

## **Estimando A Precisão Das Estimativas Das Taxas De Mortalidade Obtidas A Partir Da PNAD**

Palavras-chave: mortalidade infantil; estimação indireta; qualidade das estimativas; dados amostrais.

**Pedro Luis do Nascimento Silva**(IBGE)

**Djalma Galvão Carneiro Pessoa**(IBGE)

### **1. Introdução**

Em países com sistemas de registro civil de cobertura deficiente, como ainda é o caso do Brasil, as taxas de mortalidade e probabilidades de morte que definem as tábuas de mortalidade são estimadas usando métodos indiretos (IBGE, 2002, pag. 17). O principal destes métodos, que vem sendo utilizado pelo IBGE ao longo dos anos, requer estimativas de várias quantidades populacionais que podem ser obtidas com base em censos e/ou pesquisas por amostragem.

As quantidades populacionais requeridas para estimar as taxas de mortalidade infantil usando este método incluem o número total de mulheres, o número total de filhos tidos nascidos vivos e o número total de filhos nascidos vivos que já morreram, todos referentes a cada um dos grupos de mulheres classificadas nos grupos quinquenais de idade: 15 a 19 anos, 20 a 24 anos, ..., 45 a 49 anos. De posse de tais informações, cálculos relativamente simples permitem obter estimativas de taxas de mortalidade para idades específicas, conforme descrito em United Nations (1983, cap. III). Caso haja interesse, o mesmo método pode ser usado para obter estimativas específicas por sexo, desde que existam informações sobre os filhos nascidos vivos e mortos por sexo, como ocorre nos censos demográficos e PNADs aqui no Brasil (ver também IBGE, 2002).

O emprego do método indireto introduz uma fonte de incerteza sobre os resultados das estimativas das taxas de mortalidade. Os efeitos dessa fonte de incerteza já foram bastante estudados pelos demógrafos, que aprenderam a interpretar os resultados à luz dessa incerteza e a fazer análises corretas mesmo quando ela não pode ser evitada. Esta incerteza está presente mesmo quando os dados requeridos para a aplicação do método são obtidos através de censos ou pesquisas exaustivas.

No Brasil, entretanto, os dados disponíveis para aplicação do método indireto de

estimação das taxas de mortalidade são sempre obtidos de pesquisas amostrais. Nos anos de recenseamento, as informações são coletadas para uma amostra bem grande (cerca de 12% dos domicílios e das pessoas são pesquisados no país). Nos outros anos, a fonte principal de informações é a PNAD (Pesquisa Nacional por Amostra de Domicílios), cuja amostra é de apenas cerca de 0,2% dos domicílios e pessoas residentes no Brasil.

Parte das informações requeridas para estimar as taxas de mortalidade nas idades mais jovens são contagens de eventos raros. Estimativas recentes situam o nível da mortalidade infantil no Brasil entre 30 e 40 por mil, isto é, entre 3% ou 4%. Sabe-se também que amostras "pequenas" podem levar a estimativas imprecisas de proporções ou contagens de eventos raros.

Conseqüentemente, há uma segunda fonte de incerteza na estimação das taxas de mortalidade que não vem sendo reconhecida ou avaliada com regularidade pelos que calculam e interpretam essas estimativas. Trata-se da incerteza resultante da estimação amostral das quantidades requeridas para aplicação do método indireto, que também vai se refletir em incerteza sobre os resultados das estimativas das taxas de mortalidade. Tal incerteza pode ser grande para amostras pequenas, e se não for levada em conta, pode levar a interpretações equivocadas de variações nas estimativas, particularmente quando estas se referirem a domínios de análise "detalhados", como é o caso da estimação da mortalidade infantil por estados, etc.

Para avaliar o grau de incerteza ou de (im)precisão das estimativas indiretas de mortalidade decorrente do emprego de amostragem para estimar as quantidades populacionais requeridas, desenvolvemos neste trabalho um método aproximado para estimar os desvios padrão das estimativas das taxas de mortalidade. Sob algumas hipóteses fracas, conhecidos esses desvios padrão das taxas de mortalidade estimadas, é possível determinar intervalos de confiança para as verdadeiras taxas populacionais de mortalidade, e com isso avaliar o grau de (im)precisão das estimativas e fazer comparações com menor risco de chegar a conclusões erradas.

Na seção 2 apresentamos o método desenvolvido para estimar os desvios padrão das estimativas indiretas das taxas de mortalidade obtidas com base em informações de pesquisas amostrais. O método desenvolvido para estimar o desvio padrão das estimativas indiretas das taxas de mortalidade pode ser aplicado a dados de qualquer pesquisa por amostragem que tenha empregado amostragem probabilística, e para a qual

possam ser calculadas estimativas usuais de total e da variância das estimativas de total. Na seção 3 apresentamos os resultados obtidos com a aplicação deste método às informações coletadas na PNAD 1999. Na seção 4 apresentamos algumas conclusões e recomendações emanadas da análise dos resultados obtidos nesse trabalho.

## **2. Estimando Desvios Padrão Das Estimativas Indiretas Das Taxas De Mortalidade**

As condições necessárias para utilização do método de estimação aqui proposto para estimar os desvios padrão das estimativas indiretas das taxas de mortalidade infantil são as seguintes:

- o emprego de amostra probabilística para a qual possam ser calculadas estimativas não viciadas (ou ao menos consistentes) de total e da variância das estimativas de total;
- o emprego de amostras grandes o bastante para permitir aproximar a variância das estimativas indiretas das taxas de mortalidade usando o método de Linearização de Taylor (ver Pessoa e Nascimento Silva, 1998, cap. 3);
- a disponibilidade de informações sobre os pesos das unidades amostrais, sobre o plano amostral utilizado para obter a amostra, e de informações sobre as variáveis de estrutura do plano amostral, tais como variáveis de estratificação, de identificação de conglomerados e probabilidades de seleção das unidades de amostragem nas diversas etapas;
- a disponibilidade de programas de computador capazes de calcular estimativas de totais a partir dos microdados amostrais, e das matrizes de variância dos totais estimados, para cada um dos domínios de análise de interesse.

Todas estas condições são satisfeitas para os dados das PNADs e das amostras dos Censos Demográficos brasileiros.

Para apresentar o método, vamos primeiro introduzir alguma notação para descrever as quantidades populacionais requeridas e as quantidades resultantes da aplicação do método indireto para estimar as taxas de mortalidade. Essa notação será a mesma para qualquer domínio geográfico de interesse (um estado ou região, ou mesmo o total do país). Sendo assim, para um domínio geográfico qualquer de interesse, denotamos por:

$M_i$  o número total de mulheres na população na faixa etária  $i$ ;

$P_i$  o número total na população de filhos nascidos vivos das mulheres na faixa etária  $i$ ;

$D_i$  o número total na população de filhos nascidos vivos que já morreram, das mulheres na faixa etária  $i$ ;

$Q_i$  a taxa de mortalidade a ser estimada com base na mortalidade dos filhos das mulheres da faixa etária  $i$ ;

onde  $i=1,2,\dots,7$  representa as sete faixas etárias quinquenais para as quais se obtém dados da fecundidade e da mortalidade das mulheres pesquisadas.

Vale observar que a definição acima foi feita considerando totais de nascimentos de ambos os sexos. Entretanto, é possível aplicar o método a ser aqui descrito para estimar taxas de mortalidade específicas por sexo, desde que os valores de  $P_i$  e  $D_i$  sejam definidos da maneira adequada para representar os nascimentos e os óbitos de filhos do sexo específico para o qual são desejadas as estimativas da mortalidade.

O quadro 1 apresenta a descrição detalhada das faixas etárias, seus índices de identificação (valores de  $i$ ) e das taxas de mortalidade estimadas com dados de cada uma.

Quadro 1 - Descrição das probabilidades de morte estimadas pelo método indireto

$i$	Faixa etária da mulher	$Q_i$
1	15 a 19 anos	$Q_1=q(1)$
2	20 a 24 anos	$Q_2=q(2)$
3	25 a 29 anos	$Q_3=q(3)$
4	30 a 34 anos	$Q_4=q(5)$
5	35 a 39 anos	$Q_5=q(10)$
6	40 a 44 anos	$Q_6=q(15)$
7	45 a 49 anos	$Q_7=q(20)$

onde  $q(x)$  é a probabilidade de morte até a idade exata  $x$ , isto é, a taxa de mortalidade para a idade exata  $x$ .

