



IPARDES Instituto Paranaense de Desenvolvimento Econômico e Social

**PESQUISA DE
LOCAIS DE COMPRA
PARA O MUNICÍPIO
DE CURITIBA**

Plano de Amostragem

**CURITIBA
SETEMBRO 1993**

**I59p Instituto Paranaense de Desenvolvimento Econô-
mico e Social**
Pesquisa de locais de compra para o município
de Curitiba : plano de amostragem / Instituto
Paranaense de Desenvolvimento Econômico e Social.
- Curitiba : IPARDES, 1993.
8p.

1.Índice de preços ao consumidor. 2.Amostragem.
3.Curitiba. I.Título.

CDU 303.52:338.57

SECRETARIA DE ESTADO DO PLANEJAMENTO

CARLOS ARTUR KRÜGER PASSOS - Secretário
FERDINANDO SCHAUENBURG - Diretor Geral

INSTITUTO PARANAENSE DE DESENVOLVIMENTO ECONÔMICO E SOCIAL - IPARDES

MARIANO DE MATOS MACEDO - Diretor-Presidente
NEI CELSO FATUCH - Diretor Administrativo-Financeiro
ELVINA MARIA SOARES CHAVES - Diretora do Centro de Pesquisa
EMÍLIO CARLOS BOSCHILIA - Diretor do Centro Estadual de Estatística
EVALDO MARCOS PAVANATO - Diretor do Centro de Treinamento para o Desenvolvimento

TÉCNICO RESPONSÁVEL

Hudson Prestes dos Santos (estatístico)

SUMÁRIO

INTRODUÇÃO.....	1
1 OBJETIVO.....	2
2 O UNIVERSO DA PESQUISA.....	2
3 O MÉTODO DE AMOSTRAGEM.....	3
4 CÁLCULO DA ALOCAÇÃO DOS ELEMENTOS DA AMOSTRA EM CADA ESTRATO.....	4
5 DIMENSIONAMENTO DA AMOSTRA.....	5
6 PROCESSO DE SELEÇÃO DA AMOSTRA.....	7

INTRODUÇÃO

O presente documento delinea os procedimentos metodológicos de amostragem para a seleção dos domicílios a serem investigados sobre os locais de compra e que atendam às características das populações-alvo dos índices de preços ao consumidor a serem definidos para Curitiba, ou seja, um índice restrito referente à faixa de renda de 1 a 8 salários mínimos e um amplo, para a faixa de 1 a 40 salários mínimos.

Através dessa pesquisa será possível a definição de um cadastro de locais, isto é, de uma amostra de estabelecimentos comerciais que vendem produtos ou prestam serviços às famílias objeto do índice. A montagem desse cadastro representa uma etapa fundamental no processo de revisão do IPC-Curitiba, pois dota a pesquisa de preços de uma base de coleta de informações atualizadas e representativas das compras das famílias curitibanas.

Entre as várias fontes de dados que podem ser utilizadas para determinar-se essa amostra de estabelecimentos comerciais e de prestação de serviços, cita-se o Censo de Comércio e Serviços, realizado pelo IBGE. No entanto, para o caso de um índice de preços ao consumidor, o cadastro de informantes deve ser obtido através de pesquisa domiciliar, dado que o índice tem como objetivo medir a variação do poder de compra das famílias.

Saliente-se, porém, que a pesquisa domiciliar não é de natureza exaustiva. Existem produtos e serviços que, por possuírem características especiais de comercialização e de formação de preço, não encontram no domicílio a fonte mais adequada para a definição do local de compra. Nestes casos, o cadastro de informantes deverá ser composto pelos órgãos responsáveis pela fixação dos preços. Enquadram-se nessa categoria os serviços públicos, impostos, combustíveis, entre outros.

1 OBJETIVO

O objetivo deste plano é obter uma amostra de domicílios para Curitiba, a fim de se determinar uma relação de estabelecimentos comerciais e de prestação de serviços que represente o conjunto de locais onde as famílias curitibanas, das populações dos índices restrito e amplo, normalmente adquirem os bens e serviços que compõem a sua cesta-padrão. Pretende-se, através dos locais apontados pela amostra de domicílios, elaborar um cadastro de informantes que serão visitados mensalmente para a coleta de preços dos produtos e serviços do Índice de Preços ao Consumidor - IPC Curitiba.

2 O UNIVERSO DA PESQUISA

A população de análise é constituída por todas as unidades familiares existentes no município de Curitiba. Adotou-se como descritor do universo o cadastro de consumidores residenciais de energia elétrica da Copel em Curitiba. Este cadastro, embora não seja exaustivo, contém a relação de cerca de 98% dos domicílios de Curitiba, totalizando 373.157 domicílios ligados à rede de energia em novembro de 1992.

A atualidade deste cadastro é aspecto de suma relevância para a pesquisa, pois, torna mínima a necessidade de substituição de domicílios em campo, reduzindo assim seu custo. Outro aspecto importante é a facilidade de localização dos domicílios via cadastro da Copel, pois, em caso de dúvidas, pode-se fazer uso dos mapas com as rotas de leitura de consumo da empresa.

