

COMISSÃO "B"
DOCUMENTO OFICIAL

A APLICAÇÃO DA TÉCNICA DA AMOSTRAGEM NOS LEVANTAMENTOS
CONTÍNUOS E A PESQUISA NACIONAL POR AMOSTRA
DE DOMICÍLIOS

Autor: CARLOS MARCOS BARBOSA - Dire
tor de Levantamentos Estatísticos,
respondendo pelo Grupo Executivo
de Pesquisas Domiciliares.

"A aplicação da Técnica da Amostragem
nos Levantamentos Contínuos e
A PESQUISA NACIONAL POR
AMOSTRA DE DOMICÍLIOS"

Autor: Carlos Marcos Barbosa - Diretor
de Levantamentos Estatísti--
cos, respondendo pelo Grupo
Executivo de Pesquisas Domi-
ciliares

Rio de Janeiro, Guanabara

1968

SUMÁRIO

A Demanda de Estatísticas Básicas. Aplicação da Amostragem Probabilística. Origem das Pesquisas Domiciliares na América Latina. Implantação das Pesquisas no Brasil. Desenho da Amostra. Operações de Campo. Assuntos de Investigação. Programa de Trabalho. Aplicação da Amostragem à outros Setores. Dificuldades e Implicações.

INTRODUÇÃO

Uma das principais características do mundo de hoje é a contínua descoberta e aplicação de novas e aperfeiçoadas técnicas em quase todos os setores da vida humana, objetivando o desenvolvimento das nações e o bem-estar dos povos. No campo das ciências, por exemplo, conseguiram-se verdadeiras maravilhas; a produção industrial teve um desenvolvimento sem par. A todo momento, novas matérias primas que servirão para a elaboração dos produtos mais refinados são sintetizadas e passam a ter uma aplicação incalculável. Surgem novas pesquisas em todos os campos e, ainda assim, o homem não se dá por satisfeito na sua avidês por maiores e melhores conhecimentos. Os engenhos de guerra foram aperfeiçoados a tal ponto que passaram a constituir uma constante ameaça à conservação da espécie humana. Verifica-se, de um a outro extremo do Globo Terrestre, uma febre incontida de conhecimentos.

O desenvolvimento sócio-econômico dos povos, não se fez e não poderá ser planejado sem um profundo conhecimento dos fenômenos básicos que afetam a vida humana. É necessário que os potenciais sejam descobertos, e desenvolvidos racionalmente. E somente através desse conhecimento, poderão os planejadores, qual quer que seja sua área de ação, traçar os programas de desenvolvimento, acompanhá-lo no tempo e traçar as medidas de correção e aperfeiçoamento que se fizerem necessárias.

É nêsse particular que a Estatística, que desde as épocas mais remotas tem sido um poderoso instrumento de investigação, cada vez mais se destaca como elemento de diagnóstico, estimativa e previsão de todo e qualquer aspecto da vida moderna. Cada vez mais os levantamentos contínuos se impõem como um aliado indispensável à complementação dos levantamentos censitários, pois não se trata mais de investigar situações cuja mutação se processa lentamente, mas de acompanhar fenômenos cuja velocidade de transformação impõe um acompanhamento contínuo, a intervalos bem mais curtos do que os que se vinham utilizando.

É sumamente importante que o Brasil, país em franco desenvolvimento e que inicia a escalada que o colocará entre as nações mais desenvolvidas em futuro não muito distante, obtenha um aperfeiçoamento estatístico que acompanhe "pari-passu" o desenvolvimento dos demais setores da vida nacional.

2. A DEMANDA DE ESTATÍSTICAS BÁSICAS

A despeito dos muitos esforços que se têm empreendido, a crescente demanda de informações estatísticas básicas não tem encontrado, de forma conveniente, nos levantamentos tradicionais, a resposta adequada e isso se dá em razão de peculiaridades de nosso país, tais como, grande extensão territorial, dificuldades de transportes e meios de comunicação, grande variabilidade de situações de uma para outra região sócio-econômica, etc. o que ocasiona tempo de coleta demasiadamente longo, agravado ainda por lacunas em levantamentos de setores primordiais dado o complexo de várias situações.

Grande parte de nossas estatísticas ainda são levantadas por um sistema implantado há muitos anos, com a utilização de inquéritos anuais ou mensais, de investigação universal e, assim mesmo, com lacunas em setores de atividade que se tornaram específicos e cujo conhecimento imediato se torna indispensável. Diga-se de passagem, que esse sistema teve sua época áurea e que foi, por muitos anos o melhor que se poderia ter utilizado. Foi, a seu tempo, plenamente satisfatório no atendimento das necessidades estatísticas do país. Atualmente porém, é unanimemente reconhecida a necessidade de modificá-lo adaptando-o às necessidades do momento, que são bem maiores que as de outrora.

3. APLICAÇÃO DA AMOSTRAGEM PROBABILÍSTICA

(B1)

Há algum tempo já se vem sentindo a necessidade de uma conjugação dos esforços de todos os setores interessados, visando a obtenção de resultados que cubram o maior número possível de aspectos para os quais não existam informações atualizadas. A resposta para o problema, a exemplo do que foi conseguido em outras nações, encontra-se na aplicação da Técnica da Amostragem em larga escala e, diga-se aqui, a bem da justiça, que o IBGE tem sido o pioneiro nessa tomada de consciência, como sempre procedeu no campo estatístico.

Embora vários órgãos integrantes do Sistema Estatístico Brasileiro, já de algum tempo, venham desenvolvendo esforços no sentido da aplicação da Amostragem, os resultados até agora alcançados, embora animadores, ainda não são suficientes para o atendimento das necessidades estatísticas do Brasil. Vale sublinhar que o investimento inicial da implantação da Amostragem em larga escala é bem caro e, os lucros provenientes dessa inversão só começam a aparecer após um período não inferior a dois anos,

tempo mínimo necessário para a organização das equipes, treinamento de pessoal, testes de esquemas, preparação adequada do "frame", planejamento, revisão das amostras, preparo psicológico dos informantes, planejamento das apurações e aferição dos resultados, etc.

