

Ministério do Planejamento, Orçamento e Gestão
Instituto Brasileiro de Geografia e Estatística – IBGE
Diretoria de Pesquisas
Coordenação de População e Indicadores Sociais

Textos para discussão
Diretoria de Pesquisas
número 43

Estimadores de Diferenças Temporais e suas Variâncias: Uma Abordagem Aplicada ao Estudo de Indicadores Sociais a partir dos Dados da PNAD

Gilson Gonçalves de Matos
Ana Lucia Saboia
Leonardo Athias

Rio de Janeiro
2013

Instituto Brasileiro de Geografia e Estatística - IBGE
Av. Franklin Roosevelt, 166 - Centro - 20021-120 - Rio de Janeiro, RJ - Brasil

ISSN 1518-675X Textos para discussão. Diretoria de Pesquisas

Divulga estudos e outros trabalhos técnicos desenvolvidos pelo IBGE ou em conjunto com outras instituições, bem como resultantes de consultorias técnicas e traduções consideradas relevantes para disseminação pelo Instituto. A série está subdividida por unidade organizacional e os textos são de responsabilidade de cada área específica.

ISBN 978-85-240-4285-0

© IBGE. 2013

Impressão

Gráfica Digital/Centro de Documentação e Disseminação de Informações - CDDI/IBGE, em 2013.

Capa

Gerência de Criação/CDDI

Matos, Gilson Gonçalves de

Estimadores de diferenças temporais e suas variâncias : uma abordagem aplicada ao estudo de indicadores sociais a partir dos dados da PNAD / Gilson Gonçalves de Matos, Ana Lucia Saboia, Leonardo Athias. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2013. p. 37 - (Textos para discussão. Diretoria de Pesquisas, ISSN 1518-675X ; n. 43)

Inclui bibliografia.
ISBN 978-85-240-4285-0

1. Estimativa de parâmetros. 2. Teoria da estimativa. 3. Levantamentos domiciliares - Métodos estatísticos. 4. Indicadores sociais - Brasil. 5. Pesquisa Nacional por Amostra de Domicílios. I. Saboia, Ana Lucia. II. Athias, Leonardo. III. IBGE. Coordenação de População e Indicadores Sociais. IV. Título. V. Série.

Gerência de Biblioteca e Acervos Especiais
RJ/2013-08

CDU 519.233.5
EST

Impresso no Brasil / Printed in Brazil

Sumário

Apresentação	5
1. Introdução	7
2. Os indicadores selecionados.....	9
3. Metodologia	12
3.1. Construção de Médias, Percentuais e Razões na PNAD	12
3.2. Variância do Estimador de Médias, Percentuais e Razões na PNAD	15
3.3. Estimador de Diferenças Temporais e Variância do Estimador	17
3.4. Testes de Hipóteses	18
4. Análise dos Dados	19
5. Conclusões e sugestões	25
6. Bibliografia.....	27
Anexo 1. Exploração das Variáveis de Estrato e UPA da PNAD	29
Anexo 2. Tabelas de indicadores sociais da PNAD.....	31
Anexo 3. Box Plots dos Efeitos de Plano Amostral Ampliados (EPAs)	36

Apresentação

As análises contidas neste texto são de grande importância para o desenvolvimento dos trabalhos no âmbito das pesquisas domiciliares do IBGE. O exame detalhado das séries históricas de indicadores a partir da Pesquisa Nacional por Amostra de Domicílios - PNAD é essencial para o estudo dos fenômenos sociais. Neste particular, o presente texto contribui em termos metodológicos para o aperfeiçoamento das análises a serem realizadas, especialmente, pelo corpo técnico do IBGE.

Luiz Antônio Pinto de Oliveira

Coordenador de População e Indicadores Sociais

1. Introdução¹

As análises a partir de uma série temporal estatística são fundamentais para se conhecer a evolução de um fenômeno, especialmente, aqueles que se referem ao social. Uma série histórica nada mais é do que uma sequência de dados em intervalos de tempo durante um período específico. O primeiro passo é a observação do fenômeno em questão para descrever seu comportamento e avaliar quais os fatores que influenciaram os resultados na série para buscar as explicações. Este procedimento foi feito na análise da taxa de analfabetismo calculada a partir de uma série histórica de resultados da Pesquisa Nacional por Amostra de Domicílios - PNAD (1992 a 2009), tendo motivado a elaboração deste texto. A observação das flutuações nesse indicador exigiu um esforço de avaliar metodologicamente tais resultados, além de alguns outros indicadores com base nesta fonte.

A PNAD desde 1967 vem sendo sistematicamente utilizada como base de análise para os estudos de condições de vida. Esta fonte de dados investiga diversas características socioeconômicas, fazendo parte de seu corpo básico temas referentes à população, educação, trabalho, rendimento e condições de saneamento do domicílio. Este conjunto temático socioeconômico da PNAD faz com que seja possível avaliar e acompanhar anualmente os efeitos das políticas públicas na área social.

O presente estudo baseou-se nos dados da PNAD referentes à última década, dando enfoque à diferença de alguns indicadores em três pontos no tempo - 2001, 2005 e 2009.

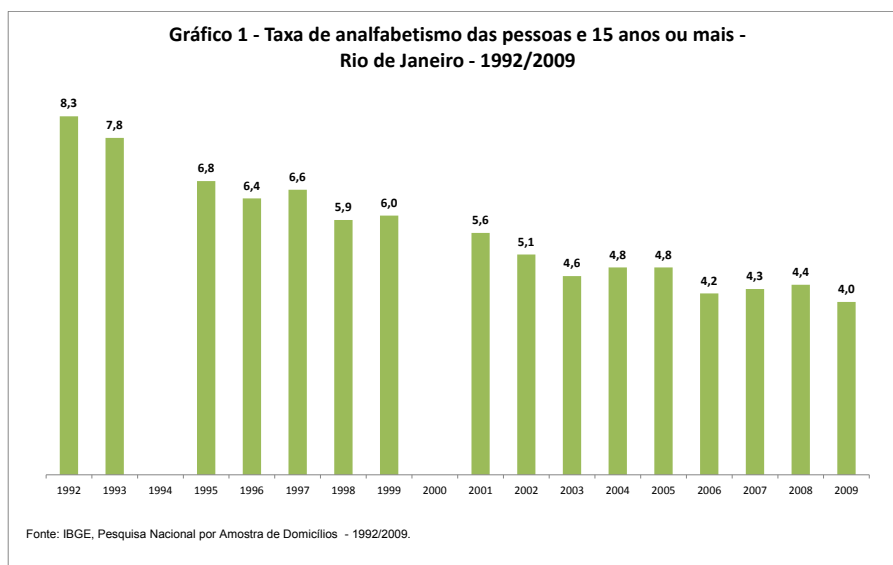
As variações e flutuações dos indicadores nas séries históricas da PNAD podem estar associadas a mudanças nos fenômenos sociais advindas de políticas públicas, como, por exemplo, a redução da taxa de analfabetismo, devido a programas específicos de erradicação do analfabetismo no país². Contudo, podem estar simplesmente relacionadas a flutuações amostrais.

