

O ESTIMADOR PONDERADO QUE COMPENSA
A AUSÊNCIA DE RESPOSTA: 666
UMA APLICAÇÃO

EP
N I L Z A N U N E S D A S I L V A

Tese apresentada à Faculdade de Saúde Pública da Universidade de São Paulo, para obtenção do título de Doutor em Saúde Pública.

Orientadora: Dra. Eunice Pinho de Castro Silva

São Paulo - 1986



AGRADECIMENTOS

- Aos doutores Eunice Pinho de Castro Silva e Jair Lício Ferreira Santos pela orientação e apoio contantes.
- Aos doutores Sabina Léa Davidson Gotlieb e Wilton de Oliveira Bussab pela cuidadosa leitura e sugestões apresentadas.
- Aos companheiros docentes da disciplina Bioestatística que me permitiram maior tempo para a execução desse trabalho.
- Aos amigos Fernão Dias de Lima e Emilia Nakamura pela paciente colaboração no processamento dos dados.
- A Ednéa Primo, Isildinha Marques dos Reis e Maria de Lurdes Vendramine que muito contribuíram para a realização desse trabalho.
- A Daisy Pires Noronha pela revisão das referências bibliográficas.

RESUMO

Como tentativa para compensar a ausência de resposta, aplicou-se o estimador ponderado por pesos amostrais nos dados de um inquérito domiciliar que registrou uma taxa global de resposta igual a 79,10%.

O procedimento procurou viabilizar a aplicação do ajuste ao nível dos elementos, usando os pesos obtidos a partir da taxa diferencial de resposta verificada na amostra composta pelas famílias. Fixado o número de cômodos do domicílio como a variável critério para a formação das classes ponderadas, calcularam-se as médias ajustadas e suas variâncias para o peso e a estatura dos menores de 19 anos, usando-se as estimativas calculadas na amostra constituída pelos elementos pertencentes às famílias que efetivamente responderam ao inquérito.

A consistência dos resultados obtidos sugere que estudos devem ser desenvolvidos a fim de buscar soluções que minimizem as pressuposições necessárias à aplicação do procedimento e determinem níveis de abrangência que possam ser considerados satisfatórios.

SUMMARY

The weighting adjustment by sample weights was applied on the data from a survey sampling with a total response rate of 79,10%, in order to compensate for the unit nonresponse.

The weighting factors were gotten by the differential response rate calculated on the household sample and they were applied on the household member level, in order to facilitate the application procedure. The variable used to define the weighting classes was the number of rooms of the dwellings. Also, the adjusted means and their variances were calculated for weight and height of persons with ages under 19 years, using the measures obtained from the same age group individuals related to the respondent households.

The consistency of the results support the suggestion that more research has to be done to search for procedures that will minimize the necessary assumptions that underlie them and that can assure wider application levels.

ÍNDICE

CAPÍTULO N°.	TÍTULO	PÁGINA N°.
1	Introdução	1
2	Objetivos	3
3	Os estimadores ponderados que compensam a ausência de resposta	4
4	Material e métodos	11
5	Resultados obtidos	16
6	Comentários finais	27
	Referências Bibliográficas	29
	Anexo n° 1	34
	Anexo n° 2	35
	Anexo n° 3	36
	Anexo n° 4	38

OBSERVAÇÕES

Este texto foi processado pela versão 1.2. R02 do Redator em um micro-computador Itautec. Fez-se necessário, portanto, introduzir algumas alterações na notação convencional da estatística matemática conforme descreve-se a seguir:

1) \sum foi substituído por S.

2) Os índices, em vez de subscritos, foram colocados após vírgulas. Por exemplo:

X_i foi substituído por (X,i)

\bar{X}_h foi substituído por (\bar{X} ,h)

CAPÍTULO I - INTRODUÇÃO

Os levantamentos por entrevistas vêm sendo amplamente utilizados como meio de produzir informações nas áreas que estudam as condições de vida das populações humanas. No Brasil, além dos inquéritos realizados pelos órgãos da administração pública, existe um grande número de pesquisas resultantes da iniciativa privada ou ligadas a interesses da área acadêmica.

A coleta dos dados, em geral, se realiza através da aplicação de questionários nos domicílios, onde os entrevistadores contam com a participação voluntária de um dos residentes, podendo ocorrer recusas e/ou a não localização das famílias cujos domicílios estiveram fechados durante o período em que se desenvolveu o trabalho de campo. Estes fatos, amplamente referidos como não-resposta ou ausência de resposta (20,21,40) são os principais determinantes da perda da informação de parte das unidades amostrais e têm sido um dos principais problemas na prática dos levantamentos, o que parece vir aumentando, notadamente, nas áreas urbanas onde registram-se taxas globais de resposta (*) inferiores a 70%.

Baixas taxas de resposta podem introduzir vício nas inferências efetuadas apenas com a informação obtida entre os respondentes, quando a ausência de resposta se concentra em alguns sub-grupos da população (não-resposta diferencial), distorcendo as distribuições que seriam observadas, para algumas variáveis, na amostra total efetivamente sorteada (10,14,23,25,33,39). E este vício poderá ser aumentado, se as variáveis para as quais as inferências são efetuadas (variáveis de interesse) também forem relacionadas com esses sub-grupos. Levantamentos realizados na área da saúde, que avaliaram o impacto da ausência de resposta nos seus resultados (6,9,11,17,19,29,31,45), têm indicado a ocorrência da não-resposta diferencial quando consideradas as variáveis sociais e demográficas básicas, comuns nesses tipos de pesquisa, destacando-se entre estas: o tamanho da família, a idade do chefe da família e a zona geográfica de residência.

* - A taxa global de resposta de um levantamento é definida (3,23) como a razão do número de questionários preenchidos nas unidades elegíveis (r) para o número total de unidades elegíveis (n) na amostra total selecionada ($TR=r/n$).

A identificação das variáveis, com seus subgrupos e respectivas taxas de resposta, permite compensar o vício introduzido. Através da aplicação de modelos ponderados (22,23) que tenta-se recuperar a representatividade da amostra sorteada, atribuindo pesos, inversamente proporcionais a essas taxas de resposta aos elementos respondentes que compõem efetivamente a amostra trabalhada.

Os procedimentos mais comuns de ponderação (4,17,23,26,33) dividem a amostra de respondentes em classes e, recorrendo à técnica de pós-estratificação, usam as informações existentes (em censos ou registros) para determinar os pesos correspondentes a essas classes e ajustar a distribuição da amostra à distribuição da população.

Quando na amostra existir informação para as unidades respondentes e para as não-respondentes, é possível calcular as taxas de resposta das classes, estimar os pesos a serem aplicados às estimativas obtidas nessas classes e, através da técnica de estratificação em duas fases (23,25), determinar a estimativa final desejada.

Este trabalho representa uma contribuição à maioria dos estudos de morbidade desenvolvidos através de amostras de domicílios, apresentando uma alternativa simples de aplicação de um procedimento de ajuste para compensar a ausência de resposta, em que os pesos aplicados às unidades elementares são estimados a partir de uma amostra, na qual a família é o conglomerado de referência básica. Tem como objetivos principais apontar algumas variáveis observáveis na família, que sejam associadas à ausência de resposta e às variáveis de interesse principal do levantamento, para as quais a informação possa ser facilmente obtida sem a participação do informante; e sugerir um procedimento de avaliação da ausência de resposta que possa ser desenvolvido paralelamente à realização das entrevistas, permitindo a determinação ágil da não-resposta diferencial e a aplicação do ajuste por ponderação, quando necessário.

CAPÍTULO 2 - OBJETIVOS

- Apresentar uma aplicação do modelo ponderado que compensa o vício introduzido pela ausência de resposta, estimando os pesos a partir da amostra total efetivamente sorteada.
- Sugerir um procedimento que avalia e compensa a ausência de resposta, fixando a família como a unidade amostral base para o cálculo das taxas diferenciais de resposta e dos fatores de ponderação; e determinando as estimativas ponderadas na amostra definida pelos elementos observados nas famílias respondentes.
- Indicar variáveis sócio-demográficas, para as quais a obtenção da informação nas famílias não-respondentes possibilite avaliar, e compensar quando necessário, o vício introduzido pela ausência de resposta.

CAPÍTULO 3 . - OS ESTIMADORES PONDERADOS QUE COMPENSAM A AU-
SÊNCIA DE RESPOSTA.

3.1. - O estimador ponderado simples

Na compensação da ausência de resposta reúnem-se os trabalhos que procuram determinar a melhor estimativa para um particular parâmetro (2,8,13,32,37,41) e aqueles cujos objetivos consistem em minimizar o vício das estimativas obtidas em levantamentos de múltiplos objetivos, desenvolvidos através de grandes amostras.

A técnica de ajuste por ponderação, amplamente utilizada na análise estatística dos resultados de levantamentos descritivos, é comumente adotada pelos autores cujas pesquisas se situam no último grupo (17,22,26,35). Consiste em assinalar pesos que refletem a importância relativa de cada elemento amostrado, em processos de estimação decorrentes de planejamentos não auto-ponderados ou em situações em que o processo equíprobabilístico de sorteio for contrariado.

Em planejamentos, a priori equíprobabilísticos, nos quais ocorreu a ausência de resposta, o estimador ponderado assinala pesos à informação obtida entre os respondentes. Concebendo a população e a amostra divididas em respondentes e não-respondentes, com os seus respectivos valores representados conforme nomenclatura apresentada na figura nº1, o modelo mais simples de ponderação expresso por:

$$\bar{x}_P = [S(P,i)(x_{r,i})] / [S(P,i)] \quad (3.1)$$

pressupõe que as médias populacionais desses grupos são iguais ($\bar{X}_R = \bar{X}_{NR}$) e assinala, como peso único, o inverso da taxa global de resposta [$(P,i) = (n/r)$] aos valores obtidos para a variável de interesse ($x_{r,i}$), em cada i -ésimo respondente da amostra.

