

Retirada da Amostra de Municípios para Previsão da Produção de Algodão, Amendoim, Arroz, Feijão e Soja no Brasil

Relatório Preliminar apresentado ao Dep.
Econômico do Ministério da Agricultura
em dezembro de 1964

Eng. Agr.^o Salomão Schattan

Das discussões mantidas com os técnicos do Ministério da Agricultura e I.B.G.E. ficou claro que no momento, os dois problemas fundamentais das estatísticas agrícolas são a pouca confiança que inspiram os dados ora publicados e o grande atraso com que é feita sua publicação.

1 — As estatísticas agrícolas publicadas pelo SEP — IBGE são obtidas pelo sistema de “estimação subjetiva”. Está claro que tais estatísticas não são necessariamente erradas, mas apresentam o inconveniente de não se poder medir sua precisão.

Em vista disso, só se pode

saber a posteriori se elas estão certas ou erradas, isto é, depois de feito o confronto com a safra efetivamente colhida (isto é possível nos casos do café e algodão).

No tempo que medeia entre a previsão e a estimativa final, as estatísticas subjetivas estão inevitavelmente sujeitas a controvérsias, pois, como ocorre frequentemente diferentes órgãos estimadores apresentam valores diferentes para o mesmo fenômeno.

O I.B.G.E. através do SEP só poderá sair deste impasse, e elaborar estatísticas universalmente aceitas, substituindo o atual método de estimação subjetiva pelo de estimação ob-

jetiva. Esta tarefa é de longo alcance e deveria ser iniciada por ocasião da safra do próximo ano.

O Estado de São Paulo tem em sua Secretaria da Agricultura longa experiência na obtenção de estatísticas agrícolas pelo método de amostragem, e ela talvez possa ser útil aos Estados irmãos.

Na Divisão de Economia Rural, não só a previsão safras mas qualquer pesquisa econômica ou sociológica é feita usando amostras, ou seja, o *método objetivo*. Assim, o sistema de amostragem torna-se o instrumento indispensável à obtenção de informações seja de previsão de safras, seja de custo de produção, estabelecimento de funções de produção, determinação do consumo nas áreas rurais, etc. O uso de amostra no Estado de São Paulo para a obtenção de informações referentes à agricultura está se generalizando porque ficou provada sua utilidade e exequibilidade. Não precisamos encarecer a importância de informações seguras em tempo hábil para a formulação de uma política agrícola ou a simples tomada de resoluções.

Nas atuais condições não se pode esperar que uma tarefa complexa e delicada como seja a obtenção de estatísticas agropecuárias através de um método objetivo, seja executado a contento para todo país, a partir de um centro único.

Por isso, sendo os Estados através de suas Secretarias de Agricultura os interessados mais diretos na obtenção des-

tas informações, parece justo que eles venham participar diretamente do esforço necessário à implantação da amostragem para estatísticas agrícolas.

Em nosso entender deveria ser estabelecida uma política de convênios entre I.B.G.E. e as Secretarias de Agricultura dos Estados, ficando assim divididos os ônus e as responsabilidades na obtenção destas estatísticas, sendo então mobilizados todos os recursos materiais e humanos disponíveis para este fim específico.

2 — Os responsáveis pelo programa de preços mínimos do Governo Federal bem como os que enfrentam o problema do abastecimento tem necessidade de conhecer a ordem de grandeza do volume de algodão, amendoim, arroz, feijão, milho e soja que serão produzidos na próxima safra.

O I.B.G.E. não tem montada a organização capaz de obter as informações objetivas e calcular com elas as previsões de safras necessárias. Tal organização não pode ser posta em funcionamento adequado em poucos meses.

Entretanto, a necessidade de informações está aí.

O Senhor Diretor do Departamento Econômico do Ministério da Agricultura nos propôs o problema da construção de u'a amostra de municípios que pudesse ser visitada por técnicos do Ministério da Agricultura e do Instituto Brasileiro de Geografia e Estatística, em um prazo limitado, a fim

de obter no local *estatísticas subjetivas revistas* da área cultivada e produção esperada para aquêles seis produtos.

