

DIMENSIONAMENTO DE AMOSTRA PARA LEVANTAMENTO DA CITRICULTURA PAULISTA¹

VERA LÚCIA FERRAZ DOS SANTOS², FRANCISCO ALBERTO PINO³,
ANTONIO AMBROSIO AMARO⁴ e FERNANDO ANTONIO DE ALMEIDA SEVER[†]

RESUMO - Dimensiona-se uma amostra para levantamento periódico da produção de citros no Estado de São Paulo. Três esquemas amostrais alternativos são comparados: 1) amostra casual simples (ACS); 2) amostra estratificada por regiões e tamanho do imóvel rural; e 3) amostra estratificada por regiões, tamanho do imóvel e grupo de municípios dentro da região. Nove níveis de variância desejada (10%, 5%, 2,5%, 1%, 0,5%, 0,4%, 0,3%, 0,2% e 0,1%) são comparados. Sugere-se o uso de uma amostra estratificada por região, tamanho e grupo, ao nível de 1%, com 583 elementos.

Termos para indexação: amostragem estratificada, alocação ótima.

SAMPLE DESIGN FOR CITRUS SURVEY IN SÃO PAULO, BRAZIL

ABSTRACT - A sample is designed for a periodic citrus production survey in the State of São Paulo, Brazil. Three alternative designs (simple random sample, stratified sample by region and farm size, stratified sample by region, farm size and group within region) and nine levels of desired variance (10%, 5%, 2.5%, 1%, 0.5%, 0.4%, 0.3%, 0.2% and 0.1%) are compared. A stratified sample by region, size and group at level 1% with 583 elements is suggested.

Index terms: stratified sample, optimal allocation.

INTRODUÇÃO

A citricultura tem ganho importância na agricultura paulista em relação a outras culturas. Sua participação no valor da produção agrícola passou de 6% em 1979 para 11% em 1984, conforme Sever & Amaro (1984), que apresentam uma avaliação da situação da citricultura. Na década de 70 a Secretaria da Agricultura do Estado de São Paulo, através do Instituto de Economia Agrícola (IEA) e da Coordenadoria de Assistência Técnica Integral (CATI) instalou um levantamento sistemático de dados para previsão e estimativa de safras de citros, realizado duas vezes por ano, numa amostra de produtores, com metodologia semelhante à descrita em Campos & Piva (1974). Tal levantamento efetuava-se apenas nas três principais Divisões Regionais Agrícolas (DIRAs) e deixou de ser feito nos últimos anos.

Dada a crescente importância da citricultura paulista, justifica-se o redimensionamento desse tipo de amostra. Assim sendo, o objetivo deste trabalho é discutir e calcular uma amostra de citricultores para levantamento de dados estatísticos sobre o desenvolvimento da citricultura em todo o Estado de São Paulo.

MATERIAL E MÉTODOS

Sistema de referência

Um primeiro cadastro de cerca de 10.000 citricultores foi fornecido pelas quatro maiores empresas fabricantes de sucos cítricos no Estado. No ano agrícola 1980/81, o IEA e a CATI aplicaram um questionário no campo para verificar tal cadastro e para incluir os citricultores que dele não faziam parte. Após um processo de depuração, obteve-se o cadastro definitivo, com 18.277 citricultores, que foi utilizado no trabalho de Sever & Amaro (1984). É este cadastro que serve de sistema de referência no presente trabalho.

Unidade amostral e variável para dimensionamento da amostra

A unidade amostral é o imóvel rural produtor de citros (laranja, limão e tangerina). Justifica-se o uso de amostra específica para citricultura por duas razões: a) a citricultura é razoavelmente concentrada, com mais de 80% dos citricultores em apenas três das dez DIRAs (Tabela 1), e b) por se tratar de cultura perene, o cadas-

¹ Aceito para publicação em 5 de fevereiro de 1986.

² Estatística, Inst. Econ. Agric. (IEA), Bolsista do CNPq. Caixa Postal 8114, CEP 04301 São Paulo, SP.

³ Eng. Agr., M.Sc., IEA, Caixa Postal 8114, CEP 04301 São Paulo, SP.

⁴ Eng. - Agr., Dr., IEA, Caixa Postal 8114, CEP 04301 São Paulo, SP.

† In memoriam.

tro é relativamente estável, não se alterando muito de um ano para outro.