Entretanto, a compensação neste caso é inócuia e de pouca utilidade prática, visto que determina a mesma estimativa e vício obtidos no grupo dos respondentes, quando se aplica o estimador não ponderado:

Figura nº1.

Nomenclatura adotada para os valores básicos usados na população e na amostra divididas em respondentes e não-respondentes.

VALORES	POPULAÇÃO			AMOSTRA		
	RESP.	NÃO-RESP.	TOTAL	RESP.	NÃO-RESP.	TOTAL
1.Nº de elementos	R	NR	N	r	nr	n
2.Valor da variável X no i-ésimo elemento.	XR,i	XNR,i	X	xr,i	xnr,i	x
3.Média por elemento para a variável X.	\bar{X}_R	\bar{X}_{NR}	\bar{X}	\bar{x}_r	\bar{x}_{nr}	\bar{x}
4.Variância para a variável X.	$(SR)^2$	$(SNR)^2$	S^2	$(sr)^2$	$(snr)^2$	s^2

$$x_r = [S(x_r, i) / r] \quad (3.2)$$

pois, $x_p = [S(n/r)(x_r, i) / S(n/r)] = [(n/r)S(x_r, i)] / n =$
 $= S(x_r, i) / r = \bar{x}_r.$

O vício expresso por:

$$Vício (\bar{x}_p) = (NR/N) (\bar{X}_R - \bar{X}_{NR})^* \quad (3.3)$$

é igual ao vício de \bar{x}_r e sua extensão depende do grau em que a pressuposição ($\bar{X}_R = \bar{X}_{NR}$) se ajusta à realidade de cada levantamento em estudo. Situação esta considerada pouco realista quando a proporção de não-respondentes na população (NR/N) é alta.

3.2. - O estimador por classes ponderadas

O modelo ponderado simples pode ser estendido à situação de estratificação (4,5,23) quando for possível identificar variáveis através das quais se divide a amostra em sub-grupos (aqui chamados classes ponderadas ou simplesmente classes) e se calculem as suas taxas de resposta. A população e a amostra, agora divididas segundo essas classes, têm os seus valores representados pela nomenclatura apresentada na figura nº2.

O estimador por classes ponderadas é mais realístico porque assume a igualdade entre as médias populacionais dos respondentes e dos não-respondentes dentro das classes que foram definidas segundo diferentes taxas de resposta calculadas por:

$$(TR, h) = (r, h) / (n, h) \quad (3.4)$$

* - A demonstração encontra-se no anexo nº1.

Figura nº2.

Nomenclatura adotada para os valores básicos usados na população e na amostra divididas segundo as classes ponderadas e em respondentes e não respondentes.

VALORES	POPULAÇÃO			AMOSTRA		
	RESP	NÃO-RESP	TOTAL	RESP	NÃO-RESP	TOTAL
1.Nº de elementos na classe h.	R,h	NR,h	N,h	r,h	nr,h	n,h
2.Valor da var.X no i-ésimo elemento da classe h.	XR,hi	XNR,hi	X,hi	xr,hi	xnr,hi	x,hi
3.Média por elemento para a var.X da classe h.	XR,h	XNR,h	X,h	xr,h	xnr,h	x,h
4.Variância para a var.X da classe h (SR,h) (SNR,h) (S,h)	2	2	2	(sr,h)	(snr,h)	(s,h)

onde o numerador é o número de respondentes e o denominador é o número total de unidades elegíveis pertencentes à classe h . Em cada classe o inverso da sua taxa de resposta é o peso aplicado aos elementos, tentando-se ajustar a distribuição da amostra de respondentes, segundo essas classes, à distribuição que deveria ocorrer para a amostra total composta por respondentes e não-respondentes. Assim, o estimador por classes ponderadas é expresso por:

$$\bar{x}_{cp} = \frac{-SS(x_{r,hi})[(n,h)/(r,h)]}{SS(n,h)/(r,h)} \quad (3.5)$$

onde $(x_{r,hi})$ é o valor da variável X observado para o i -ésimo elemento da h -ésima classe e $(n,h / r,h)$ é o inverso da taxa de resposta da classe h .

Observando, entretanto, que essa expressão quando trabalhada algébricamente pode definir um resultado igual a:

$$\begin{aligned} \bar{x}_{cp} &= \frac{-SE(n,h) / (r,h)] (x_{r,hi}) - S(n,h)S(x_{r,hi}) / (r,h)}{SE(n,h) / (r,h)} = \frac{-SE(r,h)(n,h)] / (r,h)}{SE(r,h)(n,h)] / (r,h)} = \\ &= \frac{-S(n,h)(x_{r,hi})}{n} = SE(n,h) / (n)] (x_{r,hi}) = SE(w,h) (x_{r,hi}) \end{aligned} \quad (3.6)$$

o estimador por classes ponderadas passa, por conveniência, a ser expresso por esse resultado, o qual permite o uso das técnicas de estimação da amostragem estratificada em duas fases (7,23,25). Admite-se que a amostra total foi selecionada na primeira fase e que a informação para a variável critério de estratificação foi obtida para as n unidades elegíveis, permitindo a definição das classes h e dos seus correspondentes pesos $[w,h]=(n,h)/n]$. Os respondentes de cada classe compõem uma amostra de tamanho (r,h) numa segunda fase, das (n,h) unidades elegíveis sorteadas na primeira fase; e o total de respondentes $r=[S(r,h)]<n$ é a "amostra estratificada" finalmente obtida.

A estimativa da média populacional é obtida pela expressão (3.6), onde:

$$\bar{x}_{(r,h)} = \frac{s(x_{r,h})}{(r,h)} \quad (3.7)$$

é a estimativa obtida na amostra de tamanho (r,h) da classe h e (w,h) estima os pesos populacionais (W,h) .

A sua variância é aproximada, no caso de grandes amostras e em que (r,h) for maior ou igual a 10, por:

$$v(\bar{x}_{cp}) = \frac{SE(w,h)^2}{(r,h)} + \frac{SE(w,h)^2}{n} (\bar{x}_{r,h} - \bar{x}_{cp})^2/n \quad * \quad (3.8)$$

onde

$$(s_{r,h})^2 = \frac{SE(x_{r,h})^2 - (\bar{x}_{r,h} - \bar{x}_{cp})^2}{(r,h) - 1} \quad (3.9)$$

é a variância estimada entre os respondentes de cada classe h . O primeiro termo corresponde à variância de uma amostra estratificada usual e o segundo, insignificante se n é grande, corresponde ao acréscimo na variância introduzido pela técnica de estratificação em duas fases.

Como cada conjunto (r,h) é auto-selecionado, (\bar{x}_{cp}) também é um estimador viciado para X . Entretanto, o seu vício expresso por:

$$Vício (\bar{x}_{cp}) = \frac{SE(W,h)^2}{(N,h)} E(X_{r,h} - X_{NR,h}) \quad ** \quad (3.10)$$

será minimizado na medida em que as classes definidas pelas diferentes taxas de resposta sejam eficazes no sentido de agrupar respondentes e não-respondentes semelhantes quanto às características populacionais que estão sendo estimadas.

* - A demonstração encontra-se em Kalton,G.(1983) (23) e comentários técnicos em Kish,L.(1965) (26).

** - A demonstração encontra-se no anexo-nº2.

O estimador \bar{x}_r , quando aplicado à situação de amostragem estratificada em duas fases com auto-seleção dos respondentes, também é viciado e a expressão do vício dada por:

$$\begin{aligned} \text{Vício } (\bar{x}_r) = & \text{SE}(W,h) [\bar{X}_R, h - \bar{X}_R] (R, h/N, h - R/N)] / (R/N) + \\ & + \text{SE}(W,h) [\bar{(N}_R, h)/ (N, h)] [\bar{X}_R, h - \bar{X}_{NR}, h)] \quad (3.11) \end{aligned}$$

contém no segundo termo, o vício do estimador por classes ponderadas (apresentado na expressão 3.10).

Então o vício de \bar{x}_r pode ser expresso por:

$$\text{Vício } (\bar{x}_r) = \text{SE}(W,h) [\bar{X}_R, h - \bar{X}_R] (R, h/N, h - R/N)] + \text{Vício } (\bar{x}_{cp}),$$

permitindo observar:

1º - que o vício absoluto de \bar{x}_{cp} será menor que o vício absoluto de \bar{x}_r , quando os dois termos da expressão tiverem sinais iguais, ou, no caso de sinais diferentes, quando o valor absoluto do primeiro termo for igual ou maior que duas vezes o valor absoluto do segundo termo; e

2º - que o grau de eficácia atingido pela aplicação do estimador \bar{x}_{cp} dependerá da obtenção de altos valores para o primeiro termo dessa expressão, o que será alcançado quando se definirem as classes a serem ponderadas com grandes diferenças entre as taxas de resposta e entre as médias observadas para os respondentes.

* - A demonstração encontra-se no anexo nº3.

CAPÍTULO 4 - MATERIAL E MÉTODOS.

4.1. - Material

A amostra equíprobabilística utilizada neste trabalho é composta por 1.646 famílias residentes no sub-districto do Butantã, estudadas pelo Instituto da Criança do Hospital das Clínicas da Universidade de São Paulo, nos anos de 1981 e 1982. O levantamento teve como objetivos principais a elaboração do diagnóstico sócio-econômico da população alvo do Centro de Saúde, instalado naquela área, e a avaliação do estado de saúde e nutricional dos menores de 19 anos a ela pertencentes. Foram sorteados 1.765 endereços (figura nº3), nos quais identificaram-se 1.646 unidades elegíveis, sendo efetivamente preenchidos 1.302 questionários do inquérito sócio-econômico, o que determinou uma taxa global de resposta igual a 79,10%.