Aceitamos a incumbência de construir tal amostra, embora inteiramente conscientes das limitações das estimativas subjetivas em geral, porque se nos afigura que este é no momento, a única maneira de *obter em tempo* uma idéia da ordem de grandeza da futura safra, idéia esta tão necessária aos que devem tomar decisões nas altas esferas governamentais. A nosso ver, a execução dêste trabalho revela mentalidade construtiva diante do problema das estatísticas agropecuárias.

Deve contudo ficar bem claro que a amostra proposta não se destina a ser usada permanentemente, pois, pouco se ganharia com isto. Ela deve sim, ser o passo no caminho certo que é o da construção do organismo capaz de obter informações objetivas sôbre a produção agrícola, calculá-las e publicá-las rapidamente.

3 — A fim de estudar as características estatísticas de municípios brasileiros no que se refere a área cultivada e produção dos seis produtos atrás enumerados, solicitamos ao SEP, o fornecimento dos últimos dados disponíveis.

Em meados de novembro recebemos as informações desejadas em cartões perfurados IBM. Foram feitos resumos por municípios contendo a área cultivada com os seis produtos durante o ano de 1963, a área total do município e fi-

nalmente a população rural segundo o censo.

Foi realizado um trabalho com esses dados obtendo-se as varias tabelas com o resultado da análise estatística dêstes dados por Estado da Federação. Temos aí os totais, médias, desvios padrões, matriz dos quadrados e duplos produtos e a matriz das correlações simples.

Estas tabelas foram calculadas no Centro de Cálculo Numérico da Escola Politécnica da Universidade de São Paulo, usando o computador 1620 IBM. Agradecemos a dedicação do Professor Isu Fang na execução dêste trabalho.

Essas tabelas mostram claramente quais os produtos cujas áreas cultivadas estão correlacionadas entre si, podendo em consequência ser calculado com u'a amostra única de municípios.

4 — Desta forma foram construídos os seguintes estratos de produtos:

Rio Grande do Sul — 1.º (arroz) — 2.º (feijão + milho) 3.º (soja);

Santa Catarina — 1.º (arroz) — 2.º (feijão + milho);

Paraná — 1.º (algodão) — 2.º (arroz + feijão + milho);

Mato Grosso — 1.º (algodão + amendoim + arroz + feijão + milho);

Goiás — 1.º (arroz + feijão + milho);

Minas Gerais — 1.º (algo-

ção) — 2.º (amendoim) — 3.º (arroz + feijão + milho);

Espírito Santo — 1.º (arroz + feijão + milho);

Rio de Janeiro — 1.º (algodão + arroz) — 2.º (feijão + milho).

Nos Estados de Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe e Bahia, construímos um estrato para a cultura de algodão dos nove Estados em conjunto, um estrato para a cultura de arroz nos Estados de Maranhão, Piauí, Ceará. Um estrato para arroz nos restantes Estados e um estrato para feijão e milho para todos os Estados. Ficamos assim, com quatro estratos para estes nove Estados.

5 — O passo seguinte se constitua na retirada da amostra de municípios que proporcionasse para cada estrato uma estimativa da área cultivada com erro-padrão não superior a 5%. Para isso os

municípios de cada estrato foram colocados em ordem crescente segundo a área. Dada a grande diversidade de área cultivada nos diferentes municípios, torna-se necessário sub-estratificar os estratos de produtos acima enumerados. Se quisermos uma alocação ótima da amostra, a fração de amostragem nos sub-estratos deve ser proporcional ao desvio-padrão correspondente, isto é:

$$f_1 \sim \sigma_1$$

Por outro lado, é máximo o ganho de eficiência resultante da estratificação, quando de cada sub-estrato são tirados apenas dois elementos, pois, então se constrói o número máximo de sub-estratos, todos eles tão uniformes quanto possível, distando as médias dos sub-estratos sucessivos da quantidade máxima permissível.