Escolheu-se o número total de pés plantados como variável básica para dimensionamento da amostra por ser mais estável do que a produção. É de se esperar que o erro de amostragem de qualquer item a ser levantado seja igual ou superior ao da variável básica, com raras exceções. Por outro lado, como apenas 10% dos citricultores não plantam laranja (Tabela 1), pode-se usar o número total de pés plantados (com e sem produção) com citros, ao invés de calcular amostras individuais para laranja, limão e tangerina.

c. **Estratificação por grupo de municípios** - Com base nos levantamentos subjetivos, em nível municipal, realizados pelo IEA e a CATI, no ano agrícola 1982/83 (levantamento de junho para tangerina e levantamento de novembro para laranja e limão), os municípios foram classificados quanto ao valor da produção da citricultura e divididos em dois grupos. No primeiro grupo incluíram-se os municípios de maior valor da produção citrícola até completar 50% do valor da produção de cada DIRA. No segundo grupo incluíram-se os municípios restantes (Tabelas 4 e 5). Desta maneira, o primeiro grupo contém os principais municípios de cada DIRA do ponto de vista da citricultura. -

TABELA 1. Número de citricultores por espécie e por DIRA, Estado de São Paulo, 1980/81.

DIRAs	Laranja	Limão	Tangerina	Laranja/ limão	Laranja/ Tangerina/ limão	Laranja/ tangerina limão	Total
São Paulo	7	16	318	5	27	57	590
Vale do Paraíba	0	5	15	3	24	6	122
Sorocaba	15	15	73	3	59	17	368
Campinas	1.153	47	45	218	1.114	43	4.056
Ribeirão Preto	3.352	427	62	789	1.261	86	7.604
Bauru	19	6	35	13	102	3	285
São José do Rio Preto	1.488	362	6	770	264	15	3.643
Araçatuba	21	1	2	14	44	0	324
Presidente Prudente	0	2	22	3	10	5	315
Marília	14	7	142	8	174	40	802
Estado	6.049	888	720	1.826	3.109	272	18.277

Fonte: Instituto de Economia Agrícola e Coordenadoria de Assistência Técnica Integral.

Esquemas amostrais

Compararam-se três esquemas amostrais: 1) amostragem casual simples (ACS); 2) amostragem estratificada com dois critérios de estratificação (região e tamanho) e 3) amostragem estratificada com três critérios de estratificação (região, tamanho e grupo).

Os critérios de estratificação considerados foram os seguintes:

a. **Estratificação regional** - Cada estrato é uma das dez DIRAs em que se divide o Estado (Fig. 1). Os objetivos aqui são dois: obter estatísticas por DIRA e diminuir a variância, uma vez que a citricultura não se distribui uniformemente pelo Estado.

b. **Estratificação por tamanho** - A estratificação por tamanho do cultivo pode contribuir para diminuir a variância, sendo especialmente importante para as variáveis que não foram utilizadas no dimensionamento da amostra. Utilizaram-se os mesmos estratos propostos por Sever & Amaro (1984), sendo as culturas com menos de quinhentos pés consideradas não comerciais (Tabelas 2 e 3), o que reduz a população para 13.823 elementos.

Esse critério garante a obtenção de estatísticas para todas as DIRAs. Sem ele a amostra tenderia a se concentrar nas três principais DIRAs produtoras, obtendo-se, então, melhores resultados para os totais do Estado, porém resultados inferiores por DIRA. Além disso, como o número de pés de laranja é muito superior ao de tangerina e limão, sem este critério a amostra tenderia a dar bons resultados apenas para laranja.

O grupo 1 é constituído dos seguintes municípios:

- DIRA São Paulo: Moji das Cruzes, Biritiba-Mirim e Pariquera-Açú;
- DIRA Vale do Paraíba: Jacaref, Taubaté e Pindamonhangaba;
- DIRA Sorocaba: Capela do Alto, Sorocaba, Tatuf e Itu;
- DIRA Campinas: Mogi-Guaçu, Artur Nogueira, Pirassununga, Limeira, Araras e Porto Ferreira;
- DIRA Ribeirão Preto: Bebedouro, Monte Azul Paulista, Itápolis, Taquaritinga, Barretos, Araraquara, Colina, Pirangi e Tabatinga;
- DIRA Bauru: Dois Córregos, Bauru, Getulina e Bariri;

- g) DIRA São José do Rio Preto: Olímpia, Cajobi, Tabapuã, Severínia, Paraíso e São José do Rio Preto;
- h) DIRA Araçatuba: Guaraçai, Auriflama e Andradina;
- i) DIRA Presidente Prudente: Parapuã; e
- j) DIRA Marília: Marília.