O Ministério do Planejamento, órgãos regionais, departamentos e secretarias de Estado, bem como vários outros setores da economia nacional, têm mostrado interesse quanto ao preenchimento das lacunas na obtenção de dados estatísticos básicos e também, quanto ao desenvolvimento e melhoria das estatísticas. Não menos que todos, a Fundação IBGE que recebeu o acervo de mais de trinta anos do principal produtor de estatísticas, o IBGE, não poderia deixar de demonstrar sua flexibilidade na implantação de novos métodos e técnicas que visem obter meios com os quais se possam atender a crescente demanda de informações estatísticas.

A principal característica dos levantamentos por Amostragem reside na possibilidade da obtenção de estimativas rápidas e índices de desenvolvimento, com uma precisão calculável. Podemos citar como exemplo do que acabamos de afirmar, o fato de que, tendo-se iniciado a aplicação da Amostragem às Estatísticas Industriais em 1964, quando se constatava um atraso de seis anos no fornecimento de dados, já em 1967 se possuíam resultados perfeitamente atualizados, com uma divulgação trimestral e a possibilidade da manutenção dessa atualização a períodos ainda menores. Conseguiu-se cobrir, com precisão comprovadamente boa, todo o período 1964-67. Já se dispõe, desde abril, de dados gerais para 1967 e índices atualizados até o 1º trimestre de 1967. Não se pode entretanto, deixar de levar em conta que o setor das Estatísticas Industriais já possuía uma longa tradição em levantamentos estatísticos o que facilitou a aceleração do êxito. Apesar disso os esforços desenvolvidos pela direção do IBGE, para vencer o problema, foram enormes.

4. AS PESQUISAS DOMICILIARES - Suas origens na América e principais aspectos e objetivos.

O êxito da experiência de países como os Estados Unidos e o Canadá, em pesquisas contínuas de população e habitação veio de encontro aos anseios da Aliança para o Progresso quanto ao estabelecimento e manutenção de levantamentos que proporcionassem meios para o planejamento do desenvolvimento sócio-econômico. Essa feliz associação culminou com a apresentação, por solicitação da ONU e outras organizações internacionais, de um modelo de pesquisas domiciliares aplicável à América Latina, modelo esse elaborado por técnicos do Bureau dos Censo dos E.U.A. e apresentado ao I Seminário Latino-Americano de Pesquisas Domiciliares, rea

lizado na cidade do México, em abril de 1962. Embora elaborado por solicitação da Comissão de Estatística da ONU com a finalidade de atender aos objetivos da Aliança para o Progresso, o modelo citado tem um largo campo de aplicação. Convenientemente adaptado às necessidades do Brasil poderá ser a solução para muitos dos nossos problemas estatísticos.

As pesquisas domiciliares constituem um sistema de investigação por amostra de domicílios desenhadas com propósitos múltiplos. Elas permitem a investigação de uma série de características correlacionadas com as características da população e da habitação.

O esquema padrão, conhecido com a denominação de "Estudo de Caso Atlântida" e que vem sendo adaptado ao Brasil registra como principais tópicos de investigação os seguintes assuntos:

- Características demográficas;
- Saúde e higiene;
- Consumo de alimentos e nutrição;
- Condições e meios de habitação;
- Educação e cultura;
- Mão-de-obra; e
- Nível econômico da família.

Êsses tópicos e suas subdivisões são classificados em três grupos, de acôrdo com a periodicidade aconselhada: assuntos básicos (contínuos), assuntos complementares (de periodicidade variável) e assuntos especiais (que são incluídos apenas em determinadas ocasiões).

Um outro aspecto importante das pesquisas domiciliares é que elas são levantamentos por amostragem que funcionam em combinação com os levantamentos censitários que servem para a aferição de sua precisão. A aplicação das pesquisas domiciliares permite igualmente, que em alguns casos, se reduza a profundidade de investigação dos censos demográficos, o que os torna menos dispendiosos.

Já é considerável o número de países de todo o mundo que utilizam as pesquisas domiciliares em seus levantamentos contínuos, quer adotando o esquema elaborado para a América Latina, quer utilizando processos equivalentes. Podemos citar entre êles o Canadá, os Estados Unidos, o México, a Bolívia, o Brasil, a Colômbia, Costa Rica, o Chile, o Equador, Honduras, República Dominicana, Panamá, Peru, Trinidad e Tobago, Venezuela, Alemanha, França, Índia, Paquistão, Tailândia e Turquia, cada um dêles com maior ou menor grau de experiência e profundidade, de acôrdo com os recursos disponíveis. Todos êsses e outros mais, reconhecem nas pesquisas domiciliares um poderoso instrumento de investiga--

ção das condições sócio-econômicos de suas populações.

5. IMPLANTAÇÃO DAS PESQUISAS DOMICILIARES NO BRASIL

Como foi dito anteriormente, a aplicação da Amostragem em larga escala é o único meio racional disponível para o provimento de estatísticas básicas de forma rápida, com boa precisão e a um custo razoável. Com sua aplicação, os levantamentos completos, muitas vezes calcados em critérios subjetivos, passariam gradativamente a dar lugar a investigações de melhor qualidade, proporcionando ainda, que a economia de tempo e custo que se obtivesse fosse aplicada em outras investigações ainda não iniciadas e de importância, dentro de uma escala de prioridades no plano nacional e posteriormente, no plano regional.

Os estudos preliminares para a aplicação das pesquisas domiciliares ao Brasil tiveram seu início praticamente, em princípios de 1966, quando se efetivaram os primeiros entendimentos com os técnicos do Bureau dos Censos dos E.U.A., órgão elaborador do "Estudo de Caso Atlântida". A partir daquele momento, uma equipe de técnicos da então Secretaria-Geral do Conselho Nacional de Estatística, passou, com a ajuda da assistência técnica obtida através da USAID/Brasil, a estudar todos os documentos básicos referentes ao assunto e a considerar a forma mais rápida para a adaptação do sistema às necessidades do Brasil quanto à obtenção de estatísticas básicas. Cabe aqui um pequeno parêntese para lembrar que uma das primeiras preocupações do IBGE, ontem como hoje, resume-se na definição e no estabelecimento de um programa mínimo de estatísticas básicas, no que tem sido seguido e apoiado pelos principais produtores e consumidores estatísticos.