A título de contextualizar o presente estudo, tomou-se a taxa de analfabetismo das pessoas de 15 anos ou mais para o estado do Rio de Janeiro na série que vai de 1992 a 2009 para

¹ Os autores agradecem a Pedro Luís do Nascimento Silva pelas sugestões e contribuições na parte metodológica, a Marcus Vinícius M. Fernandes pelo auxílio na construção das informações de estratos e conglomerados primários da PNAD e a Sonia Albieri pela leitura atenta.

² Neste contexto, o programa Brasil Alfabetizado pode ser citado.

mostrar as flutuações nos dados (Gráfico 1). Os valores observados indicam uma tendência clara de redução da taxa de analfabetismo no estado em pauta. Nosso primeiro questionamento foi verificar se aqueles valores estavam realmente mostrando uma redução do fenômeno e em qual dimensão. Daí a necessidade de uma avaliação mais criteriosa, que merece levar em conta a covariância implicada pela estrutura de dependência presente na amostra da PNAD ao longo das duas últimas décadas.



O outro indicador selecionado para o mesmo tipo de exame estatístico foi a “proporção de domicílios com pessoa de referência do sexo feminino, em especial, aqueles com presença de cônjuge”. Este indicador tem sido utilizado como expressão de um comportamento feminino, independentemente das limitações do conceito de pessoa de referência. Nas sociedades contemporâneas, as relações de gênero têm papel fundamental nos padrões de organização das famílias, dado que estas têm apresentado importantes mudanças tanto no âmbito privado quanto no público. A declaração por parte das mulheres como pessoas de referência das famílias ou dos domicílios pode estar expressando um novo posicionamento da mulher na sociedade brasileira, necessitando ser investigado do ponto de vista estatístico.

A essa introdução, segue uma seção (2) com uma descrição geral dos indicadores selecionados. Na seção seguinte (3) é apresentada a metodologia, subdividida em quatro subseções: (3.1) um estimador geral para indicadores de médias, proporções e razões na PNAD é descrito, em seguida (3.2) versa-se sobre a variância desse estimador em um instante de tempo,

(3.3) sobre a variância observada entre duas edições da PNAD e (3.4) detalha-se a construção do teste de hipótese levando em conta a covariância entre as pesquisas. Na quarta seção foram analisados dois indicadores originários da PNAD, taxa de analfabetismo e proporção de chefia feminina³ em domicílios particulares permanentes. Em ambos os casos, há interpretação das diferenças encontradas entre esses indicadores em dois instantes, de forma a avaliar o impacto nas conclusões a respeito dos indicadores com o uso da metodologia apresentada. Uma quinta seção corresponde a conclusões e sugestões extraídas do estudo. Nos anexos, consta uma descrição da construção das variáveis de estrato e UPA (Unidade Primária de Amostragem) na PNAD, além de tabelas adicionais de indicadores⁴ e box-plots do Efeito de Plano Amostral Ampliado, para os indicadores e níveis geográficos trabalhados.

2. Os indicadores selecionados

A alfabetização é uma condição essencial para o exercício da cidadania. Segundo a UNESCO - United Nations Educational, Scientific and Cultural Organization, uma pessoa alfabetizada é aquela que pode, com entendimento, ler e escrever uma simples declaração ou bilhete no seu dia a dia. Uma pessoa capaz de ler e escrever somente algarismos e seu próprio nome deve ser considerada analfabeta, assim como aquela que lê ou escreve uma frase ritual memorizada.

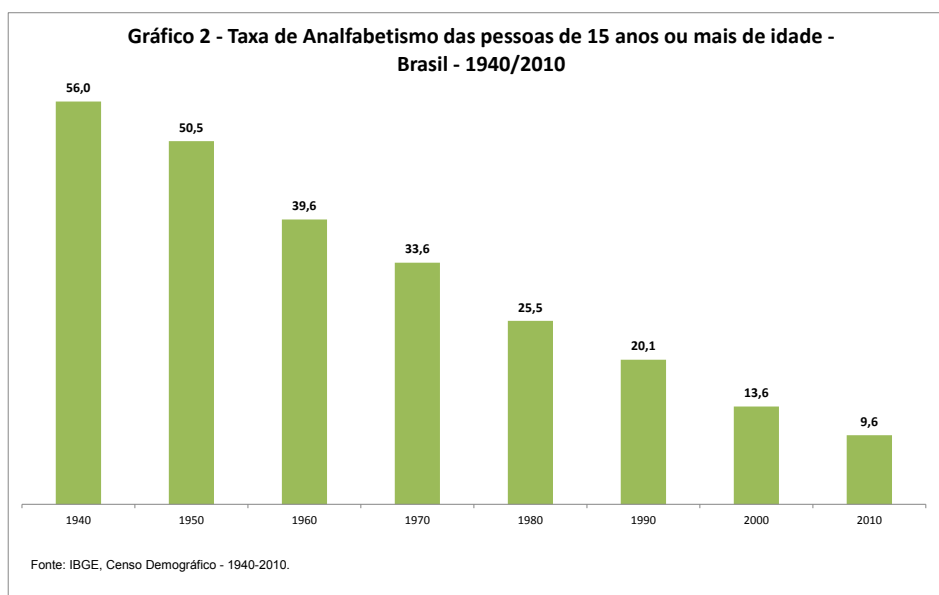
Para dimensionar a importância da informação produzida pelo IBGE, em termos internacionais, por exemplo, a UNESCO utiliza tais informações do IBGE como dados oficiais para estabelecer comparações entre os níveis educacionais dos países. A UNESCO considera que a alfabetização é pré-requisito para a maioria das formas de aprendizado independentemente do grupo etário. É uma ferramenta crucial para qualquer criança, jovem ou adulto adquirir habilidades essenciais que lhes proporcionará chances e oportunidades para vencer os desafios do cotidiano. Segundo a UNESCO, a alfabetização representa um passo essencial da educação básica, tornando-se indispensável para a efetiva participação na sociedade e nas economias do século XXI. Diante do elevado número de analfabetos no mundo, a UNESCO criou um programa (LIFE - Literacy Initiative For Empowerment) com duração de 10 anos (até 2015) para 35 países

³ O termo chefia feminina será utilizado para expressar domicílios/famílias com pessoa de referência do sexo feminino.

⁴ Tabelas completas por Grande Região, Unidades da Federação e Regiões Metropolitanas dos indicadores: taxa de analfabetismo e chefia feminina. Tabelas adicionais da média de anos de estudo das pessoas de 15 anos ou mais e a proporção de jovens de 18 a 24 anos com 11 anos ou mais de estudo.

considerados prioridade para ação e desafios na luta pela erradicação do analfabetismo. Elegeram dois grupos, tendo como base dados do período de 2000 a 2004: um primeiro onde estão 27 países cujas taxas de analfabetismo superam 50% da população, e um segundo com 11 países⁵ (Egito, Marrocos, China, Indonésia, Bangladesh, Índia, Irã, Paquistão, Etiópia e Nigéria e Brasil) cujo número absoluto de analfabetos é superior a 10 milhões.