Figura nº3.

Resumo dos resultados verificados no trabalho de campo.

```

-----
: Endereços      : 195
: sorteados=1.765 :
-----
----- -----
: Unidades não   : : Unidades      :
: elegíveis=119  : : elegíveis=1.646: 183
:                   : -----
----- -----
: Famílias respon- : : Famílias não-res :
: dentes = 1.302  : : spondentes = 344: 49
:                   : -----
----- -----
: com meno- : : sem meno- : : sub-amostra :
: res = 977: : res = 325: : de fam.=104 :
:                   : -----
----- -----
: com meno- : : sem meno- :
: res = 66:   : res = 38:
:                   :

```

A identificação das famílias que possuíam menores de 19 anos (aqui chamadas famílias com menores) efetuou-se através dos questionários do inquérito sócio-econômico já preenchidos para as 1.302 famílias respondentes, e através de exaustivos retornos realizados a uma sub-amostra equiprobabilística ($n=104$) das 344 famílias não-respondentes. Desses, 977 famílias respondentes e 66 não-respondentes que possuíam pelo menos um menor de 19 anos constituíram a amostra observada, a partir da qual a taxa de resposta diferencial foi avaliada.

As informações, referentes às condições de vida dessas famílias e consideradas possíveis indicadores da ausência de resposta, também foram retiradas dos questionários do inquérito sócio-econômico para as famílias respondentes e coletadas nos endereços da sub-amostra do segmento das famílias não-respondentes. As variáveis observadas na unidade amostral família são: tamanho da família; nº de menores de 1 ano, de 1 a 4 anos, de 5 a 9 anos e de 10 a 19 anos; idade, sexo, e instrução do chefe; tipo da habitação e número de cômodos do domicílio.

Os menores de 19 anos pertencentes às famílias respondentes constituíram a amostra de unidades elementares na qual, efetivamente, aplicou-se o modelo ponderado para compensar a ausência de resposta. Optou-se por se trabalhar com o peso e a estatura desses menores, por considerá-las as variáveis de interesse central em qualquer estudo que avalie o estado nutricional de uma população.

4.2. - Métodos

4.2.1. A formação das classes ponderadas : cálculos na amostra de famílias.

As taxas diferenciais de resposta foram calculadas, para cada variável sócio-demográfica, aplicando-se a expressão (3.4) à amostra de famílias com menores. Como, por medida de economia, as famílias não-respondentes foram observadas através de uma sub-amostra, fez-se necessário, para esse segmento, as seguintes estimativas intermediárias:

1º - o total de famílias não-respondentes com menores na amostra. Obtido pela aplicação da proporção observada na sub-amostra sobre as 344 famílias não-respondentes. Então, $\hat{n}_r = [(66/104) \cdot 344] = 218$.

2º - o número de famílias não-respondentes com menores em cada classe h . Obtido por:

$$\hat{n}_{r,h} = 218 \cdot (f_{r,h}) \quad (4.1)$$

onde $(f_{r,h})$ é a freqüência relativa simples da classe h calculada na sub-amostra desse segmento.

As taxas de resposta foram, então, calculadas por:

$$TR_{r,h} = (r_{r,h}) / [E(r_{r,h}) + (\hat{n}_{r,h})] \quad (4.2)$$

onde $(r_{r,h})$ é o número de famílias respondentes com menores observado nas distribuições de frequências simples calculadas, também, para esse segmento; e o denominador estima o total de famílias com menores em cada classe.

Com a finalidade de isolar a variável critério para a formação das classes ponderadas, avaliou-se a homogeneidade das amostras de famílias, com menores, respondentes e não-respondentes, através da estatística qui-quadrado calculada, em tabelas de contingência construídas para cada variável sócio-demográfica. Quando considerado necessário, algumas classes foram convenientemente agrupadas. A variável isolada foi aquela que apresentou maior valor para o qui-quadrado calculado em relação ao menor valor estimado para o nível de significância.

4.2.2. Obtenção das estimativas ponderadas e suas variâncias : cálculos na amostra de menores.

Em estudos antropométricos é reconhecida a conveniência de se obter as estimativas para grupos de menores separados segundo a idade e o sexo, caracterizando-se a situação em que o conglomerado família é de tamanho 1 e em que é adequado o uso dos métodos de estimação próprios à amostragem de elementos. Entretanto, na tentativa de adequar as limitações do modelo de ponderação (*) ao tamanho da amostra disponível, decidiu-se obter as estimativas ponderadas para os menores reunidos em grupos (descritos na figura nº4) cujos tamanhos médios dos conglomerados estivessem em torno de 1 menor por família respondente.

Figura nº4.

Grupos de ponderação e correspondentes médias e desvios padrão do número de menores por família respondente.

Grupo	Faixa etária em anos completos	Sexo	Média	Desvio padrão
1	1 a 4	Masc	1,136	0,356
2	1 a 4	Fem	1,163	0,385
3	5 a 9	Masc	1,217	0,434
4	5 a 9	Fem	1,198	0,442
5	10 a 19	Masc	1,409	0,676
6	10 a 19	Fem	1,312	0,573

Em cada grupo, os menores das famílias respondentes foram reunidos na classe da variável critério de estratificação em que se classificou a sua família ou o seu domicílio. As médias e as variâncias foram, respectivamente, calculadas dentro de cada classe através da aplicação das expressões (3.7) e (3.9) nas medidas do peso e da estatura registradas para cada menor, onde (x_r, h_i) é o peso ou a estatura observados no i -ésimo menor de uma família respondente

* - Como indicado no capítulo 3 página 9.

classificada na classe h ; e o denominador é substituído pelo número de menores observado nas famílias respondentes que compõem a classe h .

Admitindo-se que as famílias respondentes e as não-respondentes de uma mesma classe têm em média o mesmo número de menores, estimou-se o total de menores (n,h) em cada classe h , aplicando-se ao número de menores (r,h) , observado nas famílias respondentes, o inverso das taxas de resposta (TR,h) calculadas na amostra composta pelas famílias. Assim:

$$(n,h) = (r,h)(1/TR,h) \quad (4.3)$$

e $n = S(n,h)$ determinaram os pesos amostrais $(w,h) = (n,h)/n$.

Finalmente, as médias ponderadas e suas variâncias foram estimadas nos 6 grupos de ponderação pela aplicação dos estimadores :

$$\bar{x}_{cp} = S(w,h) \bar{(x_r,h)} \quad (4.4)$$

e

$$\bar{v}(\bar{x}_{cp}) = S[(w,h) \bar{(s_{rh})}]^2 / (r,h) + S[(w,h) \bar{(x_r,h - \bar{x}_{cp})}]^2 / n \quad (4.5)$$

anteriormente citados nas expressões (3.6) e (3.8), mas aqui reescritos a fim de facilitar a compreensão de sua aplicação à amostra de menores.

CAPÍTULO 5 - RESULTADOS OBTIDOS

As freqüências relativas simples para as variáveis sócio-demográficas segundo famílias com menores respondentes e não respondentes (tabela nº1) * evidenciam possíveis distorções nas distribuições da instrução do chefe, do tipo de construção e do número de cômodos do domicílio. A avaliação diferencial (tabela nº2) ilustra melhor este resultado, registrando as mais baixas taxas de resposta em algumas classes dessas variáveis, caracterizando-se as famílias não-responsantes como aquelas compostas por 3 ou 4 pessoas ou com chefes de maior grau de escolaridade ou cujos domicílios são casas térreas com 6 ou mais cômodos.

Ressalta-se a existência de uma possível associação entre a instrução do chefe da família e do número de cômodos do domicílio com a taxa de resposta, evidência confirmada pelo resultado dos testes de homogeneidade (tabela nº3). Entretanto, atendendo ao objetivo do trabalho de se isolar uma variável cuja informação para as famílias não-responsantes possa ser obtida com segurança e sem a participação do informante, considerou-se tomar o número de cômodos do domicílio como a variável critério para a formação das classes através das quais aplicou-se o estimador ponderado.

Os resultados dos cálculos para a obtenção dos pesos amostrais encontram-se nas tabelas números 4 e 5. O inverso das taxas diferenciais de resposta e o número de menores auto-selecionados na segunda fase da amostragem (colunas a e b) determinaram o número de menores que seriam observados caso todas as famílias com menores houvesse respondido (coluna c). Os pesos amostrais (coluna d) são as proporções que deveriam ter ocorrido na amostra preliminar e com as quais se espera corrigir as possíveis distorções introduzidas pela ausência de resposta. Apesar das taxas de resposta serem as mesmas, esses pesos são diferentes para os mesmos grupos de ponderação quando consideradas as variáveis peso e estatura, devido à diferença existente entre os números de menores auto-selecionados para uma mesma classe h em cada grupo.

* - O processamento detalhado encontra-se no anexo nº4.

Tabela nº1 - Variáveis sócio-demográficas.

Freqüências relativas simples (em %) para as famílias respondentes e não-respondentes.