O método indireto para estimar  $q(x)$  foi proposto para calcular os valores de  $Q_i$

usando dados de censos mediante a fórmula geral:

$$Q_i = \left( a_i + b_i \frac{P_1}{M_1} \frac{M_2}{P_2} + c_i \frac{P_2}{M_2} \frac{M_3}{P_3} \right) \frac{D_i}{P_i} \quad i = 1, 2, \dots, 7 \quad (1)$$

onde  $a_i$ ,  $b_i$ , e  $c_i$  são coeficientes conhecidos, considerados constantes no presente trabalho.

Na verdade, esses coeficientes podem ser diferentes dependendo da escolha do usuário interessado nas estimativas das taxas de mortalidade. Os valores selecionados para uso neste trabalho são os geralmente empregados nas análises anteriores do IBGE sobre o tema, obtidos nas tabelas encontradas em United Nations (1983, pág. 77) no trecho da tabela referente ao chamado "Modelo de Mortalidade Oeste". Para referência, os valores dos coeficientes são reproduzidos no quadro 2, incluído no anexo.

Na expressão (1), o termo  $D_i/P_i$  representa a chamada "estimativa direta" da mortalidade usando dados de um censo. Tal "estimativa direta" é a proporção de filhos nascidos vivos que já morreram para mulheres de cada uma das sete faixas etárias quinquênicas consideradas. O motivo de usar aspas aqui ao apresentar esta definição é que na verdade estas proporções não devem ser confundidas com as que seriam obtidas mediante a divisão do número de óbitos de crianças de cada idade pelas populações que iniciaram o período de análise, como seria possível calcular partindo de informações do Registro Civil, por exemplo. Por simplicidade, no restante do trabalho não usaremos mais aspas ou qualquer destaque quando fizermos referência a tais estimativas, aqui chamadas diretas para contraste com as obtidas pelo método indireto.

A expressão (1) levaria à estimativa indireta para a  $i$ -ésima taxa de mortalidade (ver quadro 1) se as informações de fecundidade e mortalidade ( $P_i, D_i$  e  $M_i$ ) viessem de um censo. Na prática, aqui no Brasil, os dados da fecundidade e da mortalidade por grupos etários das mulheres são sempre *estimados* usando dados de pesquisas amostrais. Mesmo nos anos de realização dos Censos Demográficos, as informações sobre fecundidade e mortalidade são coletadas por amostragem.

A estimação das taxas de mortalidade  $Q_i$  é usualmente feita substituindo os valores dos totais desconhecidos  $M_i$ ,  $P_i$  e  $D_i$  por estimativas amostrais denotadas por  $\hat{M}_i$ ,  $\hat{P}_i$  e  $\hat{D}_i$ , para qualquer  $i$ , sempre supondo que são empregados estimadores não viciados dos totais desconhecidos. Logo, a expressão de fato usada para obter as estimativas de taxas

de mortalidade é:

$$\hat{Q}_i = \left( a_i + b_i \frac{\hat{P}_1}{\hat{M}_1} \frac{\hat{M}_2}{\hat{P}_2} + c_i \frac{\hat{P}_2}{\hat{M}_2} \frac{\hat{M}_3}{\hat{P}_3} \right) \frac{\hat{D}_i}{\hat{P}_i} \quad i = 1, 2, \dots, 7. \quad (2)$$

Como se conclui do exame de (2),  $\hat{Q}_i$  é um estimador para a quantidade populacional desconhecida  $Q_i$ . Portanto, para avaliar a incerteza na estimação dessa quantidade devida ao emprego de amostragem, é necessário estimar a variância (ou o desvio padrão) de  $\hat{Q}_i$ . No que segue, apresenta-se uma descrição do método usado para estimar a variância de um dos valores de  $\hat{Q}_i$ , a saber, o estimador  $\hat{Q}_2$  para  $Q_2$ , que usualmente é dos componentes de  $\mathbf{Q}$  que desperta maior interesse. A estimação de variância para os outros componentes pode ser realizada de maneira análoga, e não é aqui apresentada em todo detalhe por razões de espaço.

A idéia central do método de estimação de variância utilizado aqui é que o estimador  $\hat{Q}_2$  é uma função não linear de vários estimadores de total. Não existe, portanto, expressão exata para a variância de tal estimador como função das variâncias dos estimadores de total componentes. Daí a necessidade de utilizar métodos aproximados. O método de aproximação padrão neste caso é a chamada Linearização de Taylor (ver, por exemplo, Pessoa e Nascimento Silva, 1998, cap. 3). A idéia é simples. Basta reconhecer que  $Q_2$  (e seu estimador  $\hat{Q}_2$ ) pode ser escrito como uma função de vários totais:

$$Q_2 = g(P_1, P_2, P_3, M_1, M_2, M_3, D_2) = g(\mathbf{Y}) \quad (3)$$

onde  $g(\mathbf{Y}) = g(P_1, P_2, P_3, M_1, M_2, M_3, D_2) = \left( a_2 + b_2 \frac{P_1}{M_1} \frac{M_2}{P_2} + c_2 \frac{P_2}{M_2} \frac{M_3}{P_3} \right) \frac{D_2}{P_2}$ , e

$$\mathbf{Y} = (P_1, P_2, P_3, M_1, M_2, M_3, D_2)'$$

O estimador de  $Q_2$  é dado, então, pela função  $g$  aplicada ao vetor de totais estimados correspondentes, isto é:

$$\hat{Q}_2 = g(\hat{P}_1, \hat{P}_2, \hat{P}_3, \hat{M}_1, \hat{M}_2, \hat{M}_3, \hat{D}_2) = g(\hat{\mathbf{Y}}) \quad (4)$$

onde  $\hat{\mathbf{Y}} = (\hat{P}_1, \hat{P}_2, \hat{P}_3, \hat{M}_1, \hat{M}_2, \hat{M}_3, \hat{D}_2)'$ .

O método de Linearização de Taylor recomenda estimar a variância de  $\hat{Q}_2$  por:

$$\hat{v}(\hat{Q}_2) = \Delta g(\hat{\mathbf{Y}}) \hat{v}_p(\hat{\mathbf{Y}}) \Delta g(\hat{\mathbf{Y}})^t \quad (5)$$

onde  $\hat{v}_p(\hat{\mathbf{Y}})$  é um estimador não viciado (ou ao menos consistente) para a matriz de variância do vetor de estimadores de total  $\hat{\mathbf{Y}}$ , e  $\Delta g(\hat{\mathbf{Y}})$  é o jacobiano da função  $g$ , definido como o vetor (1×7) de derivadas parciais da função  $g(\mathbf{Y})$ , avaliado no ponto  $\hat{\mathbf{Y}}$ .

O  $j$ -ésimo componente de  $\Delta g(\hat{\mathbf{Y}})$  é definido como  $\delta_j = \Delta g(\hat{\mathbf{Y}})_j = \left. \frac{\partial g(\mathbf{Y})}{\partial Y_j} \right|_{Y_j = \hat{Y}_j}$ ,  $j=1, \dots, 7$ , onde  $Y_j$  representa o  $j$ -ésimo componente do vetor  $\mathbf{Y}$  e  $\hat{Y}_j$  representa o correspondente estimador do total componente do vetor  $\hat{\mathbf{Y}}$ .

