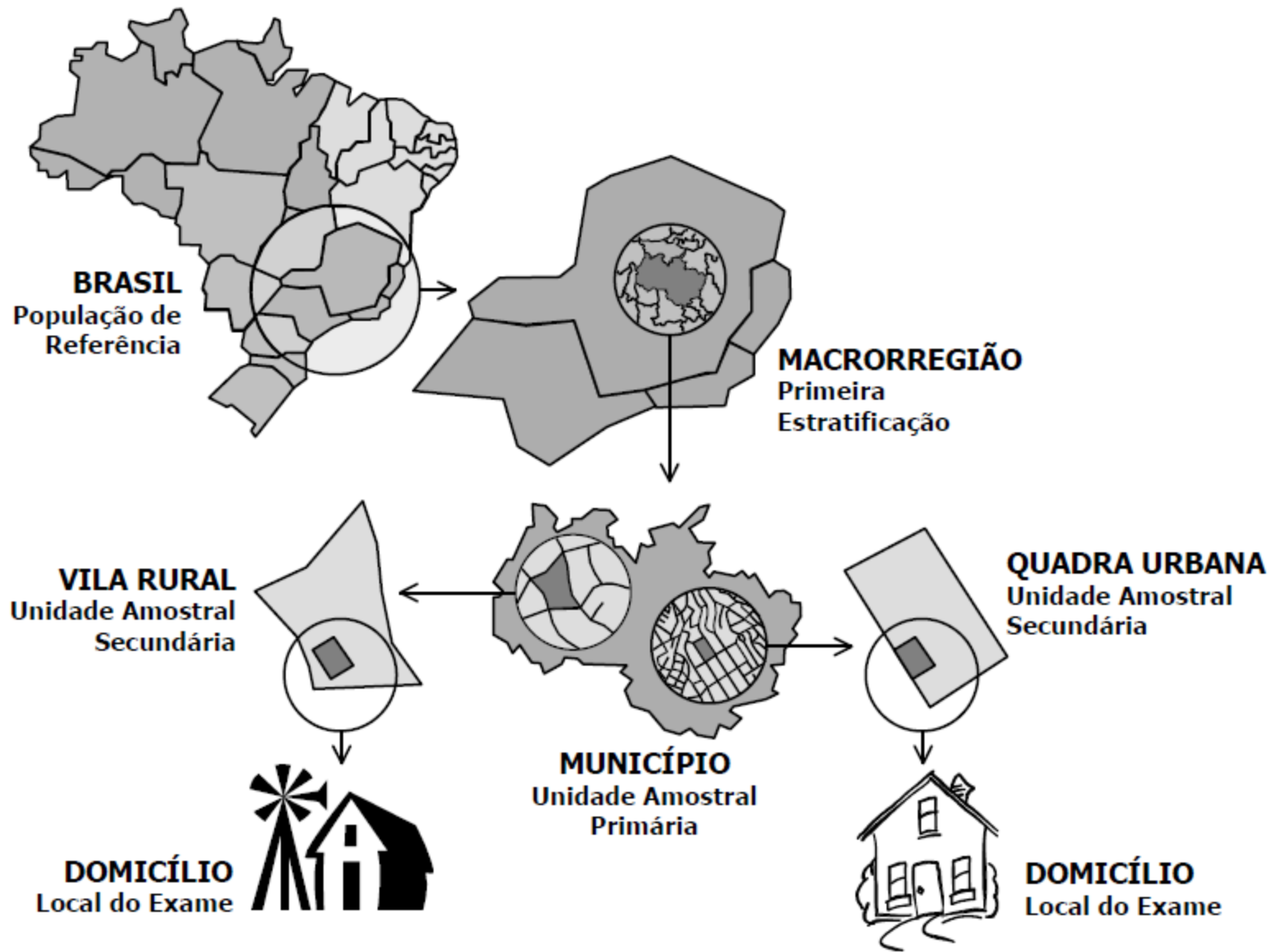


Figura 3. Esquema do plano amostral para o levantamento em domicílios em municípios até 50 mil habitantes (1º, 2º e 3º estratos).

Região Nordeste

 Município: Seringueiras

 Setor: 110150005000002

DOMDPU

157

Num	Quadra	Localização do Domicílio Rua, Nº, Complemento ou descrição do percurso	Grupos Demográficos					Domicílio Ocupado			Domicílio Vago	Observações
			5	12	65-74	15-19	35-44	participou	Recusou	Fechado		
Intervalo			3,1	3,1	3,1	10,7	8,3					
1	1	Rua Chico Mendes, 14		0				X				
2	1	Rua Chico Mendes, 16	1		1			X				
3												
4	1	Rua Chico Mendes, 20		1				X				
5	1	Rua Chico Mendes, 22	0		1			X				
6												
7	1	Rua Chico Mendes, 26		0		1						Retorno à tarde
8	1	Av. Jorge Teixeira, 33	0		0							
9	1	Av. Jorge Teixeira, 35					2	X				
10	1	Av. Jorge Teixeira, 37		0								
11	1	Av. Jorge Teixeira, 39	1		1			X				
12												
13												
14	1	Av. Jorge Teixeira, 39	0	1							X	
15	1	Av. Jorge Teixeira, 39							X			Retorno Fim de Semana

# Tamanho da amostra

- É uma questão que deve ser sempre avaliada para que eu diminua o meu erro aleatório.
- No entanto eu não posso incluir pessoas em excesso por questões de tempo/recursos e questões éticas.
- Não existe, geralmente, um número “MÁGICO” de participantes para que eu realize meu estudo

# Tamanho de amostra

- Preciso conhecer e definir alguns parâmetros anteriormente:
  - Precisão (“margem de erro”): Exemplo ao definir a prevalência de cárie em uma comunidade, eu posso considerar aceitável um erro de 3%. Ou seja, se a prevalência estiver próxima a 77%, o estudo aceitaria encontrar valores entre 74% e 80%.
  - O pesquisador define a precisão necessária

# Tamanho de amostra

- Qual é o objetivo do meu estudo????
  - *Descrever uma proporção?*
  - *Descrever uma variável quantitativa?*
  - *Comparar duas proporções?*
  - *Comparar duas médias?*
  - *Estudar determinantes de doenças em populações?*
    - *Caso-controle?*
    - *Coorte?*

# Tamanho de amostra

- Estimativas para proporções
  - Devo ter conhecimento sobre
    - Prevalência da condição a ser estudada ( $p$ )
    - Proporção de indivíduos sem a condição ( $1-p$ )
    - Precisão requerida para a estimativa ( $d$ )
    - Nível de confiança ( $1-\alpha$ ) - seu valor  $z$

# Valores de $\alpha$ ou $\beta$

Valor $\alpha$ ou $\beta$	Valores tabelados de $z$
0,001	3,291
0,01	2,576
0,05	1,960
0,10	1,645
0,20	1,290

# Tamanho de amostra

- Estimativas para proporções
  - A fórmula é:

$$n = \frac{(\text{nível de confiança})^2 (p) (1-p)}{(d) (d)}$$

# Tamanho de amostra

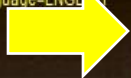
- Estimativas para proporções: Exemplos

Exemplos: A coordenação de saúde bucal de um município deseja conhecer a prevalência de crianças livres de cárie aos 12 anos de idade. Quantas crianças seriam necessárias se a precisão admissível é de 5%, com um grau de confiança de 95% e se o valor real parece não exceder 20%?

# Tamanho de amostra

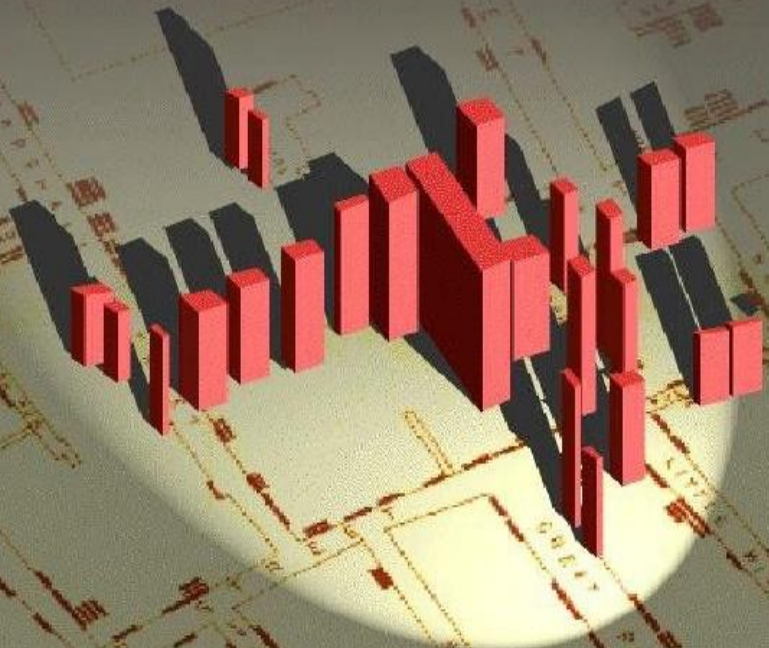
- Estimativas para proporções: Resolução
  - Proporção da população ( $p$ ) = 20% (0,20)
  - Nível de confiança ( $1 - \alpha$ ) = 95% (1,96 - *em tabela*)
  - Precisão ( $d = 15\% - 25\%$ )
  - $n = \frac{(1,96)^2 (0,2)(1-0,2)}{(0,05)^2}$
  - $n = 246$  crianças