Utilizou-se como variável de amostragem neste plano o consumo domiciliar de energia elétrica de novembro de 1992. Consideraram-se os dados referentes a esse mês porque, segundo informações da Copel, novembro é um mês representativo do padrão do consumo médio anual. Embora a variável mais apropriada para pesquisas que tratam sobre gastos familiares seja a renda

domiciliar, esta não pode ser utilizada, dada a indisponibilidade de informações atualizadas e consistentes.

Dessa forma, optou-se pela utilização do consumo de energia elétrica como proxy da renda, estabelecendo-se um critério de estratificação por faixas de consumo com o intuito de representar as características sócio-econômicas da população.

Como produto final desta pesquisa, além da relação de nomes e endereços dos estabelecimentos comerciais e prestadores de serviços, serão obtidos em cada família o rendimento bruto e a posição na ocupação principal do chefe da família. Com isso será possível a vinculação das informações às populações dos índices restrito e amplo, definidas para o IPC-Curitiba, como também correlacionar o nível de renda com as faixas de consumo inicialmente estabelecidas. Outro aspecto importante refere-se à possibilidade de uma avaliação ex-post do erro de amostragem da pesquisa, com base na renda domiciliar constatada durante os trabalhos de campo.

3 O MÉTODO DE AMOSTRAGEM

Para a seleção da amostra, adotou-se o método da amostragem estratificada, usando-se a chamada alocação proporcional ou de Bowley para a atribuição do número de elementos da amostra em cada estrato. A estratificação se justifica neste caso, pois verificou-se uma grande heterogeneidade do consumo familiar medida pelo coeficiente de variação do consumo de energia.

Com base no cadastro da Copel, classificou-se ad hoc a população em 15 faixas de consumo, com intervalos constantes de 60 kwh, com exceção da última faixa. Dentro desses estratos, constatou-se uma homogeneidade do padrão de consumo, revelada pelas variâncias do consumo de energia elétrica em cada estrato. Esse fato sugeriu a adoção da amostragem estratificada com alocação proporcional ao tamanho do estrato. Portanto, é de se supor que a

estratificação e o método de amostragem adotados trazem ganhos substanciais na precisão da amostra.

Por último convém ressaltar que, independentemente da estratificação adotada, esta sempre produz ganhos de qualidade na amostra, pois, no mínimo ela é equivalente ao método da amostragem aleatória simples.

4 CÁLCULO DA ALOCAÇÃO DOS ELEMENTOS DA AMOSTRA EM CADA ESTRATO

O cálculo da alocação proporcional para cada estrato estabelece que sejam obedecidas na amostra as mesmas proporções de elementos verificadas para o universo. Esse processo consiste, portanto, em retirar os elementos proporcionalmente ao tamanho dos estratos em que foi dividida a população.

Assim, a alocação proporcional é dada pela seguinte relação:

$$n_i = n \cdot \frac{N_i}{N}$$

onde:

n_i = número de domicílios da amostra que será extraído do estrato i ;

n = número total de domicílios na amostra;

N_i = número total de domicílios de cada estrato i da população;

N = número total de domicílios ligados à rede de energia elétrica em Curitiba.

A distribuição do número de consumidores residenciais, o consumo médio, a variância, o desvio-padrão e os coeficientes de variação do consumo domiciliar, por estrato, para Curitiba, são apresentados na tabela 1.

TABELA 1 - NÚMERO DE CONSUMIDORES RESIDENCIAIS, CONSUMO MÉDIO, VARIÂNCIA, DESVIO PADRÃO E COEFICIENTES DE VARIAÇÃO, SEGUNDO ESTRATOS DE CONSUMO, EM CURITIBA - NOVEMBRO/1992

ESTRATOS (em kwh)	NÚMERO DE CONSUMI- DORES	CONSUMO MÉDIO (em kwh)	VARIÂNCIA DO CONSUMO (em kwh ²)	DESVIO PADRÃO (em kwh)	COEF. VAR. (em %)
0 — 60	48.253	40,86	116,87	10,81	26,46
60 — 120	105.633	92,72	277,76	16,67	17,98
120 — 180	91.145	148,46	294,78	17,17	11,57
180 — 240	55.876	208,02	296,11	17,21	8,27
240 — 300	30.842	267,88	296,24	17,21	6,43
300 — 360	16.645	327,83	298,31	17,27	5,27
360 — 420	9.232	388,23	298,55	17,28	4,45
420 — 480	5.126	448,70	304,22	17,44	3,89
480 — 540	3.003	509,23	305,28	17,47	3,43
540 — 600	1.886	569,32	303,26	17,41	3,06
600 — 660	1.221	629,51	299,48	17,31	2,75
660 — 720	794	689,60	310,22	17,61	2,55
720 — 780	609	750,56	318,65	17,85	2,38
780 — 840	440	809,98	306,09	17,50	2,16
840 E MAIS	2.452	1876,27	6902115,94	2627,19	140,02
TOTAL	373.157	176,48	77063,77	277,60	157,30

FONTE DOS DADOS BRUTOS: COPEL

NOTA: Cálculos estatísticos processados pelo CEE/IPARDES.