Tais derivadas parciais são facilmente obtidas usando, por exemplo, a facilidade de computação simbólica do MAPLE ou pacote similar. Os resultados das derivadas parciais para  $Q_2$  são apresentados nas equações de (6) a (12) a seguir:

$$\frac{\partial g(P_1, P_2, P_3, M_1, M_2, M_3, D_2)}{\partial P_1} = D_2 \frac{b_2}{M_1} \frac{M_2}{P_2^2} \quad (6)$$

$$\frac{\partial g(P_1, P_2, P_3, M_1, M_2, M_3, D_2)}{\partial P_2} = -D_2 \frac{2 b_2 P_1 M_2 + a_2 M_1 P_2}{M_1 P_2^3} \quad (7)$$

$$\frac{\partial g(P_1, P_2, P_3, M_1, M_2, M_3, D_2)}{\partial P_3} = -D_2 \frac{c_2 M_3}{M_2 P_3^2} \quad (8)$$

$$\frac{\partial g(P_1, P_2, P_3, M_1, M_2, M_3, D_2)}{\partial M_1} = -b_2 D_2 \frac{P_1 M_2}{M_1^2 P_2^2} \quad (9)$$

$$\frac{\partial g(P_1, P_2, P_3, M_1, M_2, M_3, D_2)}{\partial M_2} = - \left( -b_2 P_1 M_2^2 P_3 + c_2 M_3 M_1 P_2^2 \right) \frac{D_2}{M_1 P_2^2 M_2^2 P_3} \quad (10)$$

$$\frac{\partial g(P_1, P_2, P_3, M_1, M_2, M_3, D_2)}{\partial M_3} = \frac{c_2 D_2}{M_2 P_3} \quad (11)$$

$$\frac{\partial g(P_1, P_2, P_3, M_1, M_2, M_3, D_2)}{\partial D_2} = \frac{\left( a_2 M_1 M_2 P_2 P_3 + b_2 M_2^2 P_1 P_3 + c_2 M_3 M_1 P_2^2 \right)}{M_1 P_2^2 M_2 P_3} \quad (12)$$

De posse do vetor  $\hat{\mathbf{Y}}$  das estimativas dos totais componentes do vetor  $\mathbf{Y}$ , é imediata a estimação de  $\Delta g(\hat{\mathbf{Y}})$  usando as equações (6) a (12). Para calcular a expressão (5), fica

faltando apenas a estimativa da matriz de variância  $\hat{V}_p(\hat{Y})$ . A estimação de  $\hat{V}_p(\hat{Y})$  é tarefa simples, se for empregado algum sistema próprio para estimação e tratamento de dados amostrais complexos, como é o caso do SUDAAN (Shah et al., 1997) ou similar. Neste trabalho, as estimativas de  $\hat{Y}$ ,  $\hat{V}_p(\hat{Y})$  e  $\Delta g(\hat{Y})$  foram obtidas usando uma função escrita na linguagem R, que pode ser obtida mediante solicitação aos autores.

### 3. Resultados

Nesta seção, descrevemos o processo de tratamento dos dados da PNAD-1999 requeridos para obtenção das estimativas de taxas de mortalidade.

Em primeiro lugar, consideramos importante descrever como os dados básicos foram obtidos na pesquisa. O questionário usado na coleta de dados das PNADs durante a década de 90 (1992 a 1999) incorporou duas formas alternativas de obtenção das informações de fecundidade e mortalidade, ambas através de perguntas incluídas no bloco 11 do questionário (ver IBGE, 1999). As perguntas de fecundidade e mortalidade são feitas somente para mulheres moradoras de 15 anos ou mais de idade na data da pesquisa. A primeira pergunta do bloco 11 é "*Até 25 de setembro de 1999, teve algum filho nascido vivo (ou seja, que apresentou algum sinal de vida ao nascer)?*". Mulheres respondendo "*Sim*" a esta pergunta respondiam às perguntas numeradas de 2 a 9 dentro do bloco 11. Mulheres respondendo "*Não*" a essa pergunta passavam às perguntas 10 e 11 do bloco 11, as últimas do questionário, que se referiam à existência e ao número de filhos de sete meses ou mais nascidos mortos.

A primeira forma de obtenção dos dados de fecundidade e mortalidade corresponde à investigação do número total de *filhos tidos* nascidos vivos por sexo (pergunta 2 do bloco 11) e do número total de filhos tidos nascidos vivos que ainda estão vivos por sexo (pergunta 3 do bloco 11), isto é, à pergunta direta do número de *filhos sobreviventes* por sexo. Nesta forma de investigação, a informação do número de filhos tidos nascidos vivos que já morreram (*filhos mortos*) é obtida indiretamente, calculando a diferença entre o número de filhos nascidos vivos e o número de filhos sobreviventes para cada sexo. Os valores totais por mulher são sempre obtidos por soma das informações pesquisadas separadamente para cada sexo. Esta forma de investigação foi incluída no questionário da PNAD da década de 90 para compatibilidade com a forma de investigação adotada no Censo Demográfico. Doravante, por simplicidade, chamaremos

esta forma de Modelo Censo de investigação da fecundidade e mortalidade.

A segunda forma de obtenção dos dados de fecundidade e mortalidade está baseada nas perguntas 4 a 6 do bloco 11 do questionário. Nesta forma, são feitas perguntas para cada sexo sobre o número total de filhos tidos nascidos vivos que moram no domicílio (pergunta 4), o número de filhos tidos nascidos vivos que moram em outro domicílio (pergunta 5) e o número de filhos tidos nascidos vivos que já morreram (pergunta 6). Nesta seqüência de perguntas, a obtenção da informação sobre o número de filhos mortos é direta. Novamente, os valores totais por mulher são sempre obtidos por soma das informações pesquisadas separadamente para cada sexo. Doravante, por simplicidade, chamaremos esta forma de Modelo PNAD de investigação da fecundidade e mortalidade.

Na PNAD, esta segunda forma de obtenção das informações é considerada a melhor (ver Pereira e Maluf, 1997), levando à decisão de basear as estimativas de fecundidade e mortalidade nesta seqüência de perguntas e não na primeira. Como consequência, os dados das perguntas 2 e 3 do bloco 11 não são criticados ou depurados da mesma forma que os das perguntas 4 a 6. Também não são tabulados, e sequer são incluídos nos arquivos de microdados da PNAD disseminados pelo IBGE. Um estudo com análise comparativa entre os resultados destas duas formas de investigação ao longo das PNADs de 1992, 1993 e 1995 (ver Pereira e Maluf, 1997) revelou que as estimativas por grandes regiões não apresentavam grande variação entre as duas formas de investigação, mas não investigou a variação das estimativas por estados.

Feito este esclarecimento quanto à forma de investigar a fecundidade e a mortalidade na PNAD, é importante reconhecer então dois problemas distintos relacionados com a estimação de taxas de mortalidade usando essa fonte de dados, quando se pensa em posteriormente comparar os resultados com os obtidos através do Censo Demográfico.

- Primeiro, as formas de investigar / perguntar os dados de fecundidade e mortalidade usadas na PNAD-1999 e no Censo Demográfico 2000 são distintas, podendo provocar eventualmente diferenças nos resultados devidas à mudança da pergunta. Para lidar com essa questão, decidimos fazer uso dos dados das duas formas de investigar a fecundidade e a mortalidade disponíveis na PNAD-1999 para estimar as taxas de mortalidade. A comparação destes dados internamente à mesma pesquisa permitiu avaliar o impacto da variação da forma de perguntar

sobre as estimativas, e ajudará a interpretar eventuais diferenças observadas das estimativas da PNAD-1999 para as do Censo Demográfico de 2000.

- Segundo, os dados obtidos na PNAD-1999 são baseados em amostra bem menor que a do Censo Demográfico, e portanto, sujeitos a maior incerteza na estimação dos parâmetros populacionais de interesse. A estimação dessa incerteza é justamente a maior contribuição deste relatório.

Foram selecionadas para inclusão nos cálculos somente as mulheres que tiveram código igual a 1 na variável V4735 (variável de controle da tabulação de fecundidade) e que tinham idade na faixa de 15 a 49 anos completos na data da pesquisa. O total de registros de mulheres selecionadas por este critério chegou a 97.337 mulheres. Destas foram excluídas 223 mulheres, por pertencerem a estratos de amostragem onde uma única unidade primária de amostragem da PNAD 1999 foi selecionada, o que dificultaria a estimação das medidas de precisão desejadas. Vale notar que todas as mulheres excluídas pertenciam à parte da amostra do cadastro de novas construções da PNAD. Após a exclusão destas 223 mulheres, sobraram 97.114 mulheres que tinham valores presentes em todas as variáveis relevantes (número total de filhos tidos nascidos vivos e número total de filhos nascidos vivos que já morreram).