Language=ENGLISH



- StatCalc
- Data Compare
- Table-to-View
- Visualize Data
- Epi Lock
- Compact
- Word Processor

# Epi Info <sup>TM</sup>



**M**ake View

**E**nter Data

**A**nalyze Data

**E**dit

**C**reate **M**aps

**C**reate **R**eports

**E**pi Info **W**ebsite

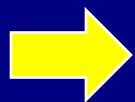
Version 3.5.1

August 13, 2008

E:\Epi\_Info\STATCALC.EXE  
EpiInfo Version 6      Statcalc      November 1993

1  
Tables (2 x 2, 2 x n)  
Sample size & power  
Chi square for trend  
E

F1-Help      F6-Open File      F10-Done



Make View

Enter Data

Analyze Data

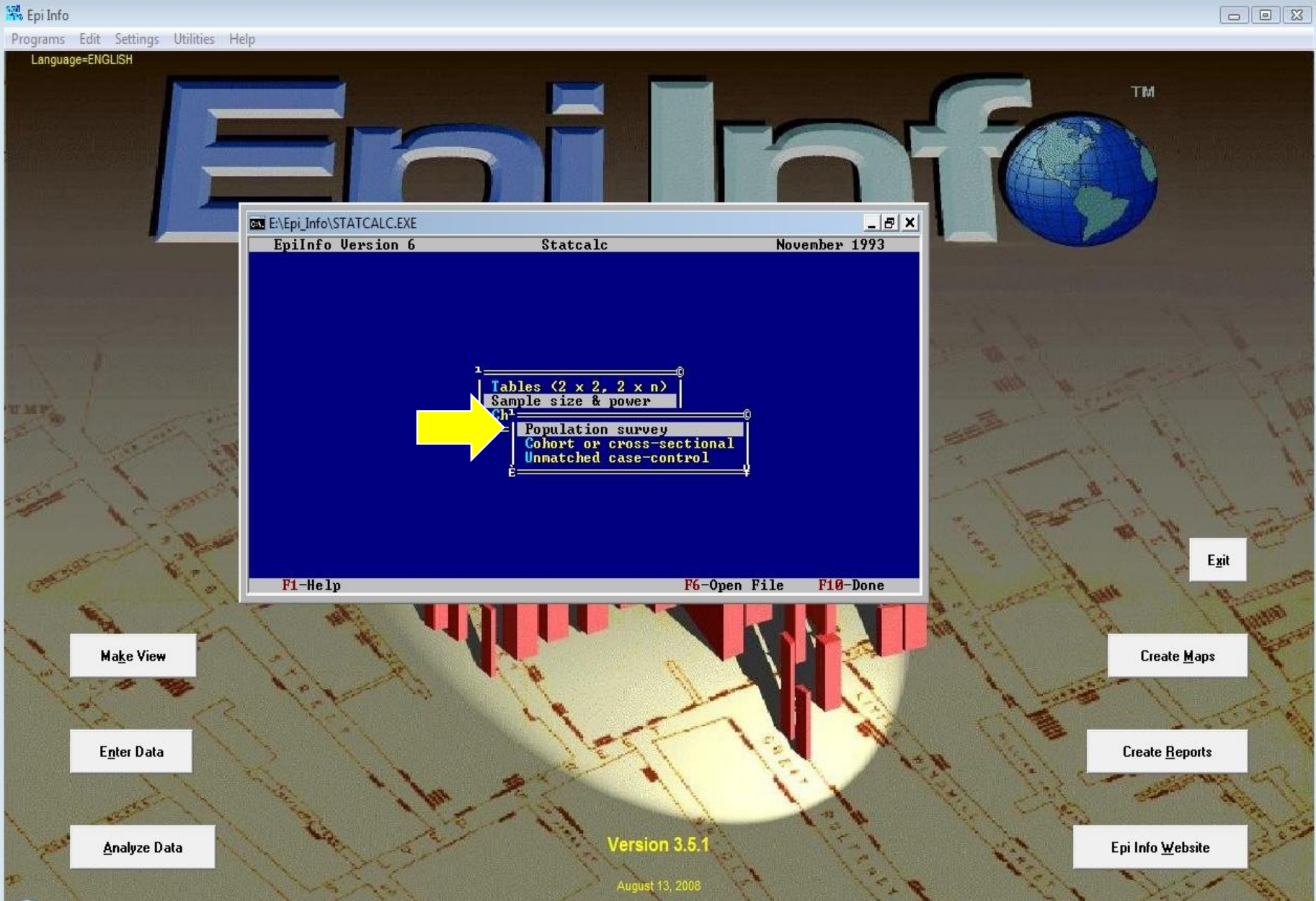
Exit

Create Maps

Create Reports

Epi Info Website

Version 3.5.1  
August 13, 2008



Epi Info  
Programs Edit Settings Utilities Help  
Language=ENGLISH

# Epi Info <sup>TM</sup>



E:\Epi\_Info\STATCALC.EXE  
EpiInfo Version 6 Statcalc November 1993

1 \_\_\_\_\_ 0  
 | Tables (2 x 2, 2 x n) |  
 | Sample size & power |  
 | \_\_\_\_\_ |  
 | Population survey |  
 | Cohort or cross-sectional |  
 | Unmatched case-control |  
 | \_\_\_\_\_ |  
 E

F1-Help F6-Open File F10-Done



Make View

Enter Data

Analyze Data

Exit

Create Maps

Create Reports

Epi Info Website

Version 3.5.1

August 13, 2008



```

E:\Epi_Info\STATCALC.EXE
EpiInfo Version 6          Statcalc          November 1993

Population Survey or Descriptive Study Using Random (Not Cluster) Sampling

Size of population from which
the sample will be selected :          999,999

Expected frequency of the factor
under study (err toward 50%) :        20.00 %

If 20.00 % is the true rate in the population, what is the result
farthest from the rate that you would accept in your sample (higher
or lower)?

Worst acceptable result :             25.00 %

F1-Help          F4-Calc          F6-Open File    F10-Done
  