Nº	VARIÁVEL	CLASSE	FAM.RESP. n=977	FAM.NÃO-RESP. nr=66
1.	Tamanho da família	<=2	2,87	1,75
		3	18,22	24,56
		4	25,49	31,58
		5	23,44	24,56
		6	14,43	7,02
		=>7	15,56	10,53
2.	Nº de menores de 1 ano	0	84,75	89,29
		1	15,25	10,71
3.	Nº de menores de 1 a 4 anos	0	56,50	57,14
		1	43,50	42,86
4.	Nº de menores de 5 a 9 anos	0	56,70	46,43
		1	29,07	37,50
		2	11,46	12,50
		=>3	2,76	8,57
5.	Nº de menores de 10 a 19 anos	0	38,69	46,43
		1	28,56	28,57
		=>2	32,75	25,00
6.	Idade do chefe	20-29	50,26	54,00
		30-39	29,58	32,00
		40-49	16,68	12,00
		50-70	3,48	2,00
7.	Sexo do Chefe	Masc	90,07	94,44
		Fem.	9,93	5,56
8.	Instrução do chefe	1º grau	59,65	41,67
		2º grau	11,91	29,17
		Superior	15,20	27,08
		Analfab.	13,24	2,08
9.	Tipo da habitação	Casa	69,43	81,82
		Apart.	7,36	7,58
		Barraco	23,21	10,61
10.	Número de cômodos	1	5,22	1,54
		2	18,83	13,85
		3	18,22	7,69
		4	21,39	18,46
		5	16,48	12,31
		6	10,03	15,38
		=>7	9,83	30,77

Tabela nº2 - Variáveis sócio-demográficas
Taxas diferenciais de resposta (TR,h)

Nº	VARIÁVEL	CLASSE	(n,h)	(TR,h)%
1.	Tamanho da família	<=2	32	87,98
		3	232	76,88
		4	318	78,34
		5	283	81,05
		6	156	90,21
		=>7	175	86,88
2.	Nºde menores de 1 ano	0	1.023	80,97
		=>1	172	86,45
3.	Nºde menores de 1 a 4 anos	0	677	81,59
		=>1	518	81,98
4.	Nºde menores de 5 a 9 anos	0	655	84,55
		1	366	77,65
		2	139	80,43
		3	35	77,62
5.	Nºde menores de 10 a 19 anos	0	479	78,88
		1	341	81,75
		=>2	375	85,45
6.	Idade do Chefe	20-29	609	80,66
		30-39	359	80,56
		40-49	189	86,17
		50-70	38	88,63
7.	Sexo do Chefe	Masc.	1.086	81,04
		Fem.	109	88,90
8.	Instrução do chefe	1º grau	672	86,48
		2º grau	180	64,59
		Superior	207	71,48
		Analfab.	134	96,60
9.	Tipo da habitação	Casa	876	76,47
		Apartam.	90	78,81
		Barraco	251	89,34
10.	Número de cômodos	1	54	93,83
		2	214	85,91
		3	195	91,39
		4	249	83,85
		5	188	85,72
		6	132	74,51
		=>7	163	58,87

Tabela nº3.

Variáveis sócio-demográficas versus grupos:respondentes e não-respondentes.

Qui-quadrados calculados (Q2), graus de liberdade (GL) e níveis de significância estimados (NS).

Nº	VARIÁVEL	(Q2)	(GL)	(NS)
1.	Tamanho da família	5,211	5	0,3907
2.	Nº de menores de 1 ano	0,855	1	0,3552
3.	Nº de menores de 1 a 4 anos	0,009	1	0,9247
4.	Nº de menores de 5 a 9 anos	2,444	3	0,4855
5.	Nº de menores de 10 a 19 anos	1,796	2	0,4074
6.	Idade do chefe	1,166	3	0,7612
7.	Sexo do chefe	1,117	1	0,2905
8.	Instrução do chefe	21,810	3	0,0001
9.	Tipo da habitação	5,736	2	0,0568
10.	Número de cômodos	32,918	6	0,0001

Tabela nº4
VARIÁVEL PESO

20

Taxas diferenciais de resposta (TR,h), nº de menores observados (r,h), nº estimado de menores (n,h) e pesos (w,h).

GRUPOS DE PONDERAÇÃO	h	a (TR,h)	b (r,h)	c (n,h)	d (w,h)
1 a 4 M	1	0,9383	19	20,249	0,062
	2	0,8591	71	82,645	0,254
	3	0,9139	48	52,522	0,161
	4	0,8385	46	54,860	0,168
	5	0,8572	37	43,164	0,133
	6	0,7451	26	34,899	0,107
	7	0,5887	22	37,351	0,115
	S		269	325,691	
1 a 4 F	1	0,9383	13	13,855	0,057
	2	0,8591	50	58,200	0,238
	3	0,9139	45	49,240	0,201
	4	0,8385	39	46,512	0,190
	5	0,8572	30	34,998	0,143
	6	0,7451	15	20,134	0,082
	7	0,5887	13	22,071	0,090
	S		205	245,010	
5 a 9 M	1	0,9383	19	20,249	0,061
	2	0,8591	59	68,677	0,206
	3	0,9139	49	53,616	0,161
	4	0,8385	61	72,749	0,219
	5	0,8572	36	41,997	0,126
	6	0,7451	26	34,899	0,105
	7	0,5887	24	40,747	0,122
	S		274	332,935	
5 a 9 F	1	0,9383	13	13,855	0,042
	2	0,8591	48	55,872	0,170
	3	0,9139	63	68,935	0,209
	4	0,8385	51	60,823	0,185
	5	0,8572	50	58,329	0,177
	6	0,7451	23	30,872	0,094
	7	0,5887	24	40,747	0,124
	S		272	329,434	
10 a 19 M	1	0,9383	13	13,855	0,023
	2	0,8591	52	60,528	0,101
	3	0,9139	100	109,421	0,183
	4	0,8385	117	139,535	0,234
	5	0,8572	94	109,659	0,184
	6	0,7451	60	80,537	0,135
	7	0,5887	49	83,192	0,139
	S			596,727	
10 a 19 F	1	0,9383	13	13,855	0,023
	2	0,8591	64	74,497	0,123
	3	0,9139	85	93,008	0,154
	4	0,8385	122	145,498	0,241
	5	0,8572	92	107,326	0,178
	6	0,7451	65	87,248	0,144
	7	0,5887	49	83,192	0,138
	S		490	604,624	

Tabela nº5
VARIÁVEL ESTATURA

21

Taxas diferenciais de resposta (TR,h); N° de menores observados (r,h), n° estimado de menores (n,h) e peso calculado (w,h).

GRUPOS DE PONDERACAO	h	a (TR,h)	b (r,h)	c (n,h)	d (w,h)
1 a 4 M	1	0,9383	19	20,249	0,061
	2	0,8591	75	87,301	0,263
	3	0,9139	49	53,616	0,162
	4	0,8385	46	54,860	0,166
	5	0,8572	37	43,164	0,130
	6	0,7451	26	34,899	0,105
	7	0,5887	22	37,351	0,113
	S		274	331,441	
1 a 4 F	1	0,9383	13	13,855	0,054
	2	0,8591	53	61,692	0,243
	3	0,9139	49	53,616	0,211
	4	0,8385	39	46,512	0,183
	5	0,8572	31	36,164	0,142
	6	0,7451	14	18,792	0,074
	7	0,5887	14	23,769	0,093
	S		213	254,401	
5 a 9 M	1	0,9383	19	20,249	0,061
	2	0,8591	59	68,677	0,207
	3	0,9139	49	53,616	0,162
	4	0,8385	60	71,556	0,261
	5	0,8572	36	41,997	0,127
	6	0,7451	26	34,899	0,105
	7	0,5887	24	40,747	0,123
	S		273	331,742	
5 a 9 F	1	0,9383	13	13,855	0,042
	2	0,8591	48	55,872	0,170
	3	0,9139	63	68,935	0,290
	4	0,8385	511	60,823	0,185
	5	0,8572	50	58,329	0,177
	6	0,7451	23	30,872	0,094
	7	0,5887	24	40,747	0,124
	S		272	329,434	
10 a 19 M	1	0,9383	13	13,855	0,023
	2	0,8591	52	60,528	0,102
	3	0,9139	100	109,421	0,184
	4	0,8385	117	139,535	0,234
	5	0,8572	93	108,493	0,182
	6	0,7451	60	80,537	0,135
	7	0,5887	49	83,192	0,140
	S		484	595,561	
10 a 19 F	1	0,9383	13	13,855	0,023
	2	0,8591	64	74,497	0,123
	3	0,9139	85	93,008	0,154
	4	0,8385	122	145,498	0,241
	5	0,8572	92	107,326	0,178
	6	0,7451	65	87,248	0,144
	7	0,5887	49	83,192	0,138
	S		490	604,624	

Os resultados dos cálculos para obtenção das médias ponderadas e suas variâncias encontram-se nas tabelas nºs.6 e 7. As médias amostrais de cada classe h (coluna a) são multiplicadas pelos pesos w_h (coluna c) cuja soma é a média ponderada, estimativa desejada da média populacional. Na composição da variância observa-se que, invariavelmente, o acréscimo devido à estratificação em duas fases (S/n da coluna e) é muito inferior ao termo da estratificação usual (coluna d), o que foi assegurado pela formação dos grupos com amostras de grandes tamanhos.

Em resumo (tabelas nºs.8 e 9), as médias ponderadas são consistentemente mais altas que aquelas obtidas entre os respondentes, refletindo, possivelmente, as correlações existentes entre a taxa de resposta, o número de cônjuges, o peso e a estatura. Entretanto, a magnitude dessa diferença pode ter sido prejudicada pela amplitude etária dos grupos de ponderação, o que determinou classes pouco homogêneas, (grande número de desvios padrão da classe são maiores que o desvio padrão do grupo) (*), e consequentemente diferenças entre as suas médias consideradas insatisfatórias. Este aspecto também interferiu nas variâncias que, igualmente às médias, são consistentemente mais altas como se vê na última coluna.