As variáveis de parturição total e mortalidade usadas na estimação das taxas de mortalidade por faixa etária da mãe foram construídas para cada mulher na amostra considerada como indicado nas expressões a seguir:

**Modelo PNAD de investigação**

$$P=(V1141+V1151+V1161)+(V1142+V1152+V1162)$$

$$D=V1161+V1162$$

**Modelo Censo de Investigação**

$$P=V1121+V1122$$

$$D=V1121+V1122-(V1131+V1132)$$

onde  $V1121$ ,  $V1122$ ,  $V1131$ ,  $V1132$ ,  $V1141$ ,  $V1142$ ,  $V1151$ ,  $V1152$ ,  $V1161$  e  $V1162$  são as variáveis contidas no arquivo de dados da PNAD-1999, conforme o dicionário de dados, correspondentes às perguntas 2, 3, 4, 5 e 6 do bloco 11 do questionário. Quando o último dígito do número da variável é igual a 1, as informações são referentes a filhos do sexo masculino, e quando é igual a 2, são referentes a filhos do sexo feminino. Além destas variáveis, a estimação das taxas de mortalidade depende ainda da contagem do

número de mulheres na faixa etária de interesse ( $i=1,2,\dots,7$ ), denotada como  $M_i$ .

Os resultados da estimação das taxas de mortalidade (por mil habitantes) e correspondentes desvios-padrão estão apresentados em quatro tabelas no anexo. Nas Tabelas 1 e 2 do anexo são apresentados os resultados para as estimativas diretas  $D_i/P_i$  (tabela 1) e indiretas  $Q_i$  (tabela 2) das taxas de mortalidade para as unidades da federação e o total nacional, junto com seus respectivos desvios padrão, obtidos com base nos dados das perguntas 4 a 6 do bloco 11 da PNAD-1999 (Modelo PNAD de investigação). Nas Tabelas 3 e 4 em anexo são apresentados os resultados para as estimativas diretas (tabela 3) e indiretas (tabela 4) das taxas de mortalidade para as unidades da federação, junto com seus respectivos desvios padrão, obtidos com base nos dados das perguntas 2 e 3 do bloco 11 da PNAD-1999 (Modelo Censo de investigação).

Começamos fazendo uma análise das estimativas indiretas das taxas de mortalidade obtidas usando o modelo PNAD. Concentramos nossa atenção na análise das estimativas para  $Q_2$ , pois é nestas que se baseia a estimação das taxas de mortalidade infantil e, posteriormente, das tábuas de mortalidade. Um exame dos valores de  $\hat{Q}_2$  na tabela 2 mostra que as taxas estimadas por estados variam de cerca de 14 por mil até cerca de 80 por mil, e os respectivos desvios-padrão variam de 4,7 por mil a 63,8 por mil, revelando grande dispersão entre os estados. Os quadros 3 e 4 no anexo apresentam os ramo-e-folhas para as estimativas de  $Q_2$  e de seu desvio padrão. No caso do estado onde é maior o desvio padrão da estimativa (Roraima), um intervalo de confiança de 95% para a taxa  $Q_2$  vai de 0 a 197 por mil! No segundo pior caso (Amapá), o intervalo de confiança de 95% para a taxa  $Q_2$  vai de 0 a 128 por mil.

Os coeficientes de variação das estimativas da taxa  $Q_2$  por estado variam entre 11% e 92%, com mediana igual a 31,6%. Assim, uma primeira conclusão a que podemos chegar com essa análise é que as estimativas das taxas de mortalidade por estados baseadas na PNAD são imprecisas, e para alguns estados, essa imprecisão é muito grande. A precisão das estimativas das taxas de mortalidade é melhor para faixas etárias mais altas, como se pode ver examinando a figura 5 do Anexo, com a faixa etária mais imprecisa sendo a das mulheres mais jovens (15 a 19 anos). Isso está de acordo com o esperado, pois as mulheres mais jovens apresentam menor incidência de eventos de fecundidade e mortalidade dos filhos tidos, tornando a estimação das taxas de

mortalidade infantil mais difícil devido tratar-se de proporções de eventos mais raros.

Uma comparação das estimativas diretas ( $D_2/P_2$ ) apresentadas na Tabela 1 com as estimativas indiretas de  $Q_2$  na Tabela 2 revela que o emprego do método indireto afeta pouco tanto as estimativas pontuais como a precisão das estimativas da taxa de mortalidade. A figura 1 do anexo mostra com clareza que os ajustes do método indireto praticamente não afetam os valores das estimativas pontuais da taxa de mortalidade, com os pontos nesse gráfico ficando todos muito próximos da linha diagonal que indicaria igualdade das estimativas diretas e indiretas. O mesmo ocorre quando comparamos as estimativas diretas e indiretas obtidas com o modelo Censo de investigação, com o gráfico tendo o mesmo comportamento que o da figura 1 mas não mostrado aqui por economia de espaço. Portanto, a introdução dos ajustes indiretos não tem grande impacto sobre as estimativas das taxas de mortalidade obtidas através da PNAD 1999.

Além disso, os desvios padrão das estimativas indiretas se aproximam bastante dos desvios padrão das estimativas diretas da proporção de filhos mortos de mães da faixa etária de 20 a 24 anos. Isto significa que uma avaliação inicial da (im)precisão das estimativas das taxas de mortalidade pode ser feita de maneira bastante simples, mediante o cálculo dos desvios padrão das proporções de filhos mortos por faixa etária das mães, que pode ser feito dentro do processo rotineiro de tabulação da própria PNAD. Todavia, o cálculo aqui realizado das estimativas dos desvios padrão das taxas de mortalidade estimadas pelo método indireto sugere que estes são relacionados com os desvios padrão das estimativas diretas por um modelo de regressão linear com intercepto igual a -0,8 e coeficiente angular igual a 1,056. Sendo assim, os valores dos desvios padrão serão subestimados em algo como 5% se não for utilizado o método aqui desenvolvido para estimar os desvios padrão das estimativas indiretas.

Uma outra análise de interesse foi feita comparando as estimativas diretas (figura 3) e indiretas (figura 4) das taxas de mortalidade quando obtidas com os dados do modelo PNAD de investigação com os dados do modelo Censo de investigação da fecundidade e mortalidade. Nessa comparação emerge um padrão diferente. As estimativas de cada tipo (diretas ou indiretas) obtidas com os dois conjuntos de dados (modelo PNAD e modelo Censo) variam bastante. Os pontos nos dois gráficos se afastam um pouco da reta diagonal que representaria a igualdade entre as estimativas obtidas com os dados coletados das duas formas. Nas duas figuras, destaca-se o ponto referente ao estado de

Roraima, cujas estimativas usando os dados obtidos segundo o modelo PNAD são superiores a 60 por mil, contra 0 se forem usados os dados obtidos segundo o modelo Censo de investigação. Essa discrepância resulta do fato de que ambas as estimativas são baseadas em amostra muito pequena, para a qual as informações básicas diferem substancialmente de uma forma de investigar para outra.

Para dar uma medida desse afastamento, ajustamos modelos de regressão linear sem intercepto entre valores das estimativas obtidas com os dados das duas fontes e estimamos a parcela da variação explicada por tais modelos (resultados dos ajustes para as estimativas diretas na Tabela 5 do anexo, e para os ajustes das estimativas indiretas na Tabela 6 do anexo, respectivamente). Como se pode ver do exame da tabela 5, a parcela da variação não explicada pelos modelos de regressão é maior justamente para as estimativas diretas de  $D_2/P_2$ , caso em que o modelo Censo de investigação leva a estimativas em média 8,5% menores que as obtidas pelo modelo PNAD. Embora modesta, a variação não explicada pelo modelo não é desprezível para as mulheres das duas faixas de idade mais jovens, e decresce substancialmente para as faixas de idade mais altas. Análise similar resulta para as estimativas indiretas obtidas segundo as duas formas de investigação.

Vale notar ainda que as duas formas de investigação da fecundidade e mortalidade levam a estimativas ligeiramente diferentes, mesmo a nível nacional (usando toda a amostra da PNAD). As estimativas para  $Q_2$  são de 40,1 por mil segundo o modelo PNAD e de 38,4 por mil segundo o modelo Censo de investigação.

#### **4. Conclusões**

Uma conclusão deste trabalho que é importante ressaltar é a revelação da grande incerteza das estimativas amostrais das taxas de mortalidade (diretas ou indiretas) obtidas com base numa fonte amostral como a PNAD, para algumas unidades da federação. Nos casos dos estados com as maiores amostras, os desvios padrão estão em torno de 5 por mil, e os respectivos coeficientes de variação são da ordem de 20% a 30%. Nos estados com as menores amostras, os desvios padrão são da ordem de 20 a 30 por mil, podendo chegar a 60 por mil (Roraima), com coeficientes de variação que atingem níveis muito elevados (de 50% a 90%).