No caso do Brasil, a taxa de analfabetismo tem sido produzida pelo IBGE há bastante tempo, sendo possível elaborar uma longa série histórica. É importante mencionar que, desde os primeiros Censos realizados pelo instituto, é possível contar com esta informação. A título de ilustração, mostra-se o comportamento deste indicador desde a década de 1940, com base nas informações censitárias. Nota-se que houve uma expressiva redução na taxa de analfabetismo entre 1940 e 2010.



Na PNAD, a pergunta que mede alfabetização é “Sabe ler e escrever?”, e as respostas são “Sim” ou “Não”, com a instrução (em itálico):

1- SIM para a pessoa capaz de ler e escrever *um recado* ou bilhete simples no idioma que conhece. **Inclua também a pessoa alfabetizada que se tornou física ou mentalmente incapacitada de ler ou escrever;** ou

⁵ Três destes 11 países estão também listados no primeiro grupo.

2 - NÃO para a pessoa que não sabe ler e escrever ou que apenas escreve o próprio nome. **Inclua** neste código a pessoa que aprendeu a ler e escrever, mas esqueceu devido a ter passado por um processo de alfabetização precário que não se consolidou.

A captação pela PNAD obedece ao mesmo critério do Censo Demográfico (conceito e instrução) desde 1992, com uma ligeira diferença na instrução.⁶

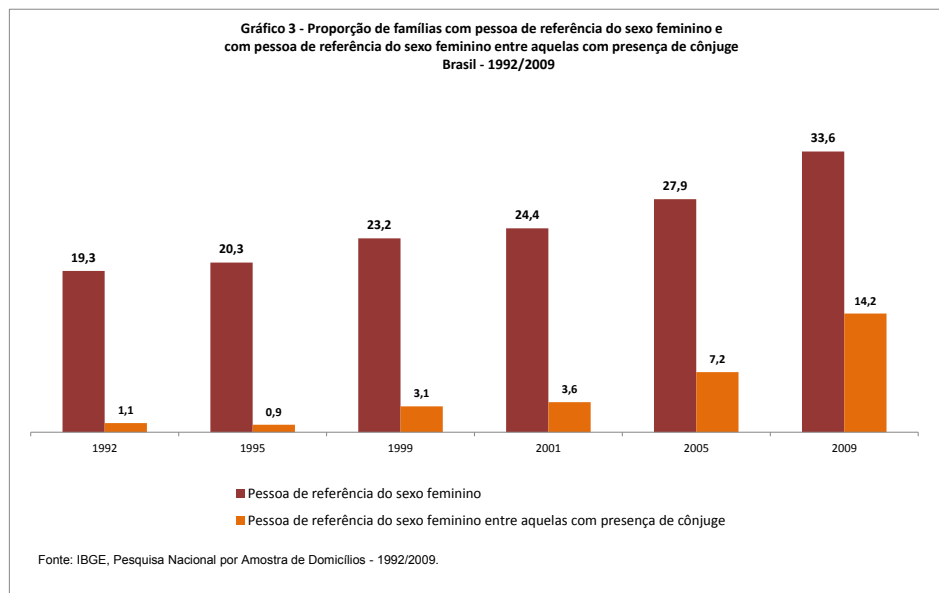
Tendo em vista os novos aportes para a análise dos fenômenos sociais, surge um desafio metodológico e conceitual de como adotar uma perspectiva de gênero — dado seu caráter transversal — ao processo de construção de indicadores sociais e de formulação de políticas públicas. Atualmente, nos estudos de gênero, a variável que analisa a questão da “chefia” da família ou domicílio tem sido amplamente utilizada, principalmente na formulação de estratégias e políticas de redução da pobreza e desigualdade.

O conceito de “chefia” ou “pessoa de referência” está associado a padrões socioculturais tais como: quem tem autoridade na família ou quem é responsável pelos negócios e pela provisão do sustento da família – papéis geralmente atribuídos às pessoas do sexo masculino. Nas pesquisas domiciliares do IBGE, a maioria das famílias apresenta pessoas do sexo masculino identificadas com a pessoa de referência ou responsável. Segundo as Nações Unidas, a definição tradicional de chefe de família pressupõe que a maioria das famílias seja constituída com uma pessoa dessa unidade familiar que tenha autoridade e responsabilidade sobre os negócios da família, e na maioria dos casos, é seu suporte econômico. Essa pessoa é então designada chefe (*head*) da família. Ainda, a ONU recomenda que sejam estabelecidas instruções claras sobre como deve ser definida a chefia de família para que sejam evitadas dificuldades decorrentes de preconceitos das pessoas que respondem aos questionários em relação a esse assunto. O procedimento para identificação do chefe, quando os membros da unidade doméstica são incapazes de fazê-lo, deve ser claro e preciso para evitar um viés de sexo. O Gráfico 3 mostra a proporção de famílias com chefia feminina e a proporção de famílias com chefia feminina entre

⁶ No Censo 2000, encontramos: Quesito 4.28 (amostra) e 4.06 (básico) - Sabe ler e escrever? Conforme o caso, assinale:

1 - SIM para a pessoa capaz de ler e escrever *pelo menos um bilhete simples* no idioma que conhece. Considere também a pessoa alfabetizada que se tornou física ou mentalmente incapacitada de ler ou escrever; e
2 - NÃO para a pessoa que nunca *aprendeu* a ler e escrever ou que, embora tenha aprendido, esqueceu. Considere também como não sabendo ler e escrever, a pessoa que só é capaz de escrever o próprio nome.

aquelas com presença de cônjuge. Os dois casos analisados apresentaram crescimento significativo nas últimas duas décadas.



3. Metodologia

Na presente seção serão apresentadas as técnicas utilizadas para a avaliação das diferenças de indicadores em dois instantes de tempo para uma pesquisa amostral complexa, neste caso, a PNAD. Os resultados metodológicos apresentados foram integralmente implementados no software SAS, através da *procedure sql* e de *data steps*.

3.1. Construção de Médias, Percentuais e Razões na PNAD

Um grande número de indicadores divulgados na Síntese de Indicadores Sociais da PNAD são médias e percentuais, sendo estes divulgados para diversos níveis de agregação e domínios de estudo. A presente subseção descreve um estimador geral para esses indicadores. Para a estimação da variância destes indicadores, podemos recorrer a pacotes como o SUDAAN, a biblioteca *Survey* do R, entre outros.

Ao estimarmos a diferença de um dado indicador para dois instantes distintos dentro da mesma década, a variância desta diferença seria estimada pela soma das variâncias nos dois instantes de tempo, em caso de independência. Acontece que a seleção de setores da amostra da PNAD é realizada no início da década, e mantida até o final, inserindo assim uma estrutura de dependência na amostra, para cada década.