Durante o ano de 1966 e sob a vigência do primeiro convênio de mútua cooperação assinado entre o IBGE, o Ministério do Planejamento, a USAID/Brasil, o Conselho Técnico da Aliança para o Progresso e, o Ponto IV, considerou-se inicialmente, o estabelecimento de um programa de pesquisas contínuas por amostra de domicílios e se decidiram as possíveis dimensões e assuntos a serem incluídos no programa. Dedicaram-se vários meses ao exame e familiarização dos técnicos brasileiros com todo o material do LASH (Latin American Household Survey), publicado sob o título "Atlântida - Um Estudo de Caso sobre Pesquisas Domiciliares - setembro de 1966".

Ao mesmo tempo em que o grupo de trabalho criado na Secretaria-Geral do Conselho Nacional de Estatística para o estudo em foco, se identificava com o processo, procediam-se as tradu

ções em português e as necessárias adaptações do plano visando atender as necessidades brasileiras. Foram considerados não só os aspectos de utilidade da pesquisa, mas também os problemas operacionais que sua implantação envolveria.

No período de outubro-novembro de 1966 realizou-se nos municípios fluminenses de Nova Iguaçu e Rio Bonito, uma pesquisa experimental tendo como finalidade o teste dos conceitos, formulários, implicações de campo e todos os demais aspectos da pesquisa que se desejava implantar. Foram testados os mapas municipais, os mapas censitários de 1960 e todos os instrumentos censitários relacionados com a localização, descrição e conteúdo dos setores censitários, a fim de se verificar se poderiam ser utilizados satisfatoriamente, como elementos básicos para a cadastração e seleção das unidades de amostra. Testaram-se, igualmente, três processos alternativos para a listagem dos domicílios e seleção da amostra. Também os manuais elaborados para o treinamento dos supervisores e pesquisadores foram submetidos ao teste.

Foram realizadas as operações de contagem-rápida, listagem e enumeração em dez setores de Rio Bonito e trinta de Nova Iguaçu. Os seguintes formulários de coleta foram utilizados:

- PNAD-1 - questionário de mão-de-obra;
- PNAD-300 - Fôlha de registro do domicílio;
- PNAD-303 - Fôlha de listagem;
- PNAD-304 - Fôlha de listagem de estruturas não residenciais;
- PNAD-305 - Fôlha de listagem em domicílio coletivo;
- PNAD-306 - Pasta do conglomerado; e
- PNAD-602 - Questionário de Migração-Interna.

A experiência obtida com o teste em muito serviu para o desenvolvimento dos trabalhos da pesquisa em caráter definitivo.

5.1 A PESQUISA NACIONAL

A partir de janeiro de 1967 passou-se a estudar os aspectos de implantação de uma pesquisa nacional de caráter contínuo. Foi necessária a tomada de decisões fundamentais estreitamente relacionadas com as áreas de estimação, o tamanho da amostra e os assuntos de investigação, além de outras medidas indispensáveis. O ideal teria sido, provavelmente, o estabelecimento de uma amostra selecionada com uma taxa de amostragem uniforme para todo o conjunto, o que proporcionaria a obtenção de estimativas regionais com o mesmo grau de precisão para um conjunto de itens. Pensava-se inicialmente na seleção de uma amostra nacional com fração de amostragem global igual a 1/1000.

Levando-se em conta que as estimativas da população projetadas para 1966 era de 84 milhões de habitantes, ou seja, aproximadamente 17 milhões de domicílios, o tamanho da amostra teria sido de cerca de 17 000 unidades domiciliares. As considerações passaram então, a versar sobre os requisitos mínimos necessários a uma amostra de tamanho mínimo para cinco agrupamentos regionais resultando em propostas de amostras que variavam de 16 a 22 mil domicílios.

Em vista porém, da extensão territorial do Brasil e da grande variabilidade de situações sócio-econômicas de uma para outra região e, tendo-se em conta que muitos itens de investigação serão incluídos em diferentes pesquisas, justifica-se plenamente que se considerasse a possibilidade de estimativas regionais e estimativas independentes para algumas unidades separadamente.

Assim, após o exame dos recursos disponíveis, dos problemas técnicos, da experiência do teste e das necessidades estatísticas, o IBGE decidiu iniciar o desenvolvimento de um programa detalhado para sete regiões (inicialmente se pensava em seis), programa esse a ser implantado gradativamente durante um período aproximado de um ano.

Em geral, cada região da pesquisa terá, inicialmente, uma amostra de no mínimo 4.000 domicílios necessários aos cálculos de variância, à determinação do erro alheio à amostragem e ao provimento da base para as futuras modificações e aumentos. As regiões definidas para a Pesquisa foram as seguintes:

<u>REGIÃO</u>	<u>UNIDADES DA FEDERAÇÃO</u>
I	Guanabara e Rio de Janeiro;
II	São Paulo;
III	Paraná, Santa Catarina e Rio Grande do Sul;
IV	Minas Gerais e Espírito Santo;
V	Maranhão, Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe e Bahia; (&)
VI	Brasília;
VII	Demais Unidades da Federação.

(&) mais conhecida como área da SUDENE.

Os estudos comparativos entre a população recenseada em 1960 e as estimativas da população projetadas para 1 de julho de 1967 revelaram os seguintes números:

REGIÃO	RECENSEAMENTO DE 1960		ESTIMATIVA PARA 1967	
	População (1000 hb)	Domicílios (1000 dom)	População (1000 hb)	Domicílios (1000 dom)
I	6 710	1 342	8 347	1 669
II	12 975	2 595	16 081	3 216
III	11 873	2 375	15 761	3 152
IV	11 372	2 274	13 354	2 671
V	22 428	4 486	25 601	5 120
VI	142	28	358	72
VII	5 467	1 093	7 078	1 416
BRASIL	70 967	14 193	86 580	17 316

Estabeleceram-se diferentes frações de amostragem, su-
jeitando-as, entretanto, a um mínimo de 1/800, a fim de se obter
uma amostra de pelo menos 4 000 domicílios para cada região. (No-
te-se que ainda não se completaram os estudos para as Regiões VI
e VII). Se a taxa de crescimento demográfico para qualquer re-
gião da Pesquisa fôr, a partir de 1960, menor ou maior que as de-
monstradas pela estimativa da população para 1967, a amostra re-
sultante será ligeiramente menor ou maior. Observe-se que se to-
mou a média de 5 pessoas por domicílio em 1967, podendo-se, por-
tanto, esperar que a amostra final seja maior que a fixada se, a
média real de pessoas por domicílio, fôr menor que 5, conforme de-
monstrou a experiência na Guanabara, Estado do Rio de Janeiro e
Estado de São Paulo.