```

População infinita

Prevalência esperada

Precisão



Make View

Enter Data

Analyze Data

Exit

Create Maps

Create Reports

Epi Info Website

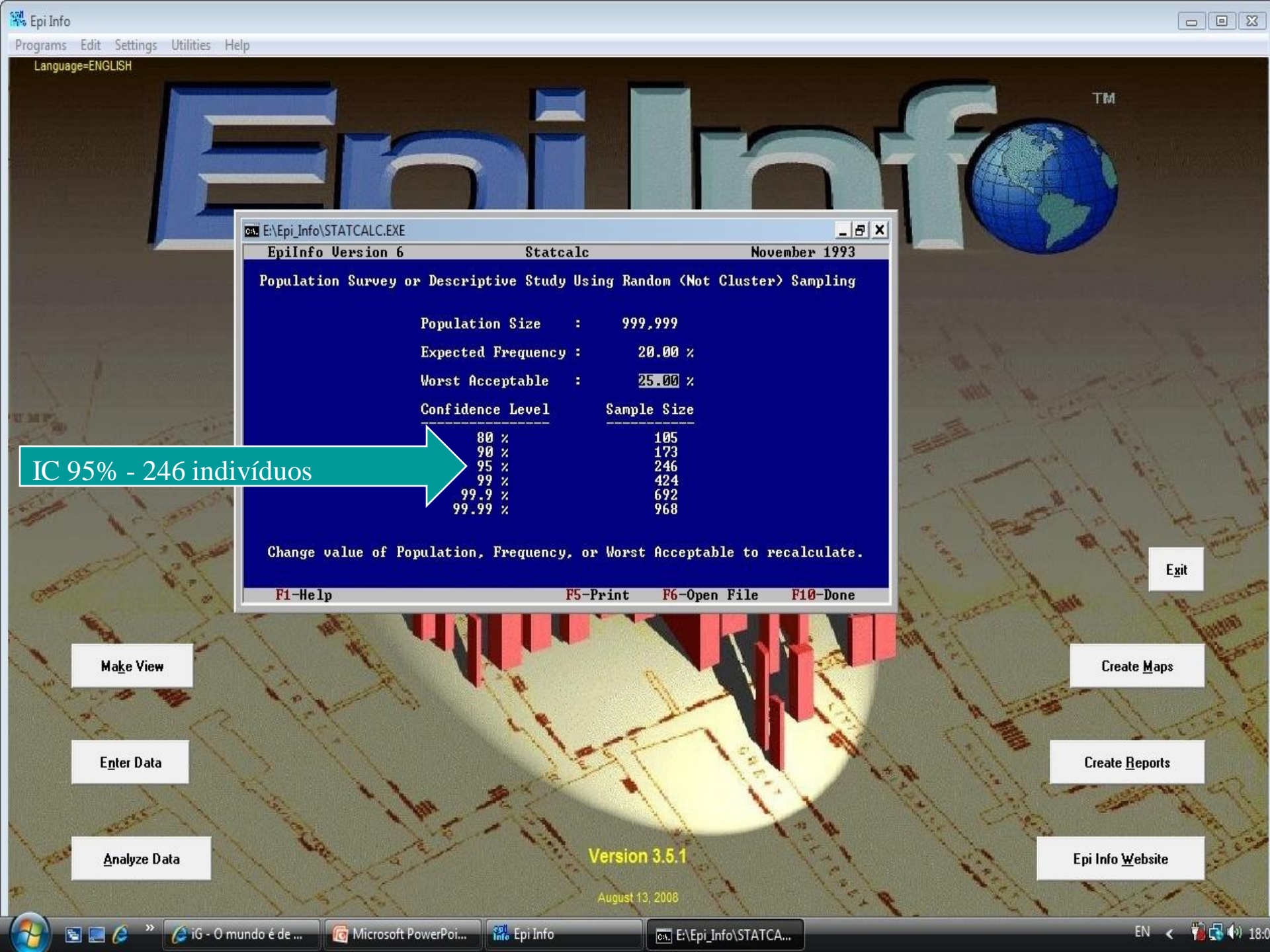
Aperte F4



Version 3.5.1

August 13, 2008





C:\Epi\_Info\STATCALC.EXE

EpiInfo Version 6 Statcalc November 1993

Population Survey or Descriptive Study Using Random (Not Cluster) Sampling

Population Size : 999,999

Expected Frequency : 20.00 %

Worst Acceptable : 25.00 %

Confidence Level	Sample Size
80 %	105
90 %	173
95 %	246
99 %	424
99.9 %	692
99.99 %	968

Change value of Population, Frequency, or Worst Acceptable to recalculate.

F1-Help F5-Print F6-Open File F10-Done

IC 95% - 246 indivíduos

Make View

Enter Data

Analyze Data

Create Maps

Create Reports

Epi Info Website

Exit

Version 3.5.1

August 13, 2008

# Tamanho de amostra

- Estimativas para proporções: Exercício
  - Você precisa fazer um cálculo amostral para conhecer a prevalência de livres de cárie dentária aos 7 anos em uma comunidade. Quantas crianças seriam examinadas se, a precisão admissível é de 5%, com um grau de confiança de 95% e se o valor real parece não exceder 40%?

# Tamanho de amostra

- Estimativas para proporções: Exemplos
  - Você precisa fazer um cálculo amostral para conhecer a prevalência de livres de cárie dentária aos 12 anos em uma comunidade. Quantas crianças seriam examinadas se, a precisão admissível é de 15 por cento do valor real ( $0,15 \times p$ ) com um grau de confiança de 95% e se o valor real parece não exceder 20%?

# Tamanho de amostra

- Estimativas para proporções: Resolução
  - Proporção da população (p) = 20% (0,20)
  - Nível de confiança = 95%
  - Precisão relativa (de 17%-23%) = 15% (de 20%)
  - $n = \frac{(1,96)^2 (0,2) (1-0,2)}{(0,15)^2 \times (0,20)}$
  - 683 crianças

# Tamanho de amostra

- Estimativas para proporções: Resolução
  - $n = 683$  crianças
  - E se meu universo for de 500 crianças?

$$- n_{final} = \frac{n}{1 + n/N}$$

Onde:

$n_{final}$  = tamanho da amostra final ajustado

$n$  = tamanho da amostra obtida no cálculo anterior

$N$  = população finita

# Tamanho de amostra

- Estimativas para proporções: Resolução

- $n_{final} = \frac{n}{1+n/N}$

- $n_{final} = \frac{683}{1+683/500}$

- $n_{final} = \frac{683}{1+1,366}$

- $n_{final} = 289$  crianças

# Tamanho de amostra

- Estimativas para médias
  - Para o cálculo eu preciso determinar
    - Desvio padrão da média ( $s$ )
    - Precisão requerida para a estimativa ( $e$ )
    - Nível de confiança ( $1 - \alpha$ ) - seu valor  $z$

# Tamanho de amostra

- Estimativas para médias

– A fórmula é a seguinte:

$$\bullet n = \frac{(\text{nível de confiança})^2 (s)^2}{e^2}$$

# Tamanho de amostra

- Estimativas para médias:
  - Exemplo: Deseja-se conhecer o número médio de dentes permanentes com experiência de cárie em uma comunidade, aos 12 anos. Em um estudo prévio, os valores encontrados foram: CPOD médio valendo 2,7 e desvio padrão 0,5. Quantas crianças deve-se examinar se eu admito uma precisão igual a 0,2 e um nível de confiança de 95%.

# Tamanho de amostra

- Estimativas para médias:
  - Desvio padrão ( $s$ ) = 0,5
  - Nível de confiança = 95%
  - Erro aceitável = 0,2

Na fórmula tem-se: 
$$\frac{(1,96) (1,96) (0,5) (0,5)}{(0,2) (0,2)} = 25$$

Deveríamos examinar **25 crianças**

# Tamanho de amostra

- Estimativas para médias: Exercício
  - Um estudo quer determinar o número médio de papilas sangrantes em pacientes com Paralisia Cerebral. O estudo piloto mostrou que em média há 12 papilas com desvio padrão igual a 6. Considerando um nível de confiança de 95%, um erro de 10% da média, quantos indivíduos deveriam participar do estudo?

# Tamanho de amostra

- Estimativas para médias: Exercício
  - Média de papilas = 12
  - Desvio padrão (s) = 6
  - Nível de confiança = 95%
  - Erro aceitável (10,8-13,2) = 10% (de 12)
  - $n = \frac{(1,96) (1,96) (6) (6)}{(1,2)^2} = 96$  indivíduos

# Tamanho de amostra

- Em um estudo laboratorial, deseja-se determinar a força de adesão em MPa de um novo sistema adesivo em dentina. Supondo em estudo piloto, o teste de tração revelou uma força foi igual a  $20,4 (\pm 3,7)$  MPa. Considerando o nível de confiança igual a 95% e erro aceitável de 10% da média, quantos corpos de prova deveriam ser testados?

# Tamanho de amostra

- Estimativas para médias: Exercício
  - Força de adesão = 20,4
  - Desvio padrão (s) = 3,7
  - Nível de confiança = 95%
  - Erro aceitável (18,36-22,44) = 10% (de 20,4)
  - $n = \frac{(1,96) (1,96) (3,7) (3,7)}{(2,04)^2} = 13$  corpos de prova

- 
- 
- 

E quando a média não for  
um bom parâmetro??

# Tamanho de amostra

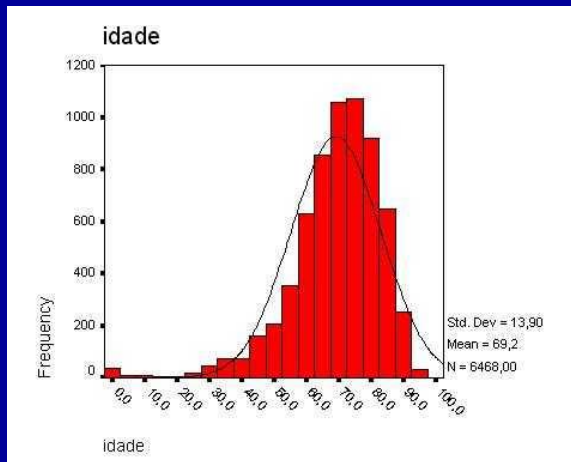
- Estimativas para variáveis quantitativas:
  - A fórmula é a seguinte:

$$n = \frac{(\text{desvio padrão})^2}{(\text{erro aceitável}) (\text{diferença DP})^2}$$

Riffenburg (2006, p.402)

•  
•  
•

Não há muito desenvolvimento de técnicas de  
cálculo amostral  
quando a distribuição não é normal



Riffenburg (2006, p.415)

•  
•  
•  
•  
•  
•  
•  
•

# Tamanho de amostra

## **ESTUDOS COMPARATIVOS: PROPORÇÃO**

Um estudo deseja comparar a eficácia de um novo selante. A eficácia da nova terapia, por estudo piloto, vem sendo de 50% enquanto que para a terapia tradicional é de 40%. Qual é a amostra necessária para se verificar a eficácia da paralisação do processo cariioso, se há uma diferença de 10% entre os grupos, com nível de significância de 5% ( $\alpha=0,05$ ) e um poder do teste de 0,9 ( $\beta=0,10$ )?

