

**ANÁLISE DE ERROS NÃO AMOSTRAIS EM LEVANTAMENTOS PARA PREVISÃO E
ESTIMATIVA DE SAFRAS DO ESTADO DE SÃO PAULO**

Francisco Alberto Pino, Denise Viani Caser

**Governo do Estado de São Paulo
Secretaria de Agricultura e Abastecimento
Coordenadoria Sócio-Econômica**

Instituto de Economia Agrícola



Governo do Estado de São Paulo
Secretaria de Agricultura e Abastecimento
Instituto de Economia Agrícola

ISSN 0101-5109
Relatório de Pesquisa
10/84

**ANÁLISE DE ERROS NÃO AMOSTRAIS EM LEVANTAMENTOS PARA PREVISÃO E
ESTIMATIVA DE SAFRAS DO ESTADO DE SÃO PAULO**

Francisco Alberto Pino
Denise Viani Caser

São Paulo
1984

ÍNDICE

1 - INTRODUÇÃO.....	1
1.1 - Levantamentos para Previsão e Estimativa de Sa- fras.....	2
1.2 - Objetivo.....	2
2 - METODOLOGIA.....	3
2.1 - Dados.....	3
2.2 - Comparações.....	4
2.3 - Tipos de Erros.....	5
3 - RESULTADOS E DISCUSSÃO.....	7
3.1 - Distribuição de Erros por Questionário.....	7
3.2 - Distribuição de Erros por Cargo ou Função do Enu- merador.....	10
3.3 - Distribuição de Erros por DIRA.....	10
3.4 - Distribuição por Tipo de Erro.....	15
3.5 - Soluções.....	19
4 - CONCLUSÕES.....	22
LITERATURA CITADA.....	24
RESUMO.....	25
SUMMARY.....	25

ANÁLISE DE ERROS NÃO AMOSTRAIS EM LEVANTAMENTOS PARA PREVISÃO E ESTIMATIVA DE SAFRAS DO ESTADO DE SÃO PAULO (1)

Francisco Alberto Pino

Denise Viani Caser

1 - INTRODUÇÃO

Nos levantamentos de dados de modo geral, e em particular nas pesquisas de opinião e nas pesquisas sócio-econômicas, tanto nos levantamentos por amostragem, como nos censitários, é inevitável a ocorrência de erros não amostrais, isto é, que não se devem ao esquema amostral, mas, ao próprio processo de levantamento dos dados. Tais erros podem ter sua origem em: a) causas técnicas, como um questionário mal formulado; b) causas circunstanciais, como diferenças regionais e dificuldades de acesso e transporte; e c) causas pessoais, tanto da parte do entrevistador como do informante, as quais podem ser resultante de qualificação inadequada, tendências pessoais, negligência, má fé, desconfiança, interesse nas respostas, vaidade, falta de registros e similares. Embora inevitáveis, tais erros podem, em geral, ser detectados e corrigidos a tempo, bem como seu estudo pode auxiliar a controlá-los e diminuí-los. O controle de qualidade das estatísticas é bem discutido por ZARKOVICH (9).

(1) Os autores agradecem a colaboração da Eng^o Agr^o Maria de Lourdes Barros Camargo.

1.1 - Levantamentos para Previsão e Estimativa de Safras

No Estado de São Paulo as estatísticas agrícolas, em especial a previsão e estimativa de safras, são elaboradas pelo Instituto de Economia Agrícola (IEA) e pela Coordenadoria de Assistência Técnica Integral (CATI). Os dados são levantados numa amostra de propriedades agrícolas, distribuídas em dez Divisões Regionais Agrícolas (DIRAs), conforme descrito em CAMPOS & PIVA (1). Antes da análise final e publicação, os dados dos questionários passam por um processo de depuração, descrito em PINO & JIMENEZ (6), que consiste em detectar possíveis erros através de testes de consistência interna e corrigi-los, se for o caso. Os resultados finais são também discutidos e comparados com dados de fontes diferentes.

Alguns estudos anteriores podem ser elementos auxiliares para o presente estudo. SCHATTAN (7) discute a questão das estatísticas agrícolas de modo geral. A qualidade das informações de preços médios recebidos pelos produtores de milho no Estado foi analisada por SENDIN & CARMO (8). PINO & CASER (5) analisaram o problema da falta de respostas nos levantamentos para previsão e estimativa de safras do Estado. Ainda quanto a tais levantamentos, existem um manual de elaboração de questionários (2), uma discussão sobre o uso das unidades de medida (3) e um estudo sobre mudanças nas unidades amostrais (4). Entretanto, embora os erros não amostrais sejam detectados e corrigidos em todos os levantamentos, até o momento não foi feito nenhum levantamento estatístico da extensão do problema nem uma análise que permitisse encontrar algumas de suas causas e estabelecer diretrizes com o intuito de controlá-los. É a isso que se propõe o presente trabalho.

1.2 - Objetivo

O objetivo deste trabalho é estudar os erros não amostrais comumente encontrados nos levantamentos por amostragem para previsão e estimativa de safras do Estado de São Paulo, estabelecendo tipos de erros, diferenças regionais e diferenças devidas aos enumeradores. Espera-se que os resultados deste estudo crítico dos levantamentos possam contribuir para o melhor conhecimento do problema dos erros não amostrais e para sua resolução.