O quadro seguinte nos fornece uma primeira idéia da
distribuição da amostra segundo as regiões da Pesquisa.

REGIÃO	ESTIMATIVA DA POPULAÇÃO (1967) (1000 hb)	FRAÇÃO GLOBAL DE AMOSTRAGEM	TAMANHO DE AMOSTRA ESPERADO (domicílios)
I	1 669	1/400	4 172
II	3 216	1/800	4 020
III	3 152	1/800	3 900
IV	2 671	1/600	4 450
V	5 120	1/800	6 400
VI	72	(-)	(-)
VII	1 416	(-)	(-)
BRASIL	17 316		28-29000

(-) Ainda não se definiram as especificações para Brasília e de-
mais Unidades da Federação. É possível que a Região VI ve-
nha a ser constituída de Brasília e algumas cidades de Goiás, e
que também, a Região VII seja limitada às principais cidades em
virtude das características especiais da mesma.

Tomando-se em conta a estimativa da população para 1967, o tamanho de amostra esperado deverá ser de 28 000 domicílios. Se, entretanto aquelas estimativas fôrem inferiores à população real, naquela data, podemos esperar uma amostra nacional de cêrca de 30 000 domicílios.

6. ESTRATIFICAÇÃO E DESENHO DA AMOSTRA

O método utilizado na estratificação e seleção da amostra baseou-se no material do Censo Demográfico de 1960 e em informações disponíveis acêrca do clima, precipitação de chuvas, tipo de agricultura, tipo e concentração da indústria e informações fisiográficas sôbre os municípios ou pequenas regiões.

A seleção da amostra é realizada em etapas, tendo o município como primeiro estágio de seleção. Em 1960 havia 2 767 municípios e, tendo em vista que os dados censitários e as outras informações utilizadas na estratificação estavam relacionados com as áreas municipais de 1960, êles foram estabelecidos como unidade primária de seleção, a despeito do aparecimento de novos municípios no período 1960-67.

As unidades secundárias de seleção são os setores censitários de 1960, considerados satisfatòriamente definidos. Havia em 1960, 57 913 setores censitários com uma média de 1 225 habitantes por setor. Os mapas dos setores censitários, as cadernetas dos setores contendo os mapas, o número de domicílios, a população e, as listagem dos setores censitários foram, na sua totalidade dos casos, localizados no Serviço Nacional de Recenseamento e provaram servir perfeitamente para os trabalhos da Pesquisa. Em face porém, da variabilidade em tamanho e do crescimento da população daqueles setores após 1960, chegou-se à conclusão de que os mesmos eram, em geral, demasiadamente grande para serem adotados como área de listagem - terceira etapa de seleção. Decidiu-se por conseguinte, adotar a operação de "contagem-rápida" que consiste na subdivisão da área dos setores da amostra em áreas menores (sub-setores) de tamanho convenientemente adequado ao terceiro estágio de seleção.

Resumindo, o plano utilizado adotou uma amostra de quatro estágios de seleção - município, setor censitário, área de listagem e domicílios. O quadro seguinte mostra a distribuição dos setores censitários de 1960, segundo as regiões da Pesquisa.

REGIÃO	NÚMERO DE MUNICÍPIOS (1960)	NÚMERO DE SETORES CENSITÁRIOS (1960)
I	62	4 843
II	503	10 754
III	414	9 651
IV	521	10 389
V	902	17 688
VI	365	4 107
VII		4 481
BRASIL	2 767	57 913

Houve necessidade, em razão da variabilidade em tamanho e composição entre os municípios e setores, de se combinarem, em alguns casos, unidades simples, visando a obtenção de pseudo-unidades baseando-se para tanto em duas condições principais: Nas regiões I e II os municípios com população inferior a 10 000 foram combinados, na maioria dos casos, com municípios contíguos; nas regiões III, IV e V, esse limite foi elevado para 12 000. Da mesma forma os municípios que apresentavam uma proporção de população urbana e de atividade não-agrícola em relação à atividade agrícola muito diferentes do conjunto em que se situavam, foram combinados para produzir uma pseudo-unidade mais homogênea e em relação às demais unidades do estrato. Em alguns casos também se procedeu ao grupamento de setores, quando esses possuíam menos de 60 domicílios.

As informações utilizadas na formação dos estratos dos municípios que não foram considerados auto-representativos (-) foram as seguintes:

- população do censo de 1960;
- proporção da população urbana em relação à população total;
- valor e tipo da principal atividade agrícola;
- classificação fisiográfica e grupamento.

Os estratos foram constituídos com tamanho igual a um nível de variação de cerca de 5% em termos da população de 1960. Houve casos, entretanto, em que se fugiu a essa norma para facilitar o grupamento de unidades muito heterogêneas. Em uns poucos casos, foram também construídos estratos de tamanhos 1,5 em relação aos demais. Nesses casos selecionaram-se 3 pseu-

(-) Ver, no final do capítulo, a definição de município auto-representativo.

do-municípios, em lugar dos 2 normalmente selecionados de cada estrato.

As etapas do desenho da amostra que vem sendo executado podem ser resumidas como se segue:

1. Designação da fração global de amostragem, dentro da região;
2. Designação dos municípios auto-representativos;
3. Estratificação dos demais municípios;
4. Seleção de 2 pseudo-municípios de cada estrato normal e 3 de cada estrato de tamanho 1,5;
5. Seleção de 5 setores em cada município não-auto-representativo e, de acordo com as necessidades, em cada município auto-representativo;
6. Seleção de 2 subsetores ou áreas de listagem em cada setor de amostra;
7. Seleção dos domicílios dentro de cada área de listagem, segundo a taxa necessária para manter a fração de amostragem pré-determinada. (O número médio esperado é 5).

No tocante ao item 7. o esquema ficou ligeiramente modificado na Região I, onde o número de domicílios esperados para os municípios não auto-representativos é 100, em vez de 50, como ocorreu nas demais regiões já trabalhadas.

A seleção da amostra nos três primeiros estágios é feita com probabilidade proporcional do tamanho, para os municípios; proporcional ao número de domicílios existentes em 1960, para os setores e; proporcional ao número de domicílios existentes em 1967, para as áreas de listagem.

Apresentamos nos dois quadros seguintes algumas informações sobre a estratificação nas regiões já iniciadas.