* - Colunas b das tabelas nºs.6 e 7.

Tabela nº6
VARIÁVEL PESO (EM GRAMAS)

23

Cálculos para a obtenção das médias ponderadas e suas variâncias.

GRUPOS DE PONDERAÇÃO	h	(a) $\bar{(xr,h)}$	(b) (sr,h)	(c) $(w,h)(\bar{(xr,h)})$	(d) 1º termo	(e) 2º termo
1 a 4 M	1	12.193,16	2.356,26	758,09	1.129,55	335.643,69
	2	14.149,86	3.419,32	3.590,56	10.603,26	34.133,59
	3	14.227,50	2.453,99	2.294,38	3.262,72	13.480,49
	4	14.682,83	3.529,16	2.473,20	7.682,14	4.653,11
	5	15.820,81	4.213,41	2.096,73	8.427,40	225.420,34
	6	15.124,62	3.138,85	1.620,68	4.351,01	39.610,79
	7	14.675,00	2.777,27	1.682,99	4.611,23	2.876,61
	r	14.483,64	3.345,45	14.516,62	40.067,29	S/n=2.013,62
1 a 4 F	1	14.461,54	2.513,31	817,77	1.553,76	220,42
	2	13.110,40	2.649,00	3.114,29	7.919,14	474.656,19
	3	14.533,56	3.652,59	2.920,81	11.974,25	18,47
	4	14.992,56	3.021,48	2.846,13	8.435,85	41.682,97
	5	15.456,67	2.702,23	2.207,86	4.966,31	124.261,70
	6	15.298,67	3.274,53	1.257,20	4.827,34	49.319,19
	7	15.096,15	4.423,77	1.359,91	12.216,04	29.492,18
	r	14.495,95	3.207,02	14.523,97	51.892,69	S/n=2.937,24
5 a 9 M	1	20.197,37	3.318,28	1.228,42	2.143,76	629.193,13
	2	20.737,43	4.417,48	4.277,64	14.073,22	1477.477,50
	3	22.458,67	4.508,16	3.616,78	10.756,66	146.894,44
	4	24.770,49	6.074,40	5.412,55	28.880,90	402.225,66
	5	22.901,39	5.066,14	2.888,84	11.344,21	33.112,15
	6	25.773,08	7.482,70	2.701,62	23.662,30	583.499,31
	7	26.864,58	7.401,89	3.287,89	34.193,70	1457.429,13
	r	23.204,50	5.830,90	23.413,74	125.054,74	S/n=14.206,50
5 a 9 F	1	22.723,08	5.184,62	955,65	3.657,25	40.639,98
	2	21.312,50	5.357,56	3.614,62	17.200,79	971.695,38
	3	23.535,71	5.996,45	4.924,93	24.991,50	6.074,83
	4	23.715,69	5.764,33	4.378,59	22.208,65	16,99
	5	24.544,00	6.484,39	4.345,74	26.363,53	124.310,47
	6	25.934,78	8.001,14	2.430,44	24.444,48	465.478,31
	7	24.708,33	6.198,99	3.056,12	24.495,32	124.241,18
	r	23.629,96	6.187,77	22.706,10	143.361,50	S/n=5.258,88
10 a 19 M	1	38.596,15	14.056,26	896,13	8.193,05	3352.553,50
	2	44.932,69	12.227,53	4.557,70	29.582,80	3272.361,00
	3	46.827,51	12.241,35	8.586,70	50.385,91	2627.058,00
	4	50.408,12	14.244,31	11.787,11	94.822,22	9.773,75
	5	55.409,57	14.885,22	10.182,50	79.601,16	4228.729,00
	6	49.220,83	13.896,49	6.643,06	58.626,83	261.417,22
	7	57.091,84	13.851,58	7.959,37	76.104,73	5852.716,00
	r	50.263,92	14.290,85	50.612,57	397.316,66	S/n=32.853,50
10 a 19 F	1	47.665,38	8.497,39	1.092,24	2.916,49	48.108,61
	2	45.533,20	10.730,24	5.610,21	27.311,21	1580.123,63
	3	46.644,12	9.465,72	7.175,17	24.943,55	938.646,00
	4	50.068,65	11.406,19	12.048,63	61.753,83	219.159,31
	5	49.528,53	9.668,33	8.791,76	32.014,98	30.454,20
	6	49.672,69	11.805,92	7.167,86	44.650,65	44.988,93
	7	52.535,20	8.985,49	7.228,46	31.194,40	1610.160,00
	r	48.911,17	10.582,04	49.114,33	224.785,06	S/n=7.395,74

Tabela nº7

VARIÁVEL ESTATURA (EM CM)

Cálculos para a obtenção das médias ponderadas e suas variâncias.

GRUPOS DE PONDERAÇÃO		(a)	(b)	(c)	(d)	(e)
	h	$\bar{(xr,h)}$	$\bar{(sr,h)}$	$\bar{(w,h)(xr,h)}$	1º Termo	2º Termo
1 a 4 M	1	85,579	10,590	5,228	0,0220313	3,7639195
	2	91,467	11,653	24,092	0,1256138	1,0131667
	3	93,265	9,539	15,087	0,0485942	0,0042800
	4	94,739	11,758	15,681	0,0823334	0,2045432
	5	97,757	9,511	12,731	0,0414681	2,4403777
	6	93,846	11,698	9,882	0,0583589	0,0184191
	7	95,182	10,812	10,726	0,0674850	0,3466442
	r	93,303	11,175	93,428	0,4458848	S/n=0,0237490
1 a 4 F	1	95,615	6,935	5,207	0,0109718	0,1279833
	2	89,755	9,833	21,766	0,1072813	4,5418629
	3	93,531	9,106	19,712	0,0751669	0,0641777
	4	95,410	11,929	17,444	0,1219631	0,3223863
	5	96,645	8,281	13,739	0,0446986	0,9336459
	6	98,000	14,507	7,239	0,0820262	1,1336725
	7	96,071	12,676	8,976	0,1001930	0,3696184
	r	93,977	10,518	94,082	0,5423007	S/n=0,0294550
5 a 9 M	1	115,105	8,582	7,026	0,0144433	2,3691236
	2	116,441	10,549	24,105	0,0800353	4,9595627
	3	119,592	9,094	19,328	0,0440886	0,4912972
	4	124,267	9,381	26,804	0,0688234	1,8535172
	5	121,056	9,462	15,325	0,0398609	0,0099039
	6	125,654	9,855	13,219	0,0413376	1,9619212
	7	126,417	9,283	15,527	0,0541649	3,1714782
	r	120,996	10,214	121,325	0,3429651	S/n=0,0446640
5 a 9 F	1	120,769	11,526	5,079	0,0180762	0,0018186
	2	117,271	10,071	19,889	0,0607800	2,3298177
	3	119,873	8,722	25,084	0,0528781	0,2551104
	4	122,196	8,917	22,561	0,0531499	0,2743271
	5	121,980	9,440	21,598	0,0558789	0,1780712
	6	122,783	12,774	11,506	0,0623069	0,3054736
	7	123,375	10,656	15,260	0,0723806	0,7111660
	r	120,835	9,926	120,977	0,3754507	S/n=0,0123110
10 a 19 M	1	144,077	17,495	3,352	0,0127419	0,5098280
	2	154,077	13,915	15,659	0,0384645	2,9523587
	3	156,290	13,048	28,715	0,0574686	1,8540189
	4	159,650	14,955	37,405	0,1049341	0,0078421
	5	163,161	13,183	29,723	0,0620136	2,4866960
	6	158,733	15,069	21,465	0,0692095	0,0727238
	7	165,714	13,132	23,148	0,0686752	5,4522447
	r	159,114	14,587	159,467	0,4135073	S/n=0,0070670
10 a 19 F	1	151,462	9,997	3,471	0,0040366	0,2850136
	2	151,359	9,902	18,649	0,0232564	1,6225124
	3	153,518	8,974	23,615	0,0224208	0,3327040
	4	155,393	9,187	37,394	0,0400644	0,0395008
	5	155,424	8,162	27,589	0,0228169	0,0336902
	6	155,431	9,503	22,429	0,0289314	0,0282617
	7	158,735	9,309	21,841	0,0334804	1,9312312
	r	154,782	9,297	154,988	0,1750069	S/n=0,0070670

Tabela nº8

PESO (GR)

Número de menores (r), médias (\bar{x}) e variâncias estimadas $v(\bar{x})$ segundo procedimentos e grupos de ponderação.

GRUPOS DE PONDERAÇÃO	r	procedimento	\bar{x}	$v(\bar{x})$
1 a 4 M	269	não ponderado	14.484,64	41.606
		ponderado	14.516,62	42.081
1 a 4 F	205	não ponderado	14.496,95	50.171
		ponderado	14.524,97	54.830
5 a 9 M	274	não ponderado	23.204,50	124.085
		ponderado	23.413,74	139.261
5 a 9 F	272	não ponderado	23.629,96	140.767
		ponderado	23.706,10	148.620
10 a 19 M	485	não ponderado	50.263,92	421.090
		ponderado	50.612,57	430.170
10 a 19 F	490	não ponderado	48.911,17	228.530
		ponderado	49.114,33	232.181

Tabela nº9.

ESTATURA (CMS)

Número de menores (r), médias (\bar{x}) e variâncias estimadas $v(\bar{x})$ segundo procedimentos e grupos de ponderação.