2 - METODOLOGIA

2.1 - Dados

É mais ou menos irrelevante para o presente estudo qual o levantamento a ser considerado, exceto se se pretendesse analisar a tendência dos erros. Assim, utilizaram-se os dados do levantamento chamado objetivo, de abril de 1978, numa amostra de 5.646 propriedades agrícolas, das quais 5.136 foram consideradas na presente análise. O questionário constava de 271 itens, que tinham a finalidade de levantar a situação da safra dos principais produtos agrícolas, mão-de-obra, composição da população residente e da população trabalhadora, insumos, máquinas e implementos. Destes, 135 itens passaram pelos testes de consistência interna, classificando-se, a seguir, os erros encontrados de acordo com seu tipo, com o cargo ou função do enumerador (pessoa que entrevista o proprietário agrícola ou administrador ou outro que possa fornecer as respostas) e com a região do Estado onde foi levantado o questionário. A percentagem de erros em cada grupo foi calculada através de:

$$p = \frac{\text{nº de erros}}{\text{nº de questionários} \times 135} \times 100$$

2.2 - Comparações

Para a análise de variância os dados em percentagem foram transformados através de:

$$y = \text{arc sen } (p/100)^{1/2},$$

a fim de se contornar problemas de desvios da normalidade e estabilizar a variância.

Para comparar a percentagem de erros entre cargos e entre DIRAs, utilizou-se o seguinte modelo cruzado, com dois fatores fixos:

$$Y_{ij} = \mu + C_i + D_j + e_{ij},$$

onde:

Y_{ij} é a percentagem de erros no cargo i da DIRA j ;

μ é a média do modelo;

C_i é o efeito do cargo ou função i ($i = 1, 2, 3, 4$);

D_j é o efeito da DIRA j ($j = 1, 2, \dots, 10$) e

e_{ij} é o erro.

Os cargos foram agrupados em quatro grupos, conforme sua semelhança, por não se dispor de observações em todas as caselas:

- a) grupo 1: engenheiro agrônomo;
- b) grupo 2: auxiliar de engenheiro agrônomo, auxiliar de campo, auxiliar de médico veterinário e auxiliar agropecuário;
- c) grupo 3: vigilante sanitário, monitor, fiscal prod. agropecuário, inspetor, escriturário e apontador; e
- d) grupo 4: trabalhador braçal, operador de máquinas, mecânico, motorista e servente.

2.3 - Tipos de Erros

Os erros foram classificados em seis tipos, os quais são definidos a seguir.

a) Resposta incompleta

Significa o não preenchimento de um ou mais itens que compõem uma pergunta. Por exemplo, é informado um valor para a produção de arroz, mas, não para a área plantada. É óbvio que, se há produção, deve haver área plantada.

b) Falta de soma

Significa o não preenchimento do item referente à soma das parcelas. Exemplo:

Item	Certo	Errado
A	1	1
B	-	-
C	2	2
Total (soma)	3	-

c) Pergunta não compreendida

Significa que o enumerador e/ou o informante não compreendeu a pergunta, fornecendo, assim, uma resposta errada. Exemplo: pergunta-se "qual o valor total do arrendamento" (isto é, o preço de um alqueire multiplicado pela área arrendada) e no questionário é informado o preço de apenas um alqueire.

d) Erro de cálculo

Significa erro numa pergunta envolvendo cálculo. O mais comum é erro de soma (não confundir com o erro do tipo "falta de soma").

e) Erro de linha

Significa que o enumerador escreveu um valor correto, mas, em linha errada, como se fosse resposta de outro item. Exemplo:

Item	Certo	Errado
Área plantada com soja	10ha	10ha
Produção de soja	350sc.	- sc.
Área plantada com arroz	- ha	350ha
Produção de arroz	- sc.	- sc.

f) Resposta fora dos limites

Significa uma resposta que se encontra fora do intervalo admitido como razoável. O caso mais comum é o da produtividade ou rendimento agrícola. Em geral este erro é devido a um zero a mais ou a menos ou a uma vírgula decimal em posição incorreta. Exemplo:

Item	Certo	Errado
Área plantada com soja	10ha	10ha
Produção de soja	350sc.	3.500sc.

A experiência mostra que uma produção de soja de 35sc.60kg/ha é razoável, mas, não uma de 350sc.60kg/ha. É óbvio que houve um erro de um zero a menos na área plantada ou de um zero a mais na produção: a comparação com a área total do imóvel ou com levantamentos anteriores poderá esclarecer qual dos dois itens está errado.

3. RESULTADOS E DISCUSSÃO

3.1 - Distribuição de Erros por Questionário

A Figura 1, elaborada a partir da última linha do Quadro 1, mostra a distribuição do número de erros por questionário. Ocorreram 2.084 erros em 1.183 questionários, isto é, 23% dos questionários continham erros passíveis de correção, numa média de 1,8 erro por questionário com erro. Verifica-se, ainda, que mais de 90% dos questionários com erros apresentavam de 1 a 3 erros.

Por outro lado, de cada mil itens levantados, apenas três apresentavam erros (quadro 1). Embora este número seja pequeno, é suficiente para causar aumento nas variâncias das estimativas, porque este aumento depende mais da magnitude do erro que do número de erros. Daí a preocupação com esse tipo de controle e correção.