5 DIMENSIONAMENTO DA AMOSTRA

O primeiro elemento para o dimensionamento da amostra é o cálculo da variância do estimador a ser utilizado. No presente caso trata-se de estimar o consumo médio mensal de energia elétrica, o qual, por hipótese, é um parâmetro que se correlaciona com a renda domiciliar e com os demais parâmetros que serão estimados na pesquisa. Essa variável possibilitará, portanto, avaliar-se a priori a precisão da amostra a ser retirada. A posteriori, a precisão dessa amostra será reavaliada com base na renda domiciliar, variável essa, mais representativa para a pesquisa de locais de compra.

Esta variância é dada por:

$$V(\hat{Y}) = \frac{1}{n} \left(\frac{\sum N_i \sigma_i^2}{\sum N_i} \right)$$

onde:

\hat{Y} = estimador do consumo médio mensal;

σ_i^2 = variância do consumo médio de todos os domicílios pertencentes ao estrato i;

N_i = conforme definição anterior.

A função acima permite relacionar-se n e $V(\hat{Y})$, ou seja, o número total de elementos na amostra e a precisão medida pela variância do estimador a ser utilizado.

Assim, com base nas informações e nos cálculos intermediários fornecidos pela Copel, infere-se que a variância do consumo médio mensal para Curitiba é:

$$V(\hat{Y}) = \frac{1}{n} (45619,48)$$

De posse dessa função pode-se determinar o erro amostral correspondente a cada número de elementos na amostra (n), supondo-se a normalidade do estimador \hat{Y} . Sabendo-se que \hat{Y} é um estimador justo de \bar{Y} (o verdadeiro consumo médio mensal), a variável:

$$z = \frac{|\hat{Y} - \bar{Y}|}{\sqrt{V(\hat{Y})}}$$

tem distribuição normal reduzida. Dividindo-se ambos os membros por \bar{Y} , obtém-se:

$$z = \frac{\frac{|\hat{Y} - \bar{Y}|}{\bar{Y}}}{\frac{\sqrt{V(\hat{Y})}}{\bar{Y}}} = \frac{e}{cv}$$

A expressão $e = \frac{|\hat{Y} - \bar{Y}|}{\bar{Y}}$ mede o erro amostral relativo que, multiplicada por 100, mede o desvio percentual do estimador (\hat{Y}) do parâmetro (\bar{Y}), em relação a este último. O denominador de z representa o coeficiente de variação do consumo de energia elétrica da população.

Adotando-se como valor de z , aquele correspondente ao nível de confiança de 95% ($z = 1,96$), obteve-se os resultados apresentados na tabela 2.

TABELA 2 - TAMANHOS DE AMOSTRA E RESPECTIVOS ERROS AMOSTRAIS PARA CURITIBA

TAMANHO DA AMOSTRA (n)	ERRO RELATIVO (%)
200	16,77
500	10,61
800	8,39
1000	7,50
1200	6,85
1500	6,12
1800	5,59
2000	5,30
2200	5,06
2500	4,74
2800	4,48
3000	4,33

FONTE: CEE/IPARDES

NOTAS: Dados de consumo de energia elétrica residencial referentes ao mês de nov./92;

$$\sqrt{n} = 2,3722.$$

6 PROCESSO DE SELEÇÃO DA AMOSTRA

Com o intuito de evitar o viés subjetivo que surge da escolha pessoal de unidades amostrais, bem como possibilitar a mesma oportunidade de escolha a cada uma destas unidades na composição das amostras para cada um dos 15 estratos de consumo de energia elétrica, será utilizada a amostragem aleatória sistemática para o sorteio dos domicílios.

Esta seleção será processada pela Copel, com base no cadastro de consumidores residenciais de energia, obedecendo à estratificação adotada e à ordenação do consumo dos domicílios dentro de cada faixa.

A título de ilustração considere-se o seguinte exemplo:

Estrato 1

Tamanho da população: $N = 48253$

Tamanho da amostra: $n = 259$

a) Cálculo do intervalo de seleção (IS)

$$IS = \frac{N}{n} = \frac{48253}{259} = 186,31$$

b) Cálculo de início aleatório (IA):

IA = x, número sorteado aleatoriamente entre 01 e IS (considerando-se apenas a parte inteira de IS).

c) Os domicílios sorteados para a amostra do estrato 1 serão:

1º domicílio = x

2º domicílio = x + IS

3º domicílio = x + 2IS

.....

.....

.....

259º domicílio = x + (n - 1)IS

Procedimento semelhante adota-se para os demais estratos.

A listagem final das unidades domiciliares selecionadas na amostra deverá incluir para cada estrato um adicional de domicílios da ordem de 25%, isto para compensar perdas provenientes de entrevistas não realizadas, seja por recusa ou por não ser o domicílio eleito para a pesquisa, pesquisa, por motivo de mudança de endereço, de inexistência ou de desocupação do domicílio, por ser de uso ocasional, ou outros entraves.