REGIÃO	TAMANHO MÉDIO DOS ESTRATOS (Pop. 1960)	NÚMERO DE ESTRATOS	NÚMERO DE MUNICÍPIOS (1960)	FRAÇÃO GLOBAL DE AMOSTRAGEM
I	334 440	5	62	1/400
II	302 996	23 (1)	503	1/800
III	299 003	30	414	1/800
IV	232 174	34 (6)	521	1/600
V	331 625	49 (6)	902	1/800

(-) O número inscrito entre parêntesis representa a quantidade de estratos de tamanho $1 \frac{1}{2}$.

REGIÃO	NÚMERO DE MUNICÍPIOS SELECIONADOS NA AMOSTRA (Pseudo-Municípios)			% DA POPULAÇÃO DE 1960 NOS MUNICÍPIOS AUTO-REPRESENTATIVOS
	Total	Auto-Representativos	Não-Auto Represent.	-
I	18	8	10	75,1
II	65	18	47	45,1
III	81	21	60	24,4
IV	90	16	74	21,8
V	137	33	104	23,1

Foram considerados como municípios auto-representativos aqueles que possuíam, em 1960, uma população acima do corte estabelecido em cada região e também, aqueles que tiveram um crescimento acentuado a partir de 1960, possuindo uma atividade econômica acima do comum da região ou também, aqueles que fazem parte de uma grande área urbana ou de área metropolitana.

7. OPERAÇÕES DE CAMPO

As operações de campo são executadas em três etapas sequenciais: contagem-rápida, listagem e enumeração.

A operação de contagem rápida consiste na subdivisão dos setores da amostra em sub-setores que, em média, possuem cerca de 100 domicílios cada um. Essa operação é procedida após reconhecimento dos limites externos e da situação interna dos setores. Procedida a subdivisão, selecionam-se, aleatoriamente, dois sub-setores que passam a constituir as áreas de listagem, passando-se em seguida à operação de listagem, que nada mais é que a cadastração completa e meticulosa de todas as unidades domiciliares existentes na área de Listagem.

A partir da Listagem os domicílios da amostra são selecionados, aleatoriamente, com uma taxa calculada para manter a fração global de amostragem estabelecida para a Região.

Passa-se, então, à operação da enumeração que consiste no processamento das entrevistas. A fim de garantir uma representação uniforme de todo o trimestre (as entrevistas são de periodicidade trimestral) o painel de informantes selecionados é distribuído aleatoriamente em 13 subconjuntos, passando-se a investigar um conjunto por semana de enumeração.

O plano prevê ainda, a substituição gradativa dos

informantes, de modo a se obter uma melhor representação do universo, além de reduzir os problemas da manutenção do mesmo corpo de informantes por tempo indeterminado.

Visando a redução do volume de trabalhos, lançou-se mão de um dispositivo de sub-amostragem para manter o tamanho médio dos conglomerados em torno de 6 unidades.

Esse dispositivo prevê a investigação de todas as unidades selecionadas se o conglomerado fôr igual ou inferior a 10 domicílios; 1/2 do número de unidades no caso de um conglomerado de 11 a 15 domicílios; 1/3 para os conglomerados de 16 a 20 domicílios e 1/4 para os de 21 a 25 domicílios.

8. SUB-UNIVERSOS ESPECIAIS

Duas sub-amostras complementares deverão ser desenhadas na intenção de melhorar a precisão das estimativas: a sub-amostra das áreas de novas construções e a sub-amostra das favelas.

Partindo-se da premissa de que a base da pesquisa repousa nos elementos de C.D. de 1960 e levando em conta o incremento domiciliar no período 1960-1967, principalmente em áreas que foram totalmente transformadas com a construção de loteamentos, vilas operárias, e outros planos habitacionais, torna-se necessário o delineamento das áreas de novas construções para a montagem desse sub-universo e o consequente desenho de uma amostra que o represente, o que infelizmente ainda não foi executado.

Outro aspecto especial, em relação à precisão da amostra é o que se refere às populações de favelas, mocambos e outros grupamentos similares, de características sócio-econômicas diferentes do conjunto das demais áreas de investigação.

Até agora o levantamento especial das favelas só vem sendo executado na Guanabara onde a disponibilidade dos elementos de estudo foi bem acessível. Espera-se porém poder desenvolver esquemas semelhantes em todas as regiões da pesquisa.

9. ASSUNTOS DE INVESTIGAÇÃO

Selecionaram-se para aplicação na Pesquisa, durante seus dois primeiros anos, quatro dos sete principais tópicos e suas divisões aconselhadas no Estudo de Caso Atlântida para pesquisas domiciliares. Investigam-se durante a implantação da

Pesquisa as características básicas da população, características básicas da habitação, mão-de-obra e migração-interna.

Como se poderá verificar no quadro de possíveis temas, anexado ao final deste documento, os assuntos de investigação foram classificados em três grupos: básicos (contínuos), suplementares e especiais, de acordo com seu grau de importância e a periodicidade necessária.

Além dos assuntos em utilização no momento, existem pelo menos dois que nos parecem merecer um grau de prioridade - maior que os demais - renda familiar e estatísticas vitais.

10. TREINAMENTO DAS EQUIPES

Como nos referimos anteriormente, o programa de implantação vem sendo aplicado em forma sequencial com intervalos de aproximadamente dois meses entre uma e outra região.

Foram implantados até agora os trabalhos da Região I, II, III e IV, estando em início de execução do treinamento a Região V que envolve nove Unidades da Federação situadas na área da SUDENE.

O treinamento se processa em três níveis diferentes, com a seleção e treinamento dos coordenadores, supervisores e pesquisadores, preparação essa executada em duas etapas consecutivas - contagem rápida e listagem, a primeira e enumeração, a segunda.

As operações de campo têm suporte na rede de coleta da Fundação IBGE suplementada eventualmente por pesquisadores que não pertencem aos quadros do Instituto. A existência da rede de coleta tem sua finalidade na execução dos levantamentos estatísticos de caráter nacional, bem como levantamentos regionais compreendidos na Convenção Nacional de Estatística. Essa rede é constituída de 2.185 Agências Municipais de Estatística que são supervisionadas e coordenadas pelas Inspetorias Regionais em número de 26, sendo uma para cada Unidade da Federação. Essa rede permanente de coleta estatística possui aproximadamente 4.000 servidores dos quais, cerca de 2.500 funcionam no trabalho de campo executado pelas agências municipais.