Para a PNAD Contínua, que sucede à PNAD no seu formato atual a partir de 2013, devido ao sistema de rotação da amostra, essa dependência permanece por um período mais curto de tempo. A amostra mestra é renovada em 2,5% a cada trimestre, sendo substituída totalmente em dez anos, ou 40 trimestres.

Na Síntese de Indicadores Sociais os indicadores são divulgados para Brasil, Grandes Regiões, Unidades da Federação e para nove Regiões Metropolitanas. Sabe-se que o peso utilizado para obter as estimativas não é o inverso da fração amostral, visto que são feitos ajustes de calibração (pós-estratificação), de tal forma que a população estimada para todos os pós-estratos seja igual à estimativa de população divulgada pelo IBGE.

Os pós-estratos da PNAD são definidos de acordo com a Unidade da Federação, a Região Metropolitana e a situação da área geográfica em análise (urbana e rural). Para as Unidades da Federação do Norte, com exceção do Pará e de Tocantins, define-se um pós-estrato para a área urbana e outro para a área rural. No caso de Tocantins, há apenas um pós-estrato, visto que sua área rural sempre foi pesquisada na PNAD, e no Pará, há dois pós-estratos para a área rural, e dois para a área urbana, e em ambos os casos separa-se a Região Metropolitana do restante da Unidade da Federação. Nas outras oito Unidades da Federação que possuem Regiões metropolitanas pesquisadas na PNAD, há um pós-estrato para a Região Metropolitana e outro para o restante da UF e nas Unidades da Federação restantes há um único pós-estrato.

O estimador do tipo razão para um dado instante de tempo t , do qual são casos particulares os estimadores de médias e proporções, pode ser expresso da seguinte forma:

$$\hat{R}^t = \frac{\sum_{k=1}^K \hat{\theta}_k^t \hat{Y}_k^t}{\sum_{k=1}^K \hat{\theta}_k^t \hat{X}_k^t} = \frac{\hat{Y}_{PS}^t}{\hat{X}_{PS}^t} \quad (1)$$

$$\hat{Y}_k^t = \sum_{h=1}^{H_k} \sum_{i=1}^{n_{kh}} \sum_{j=1}^{m_{khi}} w^t_{khij} y^t_{khij}$$

$$\hat{X}_k^t = \sum_{h=1}^{H_k} \sum_{i=1}^{n_{kh}} \sum_{j=1}^{m_{khi}} w^t_{khij} x^t_{khij}$$

$$\hat{\theta}_k^t = \frac{P_k^t}{\hat{P}_k^t}$$

P_k^t é a população projetada para o k -ésimo pós-estrato

$$\hat{P}_k^t = \sum_{h=1}^{H_k} \sum_{i=1}^{n_{kh}} \sum_{j=1}^{m_{khi}} w^t_{khij}$$

K é o número de pós-estratos

H_k é o número de estratos no k -ésimo pós-estrato

n_{kh} é o número de unidades primárias de amostragem (UPAs) selecionadas para a amostra do estrato h do k -ésimo pós-estrato

m_{khi} é o número de unidades de interesse (domicílios, famílias ou pessoas) para a amostra pertencente à UPA i , estrato h e k -ésimo pós-estrato

w^t_{khij} é o peso amostral da j -ésima unidade de interesse da UPA i , estrato h e pós-estrato k

y^t_{khij} é o valor da variável de interesse do numerador da j -ésima unidade de interesse da UPA i , estrato h e pós-estrato k . Quando estamos calculando uma proporção, y_{khij} assume os valores 0 ou 1

x^t_{khij} é o valor da variável de interesse do denominador da j -ésima unidade de interesse da UPA i , estrato h e pós-estrato k . No cálculo de médias e proporções, esta variável assume o valor 1 para todas as unidades de interesse.

Quando o indicador é obtido para um domínio de estudo d , as únicas expressões que se alteram na Eq.(1) são as de \hat{Y}_k^t e \hat{X}_k^t , que devem ser reescritas da seguinte forma:

$$\hat{Y}_k^t(d) = \sum_{h=1}^{H_k} \sum_{i=1}^{n_{kh}} \sum_{j=1}^{m_{khi}} \delta_{khij}(d) w^t_{khij} y^t_{khij}$$

$$\hat{X}_k^t(d) = \sum_{h=1}^{H_k} \sum_{i=1}^{n_{kh}} \sum_{j=1}^{m_{khi}} \delta_{khij}(d) w^t_{khij} x^t_{khij}$$

$$\delta_{khij}(d) = \begin{cases} 1, & \text{se a unidade com índice } (khij) \text{ pertence a } d \\ 0, & \text{caso contrário} \end{cases}$$

3.2. Variância do Estimador de Médias, Percentuais e Razões na PNAD

Para obter a variância do estimador dado na Eq.(1), será utilizada a técnica de Linearização de Taylor em conjunto com o Método do Conglomerado Primário (Pessoa e Nascimento Silva, 1998). Métodos como Jackknife e Bootstrap poderiam facilmente ser utilizados para a estimação das variâncias. Porém a abordagem por Linearização de Taylor é mais clara, em vista do interesse principal na estimação da covariância do estimador em dois instantes de tempo.

Seja $\hat{R}^t = f(\hat{Y}^t, \hat{X}^t, \hat{P}^t)$, onde \hat{R}^t é dado pela equação Eq.(1), $\hat{Y}^t = (\hat{Y}_1^t, \hat{Y}_2^t, \dots, \hat{Y}_k^t)$, $\hat{X}^t = (\hat{X}_1^t, \hat{X}_2^t, \dots, \hat{X}_k^t)$ e $\hat{P}^t = (\hat{P}_1^t, \hat{P}_2^t, \dots, \hat{P}_k^t)$. Pode-se então obter uma aproximação de primeira ordem para \hat{R}^t em torno do ponto (Y^t, X^t, P^t) , onde os vetores são correspondentes aos totais populacionais para os k pós-estratos referentes à unidade de análise.

$$\hat{R}^t \cong \hat{R}_L^t = R^t + \sum_{k=1}^K \left[a_k^t (\hat{Y}_k^t - Y_k^t) + b_k^t (\hat{X}_k^t - X_k^t) + c_k^t (\hat{P}_k^t - P_k^t) \right] \quad (2)$$

onde:

$$a_k = \left. \frac{\partial f(\hat{Y}^t, \hat{X}^t, \hat{P}^t)}{\partial \hat{Y}_k^t} \right|_{(\hat{Y}^t, \hat{X}^t, \hat{P}^t) = (Y^t, X^t, P^t)}$$

$$b_k = \left. \frac{\partial f(\hat{Y}^t, \hat{X}^t, \hat{P}^t)}{\partial \hat{X}_k^t} \right|_{(\hat{Y}^t, \hat{X}^t, \hat{P}^t) = (Y^t, X^t, P^t)}$$

$$c_k = \left. \frac{\partial f(\hat{Y}^t, \hat{X}^t, \hat{P}^t)}{\partial \hat{P}_k^t} \right|_{(\hat{Y}^t, \hat{X}^t, \hat{P}^t) = (Y^t, X^t, P^t)}$$

para $k = 1, 2, \dots, K$.