## DECISÃO DO JÚRI

**Inocência:** o réu não falsificou dinheiro

**Culpa:** o réu falsificou dinheiro

**Padrão para rejeitar a inocência:**  
“Dúvida além do razoável”

**Julgamento correto:** Culpar o estelionatário

**Julgamento correto:** Absolver o inocente

**Julgamento incorreto:** Condenar o inocente

**Julgamento incorreto:** Absolver o estelionatário

## TESTE ESTATÍSTICO

**H.Nula** – Não há associação entre tabaco e Ca de boca

**H.Alternativa** – Há associação

**Padrão para rejeitar uma H.Nula:**  
Nível de significância estatística ( $\alpha$ )

**Inferência correta:** concluir que há associação entre Ca e tabaco quando há essa associação na população

**Inferência correta:** concluir que não há essa associação quando não houver na população

**Inferência incorreta (erro tipo I -  $\alpha$ ):** concluir que há associação quando não houver

**Inferência incorreta (erro tipo II -  $\beta$ ):** concluir que não há associação quando houver

# UMA SÍNTESE DESSAS IDEIAS...

## Verdade na população

	Há associação entre variáveis (H1)	Não há associação (H0)
Rejeita a Hipótese Nula (H1)	CORRETA Poder do teste	<b>ERRO TIPO I</b>
Aceita a Hipótese Nula (H0)	<b>ERRO TIPO II</b>	CORRETA

Resultados na amostra do estudo

- 
- 
- 

## *Os valores...*

Valores de $\alpha$ e $\beta$	Valores tabelados (bicaudal)	Valores tabelados (unicaudal)
0.001	3.29	3.09
0.01	2.58	2.33
0.05	1.96	1.64
0.10	1.64	1.28
0.20	1.29	0.84

# FÓRMULA

$$n = \frac{\{z_{1-\alpha/2} \sqrt{2p(1-p)} + z_{\beta} [\sqrt{p_1(1-p_1) + p_2(1-p_2)}]\}^2}{d^2}$$

- I.  $p$  = média das proporções das condições comparadas
- II.  $P_1$  = proporção da condição de um grupo
- III.  $P_2$  = proporção da condição do outro grupo
- IV.  $d$  = diferença entre os grupos
- V.  $\alpha$  = nível de significância
- VI.  $1-\beta$  = poder do teste

# CÁLCULO

$$n = \frac{\{ 1,96 \sqrt{2(0,45)(1-0,45)} + 1,28[\sqrt{0,4(1-0,4)+0,5(1-0,5)}] \}^2}{(0,10)^2}$$

**n=518 dentes em cada grupo**

- I. p = média das proporções das condições comparadas (0,45)
- II. P<sub>1</sub> = proporção da condição de um grupo (0,40)
- III. P<sub>2</sub> = proporção da condição do outro grupo (0,50)
- IV. d = diferença entre os grupos (0,10)
- V. α = nível de significância (1,96)
- VI. 1-β = poder do teste (0,90 i.e. β=0,10 - 1,28)

# 1) <http://www.statpages.org/proppowr.html>

JavaStat -- Binomial Proportion Differences - Windows Internet Explorer

http://www.statpages.org/proppowr.html

Mail.com Message List E-mail de iG Mail - Entrada... Designing Clinical Researc... JavaStat -- Binomial Pr...

## Proportion Difference Power / Sample Size Calculation

(Revised 10/30/2009 - Also display results of uncorrected ("classical") calculation.)


This screen computes the sample size required to detect a difference between two proportions.

Note: Before using this page for the first time, make sure you read the [JavaStat user interface guidelines](#) for important information about interacting with JavaStat pages.

Significance Level (alpha):	<input type="text" value="0.05"/>	(Usually 0.05)
Power (% chance of detecting):	<input type="text" value="90"/>	(Usually 80)
Group 1 Population Proportion:	<input type="text" value=".4"/>	(Between 0.0 and 1.0)
Group 2 Population Proportion:	<input type="text" value=".5"/>	(Between 0.0 and 1.0)
Relative Sample Sizes Required (Group 2 / Group 1):	<input type="text" value="1.0"/>	(For equal samples, use 1.0)

Sample Size Required

	Group 1	Group 2	Total
"Classical" Calculation:	518	518	1036
With Continuity Correction:	538	538	1076



**Note:** This page incorporates a continuity correction to the usual sample-size formula based on the normal approximation to the binomial distribution. This correction increases the sample size (for each group) by an amount approximately equal to  $2/abs(p1-p2)$ , where  $p1$  and  $p2$  are the population proportions for the two groups. For a good discussion of this, see: *Statistical Methods for Rates and Proportions* by Joseph L. Fleiss (2<sup>nd</sup> ed., 1981, John Wiley & Sons, NY), chapter 3. This web page produces values consistent with those in Table A.3 of that book.

## 2) EPI INFO 6.04d – Epi Table

The screenshot shows a Windows XP desktop with the EPI INFO 6.04d software running. The software window is titled "EPI6.EXE" and displays the "Sample size, two proportions" calculation screen. The screen shows the following data:

Sample size, two proportions	
Ratio group 1/group 2	1.000
Percentage group 1	50.00
Percentage group 2	40.00
ó risk	
< > 10% < > 0.1%	< > 60% < > 90%
< * > 5% < > 0.01%	< > 70% < > 95%
< > 1%	< > 80% < > 99%
Power	: 90%
Confidence level	: 95%
Sample required in group 1	: 538
Sample required in group 2	: 538
Total #	: 1076

A red arrow points to the "Sample required in group 2" value of 538. The software interface includes buttons for "Calculate", "Reset", "Edit", and "Print". The desktop background is a blue and green abstract design. The taskbar at the bottom shows the Start button, several open applications (JavaStat, Probability of Disease, Microsoft PowerPoint, Comparação entre..., Microsoft Excel), and the system tray with the time 22:24 and date August 19, 2010.

### 3) Laboratório de Epidemiologia e Estatística – LEE – USP

[http://www.lee.dante.br/pesquisa/amostragem/di\\_2\\_pro.html](http://www.lee.dante.br/pesquisa/amostragem/di_2_pro.html)

The screenshot shows a web browser window with the following content:

- Navigation and Tabs:** The address bar shows [http://www.lee.dante.br/pesquisa/amostragem/di\\_2\\_pro.html](http://www.lee.dante.br/pesquisa/amostragem/di_2_pro.html). The browser has several tabs open, including "Lee - Amostra: Comparação de duas proporções".
- Links:** A list of links: [Proporção dos grupos.](#), [Nível de significância.](#), [Poder do teste.](#), and [Teste de hipótese.](#)
- Referências:** A reference: [Lwanga and Lemeshow \(1991\).](#)
- Instructions:** "A parte fracionária do número deve ser separada por ponto ex: 5.21"
- Form Fields:**
  - Proporção no grupo 1:  %
  - Proporção no grupo 2:  %
  - Nível de significância:
  - Poder do teste:
  - Teste de hipótese:  Monocaudal  Bicaudal
- Buttons:** "envia" and "Redefinir"
- Text:**

**Proporções dos grupos:** são as proporções estimadas sejam através de estudos anteriores, de dados da literatura ou do próprio conhecimento sobre o assunto.

**Nível de significância:** indica a probabilidade de cometer um erro do tipo I, ou seja, rejeitar a hipótese nula quando esta for verdadeira. Em outras palavras, é a chance de dizer que existe diferença entre as proporções quando na realidade elas são iguais. No exemplo considerado, é a probabilidade de concluir erroneamente que o segundo tipo de cirurgia apresenta mais complicações do que o primeiro.
- Taskbar:** Shows the Windows taskbar with the "Concluído" (Completed) notification area, the system tray with "Internet | Modo Protegido: Ativado" and "100%", and several open applications including "Lee - Amostra: Com...", "Probability of Disea...", "Microsoft PowerPo...", "Comparação entre ...", "Microsoft Excel uso ...", and "EPI6.EXE". The system clock shows "PT < 22:26".

### 3) Laboratório de Epidemiologia e Estatística – LEE - USP

Comparação de duas proporções - Windows Internet Explorer

http://www.lee.dante.br/cgi-bin/uncgi/calculo\_amostra

Mail.com Message List E-mail de iG Mail - Entrada... Designing Clinical Researc... Comparação de duas p... X

#### Valores calculados com os dados de entrada

Proporção no grupo 1: 50%  
Proporção no grupo 2: 40%  
Nível de significância: 5%  
Poder do teste: 90%  
Teste de hipótese: bicaudal  
Tamanho da amostra calculado para cada grupo: 518

Para outros valores do nível de significância e poder do teste temos:

Nív. de signif.	Poder do teste	Tam. amostra p/ grupo
5%	65%	271
5%	70%	306
5%	75%	342
5%	80%	387
5%	85%	444
5%	95%	640
0.1%	90%	1031
1%	90%	735
10%	90%	420

Concluído

Internet | Modo Protegido: Ativado 100%

Comparação de du... Probability of Disea... Microsoft PowerPoi... Comparação entre ... Microsoft Excel uso ... EPI6.EXE PT < 22:29

## 4) <http://www.cct.cuhk.edu.hk/stat/>

Sample Size Estimation - Windows Internet Explorer

http://www.cct.cuhk.edu.hk/stat/

### Sample Size Estimation



**Introduction:** In general, sample size calculation is conducted through a pre-study power analysis. The purpose is to select an appropriate sample size in achieving a desired power for correctly detection of a pre-specified clinical meaningful difference at a given level of significance. This study/tool will provide statistical procedures for determination of sample size required not only for testing *equality*, but also for testing *non-inferiority/superiority*, *equivalence* (similarity) and some *survival analysis* and *phase II clinical trials designs* that are commonly employed at various phases of clinical development. As a handful tool, it is very useful for clinical scientists and biostatisticians in the pharmaceutical industry, regulatory agencies, academia, and other scientists who are related fields of clinical development.

This program helps users to determine sample sizes for a broad range of statistical techniques, including the study of means, proportions, survival analysis, phase II clinical trials and some general cases (others).