Isto sugere que o cálculo de estimativas das taxas de mortalidade por estados com base em informações de amostras como a da PNAD é possível, mas não recomendável,

ao menos no caso dos estados com amostras pequenas. Uma estratégia mais adequada seria o cálculo e divulgação dessas estimativas para grupos de estados para os quais a amostra permita estimar com maior confiabilidade. Caso venha a prevalecer a prática atual, as estimativas por estados devem ser acompanhadas de alerta claro sobre a imprecisão das mesmas no caso de vários estados.

Recomendamos que as estimativas diretas e indiretas de taxas de mortalidade baseadas na PNAD sejam sempre acompanhadas de estimativas dos respectivos desvios padrão. A publicação das estimativas dos desvios padrão junto das respectivas estimativas das taxas de mortalidade vai para permitir aos analistas e usuários interessados incorporar em suas análises a noção da incerteza amostral envolvida na obtenção de tais indicadores, o que deve contribuir para qualificar melhor a análise de tais estimativas. Os métodos aqui desenvolvidos podem ser empregados para permitir calcular estimativas dos desvios padrão das estimativas indiretas das taxas de mortalidade.

Por último, mas não menos importante, a prática de calcular estimativas indiretas das taxas de mortalidade parece não ter grande efeito sobre as estimativas pontuais. Por outro lado, ficou evidenciado que o emprego de diferentes forma de investigação das variáveis de fecundidade e mortalidade tem efeito importante. Recomenda-se, portanto, aprofundar as análises das formas de investigação para deliberar qual das duas formas é menos sujeita a erros de medida e leva a resultados de melhor qualidade, a fim de passar a usar esta forma como padrão nas distintas pesquisas. Esta questão tem maior relevância para um projeto de pesquisa e melhoria que o eventual aprimoramento dos métodos indiretos de ajuste das estimativas de mortalidade, ao menos se a fonte de dados a ser utilizada for a PNAD.

Outra linha de investigação que deve receber maior prioridade é a que trata do desenvolvimento e aplicação de métodos capazes de estimar as taxas de mortalidade por estados com maior precisão, seja mediante a acumulação de dados de duas ou mais PNADs consecutivas, seja mediante a combinação de informações da PNAD com dados de registros administrativos ou outras fontes, seja através de técnicas de suavização de dados e estimação de tendência que aproveitem a estrutura longitudinal dada pela repetição anual da PNAD, seja por combinações destes métodos entre si ou com outros que sejam aplicáveis ao problema.

## 5. Referências

- IBGE (1999). Pesquisa Nacional por Amostra de Domicílios - PNAD 1.01 - Questionário da pesquisa.
- IBGE (2002). Censo Demográfico 2000: Fecundidade e Mortalidade Infantil, resultados preliminares da amostra.
- Pereira, N.O.M. e Maluf, K.C.F. (1997). Fecundidade e Mortalidade: Uma Avaliação das Informações Básicas nas Pesquisas Demográficas. Rio de Janeiro: IBGE, Departamento de População e Indicadores Sociais, mimeo.
- Pessoa, D.G.C. e Nascimento Silva, P.L.d. (1998). Análise de Dados Amostrais Complexos. São Paulo: Associação Brasileira de Estatística.
- R Development Core Team (2000). Introdução a R.
- Shah, B.V., Barnwell, B.G. e Bieler, G.S. (1997). SUDAAN: Software for the Statistical Analysis of Correlated Data - User's Manual, vol. 1 e 2. Research Triangle Institute.
- United Nations (1983). Manual X: Indirect Techniques for Demographic Estimation. Nova Iorque: Nações Unidas, Department of International Economic and Social Affairs, Population Studies 81.

## 6. Anexos

Quadro 2 - Coeficientes modelo Oeste para estimação indireta das taxas de mortalidade

Idades	Índice (i)	$a_i$	$b_i$	$c_i$
15-19	1	1,1415	-2,7070	0,7663
20-24	2	1,2563	-0,5381	-0,2637
25-29	3	1,1851	0,0633	-0,4177
30-34	4	1,1720	0,2341	-0,4272
35-39	5	1,1865	0,3080	-0,4452
40-44	6	1,1746	0,3314	-0,4537
45-49	7	1,1639	0,3190	-0,4435

Quadro 3 - Ramo-e-folhas das estimativas indiretas para  $Q_2$  usando modelo PNAD de investigação

Ramo	Folhas
1	578
2	36666
3	011356
4	38
5	05588
6	2679
7	2
8	0

Ramo=Dezena, Folha=Unidade, Ex.: 7|2 = 72 mortes por mil

Quadro 4 - Ramo-e-folhas dos desvios padrão das estimativas indiretas para  $Q_2$  usando modelo PNAD de investigação

Ramo	Folhas
0	55556788899
1	0122556889
2	0058
3	5
4	
5	
6	4

Ramo=Dezena, Folha=Unidade, Ex.: 3|5 = 35 mortes por mil

Tabela 1 – Estimativas diretas da mortalidade e seus desvios padrão – modelo PNAD de investigação

UF		Estimativas							Desvios padrão						
		D1/P1	D2/P2	D3/P3	D4/P4	D5/P5	D6/P6	D7/P7	D1/P1	D2/P2	D3/P3	D4/P4	D5/P5	D6/P6	D7/P7
11	Rondônia.....	0,0	24,8	14,8	55,3	17,6	62,8	79,0	0,0	18,6	7,4	17,2	12,6	25,3	23,7
12	Acre.....	0,0	50,7	129,7	63,2	30,3	32,5	203,2	0,0	25,5	47,5	28,2	17,8	20,8	55,1
13	Amazonas.....	30,3	54,9	58,7	29,7	62,2	61,8	75,1	20,9	27,9	13,2	7,5	18,5	9,4	12,3
14	Roraima.....	0,0	64,4	37,1	80,1	0,0	86,5	51,7	0,0	59,5	35,0	41,7	0,0	20,3	35,1
15	Pará.....	10,8	47,0	68,9	60,7	83,4	103,9	124,6	6,3	15,0	12,6	10,6	7,0	13,8	16,6
16	Amapá.....	0,0	57,1	97,0	29,1	56,9	31,2	64,5	0,0	33,7	20,6	14,0	30,8	15,7	28,9
17	Tocantins.....	0,0	43,6	56,8	60,2	66,9	91,2	128,6	0,0	18,8	21,5	14,9	23,1	20,7	15,5
21	Maranhão.....	52,1	80,2	71,3	98,8	133,1	167,9	193,2	22,5	11,8	11,7	13,7	21,4	15,4	16,0
22	Piauí.....	46,2	68,1	53,2	92,6	81,9	138,0	124,3	33,6	12,7	12,0	24,5	10,0	16,6	10,2
23	Ceará.....	69,9	61,6	71,0	80,5	103,4	133,4	166,2	19,0	10,9	8,0	8,7	5,8	8,9	18,1
24	Rio Grande do Norte.....	42,2	56,1	64,5	77,8	86,3	106,3	123,7	27,0	18,5	16,2	16,5	15,3	23,6	22,7
25	Paraíba.....	33,3	30,3	78,7	74,9	88,1	140,5	161,6	33,7	16,2	21,2	12,6	14,8	20,0	18,6
26	Pernambuco.....	56,8	71,5	71,4	100,6	104,9	115,8	156,7	21,9	15,2	10,0	9,3	7,6	8,6	12,4
27	Alagoas.....	14,5	58,8	100,4	139,8	142,2	151,1	183,9	14,5	20,1	25,5	22,8	18,0	19,1	23,5
28	Sergipe.....	21,3	30,5	51,9	64,3	111,8	122,5	146,5	20,1	8,7	13,2	14,4	25,9	23,1	23,3
29	Bahia.....	39,4	65,5	70,6	93,2	90,4	117,4	126,3	11,1	7,6	15,1	9,5	9,7	10,4	10,4
31	Minas Gerais.....	29,0	30,0	31,1	46,1	41,0	70,8	85,1	11,2	7,5	5,3	7,3	4,7	6,8	7,3
32	Espírito Santo.....	33,9	37,4	56,5	44,4	54,9	49,2	91,1	23,5	20,6	17,9	10,7	15,3	10,1	18,9
33	Rio de Janeiro.....	13,6	35,0	35,7	32,3	40,2	55,2	62,5	9,5	7,5	6,5	5,7	6,1	7,0	7,6
35	São Paulo.....	25,7	22,7	45,8	34,4	43,9	57,3	76,7	8,7	4,7	10,8	5,4	4,8	5,2	5,6
41	Paraná.....	35,6	18,2	39,0	38,9	46,9	53,6	75,1	15,5	5,0	7,9	6,3	8,1	5,5	9,7
42	Santa Catarina.....	12,7	26,1	42,1	68,3	26,1	45,6	60,8	11,9	8,6	12,0	20,2	5,0	6,5	10,7
43	Rio Grande do Sul.....	24,8	17,3	24,7	34,9	40,6	43,3	45,6	11,0	5,3	5,4	5,9	7,2	6,0	6,1
50	Mato Grosso do Sul.....	27,0	26,4	46,2	21,4	32,8	59,5	56,2	16,9	5,7	8,8	6,1	5,8	15,7	8,4
51	Mato Grosso.....	12,7	33,2	24,1	25,2	61,0	61,4	84,2	12,5	10,0	6,2	7,1	8,4	9,9	9,9
52	Goiás.....	29,2	14,9	33,5	42,0	39,3	69,1	113,1	17,1	4,8	5,5	5,8	5,8	7,6	16,9
53	Distrito Federal.....	0,0	27,3	37,8	35,1	26,6	43,5	51,9	0,0	9,4	11,4	7,2	5,6	9,8	11,9
br	<b>Total nacional</b>	31,2	40,3	50,9	56,6	62,6	81,8	101,4	3,5	2,3	3,0	2,2	2,1	2,3	2,7

