

OBTENÇÃO DE ESTATÍSTICAS AGRÍCOLAS PELO MÉTODO DE AMOSTRAGEM: experiências visando a criação de uma organização permanente¹

Salomão Schattan²

APRESENTAÇÃO

Trata-se aqui da republicação de um artigo³ altamente relevante, mas difícil de ser encontrado, mesmo em bibliotecas, no ano em que a introdução do método de amostragem nos levantamentos agrícolas do Estado de São Paulo completa meio século. É nesse artigo que técnicos e pesquisadores do Instituto de Economia Agrícola (IEA) têm iniciado seus conhecimentos sobre o assunto durante as últimas cinco décadas e o leitor perceberá que os problemas dos levantamentos estatísticos continuam basicamente os mesmos. O estilo do autor foi mantido, tendo sido feitas somente uma revisão técnica e alguma atualização de forma aproximando-a das regras atuais de publicação da revista. Daí a atualidade do conteúdo. A republicação é oportuna não apenas pelo seu grande valor histórico, mas pela atualidade do assunto. O prefácio do estatístico inglês Wilfred Leslie Stevens, então professor da Universidade de São Paulo, enriquece ainda mais o valioso trabalho de Schattan. Espera-se que assim novas gerações de pesquisadores possam ter acesso a um pouco da sabedoria desses mestres.

PREFÁCIO

Wilfred L. Stevens

Durante as últimas décadas, tem havido verdadeira revolução nos métodos de obtenção de informações estatísticas necessárias à administração pública, à indústria e ao comércio. A nova técnica de levantamento por amostragem se desenvolve entre o método de recenseamento, cujos custos e demoras são inaceitáveis nos dias atuais, e o método da estimação subjetiva, cujas estimativas são inseguras e passíveis de controvérsias. A fim de contornar essas dificuldades, a estatística moderna criou métodos capazes de fornecer estimativas cujos limites de erro podem ser objetivamente avaliados, sendo limitados os custos que este método envolve.

Esses progressos se registraram, especialmente, na Commonwealth Britânica e nos Estados Unidos, tendo sido fortemente estimulados pelas exigências da segunda guerra mundial. É, agora, evidente que se vai presenciar, nos anos vindouros, uma aplicação crescente do novo método, tanto no Brasil como nos outros países da América Latina. A grande obra

¹Este artigo foi resgatado e adaptado, para se aproximar o mais possível do padrão atual de Agricultura em São Paulo, pelo Pesquisador Científico Francisco Alberto Pino, a quem o Comitê Editorial agradece o empenho e a dedicação.

²Engenheiro Agrônomo, Pesquisador da Sub-Divisão de Economia Rural.

³SCHATTAN, S. **Obtenção de estatísticas agrícolas pelo método de amostragem**. São Paulo: SA/Departamento da Produção Vegetal/Divisão de Economia Rural/Sub-Divisão de Economia Rural, set. 1953. 39p. (Estudos de Economia Rural, n.7).

pioneira da Secretaria da Agricultura do Estado de São Paulo despertará, por isso, o interesse de todos os que se ocupam das questões de Estatística, no campo econômico.

Cabe-me apenas apontar algumas das lições que podem ser tiradas do trabalho apresentado pelo Engenheiro Agrônomo Salomão Schattan. Em primeiro lugar, observa-se que existe, largamente difundida no Brasil, a noção de que qualquer problema técnico pode ser melhor resolvido se for imitado a solução adotada num país estrangeiro – usualmente, os Estados Unidos. Mas, a verdade reside na verificação de que a melhor solução de um problema é sempre aquela que leva em conta todas as particularidades locais. Disso se conclui que a solução nunca pode ser copiada de outro país, onde as condições são necessariamente diferentes.

No decurso de um século, os Estados Unidos organizaram um sistema de estimativa e previsão de safras. Quando foram começadas as discussões sobre o processo a ser empregado em São Paulo também se tropeçou com a idéia de que o caminho correto poderia consistir na adaptação daquele sistema. Felizmente, foi possível estabelecer o princípio básico de que os métodos, a serem usados em nosso Estado, deveriam ser baseados no estudo minucioso das condições locais.

Ainda mais: note-se que não é questão de completar um estudo, montar um sistema e deixar o resto à rotina. Pelo contrário, o processo de criação e aperfeiçoamento é contínuo e sem fim, devendo o sistema responder, com precisão e segurança, a cada variação das condições, dos recursos e da importância relativa dos diversos assuntos que ocupam a atenção dos administradores.

Se tem sido evitado o engano de copiar os outros, deve-se lembrar aos que se propuserem aplicar novos métodos a conveniência de não copiarem, pura e simplesmente, o conteúdo deste trabalho, pois, dentro do próprio Brasil, as diferenças entre um Estado e outro são tão grandes, a ponto de fazer acreditar que, talvez, os métodos satisfatórios no Estado de São Paulo não sejam satisfatórios – ou mesmo possíveis, nos outros Estados. Assim, o trabalho do Eng. Agr. Salomão Schattan apresenta, não um modelo para imitação, mas a descrição de como o estatístico enfrenta e resolve seus problemas. A exposição inclui os princípios básicos do método de levantamento por amostragem e, em particular, chama a atenção para a necessidade imprescindível de se escolherem os elementos da amostra, **objetivamente ao acaso** (dentro das restrições impostas pela estrutura). Ressalta também que se deve fazer todo o possível para obter informações acerca de todos os elementos que constituem a amostra.

Evidentemente, o estatístico recomendará, em cada caso, o método que proporcione as informações exigidas pelo administrador, com a precisão necessária e o custo mínimo. Isso significa que a escolha do método tem que ser baseada no estudo da “contabilidade de custos” dos diversos estágios do trabalho. Como exemplo, seja considerada, a questão da escolha do “sistema de referência” (*frame*) para a seleção da amostra.

Decidiu-se usar como “sistema de referência” um cadastro das propriedades rurais do Estado de São Paulo, arquivado em cartões perfurados “I.B.M.”. Se fosse legítimo debitar ao levantamento o alto custo desse fichário, nossa decisão poderia ter sido outra. Mas, tendo o fichário outras utilidades alheias ao levantamento, só se pode debitar ao levantamento uma parte do seu custo total.

Outro exemplo é a decisão de aproveitar os Engenheiros Agrônomos Regionais nas visitas

às propriedades da amostra. De fato, não havia outra alternativa quando se iniciou o levantamento, mas, seria uma conclusão superficial supor de antemão que este método é o mais econômico, meramente pelo fato dele não necessitar novo pessoal. Deve-se observar que o uso exclusivo dos agrônomos regionais, para as visitas às fazendas, traz várias conseqüências:

- a) cerca de 50% dos Engenheiros Agrônomos Regionais não possuem condução própria, tendo sido gasta soma apreciável no aluguel de automóveis;
- b) da necessidade de distribuir o trabalho, mais ou menos, uniformemente entre os Engenheiros Agrônomos Regionais resulta certo inconveniente estatístico. Em conseqüência, a amostra tem estrutura especial, que não garante a máxima precisão de estimativa para determinado número de observações;
- c) os cálculos finais ficam consideravelmente aumentados, em conseqüência de semelhante estrutura de amostra.

Ora, essas conseqüências poderiam ter sido evitadas (em parte, pelo menos) se houvesse investigadores, independentes dos agrônomos regionais. Isso significa que uma investigação contabilística, possivelmente, mostraria que é mais econômico transferir uma parte ou todo trabalho de inquérito para turmas de investigadores independentes.

As deficiências do levantamento, tal como está organizado, resultam inteiramente das limitações de recursos. Observe-se, por exemplo, que as estimativas são baseadas nas informações fornecidas pelos fazendeiros. Embora se tenham elementos para afirmar que elas não divergem seriamente da realidade, é evidente que, no futuro, será necessário suplementar essas informações através de medições diretas. Igualmente, as previsões são baseadas nas previsões dos fazendeiros. Existem, contudo, métodos de previsão, mais objetivos, baseados na colheita de pequenas áreas e na observação de "campos testemunhas", mantidos sob rotação e cultivo permanentes.

É provável que esse trabalho especializado tenha que ser confiado a investigadores independentes.

Em conclusão, sou, porém, de opinião que o sucesso do trabalho, nos dois primeiros anos, foi impressionante.

É sabido que as repartições do Governo preferem, geralmente, ficar na rotina, tendo pouco desejo de investigar noções revolucionárias. Esta é mais uma razão para se louvar a iniciativa da Secretaria da Agricultura, por se ter proposto a criar um levantamento por amostragem sem precedentes no Brasil.

O sucesso conseguido até agora deve ser atribuído ao entusiasmo e dedicação daqueles que estão dirigindo o trabalho no centro, dos responsáveis pela apuração e cálculos, bem como dos Engenheiros Agrônomos Regionais que demonstraram notável perseverança, localizando e visitando as propriedades da amostra, freqüentemente em difíceis condições climáticas e de transporte.

Tabela 1 - Número de Propriedades Rurais no Estado de São Paulo, Classificação Segundo os Setores Agrícolas e Classes da Área (ha)

Setor	Até 2	De 3 a 9	De 10 a 29	De 30 a 99	De 100 a 299	De 300 a 999	De 1.000 a 2.999	De 3.000 a 9.999	De 10.000 a 29.999	Mais 30.000	Total
Araçatuba	577	886	4.726	4.151	1.280	528	238	75	15	2	12.478
Araraquara	497	633	1.947	2.136	984	429	96	15	1	-	6.738
Avaré	811	1.739	4.526	4.650	1.826	693	176	26	6	1	14.454
Bauru	724	882	3.376	3.439	1.271	487	115	20	-	1	10.315
Bebedouro	314	661	1.794	2.170	993	460	134	26	2	-	6.554
Bragança	4.652	5.969	5.349	2.665	650	198	12	2	-	-	19.497
Campinas	4.363	2.985	4.250	3.050	924	336	59	11	-	-	15.978
Catanduva	347	636	2.359	2.322	808	244	44	5	1	-	6.766
Itapetininga	1.668	2.586	5.016	4.397	2.029	673	156	49	11	4	16.789
Jaú	470	1.051	1.725	1.538	754	292	54	5	-	-	5.889
Marília	618	2.072	6.972	4.513	1.307	522	184	33	4	5	16.230
Piracicaba	2.000	3.416	4.893	3.181	955	365	77	10	1	-	14.898
Piraçununga	1.128	2.381	3.391	2.735	1.132	485	108	12	2	-	11.374
Presidente Prudente	824	1.867	7.827	6.006	1.738	631	257	77	20	4	19.251
Ribeirão Preto	971	1.913	3.074	3.270	1.892	964	259	41	3	-	12.387
Rio Preto	914	1.789	5.497	5.989	2.361	892	243	54	7	-	17.746
São Paulo ¹	18.675	8.296	8.759	4.385	1.004	261	66	16	3	-	41.465
Taubaté	2.828	3.596	4.756	4.436	1.924	708	142	34	1	1	18.426
Total	42.381	43.358	80.237	65.233	23.832	9.168	2.420	511	77	18	267.235

¹Foram excluídas as propriedades do Município da Capital e das Regiões Agrícolas de Miracatu, Registro e Santos.



Figura 1 - O Estado de São Paulo no Interior do Brasil.