**DIRECTIONS:**

- Choose one of techniques from "Means", "Proportions", "Survival Analysis" and "Phase II Clinical Trials" in following table
- From the table the technique appeared that you selected, please select one of specified case you are interested
- The selected sample size calculator will be appeared by a new window, then enjoy to use calculator to determine your required sample size!

<a href="#">Means</a>	<a href="#">Proportions</a>	<a href="#">Survival Analysis</a>	<a href="#">Phase II Clinical Trials</a>	<a href="#">Others</a>
-----------------------	-----------------------------	-----------------------------------	--	------------------------

Internet | Modo Protegido: Ativado 100%

Sample Size Estimat... Probability of Disea... Microsoft PowerPoi... Comparação entre ... Microsoft Excel uso ... EPI6.EXE PT < 22:36

## 4) <http://www.cct.cuhk.edu.hk/stat/>

Sample Size Estimation - Windows Internet Explorer

http://www.cct.cuhk.edu.hk/stat/Proportions.htm

### Sample Size Estimation


**Introduction:** In general, sample size calculation is conducted through a pre-study power analysis. The purpose is to select an appropriate sample size in achieving a desired power for correctly detection of a pre-specified clinical meaningful difference at a given level of significance. This study/tool will provide statistical procedures for determination of sample size required not only for testing *equality*, but also for testing *non-inferiority/superiority*, *equivalence* (similarity) and some *survival analysis* and *phase II clinical trials designs* that are commonly employed at various phases of clinical development. As a handful tool, it is very useful for clinical scientists and biostatisticians in the pharmaceutical industry, regulatory agencies, academia, and other scientists who are related fields of clinical development.

This program helps users to determine sample sizes for a broad range of statistical techniques, including the study of means, proportions, survival analysis, phase II clinical trials and some general cases (others).

**DIRECTIONS:**

- Choose one of techniques from "Means", "Proportions", "Survival Analysis" and "Phase II Clinical Trials" in following table
- From the table the technique appeared that you selected, please select one of specified case you are interested
- The selected sample size calculator will be appeared by a new window, then enjoy to use calculator to determine your required sample size!

<a href="#">Means</a>	<a href="#">Proportions</a>	<a href="#">Survival Analysis</a>	<a href="#">Phase II Clinical Trials</a>	<a href="#">Others</a>
<b>One-Sample Design</b> <a href="#">Test for Equality</a> <a href="#">Test for Non-Inferiority / Superiority</a> <a href="#">Test for Equivalence</a>				
<b>Two-Sample Parallel Design</b> <a href="#">Test for Equality</a> <a href="#">Test for Non-Inferiority / Superiority</a> <a href="#">Test for Equivalence</a>				
<a href="#">Confidence Interval – Bristol</a>				
<a href="#">Compare Two Proportions – Casagrande, Pike &amp; Smith</a>				



Internet | Modo Protegido: Ativado 100%

Sample Size Estimat... Probability of Disea... Microsoft PowerPoi... Comparação entre ... Microsoft Excel uso ... EPI6.EXE PT 22:39

## 4) <http://www.cct.cuhk.edu.hk/stat>

Sample Size Estimation - Compare Two Proportions - Windows Internet Explorer


<http://www.cct.cuhk.edu.hk/stat/proportion/Casagrande.htm>

One sided test:  $H_0: P_1 = P_2$  versus  $H_a: P_1 < P_2$  or  $P_1 > P_2$

Two sided test:  $H_0: P_1 = P_2$  versus  $H_a: P_1 \neq P_2$

Data Input: ([Help](#)) ([Example](#))

Input		Results	
$\alpha$	0.05 <input type="radio"/> one sided test <input checked="" type="radio"/> two sided test	<input type="button" value="Calculate"/>	
$\beta$	0.10	m	538
$P_1$	0.50		
$P_2$	0.40	N	1076
r	1		



Note:

Variables	Descriptions
$\alpha$	Significance level
$1-\beta$	Power of the test
$P_1$	Success proportion in arm 1
$P_2$	Success proportion in arm 2
r	Ratio of arm 2 to arm 1
m	Sample size for arm 1
N	Total sample size for arm 1 and 2

(2 item(ns) restante(s)) Aguardando <http://www.cct.cuhk.edu.hk/stat/proportion/Casagrande.htm>...

Zona desconhecida | Modo Protegido: Ativado | 100%

Sample Size Esti... Probability of D... Sample Size Esti... Microsoft Powe... Comparação en... Microsoft Excel ... cs\ EPI6.EXE PT < 22:42

# 5) <http://www.quesgen.com/SSProp.php>

Probability of Disease Calculator - Windows Internet Explorer  
<http://www.quesgen.com/SSProp.php>

Home Products Tools Blog Clients About Contact Login

Sample Size - Means  
 Sample Size - Proportion  
 Sample Size - Conf Interval for a Mean  
 Conf Interval - Conf Interval for a Proportion  
 Sample Size - Survival Analysis  
 Conf Interval - Likelihood Ratio  
 Posterior Probably of Disease  
 BMI Calculator  
 Peds BMI Calculator  
 Peds BP Calculator  
 Apache III Calculator (testing version)  
 UCSD HIPAA Tutorial

## Sample Size - Proportions

**Instructions:** Enter parameters in the Red cells. Answer will appear in the blue cells.

- Q1 = proportion of subject in group 1 (exposed)
- Q0 = proportion of subject in group 0 (unexposed)
- P1 = Risk in Group 1
- P0 = Risk in Group 0

Alpha (Two-tailed)	Beta	Q1	Q0	P1	P0
0.05	0.10	0.5	0.5	0.5	0.4
Z-Alpha	Z-Beta	P			
1.96	1.28	0.45			
Numerator1	Numerator2	Denominator			
1.95	1.27	0.10			

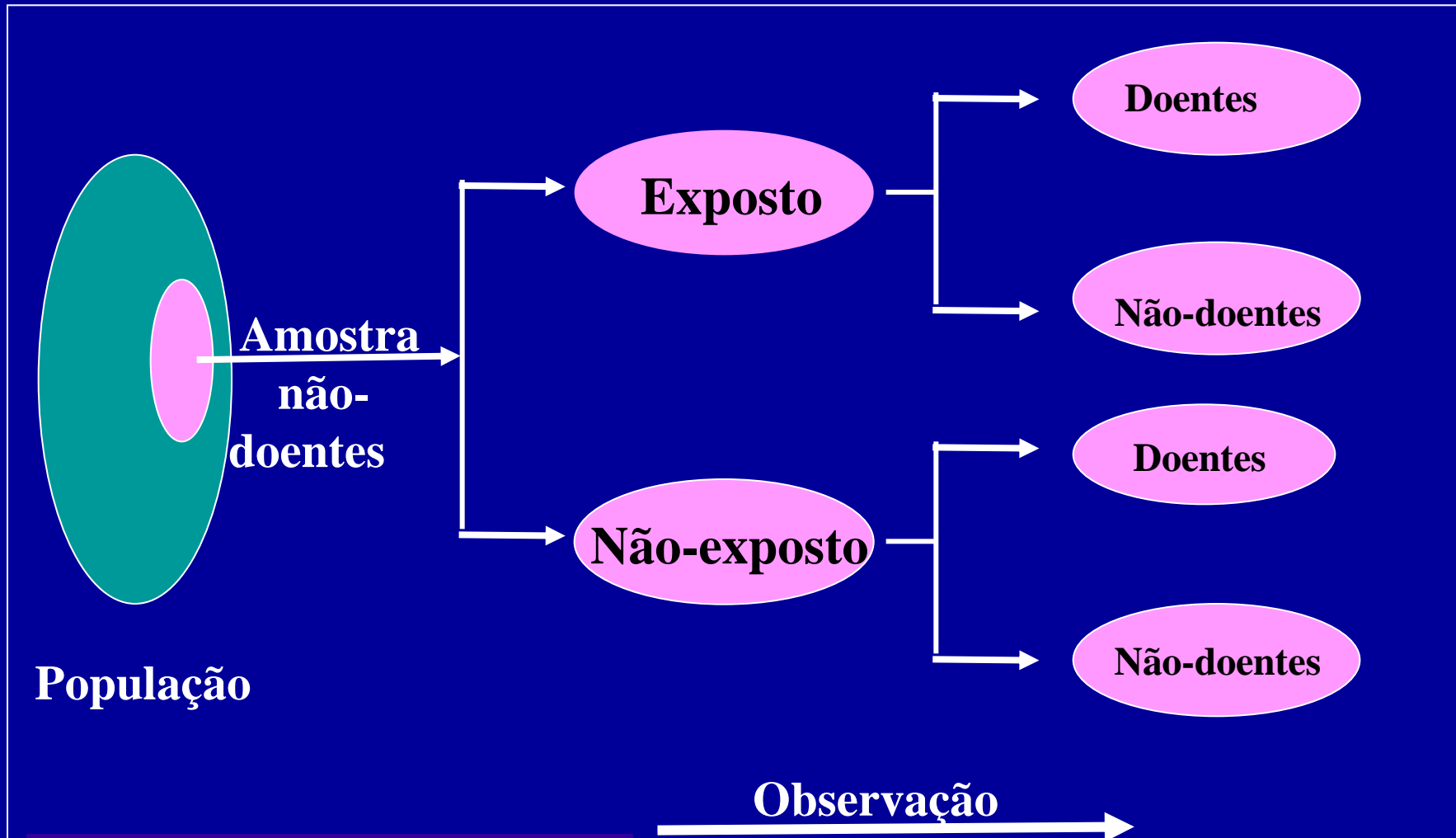
Recalc

WITHOUT THE CONTINUITY CORRECTION				WITH THE CONTINUITY CORRECTION			
	Group 1	Group 0	Total		Group 1	Group 0	Total
	518	518	1036		538	538	1076
Outcome+	259	207	466	Outcome+	269	215	484
Outcome-	259	311	570	Outcome-	269	323	592