Os trabalhos da Pesquisa são executados pelos servidores da rede, sem solução de continuidade para os encargos normais. Entretanto, naquelas áreas onde o número de servidores é insuficiente, permite-se a utilização de pesquisadores eventuais que são recrutados e treinados nas áreas de amostra.

A existência de uma rede de coleta montada e em pleno funcionamento, foi uma das principais razões, senão a primeira, que animou a equipe de técnicos a desencadear o programa que ora se implanta com um caráter muito mais ambicioso do que o recomendável para o início de um programa de pesquisas domiciliares. Seguindo-se os processos adotados em outros países iniciados nesse campo há mais de dez anos, ou as recomendações do modelo, só atingiríamos o estágio atual depois de um mínimo de quatro anos de experiência.

Convém ressaltar aqui, que nossa intenção não foi um mero rasgo de vaidade em sobrepujar outros planos, mas tão-somente, o resultado da exploração máxima dos meios disponíveis tendo em vista as prementes necessidades do país no tocante a investigações básicas para as quais não existem informações.

Ainda que não se possam apresentar, no momento, os resultados estatísticos dos primeiros experimentos, o programa de treinamento empreendido provou ser de grande utilidade para a melhoria e qualidade do pessoal da rede de coleta e também, para o seu preparo visando não só, a utilização de futuras investigações através da Pesquisa, mas também, outros tipos de levantamentos por amostragem, além de se traduzir em ótimo início de preparação psicológica e técnica para o próximo recenseamento geral de 1970.

11. PROCESSAMENTO INTERNO

Não menos intenso treinamento e, com dificuldades ainda maiores, vem sendo aplicado à equipe de processamento interno que se monta construída, paulatinamente, na sede do IBE, com o concurso de recursos humanos emprestados dos seus vários setores técnicos. O Grupo Executivo de Pesquisas Domiciliares conta, atualmente, com cerca de 80 servidores selecionados dentre o pessoal técnico da administração central, reforçado por um pequeno grupo de auxiliares sob o Regime de C.L.T., cedido pelo Serviço Nacional de Recenseamento, a título de colaboração.

As principais fases de processamento interno que vêm sendo desenvolvidas pelo Grupo são: controle da remessa e coleta do material das pesquisas, verificação da cobertura; especificação dos questionários a serem reproduzidos, em razão da utilização de sub-amostras; levantamento do número e tipo das não-entrevistas para efeito de ajustamentos possíveis (caso de não

entrevista em domicílio ocupado); crítica e codificação; revisão; perfuração de cartões, conferência de listagens e; tabulação.

A tabulação vem sendo processada em equipamento convencional IBM. Entretanto, já se iniciou o preparo das programações para um computador IBM-1401.

12. PLANO TABULAR E DE DIVULGAÇÃO

O modelo das Pesquisas Domiciliares fornece subsídios para a elaboração de cerca de 80 tabelas de apuração e divulgação, relacionadas com a investigação das características da população e habitação e, da mão-de-obra. Entretanto, o plano tabular que vem sendo adotado experimentalmente prevê a possibilidade de um grupo principal de 47 tabelas, envolvendo os seguintes aspectos principais:

1. Características da população;
2. Características da habitação;
3. Grupos familiares;
4. Mão-de-obra.

Em se tratando de uma amostra nacional com sub-amostras regionais que investigam um número muito grande de características, só se têm, a priori, uma idéia do erro de amostragem para as principais variáveis que se espera não ultrapasse a faixa de 2% de erro. No entanto, a determinação dos erros de todas as características só poderá ser obtida após um estudo completo (dos erros de amostragem e alheios a ela) a ser realizado com base nas computações obtidas.

Assim, no plano tabular experimental, procurou-se incluir o maior número de cruzamentos possíveis, levando-se em conta o interesse estatístico dessas tabulações. É fora de dúvida, entretanto, que um bom número de células se apresentará com baixa frequência e precisão, o que implicará num estudo dos fatores de controle para essas variáveis conjuntas.

Visando a melhoria da precisão das amostras que como verificam os anteriormente, foram desenhadas com um tamanho mínimo necessário à análise das variâncias.

Esses e outros estudos proporcionarão os elementos indispensáveis para o aumento do tamanho da amostra com vistas à obtenção das precisões desejadas.

O plano de divulgação, por conseguinte, será fruto de exames detalhados das tabulações obtidas.

Como nos referimos em parágrafo anterior, estão sendo processados os necessários entendimentos para a programação das tabulações, em computador IBM-1401. A fase de análise dessas tabulações envolve o exame dos dados comuns a várias tabelas no sentido da minimização do volume e complexidades das passagens na máquina, crítica de consistência e incompatibilidades.

A título de ilustração podemos citar alguns dos aspectos de grande interesse que podem ser obtidos das investigações que se vêm executando através da Pesquisa:

1. População residente;
2. Domicílios ocupados;
3. Grupos familiares;
4. Potencial de mão-de-obra (pessoas de 14 anos e mais);
5. População ocupada;
6. População desempregada ou desocupada;
7. Horas trabalhadas;
8. Pessoas ocupadas em regime de tempo parcial;
9. População inativa e reserva de trabalho.

13. PROCESSOS DE ESTIMATIVAS

Até que se disponham das distribuições atualizadas de sexo-idade da população, necessárias à análise demográfica e adoção de estimativas de razão, continuar-se-á a adotar o processo de estimativas simples.

Dessa forma os resultados iniciais apresentarão um nível de erro de amostragem superior ao que seria obtido com o auxílio de boas informações suplementares.