Considerando, para exemplificar apenas dois sub-estratos temos:

$$\frac{f_1}{\sigma_1} = \frac{f_2}{\sigma_2} \quad \therefore \quad \frac{\frac{n_1}{N_1}}{\sigma_1} = \frac{\frac{n_2}{N_2}}{\sigma_2}$$

fazendo $n_1 = n_2 = 2$

teremos $N_1\sigma_1 = N_2\sigma_2$

$$\therefore N_1^2 \sigma_1^2 = N_2^2 \sigma_2^2 \quad \therefore \quad N_1^2 \frac{\sigma_1^2}{2} = N_2^2 \frac{\sigma_2^2}{2} \quad \therefore$$

$$\therefore N_1^2 \frac{\sigma_1^2}{x_1} = N_2^2 \frac{\sigma_2^2}{x_2}$$

ou seja, são iguais as variâncias dos totais de todos sub-estratos.

Agradecemos ao Engenheiro Agrônomo Milton Nogueira de Camargo que sugeriu a construção sistemática de estratos uniformes seguindo êste caminho.

Sabemos então que os sub-estratos devem ter a mesma variância, sendo conveniente que cada um dêles contribua com dois elementos para amostra. Não sabemos contudo quantos sub-estratos serão necessários

para que a estimativa do total tenha êrro padrão igual a $(\xi.T)$ nem sabemos ainda como construir tais sub-estratos.

6 — Vejamos em primeiro lugar como devem ser construídos os sub-estratos.

Devendo o total T ser estimado com variância $(\xi.T)^2$ e havendo h sub-estratos, cada um dêles deverá contribuir para a variância do total com uma quantidade $(\xi.T)^2/h$.

Ora, a variância por unidade dentro de um sub-estrato é dada por

$$\sigma_h^2 = \frac{N_h \sum_1^{N_h} X_{hi}^2 - \left(\sum_1^{N_h} X_{hi}\right)^2}{(N_h)^2}$$

Por outro lado, a variância do total de um sub-estrato ob-

tido com u'a amostra de dois elementos é dada por

$$\sigma_{T_h}^2 = N_h^2 \frac{\sigma_h^2}{X_h} = \frac{N_h \sum_1^{N_h} X_{hi}^2 - \left(\sum_1^{N_h} X_{hi}\right)^2}{2}$$

Este valôr deverá ser igual ou menor a $(\xi T)^2/h$

Isto é:

$$N_h \sum_1^{N_h} X_{hi}^2 - \left(\sum_1^{N_h} X_{hi}\right)^2 \leq \frac{2 (\xi.T)^2}{h}$$

7 — O cálculo do número ótimo de sub-estratos, ou seja, do número de elementos necessários na amostra só poderá ser feito por tentativa, pois não se conhece a priori o ganho

de eficiência resultante da estratificação.

A quantidade de cálculos necessários a essa determinação é muito grande em nosso caso, e seria praticamente impossível executá-los a mão.

Solicitamos então a colaboração do Professor Waldemar Setzer que elaborou um programa em linguagem Fortram, com o qual foi calculado o número ótimo de sub-estratos, bem como seus limites.

O programa final usa um processo de *interação* partindo de cinco sub-estratos. Os cálculos foram feitos pelo Professor Setzer no Computador do Centro de Cálculo Numérico da Escola Politécnica da Universidade de São Paulo, e deixamos aqui consignados nossos agradecimentos ao Diretor e Técnicos deste Centro.

Obtidos os resultados da pesquisa do número ótimo de sub-estratos bem como seus limites, foi sorteada a amostra na lista dos municípios que constitui cada um dos estratos de produto.

8 — Finalmente, seguindo a técnica acima descrita foi retirada amostra baseada nos

dados de produção por Município, na safra de 1963, fornecidos pelo SEP.

Excluindo o Estado de São Paulo, fizemos o seguinte agrupamento de Estados:

a) Rio Grande do Sul, Santa Catarina e Paraná.

b) Mato Grosso, Goiás, Minas Gerais, Espírito Santo e Rio de Janeiro.

c) Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Sergipe, Alagoas e Baía.

Em cada um destes agrupamentos de Estado foi feita a ordenação dos Municípios segundo a produção de cada um dos seis produtos estudados, sendo em seguida construídos os sub-estratos que permitirão a estimativa do total da produção com erro padrão de cinco por cento.