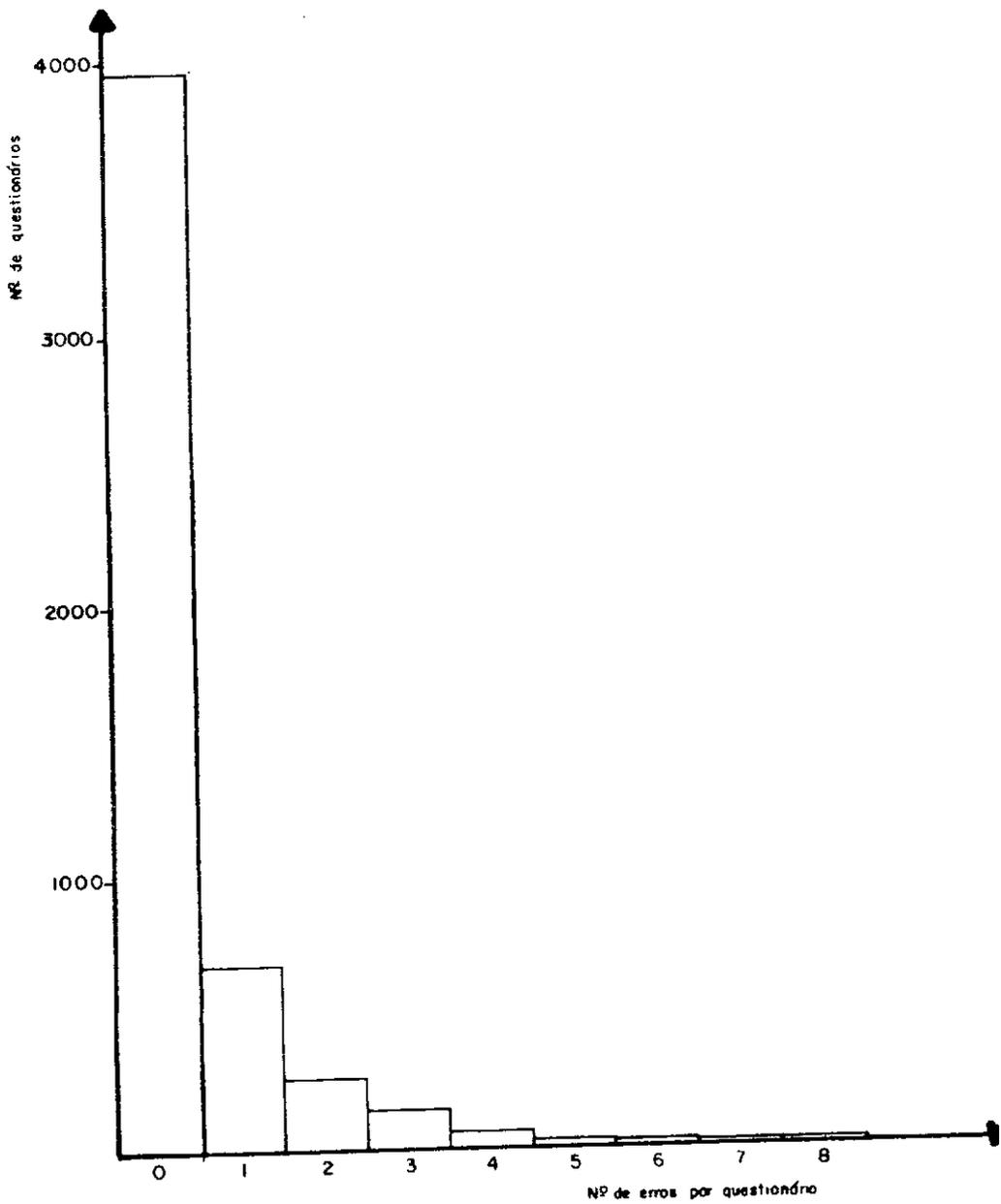


FIGURA 1.- Distribuição do Numero de Erros por Questionário.

QUADRO 1. - Distribuição do Número de Erros por Questionário de Acordo com o Cargo ou Função do Enumerador

Cargo ou função	Número de erros por questionário										Questionários		Porcentagem de erros por item
	0	1	2	3	4	5	6	7	8	Número	%		
Auxiliar de engenheiro agrônomo	2.697	444	161	79	34	14	3	2	1	3.435	66,9	0,27	
Auxiliar de campo	392	95	33	16	7	3	2	-	-	548	10,7	0,36	
Engenheiro agrônomo	289	48	31	14	6	4	-	2	1	395	7,7	0,41	
Auxiliar de médico veterinário	223	28	10	10	7	-	-	-	1	279	5,4	0,30	
Vigilante sanitário	47	6	2	1	1	-	-	-	-	57	1,1	0,22	
Monitor	31	9	3	7	2	1	-	-	-	53	1,0	0,68	
Trabalhador braçal	32	1	3	2	1	-	-	-	-	39	0,8	0,32	
Fiscal prod. agropecuário	30	4	-	1	1	-	-	-	-	36	0,7	0,22	
Operador de máquinas	27	3	1	1	-	-	-	-	-	32	0,6	0,18	
Mecânico	21	4	1	1	-	-	-	-	-	27	0,5	0,25	
Auxiliar agropecuário	9	6	3	1	1	-	-	-	-	20	0,4	0,70	
Inspetor	11	1	5	2	-	-	1	-	-	20	0,4	0,85	
Motorista	12	5	-	-	-	-	-	-	-	17	0,3	0,22	
Escriturário	14	-	-	-	-	-	-	-	-	14	0,3	0,00	
Servente	10	3	-	-	-	-	-	-	-	13	0,3	0,17	
Apontador	6	3	1	1	-	-	-	-	1	12	0,2	0,99	
Não especificado	102	22	9	3	2	1	-	-	-	139	2,7	0,33	
Total	3.953	682	263	139	62	23	6	4	4	5.136	100,0	0,30	

Fonte: Instituto de Economia Agrícola (IEA) e Coordenadoria de Assistência Técnica Integral (CATI).

3.2 - Distribuição de Erros por Cargo ou Função do Enumerador

Encontraram-se enumeradores de 16 cargos ou funções diferentes, alguns, evidentemente, não qualificados para o trabalho de enumerador (quadro 1). A maior parte dos questionários foram preenchidos por auxiliares de engenheiro agrônomo (66,9%), seguidos de auxiliares de campo (10,7%), engenheiros agrônomos (7,7%) e auxiliares de médico veterinário (5,4%).