14. PROGRAMA DE TRABALHO

O programa de trabalhos que o grupo pretende desenvolver, em seguida, envolve operações como as que registramos a seguir:

1. Prosseguimento do programa de pesquisas trimestrais para as Regiões I, II, III e IV;
2. Conclusão da fase de treinamento e início da execução contínua da Pesquisa para Região V;
3. Conclusão dos estudos e estabelecimento de um programa modificado para a Região VI;

4. Estudo e desenho de uma amostra especial para a implantação da Pesquisa em Brasília;
5. Estabelecimento de um sistema de controle de qualidade para as operações de campo e, outro para as operações de processamento-interno, incluindo a perfuração mecânica;
6. Delineamento do universo das áreas de novas construções e estabelecimento do universo das Unidades especiais (instituições), visando melhorar a eficiência da amostra e conseqüentemente a precisão das estimativas;
7. Desenvolvimento do programa para o estudo dos dados em computador-eletrônico, de maneira a garantir métodos de estimação mais precisos, tendo em vista a precisão dos resultados;
8. Cálculo das variâncias da amostra, estudo dos erros alheios à amostragem e tendenciosidade;
9. Estudo dos custos das várias fases da operação, visando o desenvolvimento de processos de estimativa para orçamento e custos futuros;
10. Revisão e preparação do conjunto de todos os manuais de instrução e operação, formulários, documentos, etc.;
11. Planejamento dos detalhes de inquéritos futuros, bem como exame dos problemas relacionados com a inclusão de itens adicionais para as investigações atuais, e ainda, o estabelecimento de uma amostra para as investigações especiais ou de caráter limitado;
12. Estabelecimento de um sistema de contato com os principais usuários, através a CONPLANE, no sentido de investigar suas necessidades, tipo de informações que necessitem, nível de precisão e outros detalhes, para os programas futuros;
13. Estudo da utilização das informações do Censo Demográfico de 1970, visando o redesenho da amostra para o período subsequente.

15. OUTRAS POSSIBILIDADES DO SISTEMA

Além das possibilidades que procuramos demonstrar nos parágrafos anteriores, há outras tantas que a implantação das pesquisas domiciliares por Amostra podem oferecer para o desenvolvimento das estatísticas básicas do país.

O estabelecimento do "frame", por exemplo, constituído da listagem dos domicílios dos setores da amostra, "frame" esse que pode ser atualizado periodicamente, fornece elementos para o desenho e seleção de amostras especiais de âmbito nacional ou restrita a áreas selecionadas para fins específicos.

A amostra das zonas rurais, por outro lado, poderá

ser aperfeiçoada e utilizada para investigações relacionadas com aspectos agro-pecuários ou planos de assistência ou de extensão rural, tendo em vista o aperfeiçoamento das estatísticas agro-pecuárias.

Os assuntos investigados através das Campanhas Estatísticas anuais, poderão ser reexaminadas, em termos da possibilidade da inclusão de certos tópicos em programas como a PNAD ou outros tipos de investigações por amostra, de sorte que, em determinado momento as investigações universais do campo contínuo fiquem reduzidas unicamente àqueles assuntos ou itens, para os quais não seja recomendável em termos custo-e-precisão a adoção de amostras.

16. APLICAÇÃO DA AMOSTRAGEM A OUTROS SETORES

Embora o presente documento tenha sido dedicado à aplicação, atualidade e possibilidades das Pesquisas Domiciliares como instrumento para as investigações das características sócio-econômicas da população, cabe ressaltar, mais uma vez, a importância e necessidade da aplicação da amostragem, para a investigação estatística de outros setores como, a indústria manufatureira, o comércio e a agricultura.

As estatísticas industriais, como nos referimos no início do documento, foi a primeira a obter êxito com a aplicação da Amostragem e, se conseguirmos manter o mesmo ritmo dos esforços que vêm sendo aplicados, teremos assegurado uma boa solução para a manutenção atualizada das informações básicas naquele setor. E, como dissemos anteriormente, a tradição de experiência de mais de duas décadas, em muito contribuiu para que rapidamente se atingisse o nível atual daquelas informações.

No setor da Agro-pecuária, vários esforços já foram desenvolvidos, através de experiências que se repetem ora com maior, ora com menor intensidade. Parece-nos chegado o momento para uma grande concentração de forças, no sentido da implantação gradativa, mas em bases seguras, da utilização da amostragem em larga escala para a obtenção das informações básicas desse setor. Só assim, se poderão substituir, de vez, os levantamentos subjetivos por outros de bases técnicas comprovadamente satisfatórias.

Um outro setor no qual as dificuldades de atualização residem principalmente, no grande volume de documentos a serem manuseados mensalmente, o que torna desaconselhável (em face

dos meios disponíveis) as apurações universais, é o do Comércio Interestadual por Vias Internas (principalmente em Minas e São Paulo). A aplicação da Amostragem, a êsse setor poderia nos conduzir à obtenção dos elementos básicos para a análise de sua situação.

Muito teríamos que escrever se tivéssemos a pretensão de desenvolver aqui os aspectos da aplicação da amostragem a êsses e outros setores de importância para as estatísticas básicas para o desenvolvimento, motivo porque, preferimos apenas dizer que todos êsses tópicos fazem parte das preocupações diárias da Fundação IBGE.

17. DIFICULDADES E IMPLICAÇÕES

Seria uma idéia totalmente falsa pensar que a Amostragem funcione como uma palavra mágica que abra a porta para a solução de todo e, qualquer problema estatístico, da noite para o dia e sem qualquer quota de sacrifício. Seria, igualmente falso pensar que ela eliminaria por completo a necessidade de levantamentos completos. É bem verdade que essa técnica possibilita a diminuição da periodicidade dos levantamentos contínuos, mas nelos se apoia para complementação, e verificação e confronto periódico, tendo em vista a melhoria da precisão e a investigação de itens de menor variação no tempo ou que não sejam passíveis de investigação por amostragem.

A grande extensão territorial do Brasil, o complexo das situações regionais, as dificuldades de meios de transportes e comunicações e, outras circunstâncias, tornam o problema ainda mais difícil.

Há necessidade do estabelecimento de um plano global pelo qual se harmonizem os levantamentos censitários e os levantamentos contínuos de investigação universal, com os levantamentos por amostragem.

A experiência de países com acentuada tradição em levantamentos por amostragem nos leva a concluir que um sistema dêsse tipo necessita de um período de implantação e consolidação de pelo menos quatro anos de aplicação em cada setor. Somente a partir daí, é que surgem realmente os resultados satisfatórios, fruto de uma boa inversão de capital-tempo-pessoal.