A Eq.(2) pode ser reescrita da seguinte forma:

$$\hat{R}^t \cong R^t + (\hat{Z}_{HT}^t - Z^t) \quad (3)$$

$\hat{Z}_{HT}^t = \sum_{k=1}^K (a_k^t \hat{Y}_k^t + b_k^t \hat{X}_k^t + c_k^t \hat{P}_k^t)$ pode ser expresso como um estimador do tipo Horvitz-Thompson;

$$Z^t = \sum_{k=1}^K (a_k^t Y_k^t + b_k^t X_k^t + c_k^t P_k^t).$$

Dessa forma, há clareza que para examinar a variância do estimador \hat{R}^t basta determinar a variância do estimador \hat{Z}_{HT}^t . As constantes a_k^t , b_k^t e c_k^t devem ser estimadas, haja vista que estas dependem de parâmetros populacionais desconhecidos. Seus estimadores são dados por:

$$\begin{aligned}\hat{a}_k^t &= \left(\frac{1}{\hat{X}_{PS}^t} \frac{P_k^t}{\hat{P}_k^t} \right) \\ \hat{b}_k^t &= - \left(\frac{1}{\hat{X}_{PS}^t} \frac{P_k^t}{\hat{P}_k^t} \right) \frac{\hat{Y}_{PS}^t}{\hat{X}_{PS}^t} \\ \hat{c}_k^t &= \left(\frac{1}{\hat{X}_{PS}^t} \frac{P_k^t}{\hat{P}_k^t} \right) \left[\left(\frac{1}{\hat{X}_{PS}^t} \right) \left(\hat{Y}_{PS}^t \frac{\hat{X}_k^t}{\hat{P}_k^t} - \hat{X}_{PS}^t \frac{\hat{Y}_k^t}{\hat{P}_k^t} \right) \right]\end{aligned}$$

A variável linearizada pode ser escrita como:

$$z^t_{khij} = \left(\frac{1}{\hat{X}_{PS}^t} \frac{P_k^t}{\hat{P}_k^t} \right) \left[\left(y^t_{khij} - \hat{R}_{Y_k}^t \right) - \hat{R}^t \left(x^t_{khij} - \hat{R}_{X_k}^t \right) \right] w^t_{khij} \quad (4)$$

As quantidades $\hat{R}_{Y_k}^t$ e $\hat{R}_{X_k}^t$ são dadas, respectivamente, por:

$$\begin{aligned}\hat{R}_{Y_k}^t &= \frac{\hat{Y}_k^t}{\hat{P}_k^t} \\ \hat{R}_{X_k}^t &= \frac{\hat{X}_k^t}{\hat{P}_k^t}\end{aligned}$$

Quando o indicador é calculado para um domínio d , a variável linearizada é dada por:

$$\begin{aligned}z^t_{khij}(d) &= \left(\frac{1}{\hat{X}_{PS}^t} \frac{P_k^t}{\hat{P}_k^t} \right) \left[\left(y^t_{khij} \delta_{khij}(d) - \hat{R}_{Y_k}^t \right) \right. \\ &\quad \left. - \hat{R}^t \left(x^t_{khij} \delta_{khij}(d) - \hat{R}_{X_k}^t \right) \right] w^t_{khij}\end{aligned} \quad (5)$$

Neste caso, $\hat{R}_{Y_k}^t$ e $\hat{R}_{X_k}^t$ são:

$$\begin{aligned}\hat{R}_{Y_k}^t &= \frac{\hat{Y}_k^t(d)}{\hat{P}_k^t} \\ \hat{R}_{X_k}^t &= \frac{\hat{X}_k^t(d)}{\hat{P}_k^t}\end{aligned}$$

onde $\delta_{khij}(d)$, $\hat{Y}_k^t(d)$, $\hat{X}_k^t(d)$ são definidos conforme a seção 3.1. As variáveis linearizadas podem facilmente ser utilizadas para a obtenção das variâncias por meio do Método do Conglomerado Primário.

As unidades primárias da PNAD, qualquer seja o estrato natural a que pertencem, são selecionadas por sorteio sistemático, com probabilidade proporcional ao tamanho (PPT), sendo a medida de tamanho o número de domicílios obtido no Censo, a menos do Universo das Novas Construções, onde se realiza sorteio sistemático simples. O estimador proposto a seguir corresponde à aproximação do plano amostral da PNAD por um plano PPT com reposição.

$$\hat{V}(\hat{R}^t) = \hat{V}(\hat{Z}_{HT}^t) = \sum_{k=1}^K \sum_{h=1}^{H_k} n_{kh} s_{kh}^2 \quad (6)$$

onde

$$s_{kh}^2 = \frac{1}{(n_{kh} - 1)} \sum_{i=1}^{n_{kh}} (Z_{khi}^t - \bar{Z}_{kh}^t)^2$$

$$Z_{khi}^t = \sum_{j=1}^{m_{khij}} z_{khij}^t$$

$$\bar{Z}_{kh}^t = \frac{1}{n_{kh}} \sum_{j=1}^{n_{kh}} Z_{khi}^t$$

n_{kh} é o número de UPAs selecionadas para a amostra do estrato h do k -ésimo pós-estrato, já definido anteriormente.

3.3. Estimador de Diferenças Temporais e Variância do Estimador

O estimador de diferenças temporais entre os instantes t e $(t-k)$, bem como sua variância, são dados respectivamente por:

$$\hat{D}_{t,k} = \hat{R}^t - \hat{R}^{t-k} \quad (7)$$

$$\hat{V}_p(\hat{D}_{t,k}) = \hat{V}_p(\hat{R}^t) + \hat{V}_p(\hat{R}^{t-k}) - 2 \cdot \widehat{Cov}_p(\hat{R}^t, \hat{R}^{t-k}) \quad (8)$$

Os dois primeiros termos do lado direito da Eq.(8) podem ser obtidos de forma independente para cada um dos instantes de tempo pela Eq.(6). Já a covariância do indicador nos dois instantes de tempo pode ser estimada pela equação abaixo.

$$\widehat{Cov}_p(\hat{R}^t, \hat{R}^{t-k}) = \sum_{k=1}^{K^*} \sum_{h=1}^{H_k^*} n_{kh}^* s_{kh}^{t, t-k} \quad (9)$$

onde:

$$s_{kh}^{t, t-k} = \frac{1}{(n_{kh}^* - 1)} \sum_{i=1}^{n_{kh}^*} (Z_{khi}^t - \bar{Z}_{kh}^t) (Z_{khi}^{t-k} - \bar{Z}_{kh}^{t-k})$$

$$Z_{khi}^t = \sum_{j=1}^{m_{khij}} z_{khij}^t$$

$$\bar{Z}_{kh}^t = \frac{1}{n_{kh}^*} \sum_{j=1}^{n_{kh}^*} Z_{khi}^t$$

n_{kh}^* é o número de UPAs que se repetem para os dois períodos, no estrato h do k -ésimo pós-estrato.