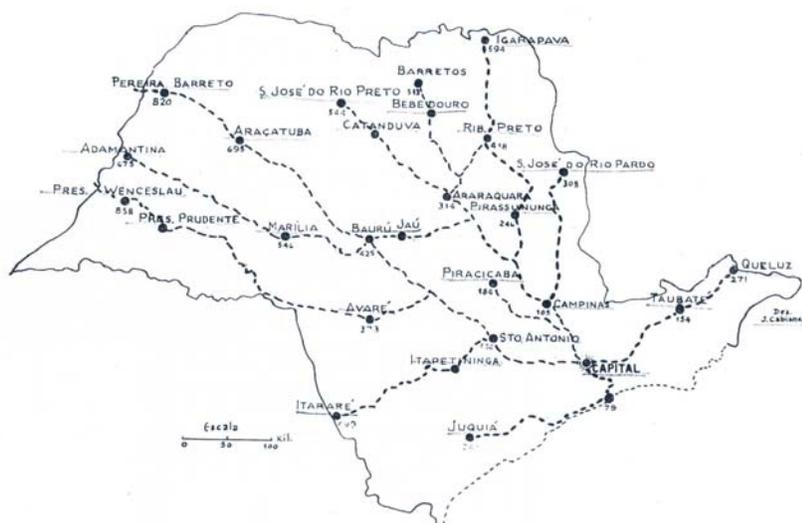


Figura 2 - Estado de São Paulo, Principais Estradas de Ferro, Distâncias da Capital em Quilômetros.

Tabela 2 - Área das Propriedades Rurais no Estado de São Paulo, Classificação Segundo os Setores Agrícolas e Classes da Área

Setor	(ha)										Mais 30.000	Total
	Até 2	De 3 a 9	De 10 a 29	De 30 a 99	De 100 a 299	De 300 a 999	De 1.000 a 2.999	De 3.000 a 9.999	De 10.000 a 29.999			
Araçatuba	492	5.583	95.070	220.834	212.558	283.395	402.870	374.671	250.355	67.189	1.913.024	
Araraquara	398	3.944	38.005	117.969	162.108	227.080	158.658	69.463	11.549	-	789.174	
Avaré	735	10.739	85.212	258.132	306.524	354.918	283.590	137.401	110.166	30.809	1.578.226	
Bauru	480	5.471	68.158	184.343	212.752	263.290	186.213	98.514	-	37.939	1.057.160	
Bebedouro	330	4.022	35.059	120.382	168.242	241.865	206.772	148.701	32.709	-	963.523	
Bragança	4.914	33.946	93.527	1.037.631	107.118	97.838	15.864	11.250	-	-	502.088	
Campinas	1.767	17.391	76.967	160.730	152.177	171.090	91.241	62.020	-	-	733.383	
Catanduba	289	4.007	45.520	125.667	133.624	121.199	68.346	21.949	18.166	-	538.767	
Itapetininga	957	16.409	93.972	257.518	341.690	341.510	256.248	274.796	178.496	212.462	1.974.058	
Jaú	496	6.417	31.580	84.501	128.809	146.882	85.949	30.926	-	-	515.560	
Marília	832	13.031	132.744	235.329	218.264	277.839	307.874	174.660	50.678	269.000	1.680.251	
Piracicaba	1.759	20.390	88.925	166.701	156.906	186.958	122.991	41.112	20.013	-	805.755	
Piracununga	1.133	13.874	62.206	148.691	194.824	250.313	172.171	49.063	21.474	-	913.749	
Presidente Prudente	772	11.898	153.246	315.185	289.314	337.778	434.927	381.041	273.918	351.985	2.550.054	
Ribeirão Preto	923	11.688	57.124	185.230	328.531	505.488	427.663	198.315	40.435	-	1.755.397	
Rio Preto	909	11.105	106.577	331.287	400.454	469.720	414.200	265.052	134.814	-	2.134.108	
São Paulo ¹	6.093	48.973	156.791	232.964	162.907	135.379	112.459	73.719	34.879	-	964.164	
Taubaté	2.436	21.089	90.339	248.188	331.940	352.587	239.499	180.005	19.726	57.981	1.543.790	
Total	25.705	259.977	1.511.019	4.008.742	3.531.282	4.765.129	3.992.976	2.592.658	1.197.378	1.027.365	22.912.231	

¹Foram excluídas as propriedades do Município da Capital e das Regiões Agrícolas de Miracatu, Registro e Santos.

1 - DESCRIÇÃO E CRÍTICA DO MÉTODO ANTERIOR

As estatísticas constituem, sem dúvida, importante elemento na orientação das atividades econômicas na sociedade moderna. Isto ainda é mais verdadeiro para o caso da agricultura, onde os preços e a produção se condicionam de forma estreita e onde a produção é levada a efeito por um sem número de empresas, espalhadas por todo o território estadual.

Não só os que se empenham diretamente em atividades agrícolas, mas, também, aqueles que desenvolvem atividades afins necessitam, em tempo oportuno, de informações estatísticas relativas aos diferentes aspectos da agricultura, para orientar suas atividades. O agricultor, o comerciante de produtos agrícolas, o diretor de estrada de ferro, os técnicos de repartições governamentais necessitam conhecer a situação da produção rural e de seus mercados a fim de programarem convenientemente a área a cultivar com cada artigo, os preços de compra e venda, a época e o número de vagões a serem destinados para casa zona, a política de amparo ou de restrições a determinados produtos ou grupo de produtores, etc.

Até o presente momento, as estatísticas oficiais brasileiras relativas à agricultura têm sido elaboradas com base no método de estimação subjetiva. As poucas estatísticas, que não são obtidas com base neste método, recorrem ao método censitário (enumeração total).

Isto ocorre, não somente no âmbito federal, como também no estadual.

No Estado de São Paulo, a Secretaria da Agricultura mantém, desde 1942, uma Seção de Previsão de Safras e Cadastro, que, a partir daquela data, começou a elaborar e a publicar regularmente estatísticas sobre a área cultivada e a produção dos principais artigos agrícolas. Como todas as estatísticas agrícolas do Brasil, as da Secretaria da Agricultura são feitas com base no método de estimação subjetiva. Por ser conveniente, descrever-se-á, em seguida, o mecanismo pelo qual são obtidas essas estatísticas.

Do ponto de vista jurídico e administrativo, o

Estado de São Paulo está dividido em 369 Municípios. Para efeito de organizar racionalmente seu trabalho a Secretaria da Agricultura agrupou esses Municípios em 153 unidades que se chamam Regiões Agrícolas. Cada uma dessas regiões está sob a responsabilidade de um Engenheiro Agrônomo, encarregado de dar orientação técnica aos agricultores de sua Região, bem como de distribuir adequadamente as facilidades proporcionadas pelo Governo aos agricultores, tais como sementes selecionadas, mudas de árvores frutíferas, produtos relacionados à lavoura (como torta de algodão, farelo de trigo), etc.

Por sua vez as 153 Regiões estão agrupadas em 19 Setores Agrícolas, cada um, sob a responsabilidade de um Engenheiro Agrônomo mais experimentado. Os Chefes de Setor são verdadeiramente os intermediários administrativos e científicos entre os Regionais e as repartições centrais da administração e de pesquisas científicas da Secretaria da Agricultura.

Para efeito de estatísticas, os Engenheiros Agrônomos Regionais, estribados no conhecimento de suas respectivas Regiões, **estimam** a área cultivada e a produção dos principais produtos agrícolas nos Municípios que constituem sua Região.

Seja para a área cultivada ou para a produção esperada e obtida de cada produto considerado, as estimativas para todo Estado se constituem da soma das estimativas feitas pelos Engenheiros Agrônomos Regionais.

Em cada ano agrícola são feitas, desta mesma forma, em seis meses consecutivos, seis estimativas de área e produção para cada artigo, sendo a última estimativa considerada como a área efetivamente plantada e a produção efetivamente colhida no Estado de São Paulo.

Em 1942, por ocasião da criação da Seção de Previsão de Safras e Cadastro, a adoção do método de estimação subjetiva, para a obtenção de estatísticas agrícolas, foi certamente o que de melhor se podia fazer com elementos existentes.

Se bem que o método de trabalho não houvesse mudado com relação ao que já vinha sendo feito, as estimativas e previsões calculadas pela Se-

ção de Previsão de Safras e Cadastro apresentaram a vantagem ponderável de serem fornecidas por técnicos de nível universitário, especializados em agricultura. Além disso, esses técnicos estavam em íntimo contato com os agricultores e a agricultura das Regiões pelas quais respondiam.

Entretanto, como não podia deixar de ser, o tempo incumbiu-se de trazer à tona as falhas das estatísticas assim elaboradas.

Um método de obtenção de estatísticas e os resultados obtidos por seu intermédio constituem um todo indivisível. Sabe-se que a eficiência de uma estatística está condicionada à veracidade e à segurança com que ela representa o fenômeno estudado; e, desde o momento em que se possa duvidar dos resultados, fica abalado todo o sistema, entrando então em discussão o próprio método de trabalho.

Ora, há hoje provas de que as estatísticas agrícolas obtidas pelo "método de estimação subjetiva" não representam de forma suficientemente precisa os fenômenos estudados. Este fato já é proverbial entre aqueles que trabalham com estatísticas ou delas se utilizam, pois há acentuada divergência entre estatísticas referentes ao mesmo assunto, mas fornecidas por fontes diferentes.

No Estado de São Paulo, as estatísticas agrícolas calculadas pelo Instituto Brasileiro de Geografia e Estatística (S.E.P.), que também usa o método de estimação subjetiva, divergem muito daquelas que foram confeccionadas pela Seção de Previsão de Safras e Cadastro da Secretaria da Agricultura, como bem o demonstra a tabela 3.

Esta grande divergência na estimação é observada com todos artigos, mesmo com aqueles que possuem certa base objetiva para o cálculo, como o algodão.

As estimativas de área cultivada com algodão deveriam ser muito próximas da área real porque a Secretaria da Agricultura, tendo o monopólio da venda de sementes de algodão no Estado, sabe com bastante precisão a quantidade de sementes vendidas para cada Município. Da quantidade de sementes vendida pode-se deduzir, com muita aproximação, a área cultivada.

Por outro lado, já é tradicional a cultura de algodão no Estado. Todo algodão produzido é vendido e beneficiado em pouco mais de duzentas máquinas de beneficiamento, sendo conhecida, com precisão, a produção total de algodão nos anos anteriores. Os dois elementos anteriormente citados, a área total cultivada e a produção total colhida nos anos anteriores, possibilitam um cálculo bastante objetivo do rendimento do algodão em caroço por unidade de área. Tendo baseado o cálculo da área cultivada e o rendimento por unidade de área em elementos objetivos, pode-se estimar a produção total fazendo o produto dessas duas quantidades.

Pois bem, apesar de todas essas bases objetivas de cálculo, as estimativas de produção de algodão elaboradas pelo Serviço de Estatística da Produção do Ministério da Agricultura divergem profundamente das da Secretaria da Agricultura, como pode ser visto na tabela 4.

Por outro lado, nota-se que, com exceção dos anos de 1945 e 1952, as estatísticas da Secretaria da Agricultura se aproximam razoavelmente da produção real de algodão em caroço entrado nas máquinas de beneficiamento.

Este fato se explica quando se sabe que a colheita é iniciada em março, sendo, portanto, conhecidas as entradas nas máquinas até maio, ocasião em que é feita a última estimativa. Isso permite correções sucessivas nas três últimas estimativas, pois há a possibilidade de um confronto muito bom entre o algodão já colhido e o por colher. Nos anos de 1948 a 1950 deve ter havido esse tipo de correções.

Com o exposto, fica provado que as estatísticas obtidas pelo método de estimação subjetiva são imprecisas. Mas, além da imprecisão das estatísticas obtidas por seu intermédio, pode-se acusá-lo de ter um campo de ação muito restrito. Ele não permite que com o seu auxílio sejam feitas pesquisas quer de natureza puramente econômica, quer de natureza sociológica.