A análise de variância não mostrou diferenças significativas entre os quatro grupos de cargos ou funções dos enumeradores quanto à percentagem de erros (Quadro 2), por causa de variações regionais dentro de cada cargo, como será visto no item 3.3 deste trabalho. Infelizmente, não foi possível analisar cada cargo ou função separadamente, devido ao número insuficiente de observações, o que, certamente, contribuiu para este resultado não significativo. Mesmo assim, classificando-os quanto à percentagem de erros, verifica-se que, em média, apontador (0,99%), inspetor (0,85%) e auxiliar agropecuário (0,70%) apresentavam os piores desempenhos, enquanto escriturários (0,00%), servente (0,17%) e operador de máquinas (0,18%) apresentaram os melhores desempenhos. O significado de tal classificação não é muito grande porque essas seis categorias de enumeradores que se encontram nos extremos foram responsáveis por menos de 2% do número de questionários, encontrando-se a maioria mais próxima da média.

3.3 - Distribuição de Erros por DIRA

A análise de variância mostra que há diferença significativa entre as DIRAs quanto à percentagem de erros (quadro 2). De fato, a percentagem de erros nos questionários das DIRAs de Campinas (0,50%), Ribeirão Preto (0,42%) e Sorocaba (0,35%) é sensivelmente maior que nas outras (Quadro 3).

QUADRO 2. - Análise de Variância da Percentagem Transformada de Erros por Cargo ou Função do Enumerador e por Divisão Regional Agrícola

Fonte de variação	Graus de liberdade	Soma de quadrados	Quadrado médio	F	H ₀	Conclusão ⁽¹⁾
Cargo	3	11,4603	3,8201	1,2749	$C_1 = \dots = C_4$	n.s.
DIRA	9	68,2121	7,5791	2,5294	$D_1 = \dots = D_{10}$	sig. 5%
Resíduo	27	80,9019	2,9964			
Total	39	160,5743				

⁽¹⁾ n.s. indica não significativo ao nível de 5%, e sig. 5% indica significativo ao nível de 5%.

QUADRO 3. - Distribuição do Número de Erros por Questionário de Acordo com a Divisão Regional Agrícola

Divisão Regional Agrícola (DIRA)	Número de erros por questionário									Questionários		Porcentagem de erros por item
	0	1	2	3	4	5	6	7	8	Número	%	
São Paulo	353	27	16	7	4	2	-	1	2	412	8,0	0,23
Vale do Paraíba	299	37	11	5	1	3	-	-	-	356	6,9	0,19
Sorocaba	454	68	32	15	15	5	1	-	2	592	11,5	0,35
Campinas	417	133	57	32	19	2	2	2	-	664	12,9	0,50
Ribeirão Preto	419	102	29	35	11	5	1	1	-	603	11,7	0,42
Bauru	425	57	21	8	1	-	1	-	-	513	10,0	0,19
São José do Rio Preto	509	87	40	12	5	1	1	-	-	655	12,8	0,26
Araçatuba	305	48	16	9	-	2	-	-	-	380	7,4	0,22
Presidente Prudente	341	54	17	9	1	2	-	-	-	424	8,3	0,22
Marília	431	69	24	7	5	1	-	-	-	537	10,5	0,22
Estado	3.953	682	263	139	62	23	6	4	4	5.136	100,0	0,30

Fonte: Instituto de Economia Agrícola (IEA) e Coordenadoria de Assistência Técnica Integral (CATI).

Por outro lado, também varia a composição do quadro de enumeradores conforme a região (quadro 4). A DIRA Marília usou o menor número de cargos ou funções diferentes de enumeradores (cinco), enquanto as DIRAs com maiores percentagens de erros foram também aquelas com maior diversidade de enumeradores: Sorocaba (nove), Campinas (nove) e Ribeirão Preto (sete). Em todas as DIRAs predominou o auxiliar de engenheiro agrônomo como enumerador: de um mínimo de 40% em Campinas a um máximo de 78% em Ribeirão Preto e em Presidente Prudente. Além dessa categoria, predominaram auxiliar de campo e engenheiro agrônomo (São Paulo, Campinas, Ribeirão Preto e Bauru); auxiliar de campo e auxiliar de médico veterinário (Sorocaba e Presidente Prudente); auxiliar de médico veterinário e vigilante sanitário (Vale do Paraíba); auxiliar de médico veterinário (São José do Rio Preto); auxiliar de campo e operador de máquinas (Araçatuba); auxiliar de campo, mecânico e auxiliar agropecuário (Marília). Somente duas DIRAs utilizaram engenheiro agrônomo em bom número, São Paulo (34%) e Campinas (20%), enquanto que duas não o utilizaram, São José do Rio Preto e Marília. A não especificação de cargo ou função do enumerador foi grande somente na DIRA São José do Rio Preto (14,5%).

Pode-se verificar, ainda, que as maiores percentagens de erros ocorreram com vigilante sanitário na DIRA Ribeirão Preto (1,85%), engenheiro agrônomo na DIRA Araçatuba (1,30%), auxiliar de campo na DIRA São José do Rio Preto (1,08%) e trabalhador braçal na DIRA Sorocaba (1,05%).