Determinação da amostra

As referências básicas são Kish (1965) e Cochran (1953).

Amostra casual simples

O tamanho da amostra é dado por $n = N/(1 + V/N S^2)$ onde n é o tamanho (número de elementos) da amostra, N é o número de elementos da população, S^2 é a variância na população e V é a variância desejada, que pode ser calculada por $V = D^2/T^2$ onde D é o valor desejado para a semi-amplitude do intervalo de confiança do total, ao nível α de probabilidade, enquanto T é o valor da estatística t de Student ao mesmo nível α . No presente trabalho utilizaram-se os seguintes valores para erro de amostragem (coeficiente de variação da estimativa do total do Estado):

$D = 10\%, 5\%, 2,5\%, 1\%, 0,5\%, 0,4\%, 0,3\%, 0,2\%$ e $0,1\%$.

TABELA 2. Estratificação quanto ao número total de pés plantados de citros.

Estrato	Número de pés plantados
0	até 500
1	501 a 5.000
2	5.001 a 12.000
3	12.001 a 20.000
4	20.001 a 40.000
5	40.001 a 100.000
6	mais a 100.00

Fonte: Sever & Amaro (1984).

Amostra estratificada

O tamanho da amostra, com H estratos é dado por

$$n = \left(\sum_{h=1}^H N_h S_h \right) / \left(V + \sum_{h=1}^H N_h S_h \right)$$

onde N_h é o número de elementos da população no estrato h e S_h é a variância da população no estrato h .



FIG. 1. Divisões regionais agr(colas (DIRAs), Estado de São Paulo.

TABELA 3. Distribuição do número de citricultores por DIRA e estrato, Estado de São Paulo, 1980/81.

DIRA	Estrato							Total
	0	1	2	3	4	5	6	
São Paulo	245	321	13	6	3	2	0	590
Vale do Paraíba	53	55	3	1	0	0	0	122
Sorocaba	130	191	16	6	15	4	6	368
Campinas	994	1.992	668	180	130	71	21	4.056
Ribeirão Preto	1.056	3.950	1.601	506	317	136	38	7.604
Bauru	282	148	21	2	4	4	2	463
São José do Rio Preto	785	1.907	637	172	96	38	4	3.643
Araçatuba	218	87	10	4	4	1	0	324
Presidente Prudente	283	31	1	0	0	0	0	315
Marília	408	357	34	2	1	0	0	802
Estado	4.454	9.039	3.004	879	570	256	75	18.277

Fonte: Instituto de Economia Agrícola e Coordenadoria de Assistência Técnica Integral.

TABELA 4. Distribuição do número de citricultores e municípios por DIRA e por grupo de municípios.

DIRA	Grupo 1	Municípios	Grupo 2	Municípios	Total	
					Grupo	Municípios
São Paulo	224	3	366	74	590	77
Vale do Paraíba	17	3	95	29	112	32
Sorocaba	54	4	314	55	368	59
Campinas	1.595	6	2.461	61	4.056	67
Ribeirão Preto	2.981	9	4.623	71	7.604	80
Bauru	101	4	362	34	463	38
São José do Rio Preto	1.586	6	2.057	78	3.643	84
Araçatuba	49	3	275	35	524	38
Presidente Prudente	25	1	290	49	315	50
Marília	225	1	577	46	802	47
Estado	6.857	40	11.420	532	18.277	572

Fonte: Instituto de Economia Agrícola e Coordenadoria de Assistência Técnica Integral.

Para alocação da amostra nos estratos usou-se a partilha de Neyman, supondo-se custos iguais em todos os estratos:

$$N_h = n \frac{N_h S_h}{\sum_{h=1}^H N_h S_h}$$

onde n_h é o tamanho da amostra no estrato h .

Depois de sorteada a amostra e levantados os dados, a expansão far-se-á da forma usual (Cochran 1953).

RESULTADOS E DISCUSSÃO

As variâncias para cálculo da amostra encontram-se na Tabela 6.

Pelos tamanhos de amostra obtidos (Tabela 7) verifica-se que a estratificação foi eficiente em relação à ACS no sentido de reduzir a variância das estimativas para um tamanho de amostra fixado (ou, equivalente, de reduzir o tamanho da amostra para uma variância das estimativas fixadas).

Entretanto, a estratificação por grupo de município só se revelou interessante com erro de amostragem menor ou igual a 1%.