Tabela 2 – Estimativas indiretas da mortalidade e seus desvios padrão – modelo PNAD de investigação

UF	Estimativas							Desvios padrão							
	q(1)	q(2)	q(3)	q(5)	q(10)	q(15)	q(20)	q(1)	q(2)	q(3)	q(5)	q(10)	q(15)	q(20)	
11	Rondônia.....	0,0	25,6	14,7	55,5	18,0	63,5	79,2	0,0	19,0	7,4	18,2	12,9	24,9	23,7
12	Acre.....	0,0	49,8	123,5	61,3	30,0	31,8	197,5	0,0	24,8	52,9	25,0	17,2	22,1	52,0
13	Amazonas.....	26,9	55,2	58,5	30,3	64,8	63,8	76,8	18,5	27,8	13,3	7,8	19,5	9,6	12,8
14	Roraima.....	0,0	69,4	38,4	83,4	0,0	90,7	53,8	0,0	63,8	35,7	44,0	0,0	20,6	36,5
15	Pará.....	10,4	47,6	68,0	61,0	85,6	105,5	125,4	6,3	15,1	12,7	10,5	7,4	14,2	16,6
16	Amapá.....	0,0	58,3	94,4	28,7	57,0	30,9	63,5	0,0	34,9	20,4	13,8	31,3	16,1	28,7
17	Tocantins.....	0,0	42,9	53,5	57,6	65,3	87,9	123,0	0,0	18,3	20,0	14,2	22,0	19,4	15,5
21	Maranhão.....	52,9	79,7	68,3	96,4	132,5	165,1	188,5	22,0	12,0	11,6	14,4	19,8	12,6	16,9
22	Piauí.....	47,3	67,2	50,7	89,8	81,0	134,8	120,6	36,1	12,4	11,1	24,4	10,3	18,4	10,4
23	Ceará.....	78,7	61,7	67,1	76,9	100,5	128,0	158,4	20,8	10,8	7,6	8,0	5,7	8,3	17,6
24	Rio Grande do Norte.....	39,2	55,4	62,8	77,4	87,7	106,9	123,4	25,4	18,0	16,4	16,6	14,7	23,1	22,6
25	Paraíba.....	31,9	30,6	77,8	75,4	90,6	142,8	163,0	33,7	16,2	21,7	13,4	16,3	20,7	18,7
26	Pernambuco.....	58,9	72,2	69,3	99,1	105,4	114,9	154,3	22,2	15,1	9,6	8,9	7,4	9,3	12,9
27	Alagoas.....	11,5	57,9	100,4	143,6	149,8	157,6	190,2	11,8	19,6	26,1	24,5	18,9	19,9	27,2
28	Sergipe.....	20,7	31,2	51,7	65,1	115,7	125,3	148,6	18,2	8,7	13,4	14,3	26,5	23,3	24,1
29	Bahia.....	37,7	65,7	69,3	93,3	92,3	118,6	126,5	11,5	7,6	15,1	9,5	10,1	10,8	10,6
31	Minas Gerais.....	26,2	30,3	31,1	47,0	42,7	72,9	86,9	9,8	7,5	5,3	7,4	4,9	7,0	7,4
32	Espírito Santo.....	26,4	36,2	55,7	45,1	57,2	50,7	93,3	18,5	19,6	18,0	10,8	16,8	10,1	19,0
33	Rio de Janeiro.....	12,6	34,8	35,0	32,3	41,1	55,9	62,6	8,8	7,4	6,5	5,8	6,3	7,1	7,7
35	São Paulo.....	23,3	22,6	45,2	34,7	45,2	58,4	77,5	7,8	4,7	10,8	5,4	4,9	5,3	5,8
41	Paraná.....	28,3	17,8	38,8	39,7	49,1	55,6	77,2	11,8	4,9	7,9	6,3	8,5	5,7	10,3
42	Santa Catarina.....	9,9	25,9	42,5	70,8	27,7	48,0	63,5	9,4	8,4	12,3	21,1	5,3	6,8	12,0
43	Rio Grande do Sul.....	22,5	17,0	24,1	34,7	41,3	43,7	45,6	9,8	5,2	5,2	5,9	7,5	6,4	6,3
50	Mato Grosso do Sul.....	24,3	25,5	44,2	21,0	33,0	59,2	55,4	14,6	5,6	8,4	6,0	6,0	15,2	8,6
51	Mato Grosso.....	12,5	32,7	23,0	24,5	60,7	60,4	82,2	12,4	10,0	6,0	7,0	8,2	10,3	10,2
52	Goiás.....	31,1	14,7	31,7	40,4	38,6	67,0	108,9	17,7	4,8	5,3	5,7	6,0	7,3	16,8
53	Distrito Federal.....	0,0	26,5	35,8	34,0	26,3	42,4	50,3	0,0	9,1	11,0	7,0	5,6	9,4	11,5
br	<b>Total nacional</b>	29,3	40,1	49,8	56,5	63,9	82,5	101,5	3,3	2,3	3,0	2,3	2,2	2,3	2,8

Tabela 3 – Estimativas diretas da mortalidade e seus desvios padrão – modelo Censo de investigação