Algumas das causas prováveis para as diferenças significativas entre as DIRAs quanto à percentagem de erros são as seguintes:

- a) quantidade e qualificação dos enumeradores. Algumas DIRAs podem ter elementos mais qualificados ou em número mais adequado para executar o levantamento no campo que outras;
- b) infra-estrutura e recursos. Algumas DIRAs podem ter melhor infra-estrutura e recursos materiais que outras;

QUADRO 4. - Distribuição dos Questionários Preenchidos (Porcentagem em Relação ao Total da DIRA) e de Erros (Porcentagem em Relação ao Número de Questionários Preenchidos) de acordo com o Cargo ou Função do Enumerador

Cargo ou função	Divisão Regional Agrícola (1)									
	São Paulo	Vale Paraíba	Sorocaba	Campinas	Rib. Preto	Bauru	S.J.R. Preto	Araçatuba	Pres. Prudente	Marília
Auxiliar de engenheiro agrônomo	51,21 (0,22)	73,87 (0,22)	58,61 (0,32)	40,51 (0,46)	78,27 (0,41)	74,46 (0,16)	77,09 (0,26)	64,73 (0,22)	78,30 (0,19)	75,98 (0,21)
Auxiliar de campo	8,25 (0,20)	-	10,97 (0,20)	16,26 (0,42)	11,44 (0,45)	9,16 (0,52)	1,67 (1,08)	22,63 (0,26)	9,66 (0,54)	16,20 (0,21)
Engenheiro agrônomo	33,98 (0,30)	1,96 (0,11)	7,77 (0,43)	20,48 (0,52)	3,64 (0,51)	7,40 (0,31)	-	1,05 (1,30)	0,47 (0,00)	-
Auxiliar de médico veterinário	2,42 (0,00)	9,27 (0,22)	15,37 (0,42)	5,27 (0,68)	3,48 (0,18)	-	6,71 (0,08)	3,68 (0,05)	7,31 (0,24)	-
Vigilante sanitário	-	9,55 (0,00)	-	3,16 (0,42)	0,33 (1,85)	-	-	-	-	-
Monitor	-	-	-	7,98 (0,68)	-	-	-	-	-	-
Trabalhador braçal	-	-	2,02 (1,05)	-	-	5,26 (0,00)	-	-	-	-
Fiscal prod. agropecuário	3,64 (0,00)	-	1,68 (0,22)	1,65 (0,54)	-	-	-	-	-	-
Operador de máquinas	-	-	-	-	0,82 (0,44)	-	-	7,10 (0,14)	-	-
Mecânico	-	-	1,18 (0,32)	-	-	-	-	-	-	3,72 (0,22)
Auxiliar agropecuário	-	-	-	-	-	-	-	-	-	3,72 (0,70)
Inspetor	-	-	-	3,01 (0,85)	-	-	-	-	-	-
Motorista	-	-	-	-	-	-	-	-	4,01 (0,22)	-
Escriturário	-	-	-	-	-	2,72 (0,00)	-	-	-	-
Servente	-	3,65 (0,17)	-	-	-	-	-	-	-	-
Apontador	-	-	2,02 (0,99)	-	-	-	-	-	-	-
Não especificado	0,48 (0,00)	1,68 (0,00)	0,33 (0,00)	1,65 (0,54)	1,99 (0,93)	0,97 (0,00)	14,50 (0,30)	0,78 (0,00)	0,23 (0,00)	0,37 (0,00)

(1) A percentagem de erros está entre parênteses.

Fonte: Instituto de Economia Agrícola (IEA) e Coordenadoria de Assistência Técnica Integral (CATI).

- c) atividades agrícolas e infra-estrutura da região. As atividades agrícolas específicas e a infra-estrutura sócio-econômica de cada região podem facilitar ou dificultar o levantamento no campo. Por exemplo, pode ser mais fácil levantar um pequeno número de grandes unidades com monocultura que um grande número de pequenas unidades em região de policultura. Além disso, as grandes distâncias e a precariedade das estradas podem dificultar o trabalho;
- d) extensão territorial. Alguns municípios de grande extensão territorial podem ter maiores dificuldades para executar o levantamento;
- e) nível de instrução e receptividade dos agricultores. As DIRAs podem diferir quanto a este item, o que pode facilitar ou dificultar o levantamento;
- f) entrosamento IEA/CATI. O entrosamento entre os técnicos das duas instituições pode, eventualmente, diferir entre as DIRAs.

O tipo de erro predominante em cada DIRA, a ser apresentado no item 3.4, também pode ajudar a esclarecer o assunto.

3.4 - Distribuição por Tipo de Erro

A distribuição por tipo de erro encontra-se no quadro 5 (por cargo ou função do enumerador) e no Quadro 6 (por DIRA). Verifica-se que 27% dos erros foram devidos a respostas incompletas, 25% à falta de soma, 17% a perguntas não compreendidas pelo enumerador e/ou informante, 14% a erros de cálculo, 9% a erros de linha e 8% a respostas fora dos limites. Os erros de respostas incompletas e de falta de soma, que em geral se devem à falta de atenção do enumerador, constituem mais da metade dos erros (algumas respostas incompletas podem ser devidas, também, ao fato de o informante não saber dar

QUADRO 5. - Distribuição do Número de Erros por Cargo ou Função do Enumerador e por Tipo de Erro

Cargo ou função	Tipo de erro						Total
	Resposta incompleta	Falta de soma	Pergunta não compreendida	Erro de cálculo	Erro de linha	Resposta fora dos limites	
Auxiliar de eng. agrônomo	360	293	186	194	97	119	1.249
Auxiliar de campo	67	67	50	43	24	13	254
Engenheiro agrônomo	63	73	30	20	19	13	188
Auxiliar de médico veterinário	28	23	22	13	14	14	154
Vigilante sanitário	4	6	4	2	1	-	17
Monitor	8	17	14	3	4	3	49
Trabalhador braçal	-	3	5	2	6	1	17
Fiscal prod. agropecuário	1	4	-	2	4	-	11
Operador de máquinas	3	1	3	1	-	-	8
Mecânico	-	4	4	-	1	-	9
Auxiliar agropecuário	3	7	9	-	-	-	19
Inspetor	1	16	4	1	1	-	23
Motorista	4	1	-	-	-	-	5
Escriturário	-	-	-	-	-	-	-
Servente	-	-	2	-	1	-	3
Apontador	4	-	10	-	-	2	16
Não especificado	19	4	10	19	4	6	62
Total	565	519	353	300	176	171	2.084

Fonte: Instituto de Economia Agrícola (IEA) e Coordenadoria de Assistência Técnica Integral (CATI).