Pode-se aquilatar adequadamente as deficiências e a incapacidade do método de estimação subjetiva, se se lembrar que para a solução dos problemas econômicos e sociológicos da agricultura não basta

Tabela 3 - Produção de Café, Arroz, Batata e Milho no Estado de São Paulo, Segundo Dados da Secretaria da Agricultura e IBGE. Comparação Percentual entre as duas Estimativas, tendo como Unidade as Estimativas da Secretaria da Agricultura

Ano	Arroz em casca (sc.60kg)			Café beneficiado (sc.60kg)		
	IBGE	Secretaria	Secretaria/ IBGE (%)	IBGE	Secretaria	Secretaria/ IBGE (%)
1943	10.495.870	10.266.290	2,2	8.187.379	9.466.717	-13,6
1944	8.242.577	9.993.067	-17,5	4.892.115	4.884.649	0,1
1945	11.298.900	11.538.651	-2,1	5.973.040	4.662.280	28,1
1946	15.816.253	12.825.799	23,3	7.399.050	7.893.346	-6,3
1947	13.974.266	10.275.346	36,0	7.365.383	7.717.198	-4,6
1948	13.877.500	8.938.616	55,2	9.152.316	11.018.172	-16,9
1949	13.031.650	9.437.827	38,1	8.124.966	8.012.582	1,4
1950	16.546.200	12.464.285	32,7	7.487.150	7.556.593	-1,0
1951	16.395.666	10.557.975	55,3	7.949.250	7.720.166	3,0

Ano	Batata (t)			Milho (sc.60kg)		
	IBGE	Secretaria	Secretaria/ IBGE (%)	IBGE	Secretaria	Secretaria/ IBGE (%)
1943	204.570	249.321	-18,0	18.175.000	20.480.418	-11,3
1944	111.021	198.097	-44,0	16.581.860	18.975.948	-12,7
1945	209.421	203.860	2,8	16.580.454	18.387.836	-9,8
1946	185.561	192.000	-3,4	21.845.145	26.634.000	-18,0
1947	232.793	261.212	-11,0	20.279.666	19.629.782	3,3
1948	196.405	202.522	-3,1	19.155.233	18.025.975	6,3
1949	275.482	225.079	22,4	18.297.833	17.088.704	7,1
1950	242.133	221.078	14,7	21.040.850	20.446.921	2,9
1951	254.733	132.019	92,9	20.824.700	17.924.799	16,2

Tabela 4 - Confronto entre as Estimativas da Produção do Algodão, nos Anos de 1943 a 1952, feitas pela Secretaria da Agricultura e o IBGE, e as Entradas de Algodão em Caroço nas Máquinas de Beneficiamento

Ano	Estimativas da Secretaria		Estimativa do IBGE	Entradas nas máquinas (t)	Secretaria/ IBGE (%)	Secretaria/ entradas nas máquinas (%)
	Final (t)	Março (t)				
1943	1.046.511	-	1.250.327	1.089.450	19,5	-4,0
1944	1.212.414	-	1.315.668	1.315.668	13,5	-7,9
1945	885.077	-	737.997	639.915	-12,6	38,3
1946	517.905	-	718.183	486.411	38,7	6,5
1947	497.109	-	577.022	491.556	16,1	-2,3
1948	438.925	522.615	518.839	416.694	18,2	5,3
1949	598.460	600.258	690.925	529.322	15,4	-5,0
1950	486.601	674.950	610.053	460.467	25,4	-5,7
1951	619.488	661.412	592.884	633.402	-4,3	-2,2
1952	863.633	833.182	-	980.000	-	-11,9

conhecer a área cultivada e a produção dos artigos mais importantes. É necessário, além disso, pesquisar um grande número de fatores, e o método de estimação subjetiva não permite pesquisas. Como exemplo de pesquisas necessárias no campo econômico citam-se as técnicas de produção dos diferentes artigos e os respectivos custos, a produtividade relativa das propriedades de diferentes tamanhos, a forma de emprego de adubos e inseticidas, etc. Entre os

problemas sociológicos cujo conhecimento é indispensável, citam-se a mobilidade das populações rurais e a sua relação com o tamanho da propriedade, para o estudo e o combate do êxodo rural; a adaptação dos diferentes grupos étnicos aos climas brasileiros, para uma atividade agrícola produtiva, para a definição de uma política imigratória; os tipos de relações de trabalho entre empresários e empregos, etc.

Foi, sem dúvida, a constatação da inconsistência das estatísticas obtidas por este processo e a sua incapacidade de fornecer elementos para a análise dos problemas econômicos e sociológicos mais complexos da agricultura que obrigou os órgãos técnicos oficiais, a partir de 1948, a encarar a possibilidade de introduzir o uso do método de amostragem em substituição ao antigo método.

2 - O QUE É A AMOSTRAGEM E EM QUE CONSISTE

O processo de amostragem consiste em se investigar as características de um todo através do estudo das características correspondentes de uma sua parte.

Trata-se de um processo historicamente estabelecido que se iniciou com populações muito uniformes e que, sistematizado à medida que se tornava mais complexo o campo abarcado, passou a constituir um ramo especial da estatística.

Uma boa idéia do que seja uma amostra é proporcionada pelos seguintes exemplos: uma moeda de um cruzeiro tirada de uma urna contendo mil moedas de um cruzeiro; um saco de cereal e um punhado de cereal tirado de forma adequada deste saco; o sangue, que circula nas veias de um homem e uma gota deste sangue. Este são exemplos de amostras em populações uniformes.

Há, entretanto, outros casos, mais freqüentes do que os acima citados, onde a população não é uniforme, isto é, onde ela é constituída de elementos sensivelmente desiguais.

Ao contrário do que sucede com as populações homogêneas, o estudo de populações heterogêneas por meio de amostragem exige técnica especial para a retirada da amostra, a fim de que, apesar da heterogeneidade, se possam inferir, para a população, os conhecimentos proporcionados pela amostra.

Métodos seguros para a seleção de amostras das populações em geral só foram criados há, relativamente, pouco tempo e o valor prático desses métodos ficou enormemente aumentado desde o mo-

mento em que se pôde fazer a determinação dos erros da amostragem e da eficiência das amostras, com base nos dados fornecidos pela própria amostra.

Devido às inúmeras vantagens que apresenta, tem sido cada vez mais freqüente o uso da amostragem nos mais diferentes campos de pesquisas. Nos estudos sociais e na auscultação da opinião pública, a amostragem tem proporcionado resultados espetaculares; na indústria, ela é usada para o controle da qualidade; sua aplicação é hoje, imprescindível em todos os ramos da pesquisa científica.

Na agricultura, além das aplicações usuais em pesquisas agrônômicas, a amostragem é particularmente indicada na obtenção de estimativas e previsões da produção dos diferentes artigos agrícolas, na pesquisa das características econômicas da produção agrícola e na pesquisa dos problemas sociológicos das populações rurais.

Se se puser de lado o método de estimação subjetiva, na elaboração de estatísticas, só restarão para tal fim o censo e a amostragem. Há quem afirme que só se deve aplicar a amostragem como último recurso, porque na opinião deles o censo proporciona sempre resultados mais seguros do que a amostragem. Entretanto, a verdade é bem diversa. Na agricultura, onde a produção se processa num grande número de propriedades, distantes umas das outras e de custoso acesso a partir da Sede de um Município, fica extremamente cara, difícil e morosa a organização e a execução de censos. A prática dos censos anteriores confirma plenamente o que ficou dito. Além disso, no censo, a fim de que o trabalho seja executado em curto prazo, há necessidade de recrutar grande número de pessoas da cidade, que não têm a menor idéia das questões concernentes à agricultura e não possuem conhecimento nem habilidade para conseguir respostas adequadas do lavrador que, em todos países do mundo, é em geral desconfiado. Estes problemas de execução do censo introduzem, em seus resultados, um erro que não pode ser calculado.

Os "estatísticos", adeptos de censos anuais de nossa agricultura, além do exemplo de alguns Esta-

dos americanos, baseiam sua argumentação no fato de que a amostragem nem sempre elimina totalmente a necessidade dos censos, e ao compararem os resultados da amostragem com os do censo partem da suposição de que os resultados do censo são perfeitos. A verdade é que os resultados do censo não servem em geral para verificar os dados obtidos pela amostragem; sua principal utilidade está em proporcionar um “sistema de referência” para a amostragem, pois as fichas do censo passam a constituir um rol adequado das unidades da população a ser estudada. Quanto à perfeição dos resultados do censo, basta lembrar o Censo Agrícola de 1946 na França. Uma pesquisa levada a efeito para verificar o valor dos dados obtidos provou que eles absolutamente não representavam a realidade, e por conseguinte não poderiam ser utilizados.

A fim de obter estimativas baratas, precisas e em tempo hábil, é preciso, pois, recorrer à amostragem.

3 - HISTÓRICO E ENSAIOS PRELIMINARES EM SÃO PAULO

Os estudos, visando ao aperfeiçoamento do sistema de previsão e estimação da área cultivada e da correspondente produção de artigos agrícolas, iniciaram-se na Subdivisão de Economia Rural da Secretaria da Agricultura com dois trabalhos⁴. O primeiro destes trabalhos encarece a necessidade de rever o método de trabalho e o esquema das estatísticas elaboradas até então; o segundo analisa a possibilidade de se adaptar às nossas condições o sistema de amostragem em uso no Agricultural Estimating and Reporting Services, do Departamento de Agricultura dos Estados Unidos.

Contudo, a execução dos trabalhos propriamente ditos só se iniciaram em 1950 com uma pesquisa preliminar que se descreve a seguir.

3.1 - Levantamento Piloto

Em 1950, influenciados ainda pelo serviço norte-americano de previsões de safras, procurou-se verificar como funcionaria aqui um sistema de levantamentos de estatísticas agrícolas por amostragem, com o envio dos Questionários pelo Correio, diretamente às propriedades da amostra. As respostas seriam voluntárias, pois o próprio agricultor se encarregaria de devolver os Questionários preenchidos à repartição responsável.

Esta pesquisa foi feita no ano agrícola de 1950/1951, e teve como sede a Região Agrícola de Campinas, constituída dos Municípios de Campinas, Indaiatuba e Monte-Mor. A Região Agrícola de Campinas, que não pode ser considerada típica do Estado, foi, contudo, escolhida para sede do Levantamento Piloto devido a sua agricultura diferenciada, proximidade da Capital, existência de boas estradas inter-municipais, etc. Além disso, trata-se de uma Região situada na zona velha do Estado onde já existe certa estabilidade da agricultura e dos agricultores.

Eis a descrição do Levantamento e em conclusão, os principais ensinamentos dele tirados: o “sistema de referência” para a retirada da amostra das propriedades rurais da Região Agrícola de Campinas se constituía das fichas Keysort confeccionadas segundo a última lista de cobrança do Imposto Territorial Rural dos três Municípios da Região.

As fichas dos três Municípios, tomadas como um todo único, foram colocadas em ordem crescente de área, fazendo-se em seguida sua divisão em cinco grupos. Para esses grupos, correspondendo as cinco classes de área, adotaram-se frações variáveis de amostragem. Às classes mais amplas, contando um menor número de propriedades de área maior, corresponderam frações de amostragem maiores. Tanto os limites das classes de área, como as frações de amostragem, foram escolhidas de forma inteiramente arbitrária.

Em seguida, com o auxílio de uma tabela especial de números ao acaso, selecionou-se, de cada um dos cinco grupos, a quantidade adequada de

⁴Auto n. 83.353 de 7 de junho de 1948. Eng. Agr. Mário Decourt Homem de Mello. Eng. Agr. Salomão Schattan.

propriedades que iriam constituir a amostra.

A cada uma das propriedades sorteadas foi endereçada uma carta onde se explicava, em linguagem acessível, as razões e a necessidade da pesquisa, o uso que se faria das informações e o seu caráter sigiloso, bem como o grande valor atribuído às respostas. Em anexo seguia um Questionário, contendo os elementos de identificação da propriedade para a qual era enviado. Ao Engenheiro Agrônomo Regional de Campinas foram entregues as cartas destinadas aos agricultores do Município de Campinas que faziam parte da amostra; o Prefeito de Indaiatuba e o Fiscal de máquinas de algodão de Monte-Mor receberam as cartas contendo os Questionários destinados a estes dois Municípios. Em cada um desses três Municípios, parte da correspondência foi entregue em domicílio aos proprietários que moravam na cidade; outra parte foi enviada através de ônibus que fazem as ligações inter-municipais e uma última parte foi depositada no Correio.