Concluído Internet | Modo Protegido: Ativado 100%

E-mail de iG M... Probability of D... Sample Size Esti... Microsoft Powe... Comparação en... Microsoft Excel ... EPI6.EXE PT 22:50

# E em estudos de coorte?



# Estudos de coorte: análise

- Incidência entre os expostos
- Incidências entre os não-expostos
- **Risco Relativo =  $IE/I\tilde{n}E$**
- Risco Atribuído =  $IE - I\tilde{n}E$

# 1) EPI INFO 6.04d e 2002 – Stat Calc

The screenshot shows the Epi Info 2002 main menu with the Statcalc window open. The Statcalc window displays the following text:

```
Statcalc
November 1993

Unmatched Cohort and Cross-Sectional Studies (Exposed and Nonexposed)

Probability that if the two SAMPLES differ this reflects a true
difference in the two POPULATIONS (Confidence level or 1- $\alpha$ ) : 95.00 %

Probability that if the two POPULATIONS differ, the two SAMPLES
will show a "significant" difference (Power or 1- $\beta$ ) : 90.00 %

Ratio (Number of Unexposed : Number of Exposed) : 1 : 1

Expected frequency of disease in unexposed group : 40.00 %

Please fill in the closest value to be detected for ONE of the following:

Risk ratio (RR) or relative risk--closest to 1.00 : 0.00
Odds ratio (OR)--closest to 1.00 : 0.00
Percent disease among exposed--closest to % for unexposed : 50.00 %
```

Four teal arrows point to the values 95.00 %, 90.00 %, 40.00 %, and 50.00 % in the Statcalc window. The main menu includes buttons for MakeView, Enter Data, Analyze Data, Epi Map, Nutrition, and Epi Info Website. The status bar at the bottom shows the taskbar with Microsoft PowerPoint, Epi Info 2002, and the Statcalc window. The system tray shows the time as 20:22 on January 30, 2003.

# 1) EPI INFO 6.04d e 2002 – Stat Calc

**Epi Info 2002**  
 Programs Edit Settings Utilities Help  
 Language=ENGLISH

**Statcalc** November 1993

Unmatched Cohort and Cross-Sectional Studies (Exposed and Nonexposed)  
 Sample Sizes for 40.00 % Disease in Unexposed Group

Conf.	Power	Unex:Exp	Disease in Exposed	Risk Ratio	Odds Ratio	Sample Size		Total
						Unexp.	Exposed	
95.00 %	90.00 %	1:1	50.00 %	1.25	1.50	538	538	1,076
90.00 %	"	"	"	"	"	442	442	884
95.00 %	"	"	"	"	"	538	538	1,076
99.00 %	"	"	"	"	"	754	754	1,508
99.90 %	"	"	"	"	"	1,052	1,052	2,104
95.00 %	80.00 %	"	"	"	"	407	407	814
"	90.00 %	"	"	"	"	538	538	1,076
"	95.00 %	"	"	"	"	660	660	1,320
"	99.00 %	"	"	"	"	924	924	1,848
"	90.00 %	4:1	"	"	"	1,340	335	1,675
"	"	3:1	"	"	"	1,071	357	1,428
"	"	2:1	"	"	"	806	403	1,209
"	"	1:2	"	"	"	404	807	1,211
"	"	1:3	"	"	"	359	1,076	1,435
"	"	1:4	"	"	"	336	1,345	1,681

Change values for inputs as desired, then press F4 to recalculate.

F1-Help F5-Print F6-Open File F10-Done

Exit

MakeView

Enter Data

Analyze Data

Epi Map

Nutrition

Epi Info Website

Revision 2  
 January 30, 2003

Microsoft PowerPoi... Epi Info 2002 C:\Epi\_Info\STATCA...

PT 20:24

## 2) Power and Sample Size Calculations

The image shows a Windows XP desktop environment with a green and blue abstract background. A central dialog box titled "Power and Sample Size Calculations" is open, featuring a logo with the letters "PS" in green and blue, overlaid on a graph of two intersecting curves. The dialog box contains the following text:

**Power and Sample Size Calculations**

Version 3.0.14

The PS program performs power and sample size calculations for many common study designs. It can also generate graphs of sample size versus power, sample size versus detectable alternative hypotheses, or power versus detectable alternative hypotheses. For further information click on the Overview button given below. To obtain interactive help, click on any blue underlined text.

Don't show this screen again

Exit Overview Continue

The desktop background is populated with various icons, including folders like "FUNDEP", "Lixeira", "Documentos", and "Saúde Coletiva UFMG", and applications like "Mozilla Firefox", "Adobe Reader 9", "Microsoft Office Word 2007", "Excel 2007", "PowerPoint 2007", "Epi Info 2002", and "Launch Internet Explorer Browser". On the right side, there are system widgets: a yellow sticky note, a clock showing approximately 10:10, a weather widget for São Paulo, SP showing 21°C, and a calendar for August 23, 2010 (segunda-feira). The taskbar at the bottom shows the Start button, several open applications (Microsoft PowerPoint, Epi Info 2002, and a command prompt), and the system tray with the date and time (20:27).

## 2) Power and Sample Size Calculations

Power and Sample Size Program: Main Window

File Edit Log Help

Survival t-test Regression 1 Regression 2 Dichotomous Mantel-Haenszel Log

[Studies that are analyzed by chi-square or Fisher's exact test](#)

**Output**

[What do you want to know?](#) Sample size

[Case sample size for uncorrected chi-squared test](#) 518

**Design**

[Matched or Independent?](#) Independent

[Case control?](#) Prospective

[How is the alternative hypothesis expressed?](#) Two proportions

[Uncorrected chi-square or Fisher's exact test?](#) Uncorrected chi-square test

**Input**

$\alpha$  0,05  $p_0$  0,40

$power$  0,90  $p_1$  0,50

$m$  1

Calculate

Graphs

**Description**

We are planning a study of independent cases and controls with 1 control(s) per case. Prior data indicate that the failure rate among controls is 0,4. If the true failure rate for experimental subjects is 0,5, we will need to study 518 experimental subjects and 518 control subjects to be able to reject the null hypothesis that the failure rates for experimental and control subjects are equal with probability (power) 0,9. The Type I error probability associated with this test of this null hypothesis is 0,05. We will use an uncorrected chi-squared statistic to evaluate this null hypothesis.