O número de elementos na amostra, como função do erro de amostragem desejado, segue uma curva semelhante a uma exponencial (Fig. 2), tornando-se evidente a comparação entre os esquemas analisados.

As alternativas mais interessantes foram as de

amostragem estratificada com grupo, com 1% ou 0,5% de erro da amostragem na estimativa do número total de pés de citros no Estado. Uma alternativa que produz uma amostra de tamanho razoável para ser levantada é a de 1%, que, ao se incluir pelo menos dois produtores em cada estrato para tornar a amostra mensurável (isto é, para se poder calcular a variância das estimativas), resulta numa amostra de 583 citricultores (Tabela 8).

TABELA 5. Número de pés de citros por estrato, por grupo e por DIRA, 1981.

Grupo 1											
Estrato	DIRA										Estado
	2	3	4	5	6	7	8	9	10	11	
1	257.024	18.113	63.210	2.416.918	3.977.451	45.263	2.011.484	9.665	31.515	214.054	9.044.697
2	54.900	0	19.290	2.790.705	5.826.287	72.828	2.391.224	32.000	9.940	97.794	11.294.968
3	44.500	0	27.100	1.482.482	3.637.356	0	1.244.169	15.000	0	16.459	6.467.066
4	39.600	0	108.400	1.653.245	4.714.489	0	1.365.053	85.000	0	21.536	7.987.323
5	107.300	0	206.670	1.976.227	4.753.711	208.412	1.307.038	40.360	0	0	8.599.718
6	0	0	570.420	1.665.031	3.269.532	110.000	460.520	0	0	0	6.075.503
Total	503.324	18.113	995.090	11.984.608	26.178.826	436.503	8.779.488	182.025	41.455	349.843	49.469.275
Grupo 2											
Estrato	DIRA										Estado
	2	3	4	5	6	7	8	9	10	11	
1	194.075	71.532	272.437	2.265.832	5.817.772	169.920	2.344.394	104.826	17.255	335.803	11.593.846
2	49.504	17.000	117.780	2.281.265	6.530.828	85.633	2.533.571	43.385	0	172.545	11.831.511
3	45.000	12.700	60.370	1.248.431	4.185.066	34.800	1.402.126	43.947	0	14.400	7.046.840
4	49.809	0	302.166	1.923.375	4.071.470	96.492	1.242.967	27.082	0	0	7.713.361
5	0	0	71.490	2.093.573	3.085.456	40.300	882.623	0	0	0	6.173.442
6	0	0	367.500	2.335.279	3.606.852	400.187	440.303	0	0	0	7.150.121
Total	338.388	101.232	1.191.743	12.147.755	27.297.444	827.332	8.845.984	219.240	17.255	522.748	51.509.121

Fonte: Instituto de Economia Agrícola e Coordenadoria de Assistência Técnica Integral.

TABELA 6. Variância do número total de pés plantados com citros, por estrato, por grupo e por DIRA, 1980/81.

DIRA	Estrato 1		Estrato 2		Estrato 3	
	Grupo 1	Grupo 2	Grupo 1	Grupo 2	Grupo 1	Grupo 2
2	$0,109 \times 10^7$	$0,835 \times 10^6$	$0,605 \times 10^7$	$0,371 \times 10^7$	$0,205 \times 10^7$	$0,267 \times 10^7$
3	$0,565 \times 10^6$	$0,145 \times 10^7$	—	$0,391 \times 10^6$	—	—
4	$0,164 \times 10^7$	$0,141 \times 10^7$	$0,536 \times 10^6$	$0,496 \times 10^7$	$0,211 \times 10^7$	$0,838 \times 10^7$
5	$0,166 \times 10^7$	$0,165 \times 10^7$	$0,341 \times 10^7$	$0,357 \times 10^7$	$0,419 \times 10^7$	$0,702 \times 10^7$
6	$0,166 \times 10^7$	$0,169 \times 10^7$	$0,403 \times 10^7$	$0,385 \times 10^7$	$0,502 \times 10^7$	$0,571 \times 10^7$
7	$0,843 \times 10^6$	$0,132 \times 10^7$	$0,362 \times 10^7$	$0,514 \times 10^7$	—	$0,160 \times 10^6$
8	$0,179 \times 10^7$	$0,164 \times 10^7$	$0,399 \times 10^7$	$0,346 \times 10^7$	$0,540 \times 10^7$	$0,517 \times 10^7$
9	$0,924 \times 10^5$	$0,921 \times 10^6$	$0,248 \times 10^7$	$0,493 \times 10^7$	—	$0,881 \times 10^5$
10	$0,975 \times 10^6$	$0,131 \times 10^7$	—	—	—	—
11	$0,145 \times 10^7$	$0,111 \times 10^7$	$0,331 \times 10^7$	$0,453 \times 10^7$	—	—
Estado	$0,117 \times 10^8$	$0,133 \times 10^8$	$0,274 \times 10^8$	$0,347 \times 10^7$	$0,187 \times 10^8$	$0,292 \times 10^8$

TABELA 6. Continuação.