QUADRO 6. - Distribuição do Número de Erros por Divisão Regional Agrícola e por Tipo de Erro

Divisão Regional Agrícola (DIRA)	Tipo de erro						Total
	Resposta incompleta	Falta de soma	Pergunta não compreendida	Erro de cálculo	Erro de linha	Resposta fora dos limites	
São Paulo	19	54	17	17	13	9	129
Vale do Paraíba	14	17	14	14	10	24	93
Sorocaba	60	83	49	39	41	12	284
Campinas	110	159	71	54	28	33	455
Ribeirão Preto	111	83	65	24	16	48	347
Bauru	45	14	28	26	15	5	133
São José do Rio Preto	71	45	16	61	19	22	234
Araçatuba	47	16	16	24	5	9	117
Presidente Prudente	42	16	29	19	17	6	129
Marília	46	32	48	22	12	3	163
Estado	565	519	353	300	176	171	2.084

Fonte: Instituto de Economia Agrícola (IEA) e Coordenadoria de Assistência Técnica Integral (CATI).

a resposta). Quase um quinto dos erros devem-se a perguntas não compreendidas, o que pode resultar de falhas na redação do questionário ou da falta de qualificação do enumerador ou da falta de registros do informante. Os erros de cálculo podem ser devidos a falhas na elaboração do questionário (solicitando cálculos complexos) ou à falta de qualificação do enumerador. Os erros de linha podem ser devidos à falta de atenção do enumerador ou a falhas na programação visual do questionário. Finalmente, as respostas fora dos limites representam menos de 10% dos erros. Estas podem se dever à falta de cooperação do informante ou ao fato de o enumerador não ter visitado a unidade amostral ou à falta de atenção (ou de qualificação) do enumerador.

Alguns esclarecimentos podem ser obtidos do exame do quadro 5. O erro por resposta incompleta foi importante no caso de auxiliar de engenheiro agrônomo, auxiliar de campo, auxiliar de médico veterinário, operador de máquinas, motorista e cargo não especificado, seguindo-se engenheiro agrônomo, vigilante sanitário e apontador. O erro por falta de soma foi importante para engenheiro agrônomo, vigilante sanitário, monitor, fiscal, mecânico e inspetor, seguidos de auxiliar de engenheiro agrônomo, auxiliar de campo, auxiliar de médico veterinário, auxiliar agropecuário e motorista. O erro devido à pergunta não compreendida foi importante no caso de operador de máquinas, mecânico, auxiliar agropecuário, servente e apontador, seguindo-se vigilante sanitário, monitor, trabalhador braçal e inspetor. Os erros de cálculo foram importantes somente no caso de cargo não especificado. Os erros de linha foram importantes para trabalhador braçal e fiscal, seguidos de servente. Finalmente, os erros devidos a respostas fora dos limites não foram importantes em nenhuma categoria.

Esclarecimentos semelhantes podem ser obtidos do exame do quadro 6. O erro por resposta incompleta foi importante nas DIRAS Ribeirão Preto, Bauru, São José do Rio Preto, Araçatuba e Presidente Prudente, seguindo-se

São Paulo, Sorocaba, Campinas e Marília. O erro por falta de soma foi importante nas DIRAs São Paulo, Sorocaba e Campinas, seguidas de Vale do Paraíba e Ribeirão Preto. O erro devido à pergunta não compreendida foi importante na DIRA Marília, seguida de Bauru e Presidente Prudente. Os erros de cálculo mostraram-se de alguma importância apenas nas DIRAs São José do Rio Preto e Araçatuba. Os erros de linha não se mostraram importantes em nenhuma DIRA, enquanto os erros devidos a respostas fora dos limites foram importantes somente na DIRA Vale do Paraíba. É interessante notar que os tipos de erro apareceram de forma nitidamente regionalizada (figura 2). É possível separar as DIRAs em quatro grupos:

- a) DIRAs onde predominaram os erros por resposta incompleta e falta de soma: São Paulo, Sorocaba, Campinas e Ribeirão Preto;
- b) DIRAs onde predominaram os erros por resposta incompleta e pergunta não compreendida: Bauru, Marília e Presidente Prudente;
- c) DIRAs onde predominaram os erros por resposta incompleta e os erros de cálculo: São José do Rio Preto e Araçatuba; e
- d) DIRA onde predominaram os erros devidos a respostas fora dos limites e por falta de soma: Vale do Paraíba.

3.5 - Soluções

Ao contrário dos erros amostrais, onde medidas de ordem técnica, como mudanças no plano amostral, podem resolver o problema, os erros amostrais apresentados neste trabalho podem ser combatidos principalmente com medidas administrativas. Do exposto até aqui fica evidente que muitos erros poderiam ser evitados, bastando para isso que se utilizassem enumerado