3.4. Testes de Hipóteses

Será utilizada a normalidade assintótica do estimador de diferenças temporais para testar a hipótese de igualdade de um dado indicador para dois instantes de tempo. As hipóteses do teste são dadas por:

$$H_0: R^t = R^{t-k}$$

$$H_1: R^t \neq R^{t-k}$$

A estatística de teste é a estatística t, denotada por:

$$t_R = \frac{\hat{D}_{t,k}}{\sqrt{\hat{V}_p(\hat{D}_{t,k})}} = \frac{(\hat{R}^t - \hat{R}^{t-k})}{\sqrt{\hat{V}_p(\hat{D}_{t,k})}} \quad (10)$$

Para grandes amostras, a distribuição da estatística de teste é a normal padrão. Rejeita-se então a igualdade do indicador em estudo para os dois instantes quando o valor da estatística t_R é significativamente diferente de zero, dada a distribuição de referência.

A extensão multivariada deste teste poderia ser feita de modo natural, e a estatística analisada seria a estatística de Wald. Essa situação seria útil para a comparação de alterações na distribuição de cor ou raça dos indivíduos, por exemplo, visto que a esta possui 5 categorias.

Para fins de comparação, fizemos o teste de hipótese univariado desconsiderando a dependência implicada pela covariância entre dois pontos distintos no tempo, a qual foi denominada estatística t viesada. Esta foi então dada por:

$$t_v = \frac{\hat{D}_{t,k}}{\sqrt{\hat{V}_p(\hat{R}^t) + \hat{V}_p(\hat{R}^{t-k})}} = \frac{(\hat{R}^t - \hat{R}^{t-k})}{\sqrt{\hat{V}_p(\hat{R}^t) + \hat{V}_p(\hat{R}^{t-k})}} \quad (11)$$

No anexo 3 são apresentados alguns box plots do Efeito de Plano Amostral Ampliado (EPA), medida obtida pela razão entre a variância corretamente especificada e aquela onde se desconsidera a covariância presente na amostra entre dois períodos. Essa medida foi obtida para cada comparação entre os dois períodos e os níveis geográficos usualmente abordados na análise de indicadores sociais da PNAD, e é designada por:

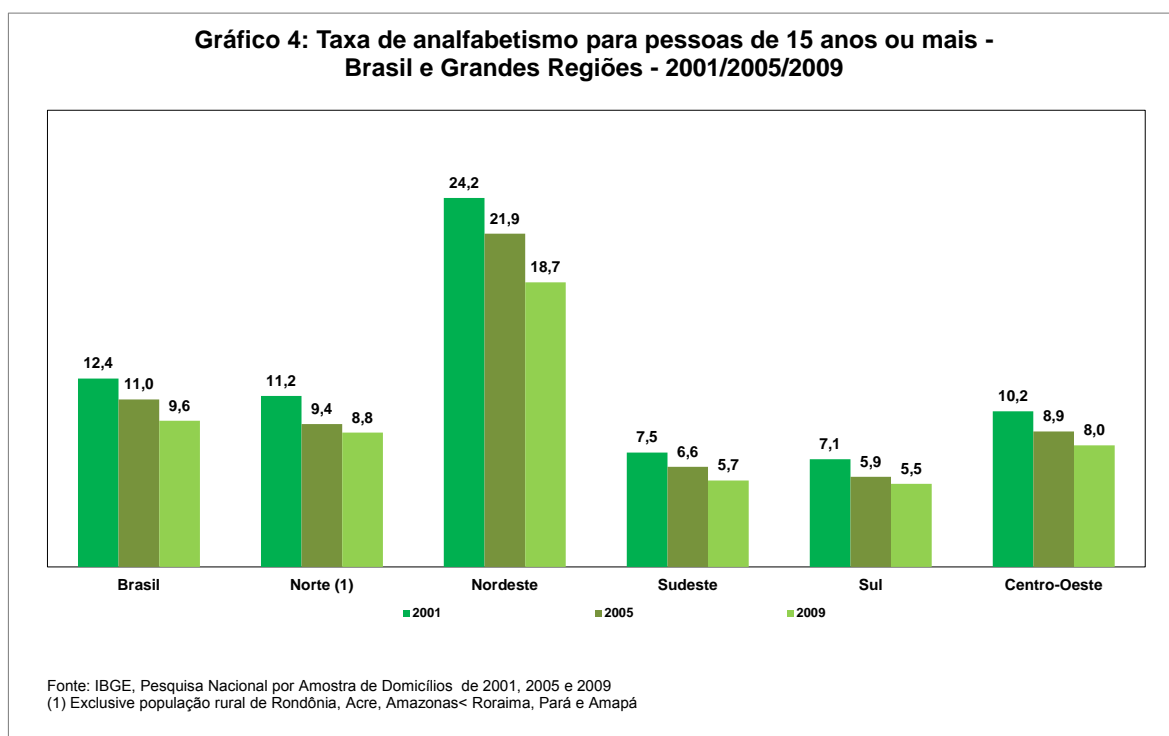
$$EPA = \frac{\hat{V}_p(\hat{R}^t) + \hat{V}_p(\hat{R}^{t-k}) - 2 \cdot \widehat{Cov}_p(\hat{R}^t, \hat{R}^{t-k})}{\hat{V}_p(\hat{R}^t) + \hat{V}_p(\hat{R}^{t-k})} = \left(\frac{t_v}{t_R}\right)^2 \quad (12)$$

4. Análise dos Dados

Após ter descrito os procedimentos técnicos para a análise, descreve-se o estudo dos indicadores da taxa de analfabetismo e da chefia feminina. Para fins de comparação entre os anos de 2001, 2005 e 2009 foram desconsideradas as áreas rurais das Unidades da Federação da Região Norte, a saber, Rondônia, Acre, Amazonas, Roraima, Pará e Amapá, pois, no ano de 2001 a amostra da PNAD não compreendia a área rural das UFs mencionadas, as quais já estavam nas amostras de 2005 e 2009. Dessa forma a taxa de analfabetismo apresentada aqui para 2009, por

exemplo, está ligeiramente distinta daquela divulgada (por 0,1%), devido à exclusão da área rural das UFs citadas.

