

ISA-CAMPINAS 2008/09
PLANO DE AMOSTRAGEM
Maria Cecília Goi Porto Alves

1. População de estudo e estratos

A população de estudo foi definida como sendo as pessoas com 10 anos de idade ou mais, residentes nos setores censitários urbanos do Município de Campinas que são constituídos por domicílios particulares ou coletivos não institucionais¹.

Foi fixado como objetivo da pesquisa o estudo de aspectos referentes a três subgrupos dessa população, os de adolescentes, adultos e idosos. Dessa forma, a população foi dividida em três estratos (de 12 a 19 anos; de 20 a 59 anos e de 60 anos ou mais), que constituíram domínios de estudo e para os quais foram sorteadas amostras independentes.

A Tabela 1 mostra a distribuição da população por esses grupos de idade. Os dados foram extraídos de arquivos do censo de 2000, constantes em CD fornecido pelo IBGE, tendo sido somados os valores referentes aos setores censitários de interesse.

Tabela 1. População residente em área urbana, segundo idade. Campinas, 2000.

Idade	População
10 a 19 anos	168759
20 a 59 anos	541414
60 anos e mais	91271
Total	801444

Optou-se pelo sorteio de amostras de tamanhos iguais, de 1000 pessoas, em cada estrato. Esperava-se que com esse número de entrevistas, o erro de amostragem estivesse entre 4 e 5 pontos percentuais em intervalos de confiança de 95%, para proporções de 0,50 e efeitos de delineamento de 2.

A estimativa do erro de amostragem foi feita por $d = z \sqrt{\frac{P \cdot (1 - P) \cdot deff}{n}}$, resultante

da expressão algébrica que determina o tamanho de amostras para estimação de proporções:

¹ Inclui setores de situação urbana: área urbanizada, não urbanizada e urbanizada isolada, e dos tipos: comum e especial subnormal (favela).

$n_o = \frac{P(1-P)}{(d/z)^2} \cdot deff$, em que P é a proporção de indivíduos a ser estimada; z é o valor na curva normal reduzida, correspondente ao nível de confiança utilizado na determinação do intervalo de confiança; d é o erro de amostragem e *deff* é o efeito do delineamento (Kish, 1965; Silva, 2001).

A partilha igual da amostra por estratos que não possuem tamanhos iguais faz com as probabilidades de sorteio (frações de amostragem) em cada estrato sejam distintas, levando, conseqüentemente, à necessidade de introdução de pesos para obter estimativas referentes ao conjunto dos estratos (Tabela 2). Os pesos são o inverso da fração amostra/população.

Tabela 2. População, amostra e pesos, segundo domínios de estudo.

Domínio	População	Amostra	Pesos
10 a 19 anos	168759	1000	168,759
20 a 59 anos	541414	1000	541,414
60 anos e mais	91271	1000	91,271
Total	801444	3000	

Em função desses pesos, o tamanho efetivo da amostra será de 1950 na obtenção de estimativas para o conjunto dos domínios (estimativa feitas para rho, a homogeneidade intraclasse, de 0,05).

O cálculo do número de domicílios foi feito a partir da média esperada de pessoas por domicílio (razão pessoas/domicílios) em cada grupo de idade, sendo que o total de domicílios existentes no Município de Campinas, considerando-se os setores censitários de interesse, era de 282147 no ano do censo. Dividiu-se, então, o tamanho da amostra (1000 pessoas) por essa média (Tabela 3).

Tabela 3. Razão população/domicílios e números de domicílios da amostra, segundo domínio de estudo.

Domínio de estudo	Razão	Domicílios amostra
10 a 19 anos	0,598124	1672

20 a 59 anos	1,918908	522
60 anos e mais	0,323487	3092

O número de setores censitários em que deveriam ser distribuídos esses domicílios foi definido considerando que o total de entrevistas por setor não deveria ultrapassar 20 (o que garantiria $deff < 2$ ($deff = 1,95$) para $\rho = 0,05$). Dessa forma, $a = 50$ (uma vez que $1000/50 = 20$).

Em função dessa decisão, foram definidos os números de domicílios a serem sorteados em cada setor (Tabela 4).

Tabela 4. Números de domicílios e setores na amostra e número de domicílios por setor, segundo domínio de estudo.

Domínio	Amostra domicílios	Setores	Amostra domicílios por setor
10 a 19 anos	1672	50	33,44
20 a 59 anos	522	50	10,44
60 anos e mais	3092	50	61,84

Prevendo-se a ocorrência de não resposta, de domicílios vagos e o aumento no número de domicílios desde o censo, foram sorteados números 20% maiores² de domicílios, 43, 14 e 78, respectivos aos três grupos de idade, elevando os tamanhos de amostra para 2150, 700 e 3900.

Processo de amostragem

Foi utilizada amostragem em dois estágios: setor censitário e domicílio (Kish, 1965). No primeiro estágio os setores censitários foram sorteados com probabilidade proporcional ao tamanho (número de domicílios). Para o sorteio, feito por

² O cálculo levou em consideração a taxa esperada de não resposta de 20%, o percentual esperado de domicílios vagos de 5% e o crescimento esperado no número de domicílios nos setores de 10%. Total: $0,8 * 0,95 * 1,10 = 0,83 \cong 0,80$

meio de sorteio sistemático, os setores foram ordenados pelo percentual de chefes que possuíam nível universitário, produzindo uma estratificação implícita por escolaridade do chefe de família.

As frações de amostragem que orientaram o sorteio em cada grupo de idade foram:

$$\text{- adolescentes} \rightarrow f = 50 \cdot \frac{M_h}{M} \cdot 43 \cdot \frac{1}{M_h} = \frac{2150}{282147};$$

$$\text{- adultos} \rightarrow f = 50 \cdot \frac{M_h}{M} \cdot 14 \cdot \frac{1}{M_h} = \frac{700}{282147};$$

$$\text{- idosos} \rightarrow f = 50 \cdot \frac{M_h}{M} \cdot 78 \cdot \frac{1}{M_h} = \frac{3900}{282147},$$

sendo $M=282147$ o número de domicílios do Município de Campinas.