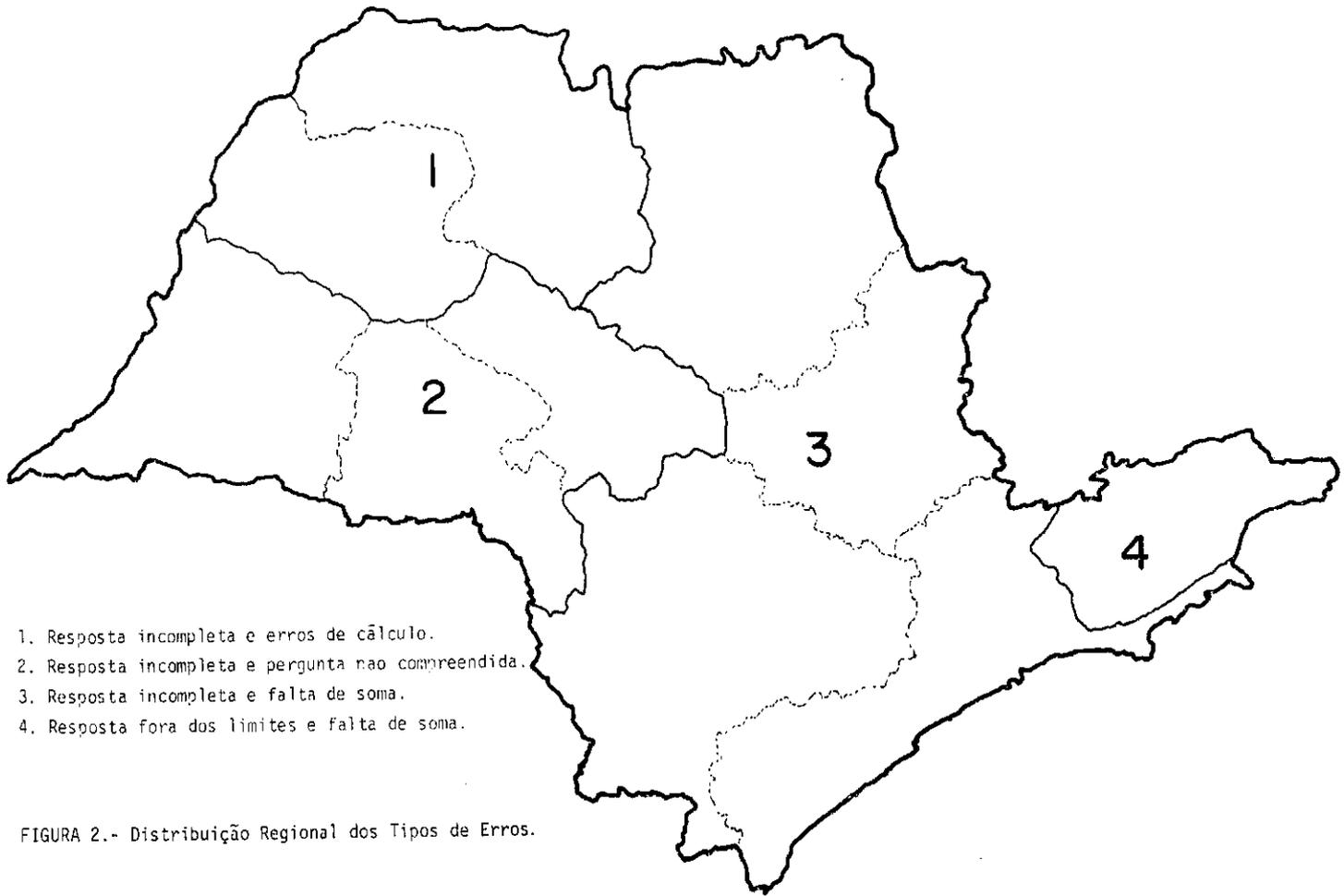


FIGURA 2.- Distribuição Regional dos Tipos de Erros.

res adequadamente qualificados e treinados. Entenda-se por qualificação do enumerador o grau e o tipo de sua instrução. Existe alguma controvérsia a esse respeito: alguns acreditam que somente pessoal de nível superior (em especial, engenheiro agrônomo) deveria ser utilizado como enumerador, enquanto outros defendem a idéia de que o pessoal de nível médio com curso técnico (em especial, técnico agrícola) é adequado, desde que receba o necessário treinamento e seja supervisionado pelo pessoal de nível superior. Um argumento a favor da primeira idéia é que o enumerador deve ter senso crítico diante das respostas do informante, bem como deve saber conseguir algumas respostas pela observação direta no local, a fim de compará-las com as dadas pelo informante. Um argumento contrário é o maior custo. Quanto ao treinamento do enumerador, entenda-se: a) um curso ou contato inicial para informá-lo a respeito das finalidades do levantamento, do significado de cada questão, da importância do levantamento e da contribuição do enumerador, de como localizar as unidades amostrais e das técnicas de abordagem a serem utilizadas durante a entrevista; e b) o contato periódico para esclarecimento de dúvidas e fluxo de informações em ambas as direções.

Uma parte das dificuldades pode ser devida à falta de recursos humanos e materiais. Entretanto, embora seja verdade que as Casas da Agricultura da CATI muitas vezes tenham dificuldades para conseguir funcionários e veículos para executar os levantamentos nas épocas devidas, também é verdade que o pessoal de escritório do IEA, que elabora e analisa os levantamentos, às vezes desconhece a realidade do campo e/ou as dificuldades encontradas para se realizar as entrevistas. Do mesmo modo, os enumeradores às vezes desconhecem os propósitos e a importância do levantamento, quando a motivação é fundamental para um serviço bem feito.

O assunto tem sido discutido em diversas ocasiões, mas, ainda se encontram muitas resistências. Em 21/11/77, por exemplo, Resolução do

Secretário de Agricultura criou um grupo de trabalho, integrado por técnicos do IEA e da CATI, para estudar os levantamentos sócio-econômicos. Este grupo recomendou a criação de uma equipe de enumeradores adrede treinados, com os necessários veículos, cuja única, ou pelo menos principal, atribuição fosse o levantamento de informações. Tal equipe seria constituída de pessoal de nível médio supervisionado por pessoal de nível superior, todos lotados no interior do Estado e subordinados ou ao IEA ou à CATI (para evitar duplicidade de comando).