A educação é um tema central na agenda do desenvolvimento e foco de políticas. A taxa de analfabetismo é um importante indicador para mensuração do nível de desenvolvimento socioeconômico de um país, bem como um indicador inicial para a mensuração das condições gerais de acesso à educação, conforme já mencionado anteriormente. Políticas para sua erradicação definitiva dependem de uma focalização baseada em estatísticas que forneçam subsídios para o controle dos ganhos/retrocessos ano a ano nos diferentes níveis geográficos. O Gráfico 4 mostra as taxas nos anos selecionados para Brasil e Grandes Regiões, onde é possível verificar uma redução em todas as regiões.



A Tabela 1 abaixo inclui as estimativas de taxa de analfabetismo e os erros-padrão associados (levando em conta a covariância entre os dois anos), ambos em porcentagem, assim como as comparações entre os anos, quer dizer, a partir da diferença em pontos percentuais. Por exemplo, a taxa de analfabetismo variou de 11,0% em 2005 para 9,6% no Brasil em 2009, ou seja, uma queda de 1,4 ponto percentual, o que, a partir do teste de hipóteses, indica diferença significativa. Além disso, esta é uma diferença substantiva, visto que o número absoluto de

analfabetos na população de 15 anos ou mais passa de 15,0 para 14,1 milhões. Ao mesmo tempo, pode-se comparar a diferença pela “estatística t”, valendo-se do controle da covariância (calculada a partir do arcabouço mostrado na seção 3) ou com sua ausência, nesse último caso, identificada como “estatística t viesada”.

Para todos os níveis de agregação apresentados, pode-se avaliar a significância das diferenças comparando as estatísticas de teste com os quantis da distribuição normal. Por exemplo, para um teste bilateral e um nível de significância de 5%, rejeita-se a igualdade (e conseqüentemente se aceita que o indicador apresenta diferença significativa entre os anos comparados) para valores observados da estatística de teste com valor absoluto superior a 1,96 (p-valor<0,05).

Tabela 1: Taxas de analfabetismo para pessoas de 15 anos ou mais, erros-padrão das estimativas e estatística t, considerando e desconsiderando a covariância implicada pela amostra, segundo Grandes Regiões - 2005 e 2009

Grandes Regiões	Taxa de Analfabetismo e Erro-Padrão (%)				Diferenças da Taxa de Analfabetismo (%), erros-padrão (%) e Estatísticas t			
	2005		2009		2009-2005			
	Estimativa	Erro Padrão	Estimativa	Erro Padrão	Diferença	Erro Padrão	Estatística t	Estatística t Viesada
Brasil (1)	11,0	0,2	9,6	0,1	-1,4	0,1	-13,1	-7,3
Norte (1)	9,4	0,5	8,8	0,5	-0,6	0,2	-2,4	-0,8
Nordeste	21,9	0,5	18,7	0,4	-3,2	0,3	-11,1	-5,5
Sudeste	6,6	0,2	5,7	0,1	-0,9	0,1	-6,4	-4,6
Sul	5,9	0,3	5,5	0,2	-0,5	0,2	-2,3	-1,4
Centro-Oeste	8,9	0,3	8,0	0,3	-0,9	0,3	-3,5	-2,4

Fonte: Pesquisa Nacional por Amostra de Domicílios 2005 e 2009.

(1) Exclusive população rural de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá.

Na comparação dos dados, para duas Grandes Regiões, Norte e Sul, o teste pela estatística t com e sem viés traz conclusões contraditórias, o que está marcado em vermelho na Tabela 1. Na Região Sul, por exemplo, a análise da diferença de cerca de -0,5 ponto percentual entre as taxas (5,9% em 2005 e 5,5% em 2009) levaria à conclusão que a variação não é significativa caso fosse usada a estatística t viesada (quer dizer, sem levar em conta a covariância), uma vez que o valor de -1,4 possui valor absoluto inferior a 1,96 (está dentro da faixa de +-1,96), enquanto a estatística t levando em conta a covariância é -2,3, significativa a 5%. Pode-se tirar a mesma conclusão em relação à região Norte.

Dois fenômenos podem estar concorrendo para essas duas regiões e não as demais apresentarem esses resultados. A região Norte, mesmo com a segunda maior proporção de analfabetos, tem a menor população total (denominador), enquanto a região Sul tem a menor taxa de analfabetismo (numerador).

As políticas educacionais contra o analfabetismo são resultado de esforços das três esferas de governo e, em muitos casos, direcionadas para os municípios (para os quais, diga-se, a PNAD não é fonte de informação) e UFs prioritárias. Dessa forma, exploramos as diferenças nas taxas nos três anos abordados (2001-2005-2009) para os estados do Nordeste. A Tabela 2 abaixo traz essa informação para UFs, e também para as regiões metropolitanas abertas nos dados divulgados pelo IBGE.

Em três estados, há resultados contraditórios (marcados em vermelho na tabela) tanto na comparação entre 2009 e 2005, quanto entre 2001 e 2009, período maior de tempo e quando se têm mudanças mais fortes de patamar das taxas de analfabetismo. Por exemplo, na Paraíba, o quantitativo de analfabetos passa de 27,2% para 25,2% e 21,6% entre 2001, 2005 e 2009. Inclusive na comparação entre 2001 e 2009, quando a taxa diminui em 5,5%, usar a estatística t viesada (-1,9) leva a considerar a diferença não significativa, enquanto se tem outra conclusão com a estatística t corretamente especificada (-2,3).

Tabela 2: Taxas de analfabetismo para pessoas de 15 anos ou mais, erros-padrão das estimativas e estatística t, considerando e desconsiderando a covariância implicada pela amostra, segundo Unidades da Federação e Regiões Metropolitanas do Nordeste - 2001, 2005 e 2009

Unidades da Federação e Regiões Metropolitanas do Nordeste	Taxa de Analfabetismo (Erro-Padrão)			Estatística t (Estatística t viesada)		
	2001	2005	2009	2005 - 2001	2009 - 2001	2009 - 2005
Maranhão	23,4 (1,5)	23,0 (2,4)	19,1 (1,6)	-0,2 (-0,1)	-2,8 (-2,0)	-2,8 (-1,3)
Piauí	29,4 (3,7)	27,4 (3,2)	23,4 (2,3)	-1,4 (-0,4)	-3,1 (-1,4)	-2,5 (-1,0)
Ceará	24,8 (1,1)	22,6 (0,9)	18,6 (0,8)	-3,2 (-1,5)	-9,4 (-4,6)	-7,1 (-3,3)
RM de Fortaleza	12,5 (2,6)	12,0 (2,1)	9,6 (1,9)	-0,3 (-0,2)	-1,9 (-0,9)	-1,7 (-0,9)
Rio Grande do Norte	24,2 (0,9)	21,5 (0,8)	18,1 (0,6)	-3,0 (-2,3)	-5,5 (-5,5)	-4,1 (-3,4)
Paraíba	27,2 (2,0)	25,2 (2,0)	21,6 (2,0)	-1,2 (-0,7)	-2,3 (-1,9)	-2,8 (-1,3)
Pernambuco	22,0 (0,5)	20,5 (0,6)	17,6 (0,7)	-3,6 (-1,9)	-8,2 (-5,2)	-4,7 (-3,4)
RM de Recife	11,0 (1,2)	9,6 (1,3)	8,2 (1,4)	-1,4 (-0,8)	-2,4 (-1,5)	-1,0 (-0,8)
Alagoas	30,6 (1,9)	29,3 (1,1)	24,6 (1,0)	-0,6 (-0,6)	-3,6 (-2,8)	-3,9 (-3,2)
Sergipe	21,4 (1,8)	19,7 (2,0)	16,3 (1,5)	-3,0 (-0,7)	-8,7 (-2,2)	-4,6 (-1,4)
Bahia	22,8 (0,7)	18,7 (0,7)	16,7 (0,5)	-8,7 (-4,2)	-12,2 (-6,7)	-4,5 (-2,3)
RM de Salvador	6,7 (2,9)	6,1 (2,5)	5,4 (2,0)	-0,3 (-0,2)	-0,7 (-0,4)	-0,4 (-0,2)