Eis, em seguida, os dados relativos ao recebimento dos Questionários no levantamento da Região de Campinas.

Com se verifica pela tabela 5, depois de um mês, só se receberam 20% mais ou menos do total de Questionários enviados e, ao final do inquérito, contou-se com um total de apenas 75 Questionários respondidos.

Tabela 5 - Recebimento dos Questionários no Levantamento Piloto na Região de Campinas

Número de propriedades com área > 5ha	2.490
Número de questionários enviados	269
Número de questionários recebidos em um mês	50
Número final de respostas	75
Número de visitas feitas a uma sub-amostra das que não responderam	10

Apesar de baixa a proporção dos Questionários devolvidos, cerca de 30%, acreditou-se que essa porcentagem não podia ser considerada típica para todo Estado porque, na pesquisa, ocorreu um fato especial. Em Monte-Mor, o Fiscal de máquinas de algodão, que é funcionário da Divisão de Economia

Rural, tendo-se interessado de modo especial pelo serviço, conseguiu 36 respostas (cerca de 50% do total recebido) dos 42 Questionários enviados para o seu Município.

Desde que a amostra, considerada como representativa da população das propriedades rurais da Região Agrícola de Campinas, se constituía de 269 propriedades, é evidente que o conjunto dos 75 Questionários recebidos não podia ser igualmente tomado como amostra ao acaso da mesma população. A utilização das informações contidas nos 75 Questionários recebidos ficou condicionada à obtenção dos dados relativos às propriedades que não responderam. A coleta desses dados complementares foi feita sorteando ao acaso dez propriedades entre as 195 que não haviam devolvido seus Questionários e visitando-as em seguida.

Em conseqüência do raciocínio anterior conclui-se: mesmo que com o correr do tempo se elevasse bastante a porcentagem de respostas e ao invés de 30%, após 2 meses, se obtivessem 50% ou 70% em um mês, seria ainda necessário fazer visitas a uma sub-amostra. A necessidade de tais visitas eliminou a maior vantagem de um levantamento por correspondência.

As informações obtidas no Levantamento Piloto e mais alguns dados do Censo Agrícola de 1950 relativos às propriedades agrícolas recenseadas em dez Municípios do Estado de São Paulo, gentilmente cedidos pelo Instituto Brasileiro de Geografia e Estatística, deram a idéia de que uma amostra ao acaso, com 1.500 propriedades, das quais se obtivesse 100% de respostas seria capaz de proporcionar estatísticas aceitáveis para os principais produtos em todo Estado.

Deve-se notar que, nas condições do Estado de São Paulo, para obter informações de todas as propriedades da amostra, seria necessário adotar como técnica de obtenção de respostas, a visita direta às propriedades e entrevistas com o responsável pela exploração.

Se se comparar o trabalho necessário à obtenção de estatísticas por amostragem, quando se adota como técnica de obtenção de respostas o Correio ou a visita direta, conclui-se em favor de amostra pe-

pequena e visitas diretas, pois se o tamanho da amostra por correspondência equivalesse a 100% da população, o trabalho de escritório necessário para ela seria muitas vezes maior do que para a amostra de 0,5%. O número de visitas a todas propriedades da amostra pequena seria também pouco superior ao das visitas à sub-amostra da amostra por correspondência. Em realidade, as visitas a sub-amostra seriam muito mais difíceis porque em cada levantamento, seria preciso localizar novas propriedades enquanto que a amostra pequena é fixa, pelo menos, durante toda uma safra.

Algumas comparações entre amostragem com respostas por correspondência e por entrevistas diretas podem ser vistas na tabela 6.

Desde este momento, não houve mais dúvidas. Aceitou-se a pequena amostra, pois ela proporcionaria, com menos trabalho e custo, resultados mais precisos e menos demorados do que a amostra por correspondência.

E ainda mais: na dependência da capacidade técnica dos entrevistadores da pequena amostra, era possível efetuar pesquisas sobre os itens mais variáveis, relativos, não só à parte econômica, mais, também, a problemas sociológicos da população rural.

Considerando todos os fatos acima citados, ficou assentado que, para o ano agrícola de 1951-1952, seria tirada uma amostra para o Estado todo, com 0,5% da população da qual se procurariam obter 100% das respostas. As entrevistas com os lavradores ficariam a cargo dos Engenheiros Agrônomos Regionais.

3.2 - Construção da Amostra do Ano Agrícola de 1951-1952

Para a retirada da amostra no Estado todo, foi utilizado como **sistema de referência** o Cadastro das propriedades rurais, organizado em 1947 na Seção de Previsão de Safras e Cadastro da Divisão de Economia Rural. Sua descrição e estudo serão feitos mais tarde.

Para fins de amostragem, a população das

propriedades rurais pode, sem nenhum inconveniente, ser estratificada em dois sentidos - geograficamente e segundo um critério de tamanho.

Quando a amostra constitui somente uma pequena porcentagem da população, o número de propriedades dentro de cada substrato da amostra será, forçosamente, muito pequeno. Todavia, para se obterem estimativas do desvio padrão, são necessários, ao menos, dois elementos de cada substrato. Conseqüentemente, a fim de se tirar o máximo proveito de uma estratificação bastante detalhada, é preciso encerrar a possibilidade de construir uma amostra contendo exatamente dois elementos de cada substrato.

Devendo as frações de amostragem ser estabelecidas de forma a proporcionar o máximo de eficiência de estimação, está claro que o tamanho dos substratos deveria ser ajustado de forma a se obterem as frações necessárias quando dois elementos são tirados de cada substrato. Teoricamente, tal ajustamento é possível - poder-se-ia, diga-se, manter fixos os limites das classes de área, variando os limites geográficos dentro de cada classe de área, de forma a obter a fração de amostragem exata, quando são tirados dois elementos de cada substrato. Está claro, porém, que tal processo seria por demais trabalhoso para ser prático, e isto conduz ao problema da determinação da perda de eficiência, quando não se obtém a fração de amostragem exata. Esta investigação⁵ mostrou que a perda é surpreendentemente pequena, mesmo quando a fração de amostragem atual diverge da fração ótima por um fator de quatro ou cinco vezes. Este fato abre caminho para a solução prática do problema, que será ilustrada a seguir.

3.2.1 - Determinação dos estratos geográficos e dos estratos dimensionais

Era natural que, por conveniência administra-

⁵Stevens, W.L. **Estimativa e previsão de safras através de um levantamento por amostragem**. São Paulo, Secretaria da Agricultura, Divisão de Economia Rural, nov. 1951. 13p. (Estudos de Economia Rural).

Nota: no original consta Stevens (1952), mas acredita-se que a referência correta seja esta apresentada acima.

Tabela 6 - Comparação entre Amostragem com Respostas por Correspondência e por Entrevistas Diretas

	Amostra por entrevista direta	Amostra por correspondência
Correspondência e questionários a enviar	A 0,5% da população	A 10% da população
Cálculos a efetuar	Com 100% dos questionários enviados	Com aproximadamente 50% dos questionários enviados
Visitas	A 0,5% da população, fixo em todos levantamentos da mesma safra	A uma sub-amostra variável em cada levantamento
Demora	30 dias, no máximo	60 dias, pelo menos
Informações possíveis	Além dos elementos econômicos básicos, podem-se efetuar pesquisas econômicas e sociológicas	Elementos econômicos mais simples

tiva, se adotasse a Região como estrato geográfico, pois, não se poderia adotar o Município porque resultaria numa amostra grande demais para os nossos recursos. Como o trabalho de campo deveria ser feito pelos Engenheiros Agrônomos Regionais, foi decidido que, no primeiro ano, a amostra seria distribuída uniformemente entre as Regiões. Está claro que isso não proporcionará obrigatoriamente as frações ótimas de amostragem para os principais produtos, mas, também, não é provável que seja tão incorreta quanto a sua natureza arbitrária poderia sugerir.

Para a estratificação dimensional, podem-se sugerir vários critérios - área total das propriedades, área cultivada, valor da propriedade, etc. Estes e outros **itens** informativos são encontrados na "declaração do contribuinte" preenchida pelo proprietário. Entretanto, só a área total aparece nos róis, sendo por isso a base obrigatória da classificação dimensional. Não há razão para supor que a área total não seja um critério eficiente para a classificação dimensional. É bem verdade que se houvesse interesse somente por produtos cultivados, a área total cultivada seria preferível.

Mas, cumpria obter outras informações, incluindo a área em mato e em pasto. Para um conjunto de dados tão divergentes, a área total da propriedade é um critério tão bom quanto qualquer outro.

Os dados para a determinação do desvio padrão dentro das classes foram obtidos do "Levantamento Piloto", anteriormente, descrito. Dados suplementares do Censo de 1950 foram gentilmente cedidos pelo Instituto Brasileiro de Geografia e Estatística.

Foram determinadas cinco classes de área com base nos dados existentes para alguns produtos importantes. A determinação foi feita de forma a que, para todo o Estado, o número de propriedades dentro de cada classe fosse inversamente proporcional ao desvio padrão dentro da classe. Foram achados os seguintes limites de classe:

5 ... 50 ... 150 ... 400 ... 1.000 ... e acima (hectares)

As propriedades com menos de cinco hectares não foram incluídas na amostra por se achar que, nelas, a produção dos artigos que deveriam ser estudados era de pouco ou nenhuma significação comercial.

Pode-se ter uma idéia mais clara da estratificação geográfica e dimensional adotada se se tomar como exemplo a tabela 7, de dupla entrada, referente a um dos 18 Setores Agrícolas em que se divide o Estado.

Cada uma das células desta tabela é um substrato formado pela interseção de uma Região com uma classe de área. Ao substrato assim definido, corresponde sempre um conjunto de fichas no Cadastro Keysort.

Para cada Região Agrícola, foram tirados ao acaso duas propriedades de cada substrato, independentemente do número total de propriedades do substrato. Desta forma, cada Engenheiro Agrônomo Regional deveria visitar, exatamente, dez propriedades, ficando a amostra para todo Estado constituída de 1.400 propriedades.

Evidentemente, o processo descrito é um tanto rudimentar. Os dados usados na determinação do desvio padrão cingiram-se a uns poucos Municípios,

Tabela 7 - Esquema da Estratificação Adotada, Setor Agrícola de Catanduva

Regiões Agrícolas	Classes de área em hectare					Número total propriedades da Região
	5 a 49	50 a 149	150 a 399	400 a 999	Acima de 1.000	
Catanduva	1.670	475	137	38	21	2.341
Novo Horizonte	740	270	96	29	19	1.154
Santa Adélia	632	219	89	15	3	958
Taquaritinga	562	244	83	17	6	912
Uchoa	695	301	93	31	0	1.120

não sendo levadas em consideração as grandes variações da distribuição por área de uma Região para outra. Entretanto, deve-se ter em mente que não preocupou muito obter uma amostra eficiente desde o primeiro ano. O trabalho, no primeiro ano, é um guia, destinado a proporcionar experiência prática, e os dados para a determinação dos desvios padrões.

Tratando-se da primeira vez que os Agrônomos Regionais da Secretaria da Agricultura iriam fazer visitas aos lavradores para obter informações estatísticas, impunha-se uma série de esclarecimentos. Esses esclarecimentos iniciais foram prestados por meio de uma carta e em folhas de instruções onde eram estudadas detalhadamente as dúvidas que poderiam surgir na localização das propriedades da amostra ou no preenchimento do Questionário.