As sugestões dos autores, baseadas no presente estudo, encontram-se nas conclusões.

4 - CONCLUSÕES

No levantamento por amostragem para previsão e estimativa de safras do Estado de São Paulo, elaborado pela Secretaria de Agricultura, em abril de 1978, ocorreram três itens com erro não amostral em cada mil itens levantados. Detectaram-se diferenças regionais quanto à percentagem de erros, mas, não entre grupos de cargos ou funções dos enumeradores. Quase 80% dos erros não amostrais deveram-se principalmente à falta de atenção. As sugestões dos autores sobre medidas de ordem administrativa para diminuir os erros não amostrais são as seguintes:

a) Formar uma equipe de enumeradores:

- devidamente qualificados e treinados;
- com a função precípua de executar este e outros levantamentos sócio-econômicos da Secretaria de Agricultura e Abastecimento;
- com a devida estrutura orgânica;

- com os recursos necessários (veículos, diárias, etc.);
- b) Reformar a equipe de elaboração e análise dos levantamentos:
 - com pessoal devidamente qualificado e treinado;
 - com os recursos necessários (computação, datilografia, desenho, divulgação dos resultados, arquivamento dos dados, etc.);
 - com métodos de trabalho mais dinâmicos que os atuais;
- c) Manter as duas equipes em permanente contato:
 - com fluxo ativo de informações em ambas as direções;
 - com permanente motivação; e
- d) Reformular, em forma e conteúdo, os questionários (recebendo, inclusive, sugestões dos enumeradores).

As alterações devem ser feitas tendo em mira:

- a) minimizar a ocorrência de erros não amostrais;
- b) facilitar ao máximo o trabalho de campo;
- c) conciliar a maximização da precisão dos dados obtidos com a minimização do número de elementos da amostra (e, portanto, do custo);
- d) facilitar ao máximo o trabalho de depuração dos dados e da elaboração das tabelas de previsão e estimativa de safras;
- e) sofisticar a análise dos resultados; e
- f) minimizar o tempo gasto entre o questionário ser preenchido e a informação final chegar aos usuários.

LITERATURA CITADA

1. CAMPOS, Humberto & PIVA, Luiz H.O. Dimensionamento de amostra para estimativa e previsão de safra no Estado de São Paulo. Agricultura em São Paulo, SP, 21(3):65-88, 1974.
2. PINO, Francisco A. Elaboração de questionário para levantamentos de campo. São Paulo, Secretaria de Agricultura e Abastecimento, IEA, 1981. 30p. (Boletim Interno, 01/81)
3. _____ . Unidades de medida na agricultura. São Paulo, Secretaria de Agricultura e Abastecimento, IEA, 1981. 13p. (Boletim Interno, 02/81)
4. PINO, Francisco A.; CAMARGO, Maria de L.B.; VIANI, Denise V. Mudanças de posse e tamanho das propriedades agrícolas do Estado de São Paulo no período de 1972-77. São Paulo, Secretaria de Agricultura, IEA, 1979. 15p. (Relatório de Pesquisa, 16/79)
5. PINO, Francisco A. & CASER, Denise V. Falta de resposta em levantamentos por amostragem: um estudo de caso. São Paulo, Secretaria de Agricultura e Abastecimento, IEA, 1984. (Relatório de Pesquisa, 08/84)
6. PINO, Francisco A. & JIMENEZ OSSIO, Julio H. Um método para a depuração de erros não amostrais em dados obtidos por levantamento em campo. In: REUNIÃO ANUAL DA SOCIEDADE BRASILEIRA DE ECONOMIA RURAL, 13., Curitiba, 1975. São Paulo, SOBER, 1977. p.409-410.
7. SCHATTAN, Salomão. Aprimoramento das estatísticas agrícolas no Brasil. Agricultura em São Paulo, SP, 18(9/10):69-84, 1971.
8. SENDIN, Paulo V. & CARMO, Maristela S. Análise da qualidade das informações dos preços médios recebidos pelos produtores de milho no Estado de São Paulo, 1969. Agricultura em São Paulo, SP, 17(7/8):1-17, 1970.
9. ZARKOVICH, S.S. Calidad de los datos estadísticos. Roma, FAO, 1968. 142p.

RESUMO

Este trabalho estuda a ocorrência de erros não amostrais nos levantamentos para previsão e estimativa de safras do Estado de São Paulo, tipificando-os e verificando diferenças regionais e devidas ao enumerador.

NON-SAMPLING ERROR IN THE SAMPLE SURVEY FOR HARVESTS ESTIMATION IN THE STATE OF SÃO PAULO

SUMMARY

This paper analyses the occurrence of non-sampling errors in the sample survey for harvests estimation in the State of São Paulo, Brazil, describing types, regional differences and differences due to interviewers.

**SECRETARIA DE AGRICULTURA E ABASTECIMENTO
INSTITUTO DE ECONOMIA AGRÍCOLA**

Comissão Editorial:

Coordenador: José Roberto Viana de Camargo

Membros: Antonio Augusto Botelho Junqueira
Celuta Moreira Cesar Machado

Elcio Umberto Gatti

Flavio Condé de Carvalho

José Luis Teixeira Marques Vieira

Rosa Maria Pescarin Pellegrini

Bibliografia: Fátima Maria Martins Saldanha Faria

Centro Estadual da Agricultura
Av. Miguel Estéfano, 3900
04301 - São Paulo - SP

Caixa Postal, 8114
01000 - São Paulo - SP
Telefone: 275-3433 r. 257



Governo do Estado de São Paulo
Secretaria de Agricultura e Abastecimento
Coordenadoria Sócio-Econômica

Instituto de Economia Agrícola

Relatório de Pesquisa
Nº 10/84