Apenas para que se tenha uma idéia da complexidade do trabalho de implantação de uma pesquisa por amostra, descreveremos, a seguir, algumas de suas principais fases:

1. Estudos iniciais envolvendo a viabilidade da pesquisa, decisões básicas, entendimentos produtor-consumidores;
2. Definição do âmbito, profundidade, assuntos de investigação e periodicidade; esquema de implantação; Recursos necessários, etc.;
3. Elaboração ou obtenção de cadastros, mapas (fotografias aéreas, principalmente no caso de amostras de áreas), de cobertura, atualização e precisão adequadas;
4. Estudos para a determinação do tamanho da amostra, sua precisão e custo. Esquema de Amostragem e desenho de amostra;
5. Elaboração dos manuais, instruções, formulários e todo o material para treinamento e trabalho, em todas as fases da operação; cálculo e impressão do material;
6. Estabelecimento de um sistema de fluxo do material órgão-central-campo;
7. Teste de campo e das fases de processamento interno;
8. Seleção e Treinamento e adaptação das equipes;
 - 8.1 de Supervisores
 - 8.2 de Pesquisadores.
9. Estabelecimento de um sistema de suporte às equipes técnicas - viaturas para trabalho, transporte de pessoal, etc.;
10. Esquema de revisão e controle das operações, dos custos e da qualidade do material coletado;
11. Plano de tabulação e programação (para equipamento convencional ou computador). Plano de divulgação e análise;
12. Cálculo e análise de variâncias;
13. Plano de revisão geral e ampliação periódica da pesquisa.

Como se pode depreender dos tópicos listados no parágrafo anterior, o período de tempo necessário ao planejamento, implantação e execução de uma pesquisa por amostra depende, em muito, das experiências, passadas e atuais, no setor específico do levantamento e em alguns aspectos, também das experiências adquiridas com pesquisas de outros setores. Há muitas vezes, possibilidade de utilização de elementos obtidos de outros levantamentos, como medida de estimação de aspectos como, tempo global de coleta, duração média da entrevista, tipo de entrevistador adequado, número de pesquisadores por área de amostra, tempo de apuração, custos globais e parciais, taxas de pagamento por tarefa, etc.

CONSIDERAÇÕES FINAIS

Foi nossa intenção ao abordar os vários aspectos apresentados neste documento, demonstrar quatro fatos que julgamos de grande importância, quais sejam:

1. As Pesquisas Domiciliares constituem o mais moderno instrumento de investigação das condições sócio-econômicas da população. Muitas das informações básicas que o país não dispõe, nesse setor, poderiam ser obtidas em caráter contínuo, sob a forma de estimativas ou índices;
2. O êxito que alguns países alcançaram com a aplicação desse novo sistema e os esforços de muitos outros no sentido de sua implantação são, por si só, motivos suficientes para que também o adotemos;
3. A fase atual de renovação do Sistema Estatístico Brasileiro, sob a orientação segura e realista da atual direção da Fundação IBGE, a constituição de colegiados técnicos como a CONPLANE, no qual participam representantes dos principais interessados nas estatísticas brasileiras, parece-nos ser o momento adequado para o início dessa aplicação;
4. A Pesquisa Nacional por Amostra de Domicílios parece-nos, dada as suas características e possibilidades, o melhor início para a implantação da Técnica da Amostragem em larga escala, como meio de desenvolvimento e aperfeiçoamento das estatísticas brasileiras no campo contínuo.

ASSUNTOS DE INVESTIGAÇÃO PARA PROGRAMAS
DE PESQUISAS DE DOMICÍLIOS

CAMPO E TEMA DE INVESTIGAÇÃO	TIPO DE PESQUISA		
	Contínua	Suple- mentar	Especial
I. Características demográficas:			
Características básicas da população (nome, parentesco, idade, sexo, estado civil) ...	X
Estatísticas vitais (nascimentos por sexo, assistência pré-natal; mortes por idade e sexo, causa da morte)	X
Natalidade (número total de crianças nascidas, e intervalo dos nascimentos)	X	...
Migração interna	X	...
II. Saúde:			
Estudo amplo de enfermidades (tipo, duração, efeito); traumatismos (tipo); seguro médico
Tipos específicos de doenças, ças, traumatismos, invalidez.	X	...
Doenças e traumatismo recentes	X	...
Vacinação contra as enfermidades contagiosas	X	...
Uso de serviços médicos	X	...
Causa da morte (veja-se "Estatísticas vitais")		
Altura, peso, atividade física para completar as necessidades de dieta (veja-se "Consumo de alimento e nutrição")..			
III. Consumo de alimentos de nutrição:			
Consumo de alimentos selecionados de alto teor proteínico (frequência e fonte)	X	...
Quantidades de alimentos consumidos, segundo tipo e fonte..	X
Requisitos dietéticos	X
Calorias e substância dos alimentos consumidos	X

CAMPO E TEMA DE INVESTIGAÇÃO	TIPO DE PESQUISA		
	Contínua	Suple- mentar	Especial
IV. Condições de habitação:			
Características básicas da habitação (classe de unidade, número de unidades de prédio, etc.)
Características do prédio e condição de ocupação (material de construção, condição da unidade, abastecimentos d'água, instalação sanitária; outras instalações de habitação, existência de indústria doméstica)	X	...
Tipo de habitação, construção nova, reforma, demolição	X
... Outras características (tipo de edifício, área do piso, espaço exterior; distância do trabalho, escolas, zonas comerciais; instalação de cozinha; banheiro; aparelhos de calefação, combustível utilizado)	X	...
V. Educação e cultura:			
Educação adquirida (assistência escolar, tipo de ensino, último grau completo, alfabetização)	X	...
Práticas de leitura, rádio e televisão	X	...
Práticas culturais e recreativas	X	...
Evasão escolar (razão pela qual não vai à escola, distância da escola)	X	...
VI. Emprego:			
Mão-de-obra (tipo de atividade, categoria, ocupação, ramo de atividade econômica, horas trabalhadas, salário duração de emprego, tipo de trabalho procurado, mão-de-obra disponível)	X
Experiência de trabalho	X	...
Indústria doméstica (veja-se "Condições de habitação")
Ocupações secundárias	X	...

CAMPO E TEMA DE INVESTIGAÇÃO	TIPO DE PESQUISA		
	Contínua	Suple- mentar	Especial
Outros temas (tamanho e tipo de estabelecimento de emprego, migração com fins de emprego)	...	X	...
VII. Nível econômico da família:			
Salários (veja-se "emprego")..
Renda anual	X	...
Gastos da família destinados a alimentação, vestuário, saúde, habitação, educação, recreação, etc.	X	...
Reforma agrária	X