Finalmente, a entrega do material referente ao primeiro levantamento em todo Estado foi feito pessoalmente aos Engenheiros Agrônomos Regionais em setembro de 1951. Para esse fim, foram promovidas reuniões que se realizaram nas cidades de São Paulo, Itapetininga, Campinas, Bebedouro, Taubaté, Lins e Lucélia. Elas tiveram como objetivo explicar àqueles que estavam incumbidos do trabalho de campo as razões da necessidade da mudança das bases do sistema de previsão de safras, a grande importância que tem para a economia dos agricultores uma boa estimativa ou previsão de safras e, finalmente, dar-lhes idéia clara e precisa do que é a amostragem e a organização do trabalho cuja execução se iniciava.

3.2.2 - Resultados obtidos com a amostra de 1951-1952

Eis, a seguir, a tabela 8 com os resultados

obtidos no Levantamento de setembro, bem como o cálculo dos respectivos desvios padrões:

Tabela 8 - Estimativas, Estado de São Paulo (excluído Setor de Araçatuba)

Cultura	Área (mil hectares)	Desvio padrão (%)
Algodão	1.328	8,7
Amendoim	50	11,0
Arroz	465	6,8
Milho	1.365	7,3
Café	1.144 milhões de pés	5,8

O setor de Araçatuba foi excluído dos cálculos porque, neste primeiro levantamento, ele só apresentou um Questionário respondido.

Os erros padrões foram calculados segundo a hipótese de haver 100% de respostas e, como se vê, em todas as culturas, eles ultrapassam o limite aceitável.

Entretanto, para um balanço real dos resultados obtidos tem-se que tomar em consideração os grandes defeitos da primeira amostra, que foi tirada sem se possuírem os conhecimentos essenciais da estrutura da população, pois, os dados em mãos eram insuficientes. A estratificação, embora baseada nos dados do recenseamento, não foi tão eficiente como deveria ter sido e o número de respostas nunca chegou muito perto de 100% porque a amostra foi tirada de um cadastro que, confeccionado em 1947 com informações de 1946, não havia sido atualizado até 1950, apresentando por isso grandes dificuldades na localização das propriedades da amostra.

4 - LEVANTAMENTO EXPERIMENTAL EM 1952-1953

Os dois anos de trabalho preparatório, isto é,

o levantamento Piloto de Campinas e o levantamento de 1951-1952 em todo Estado proporcionaram os elementos indispensáveis para aperfeiçoar, de forma decisiva, nossa amostra. No decorrer dos trabalhos, tornou-se clara a necessidade de "um sistema de referência" atualizado, uma nova estratificação, cálculo de frações de amostragem mais eficientes, etc.

Foram introduzidos, na amostra de 1952-1953, todos os melhoramentos que o trabalho feito até então indicou como necessários. É por isso que se torna interessante descrever, em detalhe, todas as características fundamentais da amostra tirada neste ano e discutir alguns resultados com ela já obtidos.

4.1 - População

Para tirar uma amostra, é necessário, em primeiro lugar, definir a população. Por sua vez, a definição da população está condicionada aos objetivos da amostragem.

A Divisão de Economia Rural precisa conhecer, em tempo oportuno, a produção dos principais artigos agrícolas e as características econômicas mais importantes da agricultura em São Paulo. Desde que se proponha obter tais informações por meio da amostragem, é preciso, automaticamente, considerar nossa população como situada nos limites dos 243.000km² ocupados pelo Estado.

Excluindo as áreas que não se podem destinar à agricultura, como as das cidades, das estradas e dos rios, e as áreas abandonadas, como as terras devolutas pertencentes ao Estado e aos Municípios, resta, dentro dos limites geográficos do Estado de São Paulo, uma área que está dividida em cerca de 300.000 propriedades.

O conjunto dessas propriedades será a população e sua definição se fará por meio do Cadastro organizado com base nas informações prestadas pelos proprietários, por ocasião do pagamento do Imposto Territorial Rural.

Anualmente, a Secretaria da Fazenda do Estado de São Paulo confecciona uma lista de todas as propriedades rurais do Estado. Nessa lista, consta o

nome do proprietário, o bairro em que está situado o imóvel e o nome da propriedade, a área em hectare e o valor venal da propriedade.

Cada uma das 369 Coletorias Municipais confecciona o rol das propriedades correspondentes ao seu Município, guardando a ordem alfabética das primeiras letras do primeiro nome do proprietário. A numeração das propriedades é feita independentemente para cada um dos Distritos que compõem o Município.

Eis um exemplo tirado de um dos róis.

MUNICÍPIO DE SÃO ROQUE - Distrito de São João Novo

1) Adão Canos

Candio

0,9075ha Cr\$2.080,00

2) Adolfina Margarida Rodrigues

São João

15,1250ha Cr\$19.165,00

Tendo aceito este cadastro como "sistema de referência" para a amostragem, é preciso analisar os fatores que o poderiam tornar inadequado. Além da análise das causas internas desfavoráveis, será indispensável fazer uma investigação objetiva capaz de esclarecer o assunto.

São dois os fatores, que podem prejudicar o valor dos róis da Secretaria da Fazenda como "sistema de referência" para a amostragem:

- a) a existência de propriedades que não estão incluídas no rol;
- b) a existência de dois ou mais registros para uma mesma gleba de terra, isto é, duas ou mais pessoas que pagam imposto sobre a mesma área de terra.

Em ambos os casos, há um prejuízo evidente para a amostra constituída à base de tal Cadastro, pois, a ausência de propriedades da lista impossibilita sua inclusão na amostra e a existência de dois registros para a mesma propriedade faz com que dobre a probabilidade de sua inclusão na amostra.

Um cadastro, onde ocorram omissões ou duplicações de propriedades em número apreciável, estará automaticamente condenado como "sistema

de referência" para amostragem.

Analisem-se em separado as possibilidades de omissão e a duplicação em nosso Cadastro.

É difícil a existência de omissões, por ser de grande interesse do proprietário o registro do imóvel em Cartório. O registro em Cartório garante legalmente a posse do imóvel. Mas, o registro só se pode processar mediante a comprovação do pagamento dos impostos, entre os quais está o Imposto Territorial Rural.

Suponha-se, porém, que não seja rigoroso o controle do pagamento dos impostos por ocasião do Registro da propriedade em Cartório.

Então, três razões podem levar um proprietário a não pagar o imposto sobre suas terras: a) ignorância da obrigação de efetuar o pagamento; b) o imposto é muito elevado, isto é, ele constitui uma parcela importante de seu orçamento; c) o imposto é desproporcionalmente elevado com relação ao valor da terra.

A ignorância da necessidade do pagamento do imposto, aparentemente, só poderia advir da existência de comunidades auto-suficientes de agricultores, que estivessem localizados em pontos inacessíveis do Estado. Não se tem conhecimento da existência de tais comunidades, pois, atualmente, todo o Estado se encontra desbravado, possuindo estradas transitáveis, e os produtores agrícolas estão intimamente ligados ao mercado, não havendo, pois, razão para tal ignorância, em escala apreciável.

O Imposto Territorial Rural é cobrado na base de 1,5% do valor venal da propriedade, valor esse, declarado pelo proprietário. As 316.000 propriedades no Estado proporcionaram à Secretaria da Fazenda, no ano fiscal de 1952, Cr\$124.373.086,80. Esta importância equivale a Cr\$392,80 por propriedade cuja área média é de 70ha, ou seja, menos de Cr\$6,00 por hectare. Para se ter idéia do que isso representa, pode-se dizer que os lavradores do Estado receberam, em junho de 1952, mês do pagamento do imposto, em média Cr\$101,00 por saca de 60 quilos de milho, Cr\$196,00 por saca de 60 quilos de arroz em casca e Cr\$299,00 por saca de 40 quilos de café em coco. A observação do número acima deixa patente

que o Imposto Territorial Rural não constitui ônus pesado para os agricultores do Estado.

Finalmente, o imposto apesar de não ser pesado poderia ser desproporcional com relação ao valor da terra. A realidade é bem o contrário disso. A especulação com imóveis, entre eles os rurais, tem sido imensa, e atualmente a desproporção entre o valor declarado das propriedades rurais para efeito fiscal e o valor real do imóvel é tão grande em favor do valor real, que a Secretaria da Fazenda está levando a efeito uma revisão geral de seus lançamentos referentes ao Imposto Territorial Rural.

Acredita-se que a sonegação do Imposto Territorial, quando ocorre, se faça, principalmente, através de declarações de área menor que a real, bem como declarações incorretas sobre a qualidade das terras, dois fatores que afetam diretamente o montante do imposto e cuja verificação se torna extremamente difícil, se não impossível.

Em resumo, todos os fatores são favoráveis ao Cadastro organizado com base nos dados do Imposto Territorial Rural, no que diz respeito à inclusão de todas propriedades rurais do Estado. Caso haja omissão, seu número deve ser muito restrito. Até o momento, só há conhecimento de uma propriedade que não foi incluída em nosso cadastro e que foi indicado por um Engenheiro Agrônomo Regional.

Outro fator que pode prejudicar a lista das propriedades obtidas através da Secretaria da Fazenda é a duplicação do registro. Este fato poderia parecer menos possível que o anterior, porém, acontece o contrário. Pode-se verificar que há casos em que duas pessoas diferentes pagam o imposto da mesma terra. Trata-se, geralmente, de terras em litígio, na fronteira noroeste do Estado, barrancas do rio Paraná, recentemente desbravadas. Estas disputas envolvem, geralmente, grandes glebas de 50.000 hectares até meio milhão de hectares, facilmente identificáveis. Há, pois, esperanças de que os Engenheiros Agrônomos Regionais, encarregados das zonas onde isso ocorre, comuniquem os fatos, a fim de se eliminarem gradativamente os casos evidentes de duplicação. Outra fonte de duplicação é a conservação, na lista, do nome de um proprietário que vendeu sua

propriedade a outro, quando o novo proprietário já consta da lista. Este tipo de duplicação é, forçosamente, muito reduzido. Afora os casos já citados de demanda e *grilo*, em nossa amostra, ainda não ocorreu caso de sorteio de dois registros de uma mesma propriedade.

Além das evidências internas de que o Cadastro, organizado no Estado de São Paulo com o auxílio das listas do Imposto Territorial Rural, é suficientemente completo e isento de duplicações, pretende-se fazer uma investigação capaz de comprovar objetivamente este fato. Ela poderá ser feita mediante a verificação da inclusão das propriedades limítrofes de nossa amostra, no Cadastro.

4.2 - Organização do Cadastro

O Cadastro, existente na Seção de Previsão de Safras e Cadastro, serviu como "sistema de referência" para a retirada de amostra do ano de 1951-1952. Organizado pelo sistema Keysort em 1947/48, não havia sido atualizado até 1951. Em consequência, figuravam em nossa amostra nomes de proprietários já completamente desconhecidos nas regiões indicadas, propriedades que haviam sido desmembradas ou, então, anexadas a outras, tornando-se, em muitos casos, bastante difícil a sua localização por parte dos Engenheiros Agrônomos Regionais.

Havia, pois, urgência em atualizar o Cadastro existente ou, então, confeccionar o novo Cadastro.

Foram comparados o trabalho e as despesas da atualização do Cadastro Keysort e o trabalho e as despesas que decorreriam da confecção de um novo Cadastro com cartões perfurados I.B.M. (Holerith). Além disso, em caso de permanência do Cadastro Keysort, foi levado em consideração o fato de que, com o pessoal disponível, seria impraticável fazer a lista das propriedades de cada Município na ordem crescente de área, e os resumos por classe de área, indispensáveis na retirada da amostra. Diante das vantagens oferecidas pelo sistema I.B.M., optou-se pela confecção de Cadastro novo, neste sistema.

Providenciou-se, então, o empréstimo dos róis da Secretaria da Fazenda; antes, porém, de se iniciar a perfuração dos cartões, foram organizados Códigos numéricos dos Municípios e Distritos.