PS version 3.0.14

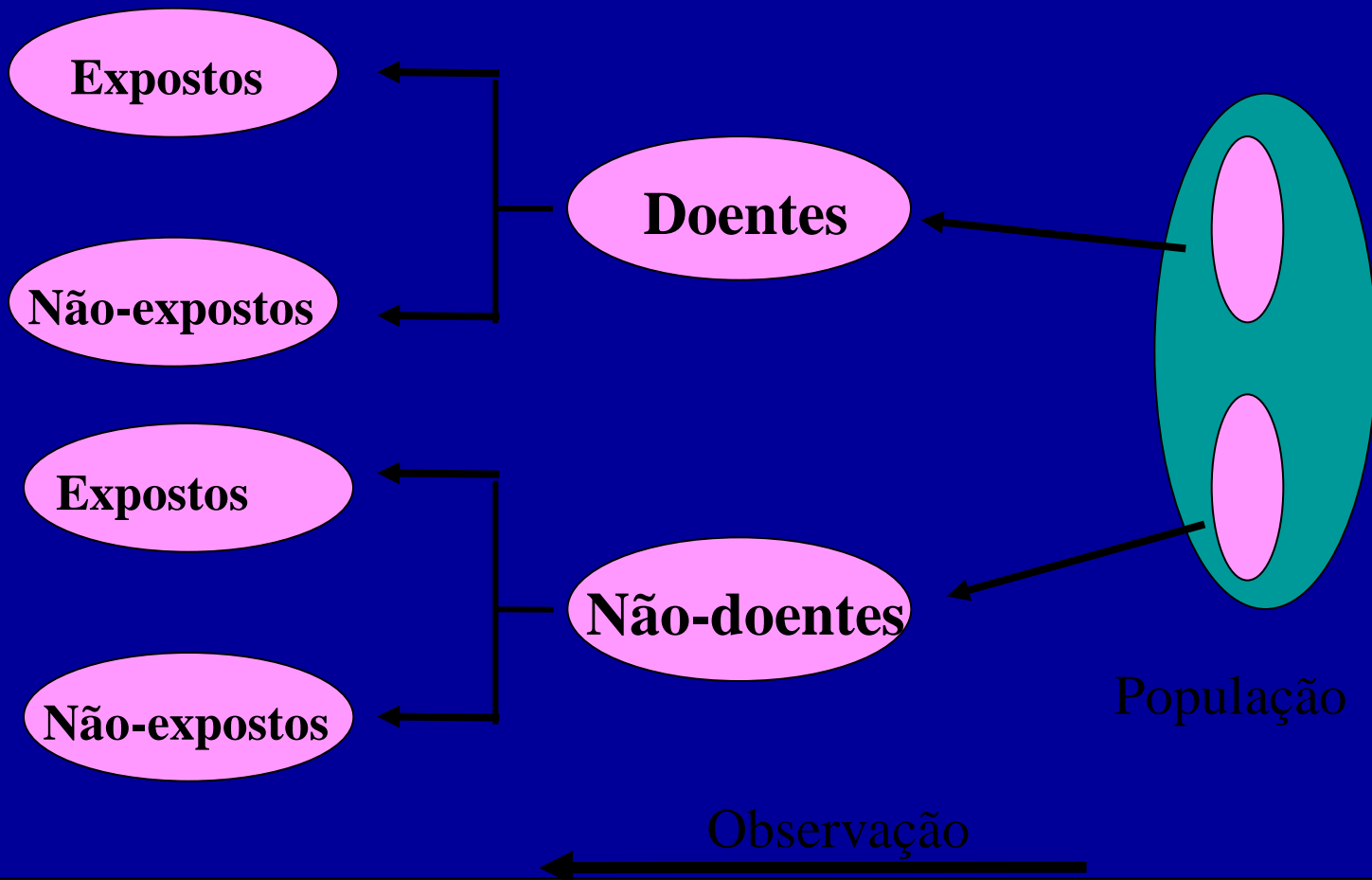
Copy to Log Exit

Logging is enabled.

Microsoft PowerPoi... Epi Info 2002 C:\Epi\_Info\STATCA... Power and Sample ...

PT 20:26

# E em estudos de caso-controle?



# 1) EPI INFO 6.04d e 2002 – Stat Calc

The screenshot shows the Epi Info 2002 main menu with the StatCalc window open. The StatCalc window displays the following data:

Conf.	Power	NOT ILL	ILL	Exposure in ILL	Odds Ratio	NOT ILL	ILL	Total
75.00 %	80.00 %	1:1		20.00 %	0.38	91	91	182
90.00 %	"	"	"			74	74	148
95.00 %	"	"	"			91	91	182
99.00 %	"	"	"			131	131	262
99.90 %	"	"	"			187	187	374
95.00 %	80.00 %	"	"			91	91	182
"	90.00 %	"	"			118	118	236
"	95.00 %	"	"			143	143	286
"	99.00 %	"	"			198	198	396
"	80.00 %	4:1				232	58	290
"	"	3:1				186	62	248
"	"	2:1				138	69	207
"	"	1:2				66	133	199
"	"	1:3				58	174	232
"	"	1:4				54	215	269

The StatCalc window also includes a text box with the instruction: "Change values for inputs as desired, then press F4 to recalculate." The main menu features buttons for "MakeView", "Enter Data", "Analyze Data", "Exit", "Epi Map", "Nutrition", and "Epi Info Website". The status bar at the bottom shows the system tray with the time 20:34 and the date January 30, 2003.

## 2) Power and Sample Size Calculations

Curso SBPqO Como calcular o tamanho da amostra na pesquisa odontológica [Modo de Compatibilidade] - Microsoft PowerPoint uso não comercial

Power and Sample Size Program: Main Window

File Edit Log Help

Survival t-test Regression 1 Regression 2 Dichotomous Mantel-Haenszel Log

[Studies that are analyzed by chi-square or Fisher's exact test](#)

**Output**

[What do you want to know?](#) Sample size

[Case sample size for uncorrected chi-squared test](#) 81

**Design**

[Matched or Independent?](#) Independent

[Case control?](#) Case-Control

[How is the alternative hypothesis expressed?](#) Two proportions

[Uncorrected chi-square or Fisher's exact test?](#) Uncorrected chi-square test

**Input**

$\alpha$  0,05  $p_0$  0,20

power 0,80  $p_1$  0,40

$m$  1

Calculate

Graphs

**Description**

We are planning a study of independent cases and controls with 1 control(s) per case. Prior data indicate that the probability of exposure among controls is 0,2. If the true probability of exposure among cases is 0,4, we will need to study 81 case patients and 81 control patients to be able to reject the null hypothesis that the exposure rates for case and controls are equal with probability (power) 0,8. The Type I error probability associated with this test of this null hypothesis is 0,05. We will use an uncorrected chi-squared statistic to evaluate this null hypothesis.

PS version 3.0.14

Copy to Log Exit

Logging is enabled.

Slide 83 de 91 "Contemp" Português (Brasil)

Curso SBPqO Co... Aula de desenho ... Epi Info 2002 C:\Epi\_Info\STAT... KINGSTON (E) Power and Sampl... PT 79% 20:37

## 2) Power and Sample Size Calculations

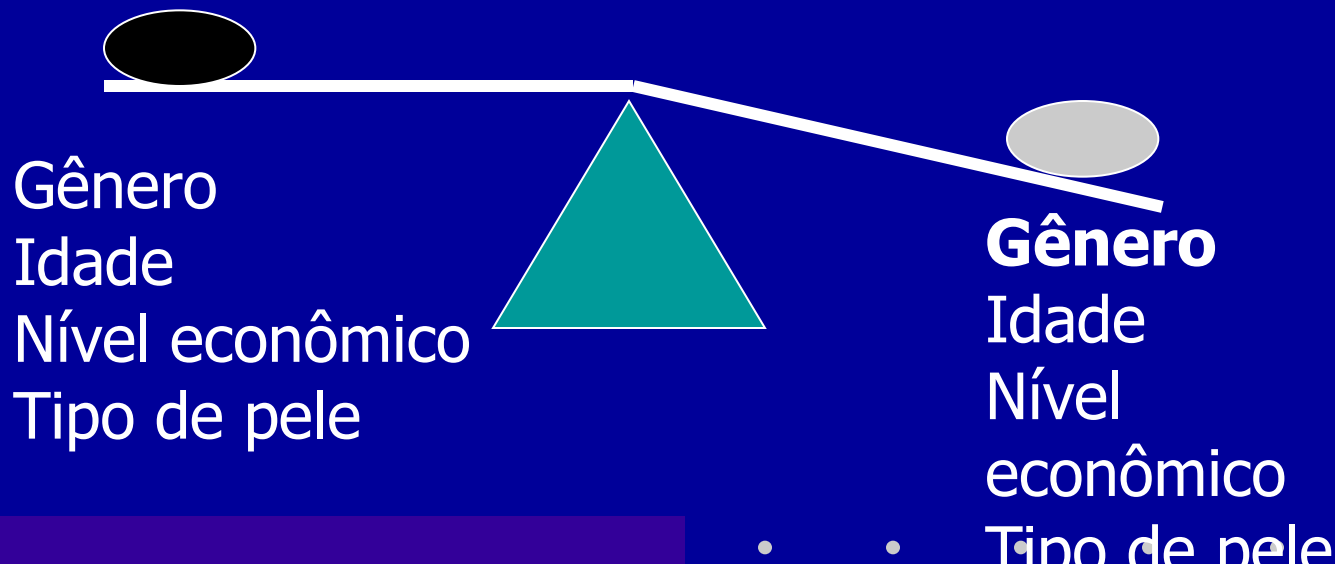
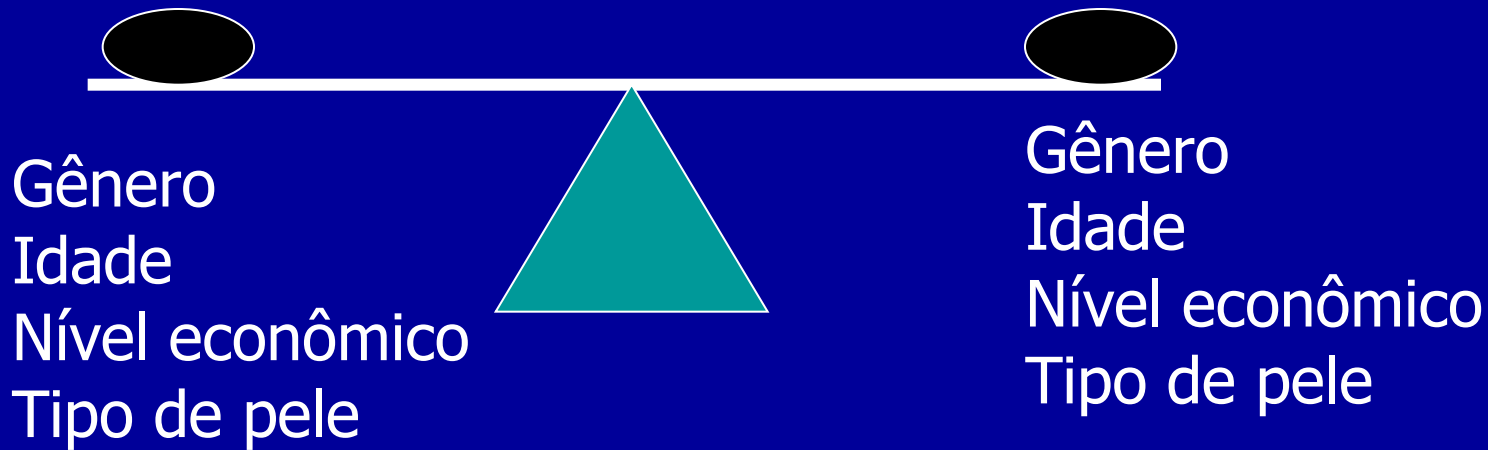
The screenshot displays the 'Power and Sample Size Program: Main Window' interface. The 'Output' section is titled 'Studies that are analyzed by chi-square or Fisher's exact test'. Under 'What do you want to know?', the 'Sample size' dropdown is set to '91'. The 'Design' section includes 'Matched or Independent?' set to 'Independent', 'Case control?' set to 'Case-Control', 'How is the alternative hypothesis expressed?' set to 'Two proportions', and 'Uncorrected chi-square or Fisher's exact test?' set to 'Fisher's exact test'. The 'Input' section shows  $\alpha$  (0,05),  $p_0$  (0,20),  $p_1$  (0,40),  $m$  (1), and  $\text{power}$  (0,80). A 'Calculate' button is visible. The 'Description' section contains a text box with the following text: 'We are planning a study of independent cases and controls with 1 control(s) per case. Prior data indicate that the probability of exposure among controls is 0,2. If the true probability of exposure among cases is 0,4, we will need to study 91 case patients and 91 control patients to be able to reject the null hypothesis that the exposure rates for case and controls are equal with probability (power) 0,8. The Type I error probability associated with this test of this null hypothesis is 0,05. We will use a continuity-corrected chi-squared statistic or Fisher's exact test to evaluate this null hypothesis.'