DIRA	Estrato 4		Estrato 5		Estrato 6		Total
	Grupo 1	Grupo 2	Grupo 1	Grupo 2	Grupo 1	Grupo 2	
2	—	$0,362 \times 10^7$	$0,540 \times 10^8$	—	—	—	$0,113 \times 10^9$
3	—	—	—	—	—	—	$0,166 \times 10^7$
4	$0,421 \times 10^8$	$0,185 \times 10^8$	$0,396 \times 10^9$	—	$0,342 \times 10^{10}$	$0,354 \times 10^9$	$0,337 \times 10^{10}$
5	$0,292 \times 10^8$	$0,347 \times 10^8$	$0,301 \times 10^9$	$0,994 \times 10^8$	$0,111 \times 10^{11}$	$0,129 \times 10^{11}$	$0,124 \times 10^{11}$
6	$0,309 \times 10^8$	$0,249 \times 10^8$	$0,242 \times 10^9$	$0,229 \times 10^9$	$0,133 \times 10^{11}$	$0,133 \times 10^{10}$	$0,130 \times 10^{11}$
7	—	$0,741 \times 10^7$	$0,225 \times 10^8$	—	—	—	$0,212 \times 10^{11}$
8	$0,248 \times 10^8$	$0,349 \times 10^8$	$0,239 \times 10^9$	$0,237 \times 10^9$	$0,147 \times 10^9$	$0,859 \times 10^8$	$0,406 \times 10^9$
9	$0,402 \times 10^8$	—	—	—	—	—	$0,355 \times 10^8$
10	—	—	—	—	—	—	$0,112 \times 10^7$
11	—	—	—	—	—	—	$0,643 \times 10^7$
Estado	$0,167 \times 10^9$	$0,124 \times 10^9$	$0,125 \times 10^{10}$	$0,566 \times 10^9$	$0,280 \times 10^{11}$	$0,246 \times 10^{11}$	$0,507 \times 10^{11}$