Devido ao grande número de informações alfabéticas, não passíveis de codificação, e que deveriam ser colocadas nos cartões, foi-se obrigado a fazer dois cartões para cada propriedade do rol.

Foi o seguinte o esquema da perfuração (Figura 3).

Terminada a perfuração e a sua conferência, os cartões de cada Município foram postos em ordem crescente de área. Em seguida, foram intercalados cartões mestres, separando os grupos de cartões pertencentes às classes de área.

Foi o seguinte o esquema da listagem, feita com o auxílio dos cartões, assim organizados (Figura 4).

Ao mesmo tempo que se fez a listagem por Município, foram tirados resumos por classe de área dentro do Município. O número de propriedade, dentro de cada classe, e a área total dessas propriedades são os principais elementos contidos no cartão resumo.

A cada Engenheiro Agrônomo Regional, foram enviadas cópias das listas das propriedades dos Municípios de sua Região, para o fim de melhor familiarizá-lo com as propriedades agrícolas dela.

4.3 - Determinação das Classes de Área e das Frações de Amostragem

Com os dados obtidos no primeiro ano, pôde-se tentar, de forma conseqüente, a construção de uma amostra capaz de proporcionar estimativas eficientes.

Se por um lado a estratificação geográfica devia permanecer a mesma do ano anterior, porque a coleta dos dados continuaria sendo feita pelos Engenheiros Agrônomos Regionais, o mesmo não se dava com a questão dos limites das classes de área e o sistema de frações variáveis de amostragem.

Quanto aos limites das classes de área, po-

Coluna	Cartão 1	Cartão 2
1	Dígito 1	Dígito 2
2-5	Número de ordem da propriedade no rol	
6-48	Nome do proprietário	Nome da propriedade Nome do bairro
49-51	Três primeiras letras do segundo nome	
52-54	Três primeiras letras do último sobrenome	
55-59	Área em hectares	
60-62	Código numérico do município ¹	
63	Código numérico do distrito ¹	
64-67	Quatro primeiras letras do nome	
68	Código de classe de tamanho ²	

¹Duplicação.

²Multiperfuração após a classificação por área.

Figura 3 - Esquema de Perfuração.

Número da ordem dentro da classe de área	Nome do proprietário Nome da propriedade Nome do bairro	Número da propriedade e rol	Código do município e distrito	Área em hectares	Número de proprietário em cada classe de área	Código de classe de área

Figura 4 - Esquema de Listagem.

dem-se fazer as seguintes considerações: o exame das tabelas estatísticas de distribuições altamente assimétricas, tais como área das propriedades agrícolas, salários individuais, etc., mostra que a escolha dos limites das classes é feita de forma arbitrária e irregular. A pessoa que organiza a tabela nota, corretamente, que a amplitude dos intervalos deve aumentar à medida que se atingem pontos mais altos da escala. Mas, ao escolher séries como 1 ... 2 ... 3 ... 4 ... 5 ... 10 ... 20 ... 30 ... etc., ela provoca saltos indesejáveis nas frequências, os quais tendem a obscurecer as informações que a tabela deveria proporcionar. Pareceria mais apropriado fazer com que os limites das classes se desenvolvessem segundo uma lei regular.

Sugestão óbvia para isso é uma progressão geométrica, como 3 ... 9 ... 27 ... 81 ... 343 ... etc. Pode-se notar, também, que esta série de limites de classes se torna aritmética ao se fazer a transformação logarítmica, facilitando assim, em muitos casos, o ajustamento de curvas de frequência, especialmente, pela curva normal logarítmica. Por outro lado, a prática demonstra que os homens consideraram mais exequível o trabalho com números redondos, tais, por exemplo, números divisíveis por potências de dez.

Diante dessas considerações, adotou-se a série 3 ... 10 ... 30 ... 100 ... 300 ... 1.000 ... 3.000 ... 10.000 ... 30.000 ... 100.000 hectares para limites de nossas classes de área. Observa-se que, sendo 3 a raiz quadrada aproximada de 10, esta série é a aproximação grosseira de uma progressão geométrica.

É interessante notar que, no ano anterior, se tentou arranjar os intervalos de classe de forma que o mesmo número de elementos fosse tirado de cada estrato de área. Constatou-se, felizmente, que a série de intervalos acima proposta possibilita a construção de uma amostra satisfatória com, aproximadamente, o mesmo número de propriedades em cada estrato de área, exceção feita ao primeiro estrato (3-10 hectares).

Quanto às frações variáveis de amostragem, deve-se lembrar, em primeiro lugar, que não se poderia continuar retirando para a amostra um número constante dentro de cada substrato, sem levar em consideração o número de propriedades dentro do substrato e nem haveria razões para supor que, para todas as Regiões do Estado, seria ótimo um mesmo conjunto de frações de amostragem. Ao contrário, não há dúvida de que seria desejável determinar conjuntos de frações de amostragem para cada Re-

gião, ou, ao menos, para cada Setor. Mas, nem mesmo para o Setor foi possível determinar o desvio padrão com suficiente precisão. O melhor que se pôde fazer foi reunir as Regiões em quatro grupos, com base em certa semelhança da agricultura, o que não imprimiu, contudo, a esses grupos a desejada uniformidade. Ao agrupamento de Regiões Agrícolas chamou-se ZONA, sendo, em seguida, calculadas em separado as frações de amostragem para cada uma das quatro zonas, em que ficou dividido o Estado. Em segundo lugar, deve-se lembrar que não se poderia esperar a determinação de frações de amostragem simultaneamente ótimas para todas observações (itens do Questionário). Neste caso, poder-se-ia encarar o interessante problema de construir uma amostra de tamanho mínimo que, em nenhum caso, proporcionasse estimativas cujo desvio padrão fosse superior aos limites tolerados.

Deixando, porém, de lado tais requintes técnicos, está claro que o problema prático consistia em determinar frações de amostragem que fossem razoavelmente boas para as principais culturas.

É inevitável que algumas culturas secundárias, cuja produção se localiza em áreas limitadas, tenham estimativas com desvio padrão superior ao aceitável; a situação pode todavia ser remediada desde que se tenham os recursos necessários, aumentando as frações de amostragem nas Regiões onde tais culturas se desenvolvem.

A amostra mais eficiente é aquela que tem frações de amostragem proporcionais aos desvios padrões dentro de cada cadastro.

Ora, os desvios padrões foram calculados para cada produto e dentro de cada classe de área. Entretanto, sendo a estratificação, proposta em 1952-1953, diferente da usada no ano anterior, surgiu o problema da maneira de determinar estimativas do desvio padrão dentro dos novos estratos, a partir dos dados obtidos para os antigos estratos. A reclassificação dos dados originais, segundo os novos estratos, era impraticável, porque, em muitos casos, não se obteriam os mínimos necessários de duas observações por substrato, capazes de proporcionar uma contribuição à soma dos quadrados. Poder-se-

ia, naturalmente, (reclassificando os dados) fazer estimativas para grupos de Regiões; entretanto, tais estimativas estariam inflacionadas por variações entre regiões, não sendo, pois, adequadas ao problema a resolver.

O método, efetivamente, utilizado foi o de interpolação gráfica (Figura 5) e será ilustrado com os dados da Quarta Zona (Setores de Araçatuba, Bebedouro, Presidente Prudente, Ribeirão Preto e São José do Rio Preto).

A estimativa da variância média dentro do substrato é dada pela fórmula:

$$S^2 = \sum \frac{(X_1 - X_2)^2}{2n}$$

onde X_1 e X_2 são as duas observações do substrato, tendo lugar a soma para os n substratos de cada classe de área dentro da zona para a qual existem **ambas** observações. Foram feitas estimativas da variância para a área plantada com os quatro principais produtos - algodão, arroz, milho e café.

Naturalmente, a variância dependerá dos limites inferior e superior da classe; mas, para simplificar o problema, supõe-se que ela depende, principalmente de:

$$\frac{1}{2} \log (I S) = \frac{1}{2} (\log I + \log S)$$

onde I e S são, respectivamente, os limites inferior e superior da classe (a maior classe, aberta à direita, foi tomada como sendo de 3.000 a 10.000).

Na figura 5, há os logaritmos das estimativas dos desvios padrões no eixo das abcissas e $\log(IS)/2$ no eixo das ordenadas.

Pode-se observar que as curvas da figura, representativas das quatro culturas, são, aproximadamente, paralelas. Esta verificação é bastante confortadora, pois, significa que as frações de amostragem serão eficientes para as quatro culturas. Finalmente, calcularam-se e assinalaram-se, na figura, as médias de $\log(s)$ para as quatro culturas. No caso da Quarta Zona, foi consignado ao algodão peso dobrado, porque ele é o produto mais importante da zona. A regressão foi ajustada visualmente por meio de duas linhas retas. Os pontos dessas linhas de regressão, correspondendo às novas classes de área, deram, à

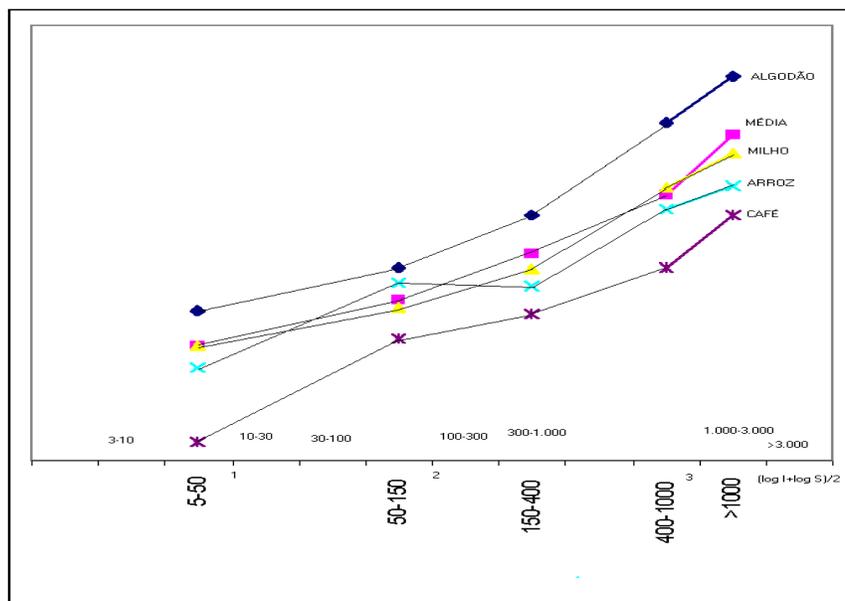


Figura 5 - Logaritmos dos Desvios Padrões para Interpolação.

parte uma constante adicional arbitrária, os valores de $\log(f)$ onde f é a fração de amostragem. O resultado encontra-se na tabela 9 e figura 6.

A tabela 10 mostra, para as quatro zonas, as quantidades proporcionais às frações de amostragem, onde, para efeito de comparação, à classe central foi atribuído o valor 100. Para as propriedades inferiores a 3 hectares, a fração foi zero, porque apesar de constituírem 15,7% do número de propriedades sua área é de apenas 0,2% da área total, das propriedades.

4.4 - Coleta de Dados

4.4.1 - Descrição da organização

O trabalho de amostragem constitui-se de duas partes que, conquanto se realizem independentemente no espaço e no tempo, estão indissolivelmente ligadas.