GRUPOS DE PONDERAÇÃO	r	procedimento	\bar{x}	$v(\bar{x})$
1 a 4 M	274	não ponderado	93,30	0,4557
		Ponderado	93,43	0,4696
1 a 4 F	213	não ponderado	93,98	0,5194
		Ponderado	94,08	0,5718
5 a 9 M	273	não ponderado	120,99	0,3822
		Ponderado	121,34	0,3876
5 a 9 F	272	não ponderado	120,83	0,3622
		Ponderado	120,98	0,3878
10 a 19 M	484	não ponderado	159,11	0,4396
		Ponderado	159,47	0,4443
10 a 19 F	490	não ponderado	154,78	0,1764
		Ponderado	154,99	0,1821

CAPÍTULO 6 - COMENTÁRIOS FINAIS

Infelizmente, a natureza empírica deste trabalho não permite alcançar a concreta avaliação da eficácia do ajuste efetuado, quanto à redução do vício potencialmente atribuído à ausência de resposta. A utilidade dos resultados, apesar do acréscimo na variância, reside no reconhecimento de que o vício se reduziu na medida em que foram adotadas as mais diferentes taxas de resposta para a determinação dos pesos amostrais. O delineamento, o tamanho da amostra inicial e a amplitude da idade foram as principais limitações impostas aos resultados obtidos. Face à consistência dos resultados verificada numa situação em que a taxa global de resposta foi igual a 79,10% e considerando o procedimento técnico como extremamente simples e de baixo custo dentro do contexto operacional de um levantamento, aconselha-se a aplicação do estimador por classes ponderadas, principalmente nas situações em que taxas globais de resposta inferiores a 80% sejam alcançadas.

Como foi verificado, a identificação da variável (ou das variáveis) critério para a formação das classes ponderadas é fundamental no procedimento e requer a existência da informação, na amostra, para as unidades respondentes e para as não respondentes. O número de cômodos do domicílio foi, no inquérito do Instituto da Criança, a variável com maior potencial para caracterizar os não-respondentes. Como se trata de uma variável cuja informação pode ser obtida com segurança e agilidade, sem a participação do informante, evidencia-se a importância de que tentativas sejam feitas no sentido de se conseguir a informação para as famílias não-respondentes durante o desenvolvimento do trabalho de campo, afastando-se assim a necessidade de sub-amostrar esse segmento e de calcular as estimativas que se fizeram necessárias para a aplicação do procedimento.

Tomar a família ou o domicílio como a unidade amostral base para o cálculo das taxas de resposta mostrou-se prático e além de identificar pontos de resistência durante a realização das entrevistas, facilitou o acesso à informação para as unidades não respondentes, viabilizando a aplicação do estimador por classes ponderadas. Para a obtenção das estimativas ponderadas ao nível dos menores de 19 anos, aplicando as taxas de resposta calculadas na amostra das famílias, admitiu-se que em cada grupo de ponderação o impacto da

correlação intra-classe é desprezível, permitindo a aplicação da análise estatística para uma amostra de elementos estratificada em duas fases.

Casos mais gerais, em que a homogeneidade dos componentes das famílias, quanto às variáveis para as quais as estimativas são desejadas, seja considerada relevante, podem ser tratados pela aplicação do estimador razão estratificado (10,25) em que cada média ($\bar{x}_{r,h}$) calculada na amostra de elementos é uma razão estimada para o estrato composto pelo conjunto das famílias respondentes. O estudo das propriedades e da viabilidade de aplicação desse procedimento deverá, portanto, ser objeto de investigações posteriores.

REFERÊNCIAS BIBLIOGRÁFICAS

- 01 - AMERICAN STATISTICAL ASSOCIATION. -Report on the ASA conference on survey of human populations. *Amer.Stat.*, 8 (1):30-4, 1974.
- 02 - AUSTIN, M.A. et al. - The effect of response bias on the odds ratio. *Amer.J.Epidem.*, 114: 137-43, 1981.
- 03 - BAILAR, B.A. & LANPHIER, C.M. - A pilot study to develop survey methods to assess survey practices. *Amer.Stat.*, 32 (4): 130-2, 1978.
- 04 - BANKS, M.J. - An indication of the effect of noninterview adjustment and post-stratification on estimates from a sample survey. In:MEETING OF THE AMERICAN STATISTICAL ASSOCIATION, Chicago, 1977. *Proceedings_of_the_Social_Statistics_Section*. Washington, ASA, c1978. p.291-95.
- 05 - BUREAU of the CENSUS. - *The current population survey-a report_on_methodology*. Washington, D.C., 1963. (Technical Paper, 7).
- 06 - CARLSSON, G.S. et al. - Comparison of participants and non-participants in a population study of injuries. The study of men born in 1913 and 1923 . *Scand.J.soc.Med.*, 13(1): 15-22, 1985.
- 07 - CENTRAL BUREAU OF STATISTICS. - *Indonésia_fertility_survey_1976*. Jakarta, Indonésia, 1978.v.i.
- 08 - CHAVANCE M. - Essai de correction du biais du aux non-réponses. *Rev.Épidém.Santé_publ.*, 31:39-44, 1983.
- 09 - COBB, S. et al. - Differences between respondents and non-respondents in a morbidity survey involving clinical examination. *J.chron.Dis.* 6:95-108, 1957.

- 10 - COCHRAN, W.G. - Sample techniques. 3rd.ed. New York, John Wiley, 1965.
- 11 - COMSTOCK, G. W. et al. - Characteristics of respondents and nonrespondents to a questionnaire for estimating community mood. Amer.J.Epidem., 92: 233-39, 1973.
- 12 - CRAWFORD, A. - Response rates in british general population surveys of alcohol consumption: a comparison. Alcohol & Alcohol, 19 (2): 141-5, 1984.
- 13 - CRIQUI, M.H. - Response bias and risk ratios in epidemiologic studies. Amer.J.Epidem., 109: 394-9, 1979.
- 14 - CRIQUI, M.H. et al. - The effect of non-response on risk ratios in a cardiovascular disease study. J.Chron.Dise., 32:633-38, 1979.
- 15 - DUNKELBERG, C.W. & DAY, G.S. - Nonresponse bias and callbacks in sample surveys. J.Market.Res., p.160-8, May, 1973.
- 16 - FELLER, W. - Introdução à teoria das probabilidades e suas aplicações - Parte I - Espaços amostrais - discisos. São Paulo, Ed.Edgard Blucher, 1976.
- 17 - FORTHOFER, R.N. - Investigation of nonresponse bias in NHANES II. Amer.J.Epidem., 112:507-15, 1983.
- 18 - HAWKINS, D.F. - Estimation of nonresponse bias. Sociol.-Meth.Res., 3:461-85, 1975.
- 19 - ISKRRANT M.A. et al. - Effect of incomplete information on estimating prevalence of disease. Publ.Hlth.Rep., 62:384-89, 1982.
- 20 - KALSSBEEK W.D.- A conceptual review of survey error due to nonresponse. In:MEETING OF THE AMERICAN STATISTICAL

- ASSOCIATION, Houston, Texas, 1980. Proceeding_of_the_Section_on_Survey_Research_Methods, Washington, ASA, 1980. p.131-6.
- 21 - KALTON, G. et al. - Some problems of nonresponse and nonresponse adjustment in the survey of income and program participation. In: MEETING OF THE AMERICAN STATISTICAL ASSOCIATION, Houston, Texas, 1980. Proceeding_of_the_Section_on_Survey_Research_Methods, Washington, ASA, 1980. p.501-6.
- 22 - KALTON, G. - Conceptual and theoretical framework for survey sampling. In: IASS MEETINGS, 4th, Buenos Aires, 1981. Proceedings. Buenos Aires, 1981. p.514-30.
- 23 - KALTON, G. - Compensating_for_missing_survey_data. Ann Arbor, Institute for Social Research. The University of Michigan, 1983. (Research Report Series).
- 24 - KALTON, G. - Introduction_to_survey_samploing. Beverly Hills, Sage University, 1983. (Paper Series on Quantitative Application in the Social Sciences, n.07-035).
- 25 - KISH, L. - Survey_samploing. New York, John Wiley, 1965.
- 26 - LLANDIS J.R. et al. - A statistical methodology for analyzing data from a complex survey: the first national health and nutrition examination survey. Vital_Hltb_Stat, Ser.2 (92): 1-52, 1982.
- 27 - MACRAE, F.A. et al. - Colorectal cancer:Knowledge and attitudes of doctors in Victoria. Aust_N.Z.J.Med., 12: 278-83, 1982.
- 28 - MANGIONE, T.W. et al. - Collecting sensitive data- a comparison of three survey strategies. Sociol_Meth_Res. 10:133-7-46, 1982.
- 29 - MEHEUS, A. - Response raate in prevalence studies. In: HOLLAND W.W. & KARHAUSEN,L., eds. Health_care_and_epidemiology. London, H.Kimpton, 1978.p.105-7.

- 30 - MOOD, A.M. et al.- *Introduction to the theory of statistics*. 3rd.ed. Tokyo, McGraw-Hill Kogakusha,Ltd., 1963.
- 31 - NAPIER, J.A. - Field methods and response rates in the Tecumseh Community Health Study. *American Public Health*, 52:208-16, 1962.
- 32 - RUBIN, D.B. - Formalizing subjective notions about the effect of nonrespondents in sample survey. *J.Amer.Stat.Ass.*, 35: 538-43, 1977.
- 33 - SCHAIBLE, W.L. - Quality control in a national health examination survey. *Vital Health Stat. Ser.2*, (44): 1-21, 1972.
- 34 - SILMAN, A.J. & LOCKE, C.M. - Blood pressure distribution in responders and initial non-responders in a population screening study. *J.Epidem.Community Health*, 36: 248-50, 1982.
- 35 - SILVA, E.P.C. - Encuesta demográfica sobre el bajo registro de nacimientos y muertes. In: REUNION LATINOAMERICANA DE MUESTREO, 1º, Mexico, 1978. Memoria. Mexico, 1979.p.165-70.
- 36 - SILVA, N.N. - Ausência de resposta em pesquisas por amostragem. São Paulo, 1979. [Dissertação de Mestrado-Faculdade de Saúde Pública da USP].
- 37 - SINGH, B. & SEDRANSK, J. - Sample size selection in regression analysis when there is nonresponse. *J.Amer.Stat.Ass.*, Z3 (362): 362-5, 1978.
- 38 - STEEH, C.G. - Trends in nonresponse rates, 1952-1979. *Publ.Opin.Quart.*, 45:40-57, 1981.
- 39 - TIBBLIN, G. - A population study of 50-year-old men: an analysis of the non-participation group. *Acta Med.Scand.*, 128 (4): 453-9, 1965.