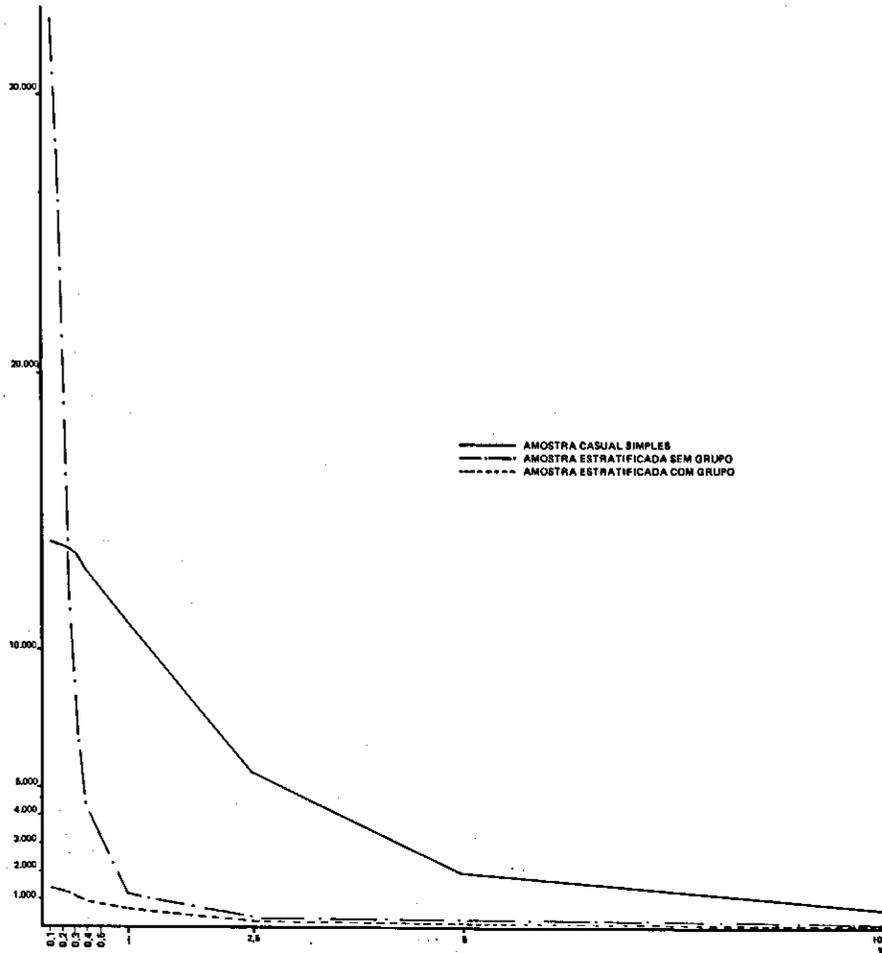


FIG. 2. Tamanhos de amostra obtidos.

TABELA 7. Tamanhos de amostra obtidos em função dos vários erros de amostragem.

Erro de amostragem (%)	Amostra casual simples	Amostra estratificada	
		Sem grupo	Com grupo
10,0	556	11	11
5,0	1.983	44	41
2,5	5.542	177	150
1,0	11.142	1.105	580
0,5	13.021	4.418	985
0,4	13.291	6.903	1.075
0,3	13.508	12.270	1.158
0,2	13.668	27.596	1.225
0,1	13.765	110.112	1.252

balho presta-se, assim mesmo, como base para essa revisão e a amostra aqui calculada presta-se como amostra de transição.

Uma crítica à estratificação utilizada é a presença de caselas vazias, isto é, estratos sem elementos.

CONCLUSÕES

1. A estratificação mostrou-se eficiente em relação à amostra casual simples.

2. A estratificação por grupo só foi eficiente com erro de amostragem menor ou igual a 1%. Recomendou-se uma alternativa de amostra com 583 elementos e com erro de amostragem de 1% para a variável número total de pés plantados com citros. Sugeriu-se a atualização do cadastro.

TABELA 8. Tamanho da amostra estratificada com 1% de erro de amostragem da estimativa do número total de pés de citros dentro da DIRA, do estrato e do grupo no Estado de São Paulo.

DIRA	Estrato												Total
	1		2		3		4		5		6		
	G1	G2	G1	G2	G1	G2	G1	G2	G1	G2	G1	G2	
São Paulo	3	2	—	—	—	—	—	—	—	—	—	—	5
Vale do Paraíba	2	2	—	—	—	—	—	—	—	—	—	—	4
Sorocaba	2	3	—	—	—	—	—	—	—	—	—	—	5
Campinas	22	23	12	10	3	4	6	7	10	7	17	24	145
Ribeirão Preto	34	56	26	30	9	11	16	13	22	15	40	33	305
Bauru	2	2	—	—	—	—	—	—	—	—	—	—	4
São José do Rio Preto	20	24	11	11	3	4	4	5	6	4	2	2	96
Araçatuba	—	2	—	—	—	—	2	—	—	—	—	—	4
Presidente Prudente	2	2	—	—	—	—	—	—	—	—	—	—	4
Marília	3	4	2	2	—	—	—	—	—	—	—	—	11
Estado	90	120	51	53	15	19	28	25	38	26	59	59	583

Cabem aqui algumas considerações sobre o cadastro utilizado. É óbvio que, num levantamento periódico, quanto mais rapidamente um cadastro for atualizado, tanto melhores serão os resultados. Na impossibilidade de atualização mais freqüente, num caso como este da citricultura, é recomendável uma a cada cinco anos ou sempre que se notar alteração mensurável. É o caso presente, em que o cadastro de 1980/81 já não reflete totalmente a realidade de 1984/85, sendo recomendável sua atualização antes do sorteio da amostra e a instalação do levantamento. Entretanto, o presente tra-

REFERÊNCIAS

CAMPOS, H. & PIVA, L.H.O. Dimensionamento de amostra para previsão e estimativa de safra no Estado de São Paulo. *Agric. SP*, 21(3):65-88, 1974.

COCHRAN, W.G. *Sampling techniques*. New York, J. Wiley, 1953.

KISH, L. *Survey sampling*. New York, J. Wiley, 1965.

SEVER, F.A.A. & AMARO, A.A. A citricultura em São Paulo; uma análise estrutural. s.l., s.e.d., 1984. Trabalho apresentado no "International Citrus Congress", São Paulo, SP, 1984.