UF	Estimativas							Desvios padrão						
	D1/P1	D2/P2	D3/P3	D4/P4	D5/P5	D6/P6	D7/P7	D1/P1	D2/P2	D3/P3	D4/P4	D5/P5	D6/P6	D7/P7
11 Rondônia.....	0,0	17,0	22,2	85,7	22,3	67,0	81,1	0,0	12,2	10,0	27,5	10,4	24,7	24,4
12 Acre.....	0,0	70,5	129,7	63,2	30,3	31,7	203,2	0,0	27,5	47,5	28,2	17,8	19,7	55,1
13 Amazonas.....	15,4	36,2	43,2	48,9	64,1	65,6	76,0	15,1	10,6	13,5	12,3	18,2	10,5	12,3
14 Roraima.....	0,0	0,0	36,7	80,1	0,0	108,5	51,7	0,0	0,0	34,6	41,7	0,0	25,1	35,1
15 Pará.....	20,3	51,8	72,7	63,6	95,5	106,3	127,9	11,2	16,3	12,6	10,5	13,0	14,0	16,7
16 Amapá.....	0,0	57,1	97,0	29,1	56,9	31,2	64,5	0,0	33,7	20,6	14,0	30,8	15,7	28,9
17 Tocantins.....	0,0	34,2	42,7	54,1	66,9	80,3	134,2	0,0	16,7	20,0	14,0	23,1	18,0	15,7
21 Maranhão.....	73,7	86,2	82,7	93,1	128,0	164,4	184,2	11,6	14,0	14,2	18,0	21,7	15,0	19,6
22 Piauí.....	46,2	72,3	53,2	88,2	86,2	138,7	125,2	33,6	13,5	10,6	24,8	10,8	16,6	10,0
23 Ceará.....	81,9	60,3	73,2	82,0	101,7	135,7	169,5	20,2	11,2	9,0	8,2	6,0	9,2	17,8
24 Rio Grande do Norte.....	42,2	57,1	56,7	81,0	89,5	104,7	121,3	27,0	18,8	15,0	17,1	16,9	25,8	21,3
25 Paraíba.....	33,3	30,3	81,0	74,9	88,6	145,2	176,0	33,7	16,1	20,8	12,6	13,1	19,5	17,9
26 Pernambuco.....	56,8	58,5	68,1	92,4	114,4	116,2	159,0	22,0	8,2	7,4	8,1	9,2	8,1	11,8
27 Alagoas.....	14,5	58,8	110,5	137,9	147,4	153,7	187,3	14,5	20,1	26,7	22,3	17,2	19,4	22,9
28 Sergipe.....	21,3	36,4	56,9	74,1	130,9	119,3	145,5	20,1	16,1	13,7	10,8	35,1	26,6	21,9
29 Bahia.....	36,5	63,2	59,1	94,1	88,1	115,7	127,4	11,0	6,6	7,3	9,5	10,2	10,6	10,7
31 Minas Gerais.....	25,2	25,1	30,8	37,5	40,4	70,9	84,8	10,6	5,6	5,3	5,3	4,3	7,5	7,1
32 Espírito Santo.....	65,6	38,0	42,5	53,6	57,1	49,2	92,9	46,5	20,9	12,7	10,3	15,0	10,1	20,7
33 Rio de Janeiro.....	20,2	35,5	37,1	42,1	44,2	54,9	72,0	11,3	7,9	11,0	9,2	9,5	9,0	8,3
35 São Paulo.....	22,9	20,3	32,8	31,0	44,0	54,6	77,8	8,2	4,4	4,2	3,9	4,4	4,6	5,6
41 Paraná.....	26,4	19,4	33,3	40,0	50,1	51,2	77,1	14,6	5,3	6,7	6,5	7,5	5,6	9,7
42 Santa Catarina.....	12,8	30,4	31,5	65,4	38,9	42,1	63,4	12,0	9,7	9,2	18,9	13,1	5,8	11,0
43 Rio Grande do Sul.....	24,8	18,7	27,5	39,6	38,2	48,6	45,9	11,0	5,4	5,8	7,6	5,0	7,7	6,1
50 Mato Grosso do Sul.....	40,0	30,2	46,4	21,6	33,2	44,2	59,4	20,0	6,6	8,8	6,0	5,9	15,0	12,6
51 Mato Grosso.....	25,7	22,4	26,3	25,4	61,6	51,6	84,2	17,4	7,5	6,6	5,5	8,5	9,8	10,2
52 Goiás.....	37,0	16,6	31,7	46,0	36,9	69,4	101,5	18,7	5,0	5,1	6,9	5,7	7,4	12,9
53 Distrito Federal.....	0,0	24,9	31,1	30,9	27,4	42,3	48,1	0,0	9,2	9,3	6,5	5,8	8,7	9,8
<b>br Total nacional</b>	<b>33,0</b>	<b>38,6</b>	<b>47,0</b>	<b>56,2</b>	<b>64,2</b>	<b>81,0</b>	<b>102,7</b>	<b>3,5</b>	<b>2,1</b>	<b>2,1</b>	<b>2,2</b>	<b>2,2</b>	<b>2,3</b>	<b>2,8</b>

Tabela 4 – Estimativas indiretas da mortalidade e seus desvios padrão – modelo Censo de investigação

UF	Estimativas							Desvios padrão						
	q(1)	q(2)	q(3)	q(5)	q(10)	q(15)	q(20)	q(1)	q(2)	q(3)	q(5)	q(10)	q(15)	q(20)
11 Rondônia.....	0,0	17,2	21,6	84,4	22,4	66,4	79,7	0,0	12,4	9,6	27,5	10,4	23,6	24,0
12 Acre.....	0,0	70,0	126,9	63,1	30,9	32,0	203,6	0,0	27,7	49,9	27,8	17,8	20,1	54,4
13 Amazonas.....	13,6	36,3	43,1	49,8	66,7	67,6	77,7	13,1	10,3	13,3	12,6	19,0	10,5	12,8
14 Roraima.....	0,0	0,0	38,4	84,6	0,0	115,6	54,6	0,0	0,0	35,7	44,8	0,0	27,6	37,1
15 Pará.....	19,8	52,2	71,3	63,5	97,3	107,2	127,9	10,6	16,2	12,9	10,5	13,2	14,3	16,6
16 Amapá.....	0,0	58,3	94,4	28,7	57,0	30,9	63,5	0,0	34,9	20,4	13,8	31,3	16,1	28,7
17 Tocantins.....	0,0	33,5	40,1	51,7	65,2	77,3	128,3	0,0	16,3	18,6	13,3	22,0	16,8	15,9
21 Maranhão.....	76,2	85,4	78,8	90,2	126,5	160,5	178,5	15,3	14,2	14,5	18,2	20,0	12,7	18,6
22 Piauí.....	47,3	71,4	50,7	85,5	85,3	135,5	121,4	36,0	13,2	9,9	24,7	10,6	18,5	10,3
23 Ceará.....	92,3	60,4	69,0	78,3	98,8	130,2	161,4	22,1	11,1	8,5	7,6	6,1	8,6	17,2
24 Rio Grande do Norte.....	38,6	56,3	55,2	80,7	91,2	105,5	121,2	24,9	18,2	15,2	17,0	16,0	25,4	21,0
25 Paraíba.....	31,8	30,6	80,1	75,4	91,1	147,7	177,5	33,7	16,1	21,2	13,4	14,3	20,3	18,1
26 Pernambuco.....	57,7	59,1	66,4	91,5	115,5	116,0	157,5	21,8	8,3	7,4	7,9	9,2	8,9	12,1
27 Alagoas.....	11,5	57,9	110,4	141,5	155,1	160,1	193,4	11,9	19,6	27,3	23,9	18,3	20,1	26,6
28 Sergipe.....	20,3	37,3	57,1	75,6	136,3	122,8	148,6	17,7	16,2	14,0	10,9	35,7	27,2	22,6
29 Bahia.....	35,2	63,1	57,7	93,6	89,5	116,2	127,0	11,2	6,6	7,4	9,5	10,4	10,9	11,1
31 Minas Gerais.....	22,8	25,3	30,7	38,2	42,0	73,1	86,6	9,4	5,6	5,3	5,3	4,4	7,9	7,2
32 Espírito Santo.....	48,6	36,5	41,9	54,6	59,7	50,9	95,4	34,8	19,7	12,7	10,2	16,3	10,2	20,9
33 Rio de Janeiro.....	18,5	35,2	36,3	42,1	45,2	55,5	72,3	10,4	7,8	10,9	9,2	9,7	9,1	8,4
35 São Paulo.....	20,9	20,2	32,3	31,2	45,3	55,6	78,5	7,5	4,4	4,1	3,9	4,5	4,6	5,7
41 Paraná.....	21,3	19,0	33,0	40,8	52,4	53,0	79,1	11,3	5,2	6,7	6,4	7,8	5,7	10,3
42 Santa Catarina.....	10,2	30,2	31,6	67,5	41,0	44,1	65,7	9,7	9,5	9,2	19,7	13,9	5,9	12,1
43 Rio Grande do Sul.....	22,4	18,5	26,8	39,5	39,0	49,1	45,9	9,7	5,3	5,7	7,7	5,3	8,1	6,3
50 Mato Grosso do Sul.....	35,7	29,1	44,4	21,2	33,3	43,9	58,5	17,6	6,4	8,5	5,9	6,1	14,6	12,8
51 Mato Grosso.....	25,3	22,1	25,2	24,8	61,5	50,9	82,5	17,1	7,4	6,3	5,4	8,4	9,9	10,4
52 Goiás.....	39,3	16,5	30,1	44,4	36,3	67,4	97,9	19,5	5,0	5,0	6,8	5,8	7,1	12,7
53 Distrito Federal.....	0,0	24,1	29,3	29,8	27,0	41,2	46,5	0,0	8,9	8,9	6,3	5,8	8,4	9,5
<b>br Total nacional</b>	<b>31,0</b>	<b>38,4</b>	<b>45,9</b>	<b>56,1</b>	<b>65,4</b>	<b>81,6</b>	<b>102,7</b>	<b>3,3</b>	<b>2,1</b>	<b>2,1</b>	<b>2,2</b>	<b>2,3</b>	<b>2,3</b>	<b>2,8</b>