A precisão das estatísticas, obtidas por amostragem, depende, igualmente, da perfeição do trabalho de escritório e da veracidade das informações

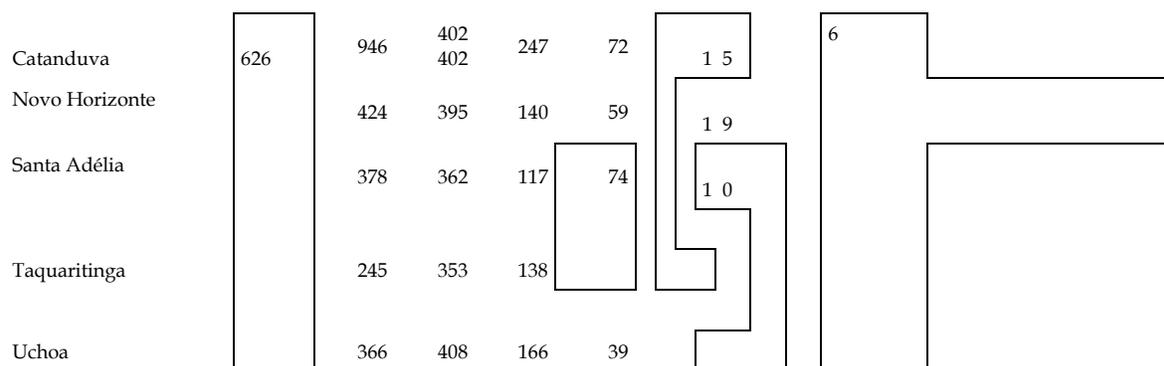
obtidas junto dos lavradores. Dessas estatísticas, o trabalho de campo é, inegavelmente, o que custa mais caro, o que necessita de mais gente e está mais sujeito a erros. Os erros são de controle mais difícil do que os erros de escritório.

A Secretaria da Agricultura, ao estabelecer o serviço de amostragem, visou não somente divulgar dados estatísticos, mas, também, efetuar pesquisas e proporcionar aos economistas conhecimentos básicos sobre a agricultura de São Paulo. Por isso, os Questionários deveriam conter, forçosamente, itens complexos que, para serem respondidos, demandariam conhecimentos técnicos especializados. Ora, apesar de ser objetivado o estabelecimento de um sistema permanente de amostragem, não se contava, até o momento, com recursos suficientes para manter um corpo de pessoas capazes de fazer as visitas periódicas a todas propriedades da amostra. Por isso, foi solicitada a cooperação de um organismo já estruturado na Secretaria da Agricultura, a fim de que ele se encarregasse de coletar informações junto aos lavradores.

Felizmente, a Divisão de Fomento Agrícola conta com a Seção de Regiões, que cobre todo o Es-

Tabela 9 - Estrutura da Amostra no Setor Agrícola de Catanduva

	<3	3-10	10-30	30-100	100-300	300-1.000	1.000-3.000	3.000-10.000	10.000-100.000	>100.000
Número de propriedades por estrato (ha)										
Catanduva	147	276	946	804	247	72	10	1	0	0
Novo Horizonte	55	97	424	395	140	59	19	3	1	0
Santa Adélia	32	78	378	362	117	37	3	0	0	0
Taquaritinga	16	71	245	353	138	37	5	0	0	0
Uchoa	97	114	366	408	166	39	7	1	0	0
Fração de amostragem	0	0,11	0,22	0,51	1,00	2,10	4,40	8,90	0	0
Número de propriedades multiplicado pela fração amostral										
Catanduva	0	30,36	208,12	410,04	247,00	151,20	44,00	8,90	0	0
Novo Horizonte	0	10,67	93,28	201,45	140,00	123,90	83,60	26,70	0	0
Santa Adélia	0	8,58	83,16	184,62	117,00	77,70	13,20	0	0	0
Taquaritinga	0	7,81	53,90	180,03	138,00	77,70	22,00	0	0	0
Uchoa	0	12,54	80,52	208,08	166,00	81,90	30,80	8,90	0	0
Número de propriedades agrupadas										



tado com uma rede de Casas da Lavoura. São estas, ao todo, 153 organismos regionais, cada um sob a chefia de um Engenheiro Agrônomo Regional.

Analisando as qualidades do Engenheiro Agrônomo Regional sob o ângulo do interesse de nosso serviço de amostragem, verificou-se que apresentava as seguintes vantagens:

- a) é um funcionário bem remunerado e de categoria;
- b) tem, geralmente, veículo próprio (Jeep ou automóvel);
- c) conhece bem sua Região, assim como grande número

de seus agricultores;

- d) está habituado a lidar e a conversar com agricultores;
- e) seus conhecimentos lhe permitem ajuizar da veracidade das respostas que recebe, discuti-las com o lavrador, consignando no Questionário o dado mais correto.

Por outro lado, o Engenheiro Agrônomo Regional apresenta, como fator negativo para o nosso serviço, o fato de sua função primordial na Secretaria da Agricultura ser a do fomento da técnica agrícola

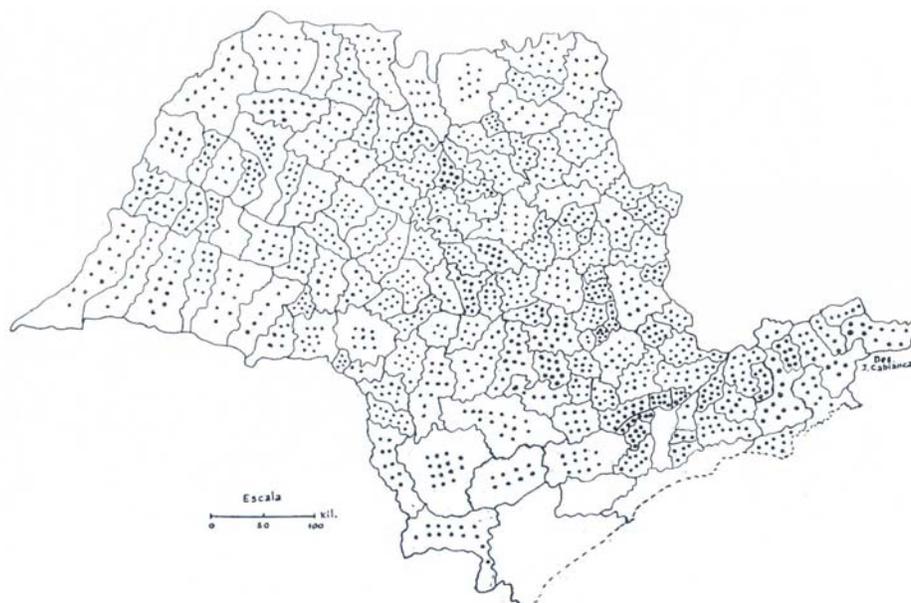


Figura 6 - Distribuição da Amostra por Regiões, cada Ponto Equivale a uma Propriedade.

Tabela 10 - Frações de Amostragem Proporcionais

Classes de tamanho (ha)	Zonas			
	1	2	3	4
3-10	26	15	11	14
10-30	42	29	22	25
30-100	64	52	51	45
100-300	100	100	100	100
300-1.000	300	280	210	220
1.000-3.000	1.300	810	440	710
3.000-10.000	5.300	2.600	890	2.200

mais avançada e da produção em suas respectivas Regiões, dedicando apenas uma pequena parte de seu tempo às entrevistas com os agricultores para o preenchimento de nossos Questionários. Ora, nada mais plausível que esperar do agrônomo que, ao tentar ganhar a confiança do agricultor, por ocasião da entrevista, faça, ao mesmo tempo, "fomento", abordando os problemas do agricultor e pondo à sua disposição as facilidades da Casa da Lavoura bem como seus próprios conhecimentos técnicos.

Pode-se admitir que essa atitude do Regional influa, até certo ponto, nos resultados obtidos pelos agricultores da amostra, viciando desta forma as estimativas finais. Mas, se for admitida a possibilidade dessa influência, pois, não se pode negar o efeito da educação, deve-se acreditar que ela não se

faça sentir em curto período de tempo e, muito menos, de forma ponderável em nossa amostra. Isto porque a maioria das propriedades e dos proprietários da amostra é, inicialmente, desconhecida dos Regionais; e levando em consideração a circunstância de que o agricultor tem, em geral, mentalidade desconfiada e conservadora, não se pode aceitar que ele modifique suas práticas tradicionais, em virtude das simples afirmativas de um estranho, por mais lógicas que sejam e por mais convincente a forma de fazer essas afirmativas.

Mas, se apesar dessas considerações, for julgado conveniente reduzir ainda mais as possibilidades de vícios introduzidas pelos Engenheiros Agrônomos Regionais, então, pode-se adotar como norma substituir as propriedades da amostra em um

ano, só no fim da época de plantio das culturas anuais mais importantes. Em São Paulo, tal substituição se daria em dezembro.

4.4.2 - Crítica ao sistema

Já foi dito que a técnica de obtenção de respostas consiste na entrevista entre o Engenheiro Agrônomo Regional e o responsável pela exploração agrícola. Para cada levantamento são enviados aos Regionais os Questionários em duplicata, contendo cada um o nome do proprietário, o nome da propriedade e o Bairro, o Distrito e o Município em que se localiza o imóvel, e sua área, segundo o Cadastro. O Regional visita, então, o empresário na cidade, se ele reside aí, ou na fazenda se ele reside no campo, preenchendo o Questionário com base nas informações prestadas pelo responsável. É inegável que ao ser inquirido, o empresário **estima** a área cultivada e a produção provável, podendo haver enganos nesta estimativa. Se se trata de pesquisas referentes à técnica de produção, onde são perguntados fatos ocorridos há algum tempo, as possibilidades de erro são maiores, porque são poucas as propriedades agrícolas que mantêm escrituração adequada.

Entretanto, pode-se afirmar que, excetuando os casos em que as respostas são dadas com má fé, o erro possível numa estimativa feita pelo empresário, relativa à sua exploração agrícola, é forçosamente pequeno.

São, provavelmente, os pequenos empresários que têm menos idéia das áreas que ocupam com cada uma de suas culturas e da produção total colhida, porque o conhecimento exato das áreas lhe é de pouco interesse, considerando-se que grande parte da produção é consumida pela própria família. Já nas propriedades médias e grandes, onde o trabalho é feito com assalariados, o conhecimento da área cultivada com um produto decorre da necessidade do parcelamento dessa área para efeito de pagamento das **tarefas**. As estimativas acertadas de produção, nestes casos, também interessam bastante ao empresário.

O fato de não serem grandes os erros de esti-

mação do empresário com relação à sua exploração agrícola não elimina, em absoluto, a necessidade de se levar a efeito uma pesquisa capaz de determinar os erros a que estão sujeitas as estimativas consignadas pelos lavradores. Esta pesquisa ainda não foi realizada aqui, porque a organização criada não possui, até o momento, os meios indispensáveis.

Outra dificuldade, que surge na coleta dos dados, relaciona-se à maneira pela qual foi definida a população.

Esta dificuldade se origina do fato de que, sendo a propriedade rural nossa unidade de amostragem, a ela nem sempre corresponde uma única exploração agrícola. Isto ocorre toda vez que o proprietário arrenda ou dá em parceria suas terras a diversas pessoas. Nestes casos, é o empresário de cada uma das explorações, e não o proprietário, a pessoa mais habilitada a prestar informações. A solução consiste em obter os dados junto de cada um dos empresários, que fazem parte das propriedades da amostra. Houve casos, em que era grande o número de empresários em uma propriedade.

Mas, desde que se obtenham as informações sobre todas as explorações, situadas nos limites da propriedade da amostra, não haverá prejuízo para os resultados finais.

5 - RESULTADOS OBTIDOS NO SEGUNDO ANO

A estimativa final da área cultivada com algodão, arroz, milho e o número de cafeeiros foi obtida a partir do Questionário de março, enquanto a produção se refere ao Questionário de julho.

A tabela 11 dá as áreas, produções e respectivos desvios padrões daqueles produtos no ano agrícola de 1952/1953.

Dos 1.360 Questionários enviados no mês de julho, voltaram devidamente preenchidos 1.170, ou seja, 86% do total. Entretanto, os cálculos foram feitos como se todos houvessem respondido, preenchendo-se as falhas com as informações contidas nos Questionários recebidos.