**Teste Exato de Fisher**

PS version 3.0.14  
Copy to Log Exit  
Logging is enabled.

•  
•  
•

# E se o caso-control for pareado?



Desktop icons and folders:

- FUNDEP
- Lixeira
- Mozilla Firefox
- Adot
- Microsoft Office
- Saúde Coletiva UFMG
- EPI6 - Atalho
- Computador
- resultad\_final03
- Iwanga limeshow 1-22
- Rede
- Disertacao
- Painel de Controle
- revista-idpc-2000
- Aula caso-control...
- itens-bio
- Epi Info 2002
- Launch Internet Explorer Browser

Power and Sample Size Program: Main Window

File Edit Log Help

Survival | t-test | Regression 1 | Regression 2 | Dichotomous | Mantel-Haenszel | Log

[Studies that are analyzed by chi-square or Fisher's exact test](#)

**Output**

[What do you want to know?](#) Sample size

[Required number of case patients](#) 118

---

**Design**

[Matched or Independent?](#) Matched or paired

[Case control?](#) Case-Control

---

**Input**

$\alpha$  0,05      $p_0$  0,60     Calculate

[power](#) 0,80      $\phi$  0,68     Graphs

$m$  1      $\psi$  4,20

---

**Description**

We are planning a study of matched sets of cases and controls with 1 matched control (s) per case. Prior data indicate that the probability of exposure among controls is 0,6 and the correlation coefficient for exposure between matched cases and controls is 0,68. If the true odds ratio for disease in exposed subjects relative to unexposed subjects is 4,2, we will need to study 118 case patients with 1 matched control(s) per case to be able to reject the null hypothesis that this odds ratio equals 1 with probability (power) 0,8. The Type I error probability associated with this test of this null hypothesis is

PS version 3.0.14

Copy to Log     Exit

Logging is enabled.

System tray area:

- Yellow sticky note
- Clock: 10:10
- Weather: 21° São Paulo, SP
- Calendar: agosto 2010, 23 segunda-feira

Taskbar:

- Windows Start button
- Taskbar icons: Curso SBPQO..., Aula de dese..., Epi Info 2002, C:\Epi\_Info\S..., KINGSTON (E:), Power and Sa..., Help
- System tray: PT, network, volume, clock 20:47

### Input variables on the *Dichotomous* dialog box

$\alpha$

The Type I error probability for a two sided test. This is the probability that we will falsely reject the null hypothesis.

$n$

For case-control studies  $n$  is the number of case patients. For prospective studies  $n$  is the number of patients receiving the experimental treatment.

*power*

the probability of correctly rejecting the null hypothesis that the relative risk (odds ratio) equals 1 given  $n$  case patients,  $m$  control patients per experimental patient, and a Type I error probability  $\alpha$ .

$p_0$

For case-control studies,  $p_0$  is the probability of exposure in controls. In prospective studies,  $p_0$  is the probability of the outcome for a control patient.

$p_1$

For case-control studies,  $p_1$  is the probability of exposure in cases. In prospective studies,  $p_1$  is the probability of the outcome in an experimental subject.

$m$

For independent prospective studies  $m$  is the ratio of control to experimental subjects. For matched prospective studies  $m$  is the number of controls subjects matched to each experimental subject. For independent case-control studies  $m$  is the ratio of control to case patients. For matched case-control studies  $m$  is the number of control patients matched to each case.

$R$

Relative risk of failure for experimental subjects relative to controls.

$\phi$

For case-control studies  $\phi$  is the correlation coefficient for exposure between matched cases and controls. For prospective studies  $\phi$  is the correlation coefficient for the outcome between matched experimental and control subjects.

$\psi$

Odds ratio of exposure in cases relative to controls. For a matched study design  $\psi$  is the odds ratio adjusted for the matching variable(s). For an independent design it is the unadjusted odds ratio. If the cases and controls are unbiased samples from respective target populations then  $\psi$  also equals the odds ratio for disease in exposed subjects relative to unexposed subjects. The odds ratio approximates the relative risk in case-control studies when the disease is rare in the general population.

**Coefficiente de correlação para exposição entre os pares casos e controles**

**Odds Ratio (Razão das chances)**



Curso SBPqO ...

Aula de dese...

Epi Info 2002

C:\Epi\_Info\S...

KINGSTON (E:)

Power and Sa...

Help

PT

System tray icons: network, volume, power, and system clock showing 20:50

# *RISCO* versus *ODDS*

**RISCO** = a possibilidade de que algo aconteça  
a possibilidade de que **tudo** aconteça

**ODDS** = a possibilidade de que algo aconteça  
a possibilidade de que algo **não** aconteça

Exemplo: Entre 100 pessoas estudadas, 20 desenvolvem câncer de boca em um ano.

O **RISCO** é 1 em 5 (Ex: 20 em 100).

A **ODDS** é de 1 em 4 (Ex: 20 comparados a 80).

# Estudos de caso-controle

- Análise dos dados

ODDS RATIO = **CHANCE/ODDS** = possibilidade de ocorrência de um evento  
possibilidade da não ocorrência do evento

$$\frac{\frac{a}{c}}{\frac{b}{d}}$$



$$\text{OR} = \frac{a \times d}{b \times c}$$

Exposição	Doença / evento	
	Presente	Ausente
Presente	a	b
Ausente	c	d

## TAMANHO DE AMOSTRA

### **ESTUDOS COMPARATIVOS: MÉDIA**

Uma pesquisadora deseja verificar a existência de diferenças entre o CPOD de crianças de níveis econômicos distintos. Estudos anteriores indicam que o CPOD das crianças mais ricas é igual a 2,5 ( $s = 0,5$ ) enquanto que nas mais pobres a média é igual a 4,0 ( $s = 0,7$ ). Quantas crianças deveriam participar do estudo se é desejado detectar um diferença entre os dois grupos de 0,5 unidade, com nível de significância de 5% e um poder do teste de 80%?



- 
- 
- 

$$n = \frac{(1,96+1,29)^2 (0,5^2 + 0,7^2)}{0,5^2}$$

n= 32 crianças

# Falhas que devem ser evitadas (1)

- *“A entrada de pacientes na clínica é por volta de 50 pacientes no ano, dos quais 10% podem não querer participar do estudo. Assim, ao longo de 2 anos do estudo, a amostra será de 90 pacientes.”*
- Embora a maioria dos estudos precise balancear a viabilidade com o poder, o tamanho da amostra não pode ser decidido no número de pacientes disponíveis apenas.
- Em estudos nos quais o número de pacientes disponíveis é um conhecido fator limitante, o cálculo amostral deverá, também ser feito, para indicar se:
- O poder que o estudo tem para detectar as diferenças de importância clínica/epidemiológica, OU
- A diferença que será detectada quando o poder desejado é aplicado.

# Falhas que devem ser evitadas (2)

- *“O cálculo amostral não foi realizado pois não havia qualquer informação anterior para se basear”*
- Faça um estudo piloto;
- Utilize a proporção que garanta o maior cálculo (50%);

# Falhas que devem ser evitadas (3)

- “Um estudo prévio na mesma região que eu estou estudando, recrutou 150 indivíduos e encontrou resultados estatisticamente significativos ( $p=0,014$ ), assim uma amostra similar deve ser suficiente no meu estudo.”
- Estudos prévios podem ter tido “sorte” de encontrar tais diferenças, devido à variação amostral aleatória.



[maurohenrique@ufmg.br](mailto:maurohenrique@ufmg.br)