Figura 1 – Comparação das estimativas diretas e indiretas para  $Q_2$  obtidas segundo modelo PNAD de investigação da mortalidade

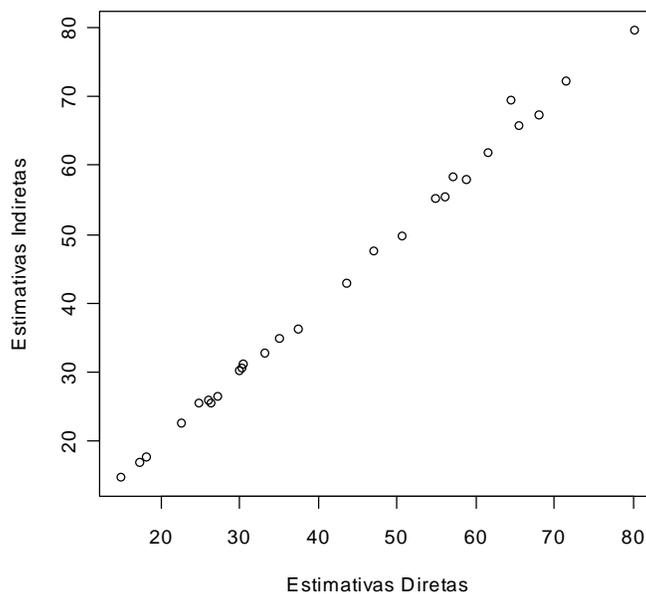


Figura 2 – Comparação das estimativas dos desvios padrão das estimativas diretas e indiretas para  $Q_2$  obtidas segundo modelo PNAD de investigação da mortalidade

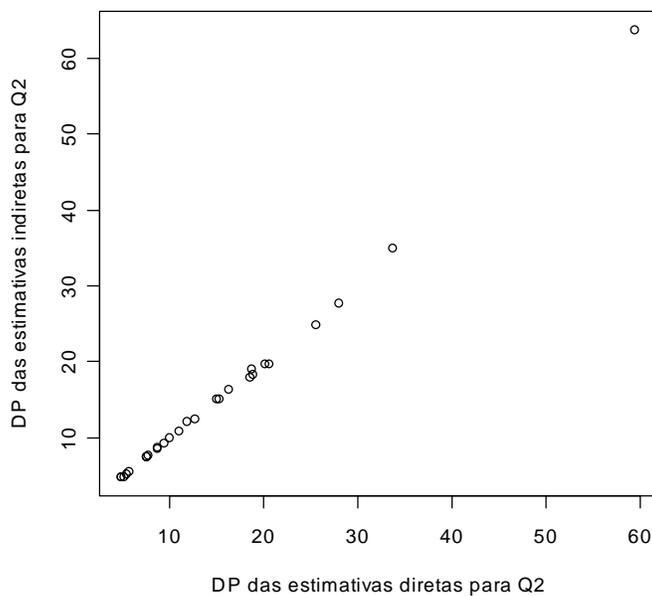


Figura 3 – Comparação das estimativas diretas da taxa de mortalidade  $D_2 / P_2$  obtidas usando as diferentes formas de investigação

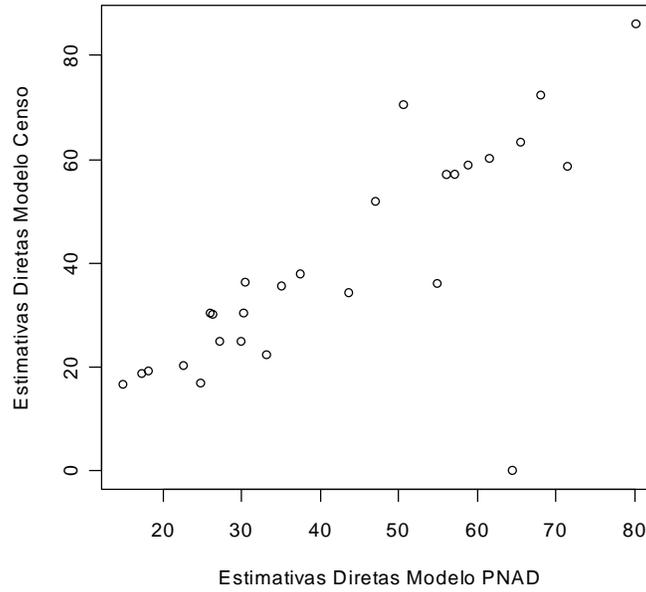


Figura 4 – Comparação das estimativas indiretas da taxa de mortalidade  $Q_2$  obtidas usando as diferentes formas de investigação

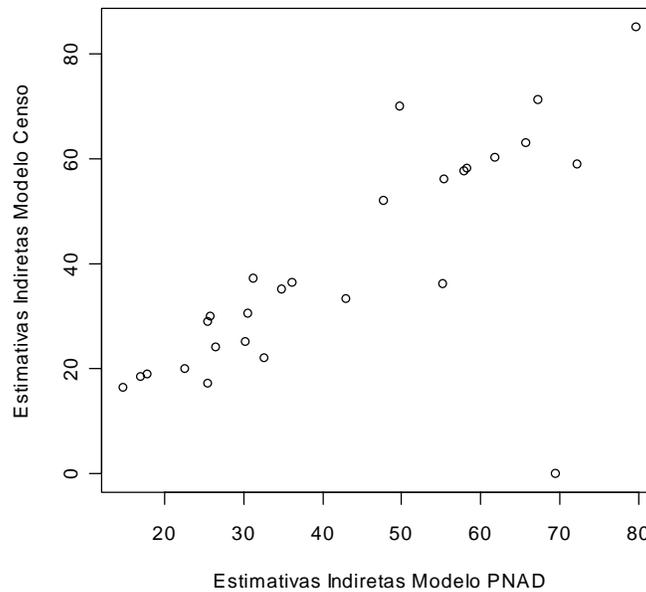


Figura 5 – Boxplots dos coeficientes de variação das estimativas indiretas de taxas de mortalidade obtidas segundo modelo PNAD de investigação

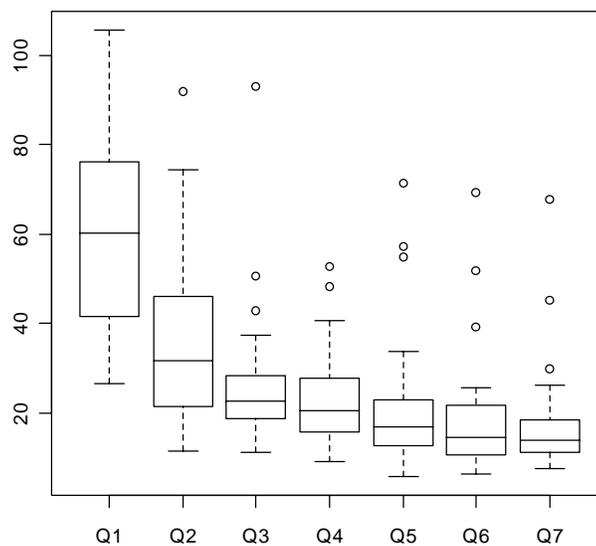


Tabela 5 - Estimativas dos coeficientes angulares e  $R^2$  em modelos de regressão para estimativas diretas da mortalidade segundo modelo Censo usando como preditores as estimativas diretas segundo modelo PNAD

Parâmetro	Coefficiente estimado	$R^2$
$D_1/P_1$	1,11306	0,934
$D_2/P_2$	0,91425	0,906
$D_3/P_3$	0,97381	0,985
$D_4/P_4$	1,01123	0,985
$D_5/P_5$	1,03752	0,995
$D_6/P_6$	1,00014	0,995
$D_7/P_7$	1,00839	0,998

Tabela 6 - Estimativas dos coeficientes angulares e  $R^2$  em modelos de regressão para estimativas indiretas da mortalidade segundo modelo Censo usando como preditores as estimativas indiretas segundo modelo PNAD

Parâmetro	Coefficiente estimado	$R^2$
$Q_1$	1,1130	0,950
$Q_2$	0,9015	0,895
$Q_3$	0,9771	0,985
$Q_4$	1,0123	0,985
$Q_5$	1,0387	0,995
$Q_6$	1,0012	0,995
$Q_7$	1,0116	0,998