Tabela 11 - Estimativa Final de Área e Produção de Algodão, Arroz, Milho e Café e os Respectivos Desvios Padrões

Produto	Área (ha)	Desvio padrão (%)	Produção	Desvio padrão (%)
Algodão	837.000	13	655.000 t em caroço	12
Arroz	543.000	3	7.200.000 sc.60kg	5
Milho	1.162.000	5	21.000.000 sc.60kg	5
Café	1.137.000.000 pés	5	8.700.000 sc.60kg	9

Pode-se verificar na tabela 11 que os desvios padrões da área cultivada e da produção de algodão e da produção de café são superiores aos aceitáveis. Entretanto, o fato de em cada um dos casos citados alguns substratos contribuírem até com 70% do total da variância leva a crer que o fenômeno seja atribuível à estimação pouco precisa da variância.

Esta hipótese da baixa precisão da variância é reforçada quando se compara o número de sacos de semente de algodão comprados pelos lavradores e estimado no Questionário de janeiro, com o número de sacos de semente de algodão efetivamente vendidas pela Secretaria da Agricultura; ou então quando é comparada a estimativa da produção de algodão feita em julho com a produção efetiva, obtida através das máquinas de benefício de algodão (Tabela 12).

Tabela 12 - Comparação entre Dados Estimados e Efetivos

	Estimado	Efetivo
Sacas de sementes compradas pelos lavradores	800.000	787.358
Produção de algodão safra de 1952/53	655.000	653.651

Vê-se que as estimativas estão muito próximas dos parâmetros.

A baixa precisão da variância parece residir no fato de a amostra não se distribuir pelos diferentes Setores Agrícolas da mesma forma que se distribui a produção. É sabido que as zonas novas do Estado com maior produção são as que têm Regiões Agrícolas de maior área. Ora, sendo mais ou menos constante o número de visitas a se fazerem em cada Região Agrícola, conclui-se que o número de elementos da amostra em cada Setor Agrícola não é proporcional a sua produção. Torna-se, pois, neces-

sário aumentar substancialmente o número de elementos da amostra nos Setores onde há acentuada divergência entre a densidade da amostra e da produção.

Tem-se um prazo de 35 a 40 dias entre o envio do Questionário e o início dos cálculos. É de 10 dias a duração dos cálculos que, no momento, são feitos manualmente.

Para cada levantamento o custo das visitas dos Engenheiros Agrônomos Regionais às propriedades da amostra é de Cr\$180.000,00 em despesas de transporte e diárias.

6 - CONCLUSÃO

Da descrição dos trabalhos já realizados e dos resultados obtidos, pode-se admitir a viabilidade do emprego do "método de amostragem" na obtenção de dados estatísticos relacionados com a agricultura do Estado de São Paulo.

Comparado ao método anterior (o subjetivo), o da amostragem fornece estatísticas mais fidedignas, pois, ele permite a medida do erro de estimação. Além disso, há considerável ampliação do número de informações que se podem obter utilizando o novo método; assim, os Questionários dos diferentes levantamentos incluíram itens alheios às estatísticas atualmente fornecidas pelos Engenheiros Agrônomos Regionais ou os Agentes Municipais de Estatísticas. Os novos elementos devem servir de base para estudos e análises das outras Seções da Subdivisão de Economia Rural.

Os processos de comercialização de café no Estado de São Paulo e a renda das propriedades agrícolas são agora estudados em uma pesquisa

levada a efeito pela Subdivisão de Economia Rural. As propriedades agrícolas visitadas nesta pesquisa constituem uma sub-amostra da amostra utilizada nos cálculos da produção.

Vê-se, pois, que a amostra não é simplesmente um instrumento eficiente para a previsão e estimação das safras, mas, pode, além disso, ser o alicerce sólido sobre o qual repousam as pesquisas econômicas que levarão ao melhor conhecimento de nossa agricultura.

A discussão dos resultados obtidos com o último Questionário de julho tornou clara a necessidade de se continuar o aperfeiçoamento da amostra.

Nas atuais condições, essa melhoria deve ser conseguida por dois meios: aumento do número de elementos da amostra nos Setores Agrícolas de maior produção e correção das imperfeições do Cadastro.

A duplicação do número de propriedades da amostra nos Setores Agrícolas de Presidente Prudente, Marília, Araçatuba e Rio Preto não pode ser resolvida duplicando o número de visitas a serem feitas pelos Engenheiros Agrônomos Regionais destes Setores, porque, na maioria dos casos, os trabalhos se tornariam excessivos.

A solução parece estar ou no aumento do número de Engenheiros Agrônomos Regionais nos Setores considerados ou, então, na localização de, pelo menos, um elemento especializado que se incumbiria de visitar grande número de propriedades em cada um dos quatro Setores. Destas soluções, a última parece ser mais viável no momento porque implica em menores despesas.

Depois de aumentar o número de elementos da amostra nos Setores de maior produção, será necessário verificar a precisão das informações dos lavradores, sobre a área cultivada. Esta verificação poderá ser feita por meio do levantamento trigonométrico das áreas cultivadas em uma sub-amostra adequada, tornando-se então possível a correção dos totais de área.

Esta correção redundará em maior precisão das estimativas das áreas cultivadas.

Para se melhorar a previsão da produção, deverá ainda ser resolvido o problema da determina-

ção do rendimento médio de cada uma das culturas. Dois caminhos podem ser seguidos na solução deste segundo problema. O primeiro liga-se ao estabelecimento de certo número de pequenas "áreas de prova" em todo Estado, as quais, cultivadas de forma adequada, servirão de "termômetro" do efeito das condições mesológicas sobre os diferentes cultivos. O segundo caminho consiste numa pesquisa para a determinação da técnica a ser adotada no cálculo do rendimento dos campos de cultura de uma sub-amostra escolhida adequadamente.

O aperfeiçoamento da amostra e o estabelecimento de controles objetivos das informações obtidas constituem a linha segundo a qual se desenvolverão os trabalhos futuros do "serviço de amostragem".

**OBTENÇÃO DE ESTATÍSTICAS AGRÍCOLAS PELO MÉTODO DE AMOSTRAGEM:
experiências visando a criação de uma organização permanente**

Anexo 1

DEPARTAMENTO DA PRODUÇÃO VEGETAL
DIVISÃO DE ECONOMIA RURAL
SECRETARIA DA AGRICULTURA

LEVANTAMENTOS ECONÔMICOS
DA SUBDIVISÃO DE ECONOMIA RURAL

PREVISÃO DE SAFRAS
Questionário de Janeiro

código							Área total em alqueires
região							
Distância da Sede em quilômetros							
	ALGODÃO arobas	AMENDOIM sacas 25 quilos	ARROZ em casca - sc.60kg	FEIJÃO das águas - sc.60kg	MILHO em sacas 60kg	CAFÉ	
Produção na Safra 1951-1952						Produção 1951-1952 sc.40kg - coco	
Área plantada Safra 1951-1952	Alqueires					N. de pés	até 3 anos sem colheita
Área plantada Safra 1952-1953							de 3 anos a 8 anos
							com mais de 8 anos
Número de sacas de sementes de algodão compradas							
ÁREA FLORESTAL				PASTOS			
MATAS naturais em alqueires	Matas formadas em alqueires			CAMPOS em alqueires	JARAGUÁ em alqueires	COLONIÃO em alqueires	GORDURA em alqueires
	EUCALIPTOS	PINHEIROS					
código							Área total em alqueires
região							
Distância da Sede em quilômetros							
	ALGODÃO arobas	AMENDOIM sacas 25 quilos	ARROZ em casca - sc.60kg	FEIJÃO das águas - sc.60kg	MILHO em sacas 60kg	CAFÉ	
Produção na Safra 1951-1952						Produção 1951-1952 sc.40kg - coco	
Área plantada Safra 1951-1952	Alqueires					N. de pés	até 3 anos sem colheita
Área plantada Safra 1952-1953							de 3 anos a 8 anos
							com mais de 8 anos
Número de sacas de sementes de algodão compradas							
ÁREA FLORESTAL				PASTOS			
MATAS naturais em alqueires	Matas formadas em alqueires			CAMPOS em alqueires	JARAGUÁ em alqueires	COLONIÃO em alqueires	GORDURA em alqueires
	EUCALIPTOS	PINHEIROS					

DEPARTAMENTO DA PRODUÇÃO VEGETAL
 DIVISÃO DE ECONOMIA RURAL
 SECRETARIA DA AGRICULTURA

LEVANTAMENTOS ECONÔMICOS
 DA SUBDIVISÃO DE ECONOMIA RURAL

PREVISÃO DE SAFRAS
 Questionário de Março

					código	
					região	
Safra 1952/53	ALGODÃO	AMENDOIM DA SECA	ARROZ EM CASCA	BATATA DA SECA	CANA-DE-AÇÚCAR	
Área plantada - alq.						
Produção provável	arrobas	sacos 25kg	sacos 60kg	sacos 60kg	toneladas	
	FEIJÃO DA SECA	MAMONA	MILHO	CAFÉ	LARANJAS PRECOSES (baianinha, lima, etc.)	
Área plantada - alq.			$\frac{q}{n}$	$\frac{q}{n}$		
Produção provável	sacos 60kg	sacos 50kg	sacos 60kg	sacos 40kg em coco	caixas de colheita	
	AMENDOIM DAS ÁGUAS	BATATA DAS ÁGUAS	FEIJÃO DAS ÁGUAS	BANANA NANICA	BANANA MAÇÃ	LARANJAS TARDIAS (pêra, natal, etc.)
Área plantada - alq.				$\frac{n}{\text{touceira}}$	$\frac{n}{\text{touceira}}$	$\frac{n}{\text{pês}}$
Produção colhida	sacos 25kg	sacos 60kg	sacos 60kg	cachos nos 2 últimos meses	cachos nos 2 últimos meses	caixas de colheita

OBSERVAÇÕES _____

DEPARTAMENTO DA PRODUÇÃO VEGETAL
 DIVISÃO DE ECONOMIA RURAL
 SECRETARIA DA AGRICULTURA

LEVANTAMENTOS ECONÔMICOS
 DA SUBDIVISÃO DE ECONOMIA RURAL

PREVISÃO DE SAFRAS
 Questionário de Julho

										código					
										região					
ALGODÃO			ARROZ				MILHO			CAFÉ					
Produção total 1952-53 - arrobas		Área a plantar na próxima safra alqueires	Produção total 1952-53 sacos 60kg em casca		Área a plantar na próxima safra alqueires		Produção total 1952-53 sacos 60kg		Área a plantar na próxima safra alqueires	colhido sacos de 40kg em coco		a colher sacos de 40kg em coco			
(1) TRABALHADORES DA PROPRIEDADE							MÁQUINAS AGRÍCOLAS DE TRACÇÃO ANIMAL								
(2) PROPR.		(2) ARREND.		CAMARADAS		COLONOS		PARCEIROS		(4) N. TOTAL DE ARADOS		(4) N. TOTAL DE GRADES		(4) N. TOTAL DE CULTIV.	
(3) n. de enxadas	(3) n. de enxadas	n. de diaristas	n. de mensal.	n. de famílias	n. de enxadas	n. total	n. de enxadas	aiveca	disco	disco	dente	(5) carpideira		(6) planets	

- (1) Mulheres que trabalham parte do tempo e crianças menores de 15 anos, consideram-se 1/2 (meia) enxada.
- (2) Quando o proprietário arrenda parte da propriedade, preencher separadamente para cada um deles - arrendatário e proprietário.
- (3) Incluir somente as pessoas da família do empresário que fazem trabalho braçal na agricultura.
- (4) Não incluir os implementos de trator.
- (5) Carpideira - máquina agrícola para uma só enxada.
- (6) Planets - comporta geralmente 3 ou 5 enxadinhas.

OBSERVAÇÕES _____

Assinatura do Engenheiro Agrônomo: _____
 Data: _____