- 40 - UNITED NATIONS - National household survey capability programme. Non-sampling errors in household surveys: Sources, Assessment and control; Preliminary version. New York, 1982.
- 41 - WAYNE, W.D. et al. - An adjustment for nonresponse in sample surveys. Educ.psychol.Meas., 42: 57-67, 1982.
- 42 - WILHELMSEN, L. et al. - A comparison between participants and non-participants in a primary preventive trial. J.chron.Dis., 29: 331-39, 1976.

ANEXO Nº 1

$$\text{Vício } (\bar{x}_r) = (NR/N) (\bar{X}_R - \bar{X}_{NR})$$

$E(\bar{x}_r) = E_1 E_2(\bar{x}_r)$ onde E_2 é a esperança condicional (16.30) para um fixado número de respondentes ($r > 0$) na amostra, e E_1 é a esperança calculada sobre todos os possíveis diferentes valores que r pode assumir.

Então:

$$E(\bar{x}_r) = E_1 \left(\sum E_2(\bar{x}_r, i) / r \right) = E_1(\bar{X}_R) = \bar{X}_R$$

Considerando a população dividida em respondentes (R) e não-respondentes (NR), pode-se definir a média populacional pela expressão:

$$\bar{X} = (R/N) \bar{X}_R + (NR/N) \bar{X}_{NR} \quad (A1.1)$$

Então:

$$\begin{aligned} \text{Vício } (\bar{x}_r) &= E(\bar{x}_r) - \bar{X} = \bar{X}_R - [(R/N) \bar{X}_R + (NR/N) \bar{X}_{NR}] \\ &= [(1 - R/N) \bar{X}_R] - [(NR/N) \bar{X}_{NR}] \\ &= (NR/N) (\bar{X}_R - \bar{X}_{NR}) \end{aligned}$$

ANEXO N°2

$$\text{Vício } (\bar{x}_{cp}) = S\bar{c}(W,h) (NR,h/N,h) (\bar{X}_R,h - \bar{X}_{NR},h)]$$

$E(\bar{x}_{cp}) = E_1 E_2 (\bar{x}_{cp})$, onde E_2 é a esperança condicional (16,30) para fixado número de respondentes ($r,h \geq 0$) da classe h , na amostra, e E_1 é a esperança calculada sobre todos os possíveis diferentes valores de r,h .

Então:

$$\begin{aligned} E_1 E_2 (\bar{x}_{cp}) &= E_1 E_2 [S\bar{c}(w,h)(\bar{X}_r,h))] = E_1 [S E_2 (w,h)(\bar{X}_r,h)] \\ &= E_1 [S\bar{c}(w,h)(\bar{X}_R,h)] \\ &= S\bar{c}(W,h)(\bar{X}_R,h) \end{aligned} \quad (\text{A2.1})$$

Considerando a população dividida em estratos h nos quais existem os respondentes e os não-respondentes, define-se a média populacional por:

$$\begin{aligned} \bar{X} &= S(W,h) \bar{X}_h \\ &= S\bar{c}(W,h) [(R,h/N,h)(\bar{X}_R,h) + (NR,h/N,h)(\bar{X}_{NR},h)] \\ &= S\bar{c}(W,h)(\bar{X}_R,h) - S\bar{c}(W,h)(NR,h/N,h)(\bar{X}_R,h - \bar{X}_{NR},h) \end{aligned} \quad (\text{A2.2})$$

Como $\text{Vício } (\bar{x}_{cp}) = E(\bar{x}_{cp}) - \bar{X}$ e substituindo esses termos pelas expressões (A2.1) e (A2.2), tem-se:

$$\begin{aligned} \text{Vício } (\bar{x}_{cp}) &= S\bar{c}(W,h)(\bar{X}_R,h) - S\bar{c}(W,h)(\bar{X}_R,h) + \\ &\quad + S\bar{c}(W,h)(NR,h/N,h)(\bar{X}_R,h - \bar{X}_{NR},h) \\ &= S\bar{c}(W,h)(NR,h/N,h)(\bar{X}_R,h - \bar{X}_{NR},h) \end{aligned}$$

ANEXO Nº3

$$\begin{aligned} \text{Vício } (\bar{x}_r) &= \text{SE}(W, h)(\bar{X}_R, h - \bar{X}_{NR}, h) (R, h/N, h - R/N)/(R/N)] + \\ &+ \text{SE}(W, h)(N_R, h/N, h)(\bar{X}_R, h - \bar{X}_{NR}, h)] \end{aligned}$$

Considerando a população dividida em estratos h com diferentes taxas de resposta $(R, h/N, h)$, a média populacional calculada para o estrato dos respondentes poderá ser expressa por:

$$\begin{aligned} \bar{X}_R &= \text{SE}(R, h)(\bar{X}_R, h)] / \text{S}(R, h) \\ &= \text{SE}(R, h)(N, h/N, h)(\bar{X}_R, h)] / \text{SE}(R, h)(N, h/N, h) \\ &= \text{SE}(R, h/N, h)(N, h/N)(\bar{X}_R, h)] / \text{SE}(R, h/N, h)(N, h/N)] \\ &= \text{SE}(R, h/N, h)(W, h)(\bar{X}_R, h) / (R/N) \end{aligned} \quad (\text{A3.1})$$

Sendo $\text{Vício } (\bar{x}_r) = \bar{X}_R - \bar{X}$ e substituindo esses termos pelas expressões (A3.1) E (A2.2) tem-se:

$$\begin{aligned} \text{Vício } (\bar{x}_r) &= \text{SE}(R, h/N, h)(W, h)(\bar{X}_R, h)] / (R/N) - \text{SE}(W, h)(\bar{X}_R, h)] \\ &\quad + \text{SE}(W, h)(N_R, h/N, h)(\bar{X}_R, h - \bar{X}_{NR}, h)] \\ \\ &= \text{SE}(W, h)(\bar{X}_R, h)] \text{E}(R, h/N, h)/(R/N) - 1)] \\ &\quad + \text{SE}(W, h)(N_R, h/N, h)(\bar{X}_R, h - \bar{X}_{NR}, h)] \\ \\ &= \text{SE}(W, h)(\bar{X}_R, h)] \text{E}(R, h/N, h) - (R/N)]/(R/N)] \\ &\quad + \text{SE}(W, h)(N_R, h/N, h)(\bar{X}_R, h - \bar{X}_{NR}, h)] \end{aligned}$$

Desde que $SE(W,h)(R,h/N,h) - (R/N) = 0$, então:

$$\begin{aligned} Vício(\bar{x}_{cp}) &= SE(W,h)(\bar{x}_R, h - \bar{x}_R)(R, h/N, h - R/N)/(R/N) \\ &\quad + SE(W,h)(NR, h/N, h)(\bar{x}_R, h - \bar{x}_{NR}, h) \end{aligned}$$

ANEXO N° 4

Variáveis sócio-demográficas, processamento detalhado:
Distribuições de freqüências, estatísticas qui-quadrado e ta-
xas diferenciais de resposta.

A4.1. TAMANHO DA FAMÍLIA

SAS
TABLE OF NPES BY RESP
NPES

FREQUENCY COL PCT	NÃO		TOTAL
	RESPOND	RESPOND	
<= 2	1	28	29
	1.75	2.87	
3	14	178	192
	24.56	18.22	
4	18	249	257
	31.58	25.49	
5	14	229	243
	24.56	23.44	
6	4	141	145
	7.02	14.43	
>= 7	6	152	158
	10.53	15.56	
TOTAL	57	977	1034

STATISTICS FOR 2-WAY TABLES

CHI-SQUARE 5.211 DF= 5 PROB=0.3907
 PHI 0.071
 CONTINGENCY COEFFICIENT 0.071
 CRAMER'S V 0.071
 LIKELIHOOD RATIO CHISQUARE 5.650 DF= 5 PROB=0.3418
 TABELA REFERENTE A NPES COM AJUSTAMENTO PARA M=216

OBS	NPES	COUNT	COUNTL	TOTAL	PR
1	2	29	3.8246	31.825	0.879824
2	3	173	53.5439	231.544	0.768753
3	4	249	68.8421	317.842	0.783408
4	5	229	53.5439	282.544	0.810494
5	6	141	15.2982	156.298	0.902121
6	7	152	22.9474	174.947	0.868833

A4.2. NÚMERO DE MENORES DE 1 ANO

SAS
TABLE OF M1 BY FESP

M1

FREQUENCY		NÃO		TOTAL
COL	PCT	RESPOND.	RESPOND.	
0		50	828	878
		89.29	84.75	
>= 1		6	149	155
		10.71	15.25	
TOTAL		56	977	1033

STATISTICS FOR 2-WAY TABLES

CHI-SQUARE	0.855	DF =	1	PROB=0.3552
PHI	0.029			
CONTINGENCY COEFFICIENT	0.029			
CRAMER'S V	0.029			
LIKELIHOOD RATIO CHISQUARE	0.929	DF =	1	PROB=0.3350
CONTINUITY ADJ. CHI-SQUARE	0.536	DF =	1	PROB=0.4641
FISHER'S EXACT TEST (1-TAIL)				PROB=0.2381
(2-TAIL)				PROB=0.4437

TABELA REFERENTE A M1 COM AJUSTAMENTO PARA M=218

OBS	M1	COUNT	COUNTL	TOTAL	PR
1	0	828	194.643	1022.64	0.809667
2	1	149	23.357	172.36	0.864484

A4.3. NÚMERO DE MENORES DE 1 A 4 ANOS

SAS
TABLE OF M14 BY RESP
M14

FREQUENCY	NÃO	COL PCT	RESPOND.	RESPOND.	TOTAL
0	32	57.14	552	56.50	584
1	24	42.86	425	43.50	449
TOTAL	56		977		1033

STATISTICS FOR 2-WAY TABLES

CHI-SQUARE	0.009	DF= 1	PROB=0.9247
PHI	0.003		
CONTINGENCY COEFFICIENT	0.003		
CRAZER'S V	0.003		
LIKELIHOOD RATIO CHISQUARE	0.009	DF= 1	PROB=0.9247
CONTINUITY ADJ. CHI-SQUARE	0.002	DF= 1	PROB=0.9648
FISHER'S EXACT TEST (1-TAIL)			PROB=0.5197
(2-TAIL)			PROB=1.0000

TABELA REFERENTE A M14 COM AJUSTAMENTO PARA M=218

OBS	M14	COUNT	COUNTL	TOTAL	PE
1	0	552	124.571	676.571	0.815878
2	1	425	93.429	518.429	0.819785

A4.4. NÚMERO DE MENORES DE 5 a 9 ANOS

SAS
TABLE OF M59 BY RESP

M59

		FREQUENCY NÃO			
	COL PCT	!RESPOND.	!RESPOND.		TOTAL
0	1	26	554	1	580
		46.43	56.70		
1	1	21	284	1	305
		37.50	29.07		
2	1	7	112	1	119
		12.50	11.46		
>= 3	1	2	27	1	29
		3.57	2.76		
TOTAL		56	977		1033

STATISTICS FOR 2-WAY TABLES

CHI-SQUARE	2.444	DF= 3	PROB=0.4855
PHI	0.049		
CONTINGENCY COEFFICIENT	0.049		
CRAMER'S V	0.049		
LIKELIHOOD RATIO CHISQUARE	2.397	DF= 3	PROB=0.4942
TABELA REFERENTE A M59		COM AJUSTAMENTO PARA M=216	

OBS	M59	COUNTI	COUNTL	TOTAL	PR
1	0	554	101.214	655.214	0.845525
2	1	284	81.750	365.750	0.776487
3	2	112	27.250	139.250	0.804309
4	3	27	7.786	34.786	0.776181

A4.5. NÚMERO DE MENORES DE 10 a 19 ANOS

S A S
TABLE OF M1019 BY RESP
M1019

FREQUENCY	NÃO		TOTAL
	COL PCT	RESPOND.	
0	26	378	404
	46.43	38.69	
1	16	279	295
	28.57	28.56	
>= 2	14	320	334
	25.00	32.75	
TOTAL	56	977	1033

STATISTICS FOR 2-WAY TABLES

CHI-SQUARE	1.796	DF = 2	PROB=0.4074
PHI	0.042		
CONTINGENCY COEFFICIENT	0.042		
CHAMER'S V	0.042		
LIKELIHOOD RATIO CHISQUARE	1.832	DF = 2	PROB=0.4002
TABELA REFERENTE A M1019 COM AJUSTAMENTO PARA M=218			

OBS	M1019	COUNT	COUNTL	TOTAL	PR
1	0	378	101.214	479.214	0.788791
2	1	279	62.286	341.286	0.817497
3	2	320	54.500	374.500	0.854473

A4.6. IDADE DO CHEFE DA FAMÍLIA

S A S
TABLE OF IDADE BY RESP

IDADE

FREQUENCY

COL PCI	NÃO	RESPOND.	TOTAL
<= 30	27	491	518
	54.00	50.26	
40	16	289	305
	32.00	29.58	
50	6	153	169
	12.00	16.68	
>= 60	1	34	35
	2.00	3.48	
TOTAL	50	977	1027

STATISTICS FOR 2-WAY TABLES

CHI-SQUARE	1.166	DF= 3	PROB=0.7612
PHI	0.034		
CONTINGENCY COEFFICIENT	0.034		
CRAMER'S V	0.034		
LIKELIHOOD RATIO CHISQUARE	1.275	DF= 3	PROB=0.7351
TABELA REFERENTE A IDADE COM AJUSTAMENTO PARA M=218			

OBS	IDADE	COUNT	COUNTL	TOTAL	PR
1	30	491	117.72	608.72	0.806611
2	40	289	69.76	358.76	0.805552
3	50	163	26.16	189.16	0.861704
4	60	34	4.36	38.36	0.886340

A4.7. SEXO DO CHEFE DA FAMÍLIA

S A S
TABLE OF SEXF BY RESP

SEXF

FREQUENCY	NÃO		TOTAL
	COL PCT	RESPOND.	
1	51	880	931
	94.44	90.07	
2	3	97	100
	5.56	9.93	
TOTAL	54	977	1031

STATISTICS FOR 2-WAY TABLES

CHI-SQUARE	1.117	DF= 1	PROB=0.2905
PHI	0.033		
CONTINGENCY COEFFICIENT	0.033		
CRAMER'S V	0.033		
LIKELIHOOD RATIO CHISQUARE	1.291	DF= 1	PROB=0.2559
CONTINUITY ADJ. CHI-SQUARE	0.674	DF= 1	PROB=0.4118
FISHER'S EXACT TEST (1-TAIL)			PROB=0.2114
(2-TAIL)			PROB=0.4747

TABELA REFERENTE A SEXF COM AJUSTAMENTO PARA M=218

OBS	SEXF	COUNT	COUNTL	TOTAL	PR
1	1	880	205.889	1085.89	0.810396
2	2	97	12.111	109.11	0.889002

4.8. INSTRUÇÃO DO CHEFE DA FAMÍLIA

S A S
TABLE OF INST BY RESP
INST

FREQUENCY COL PCT	NÃO RESPOND.		TOTAL
	1	20 41.67	581 59.65
1	20	581	601
2	14	116	130
3	13	148	161
>= 4	1	129	130
TOTAL	48	974	1022

STATISTICS FOR 2-WAY TABLES

CHI-SQUARE	21.810	DF= 3	PROB=0.0001
PHI	0.146		
CONTINGENCY COEFFICIENT	0.145		
CRAMER'S V	0.146		
LIKELIHOOD RATIO CHISQUARE	20.958	DF= 3	PROB=0.0001

TABELA REFERENTE A INST COM AJUSTAMENTO PARA M=218

OBS	INST	COUNT	COUNTL	TOTAL	PR
1	1	581	90.8333	671.833	0.864798
2	2	116	63.5833	179.583	0.645940
3	3	148	59.0417	207.042	0.714832
4	4	129	4.5417	133.542	0.965991

A4.9. TIPO DA HABITAÇÃO

SAS
TABLE OF CONDHAB BY RESP
CONDHAB

		NÃO		TOTAL
COL	PCT	RESPOND.	RESPOND.	
1	1	54	670	724
		81.82	69.43	
2	1	5	71	76
		7.58	7.36	
3	1	7	224	231
		10.61	23.21	
TOTAL		66	965	1031

STATISTICS FOR 2-WAY TABLES

CHI-SQUARE	5.736	DF= 2	PROB=0.0568
PHI	0.075		
CONTINGENCY COEFFICIENT	0.074		
CRAMER'S V	0.075		
LIKELIHOOD RATIO CHISQUARE	6.672	DF= 2	PROB=0.0356
TABELA REFERENTE A CONDHAB COM AJUSTAMENTO PARA N=252			

OBS	CONDHAB	COUNT	COUNT L	TOTAL	PR
1	1	670	206.182	876.182	0.764681
2	2	71	19.091	90.091	0.788093
3	3	224	26.727	250.727	0.893401

A4.10. NÚMERO DE CÔMODOS DO DOMICÍLIO

SAS
TABLE OF NCOM BY RESP
NCOM

FREQUENCY COL PCT	NÃO RESPOND.		TOTAL
	1	51	
	1.54	5.22	
2	9	184	193
	13.85	18.83	
3	5	178	183
	7.69	18.22	
4	12	209	221
	18.46	21.39	
5	8	161	169
	12.31	16.48	
6	10	98	108
	15.38	10.03	
>= 7	20	96	116
	30.77	9.83	
TOTAL	65	977	1042

STATISTICS FOR 2-WAY TABLES

CHI-SQUARE 32.918 DF = 6 PROB=0.0001
 PHI 0.178
 CONTINGENCY COEFFICIENT 0.175
 CRAMER'S V 0.178
 LIKELIHOOD RATIO CHISQUARE 27.086 DF = 6 PROB=0.0001
 TABELA REFERENTE A NCOM COM AJUSTAMENTO PARA M=218

OBS	NCOM	COUNT	COUNTL	TOTAL	PR
1	1	51	3.3538	54.354	0.938296
2	2	184	30.1846	214.185	0.859072
3	3	178	16.7692	194.769	0.913902
4	4	209	40.2462	249.246	0.838528
5	5	161	26.8308	187.831	0.857155
6	6	98	33.5385	131.538	0.745029
7	7	96	67.0769	163.077	0.588679