Fonte: Pesquisa Nacional por Amostra de Domicílios 2005 e 2009.

Nota: Os valores em vermelho indicam inversão na decisão ao se utilizar a estatística t viesada.

Em relação a esse indicador, a correta especificação da estatística t trouxe maior precisão à análise e conclusões diferentes da ausência da especificação. Na sequência, observar-se-á como isso se comporta no estudo da chefia feminina.

Os resultados discutidos a seguir se referem às comparações das proporções de famílias com pessoas de referência do sexo feminino, no total e também quando há presença de cônjuge (casais). Deve-se ressaltar que, desde o início da década de 1990, houve um crescimento do fenômeno em estudo. Isso não parece ser resultado de mudança metodológica, pois a PNAD utiliza o mesmo questionário desde 1992.⁷

Nos três anos selecionados, como visto no Gráfico 3 acima, as diferenças são fortes, tanto para o total de família quanto nas famílias com presença de cônjuge. Nas Tabelas 3 e 4 abaixo, traz-se a informação para Brasil e Grandes regiões nos dois casos.

Tabela 3: Proporção de famílias com chefia feminina, erros padrão das estimativas e estatística t, considerando e desconsiderando a covariância implicada pela amostra, segundo Grandes Regiões - 2001, 2005 e 2009

Grandes Regiões	Proporção de Famílias com Chefia Feminina (Erro-Padrão)			Estatística t (Estatística t viesada)		
	2001	2005	2009	2005 - 2001	2009 - 2001	2009 - 2005
Brasil (1)	27,3 (0,2)	30,8 (0,2)	35,2 (0,2)	15,6 (13,0)	30,7 (26,2)	17,7 (15,0)
Norte (1)	29,1 (0,6)	33,5 (0,9)	38,7 (0,7)	6,0 (3,9)	11,1 (10,1)	4,5 (4,4)
Nordeste	27,8 (0,3)	31,3 (0,4)	34,8 (0,5)	9,0 (7,0)	17,8 (12,4)	8,4 (6,0)
Sudeste	28,0 (0,3)	31,4 (0,3)	36,0 (0,4)	9,0 (8,0)	17,5 (16,2)	10,2 (9,5)
Sul	24,4 (0,4)	27,6 (0,5)	33,0 (0,6)	6,2 (4,9)	14,4 (11,8)	10,0 (7,3)
Centro-Oeste	26,7 (0,6)	30,3 (0,5)	36,2 (0,6)	6,0 (4,7)	13,6 (11,5)	9,3 (7,2)

Fonte: Pesquisa Nacional por Amostra de Domicílios 2001, 2005 e 2009.

(1) Exclusive população rural de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá nos três anos de estudo.

Tabela 4: Proporção de famílias com chefia feminina entre as famílias com responsável e cônjuge, erros padrão das estimativas e estatística t, considerando e desconsiderando a covariância implicada pela amostra, segundo Grandes Regiões, Unidades da Federação e Regiões Metropolitanas - 2001, 2005 e 2009

Grandes Regiões, Unidades da Federação e Regiões Metropolitanas	Proporção de Famílias com Chefia Feminina entre aquelas com Cônjuge (Erro-Padrão)			Estatística t (Estatística t viesada)		
	2001	2005	2009	2005 - 2001	2009 - 2001	2009 - 2005
Brasil (1)	3,6 (0,1)	7,2 (0,2)	14,3 (0,3)	20,7 (20,6)	38,6 (38,0)	24,9 (23,6)
Norte (1)	5,8 (0,3)	10,0 (0,9)	18,0 (0,9)	4,6 (4,4)	13,1 (12,7)	6,1 (6,3)
Nordeste	2,9 (0,2)	7,0 (0,3)	11,9 (0,4)	14,2 (14,2)	23,3 (21,2)	11,5 (10,2)
Sudeste	3,7 (0,2)	6,9 (0,2)	14,9 (0,5)	11,2 (11,2)	21,5 (21,8)	15,8 (15,1)
Sul	3,4 (0,2)	6,8 (0,4)	14,5 (0,5)	8,4 (8,3)	20,6 (19,2)	12,7 (12,0)
Centro-Oeste	4,4 (0,3)	8,8 (0,5)	16,2 (0,7)	7,8 (7,5)	15,4 (15,3)	9,2 (8,6)

Fonte: Pesquisa Nacional por Amostra de Domicílios 2001, 2005 e 2009.

(1) Exclusive população rural de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá nos três anos de estudo.

⁷ Isso difere do Censo Demográfico, que utilizou o termo “chefe do domicílio” em 1991, mas “pessoa responsável pelo domicílio” em 2000 e 2010.

Nas duas tabelas, as conclusões são as mesmas ao se utilizar a estatística t corretamente especificada e aquela com a hipótese de independência na amostra da pesquisa entre os dois períodos, ou seja, rejeita-se a igualdade dos indicadores em todas as comparações realizadas. Isso ocorre pelo aumento acentuado do fenômeno no período em análise. Embora não foram reportados, os p-valores associados aos valores observados das estatísticas t são pequenos, e dessa forma, mesmo adotando níveis de significância bastante conservadores não haveria alteração nas conclusões apresentadas.

Não se tem, no entanto, a mesma conclusão, apesar de ser um fenômeno com grande crescimento generalizado, na comparação sub-regional. As tabelas 5 e 6 abaixo trazem apenas os cortes regionais (UFs e RMs) que mostram diferenças na significância entre estatísticas t corretamente especificadas (levando em conta a dependência temporal na amostra da PNAD) e quando se desconsidera essa correção.

Tabela 5: Proporção de famílias com chefia feminina, erros padrão das estimativas e estatística t, considerando e desconsiderando a covariância implicada pela amostra, segundo Unidades da Federação e Regiões Metropolitanas selecionadas - 2001, 2005 e 2009

Unidades da Federação e Regiões Metropolitanas selecionadas	Proporção de Famílias com Chefia Feminina (Erro-Padrão)			Estatística t (Estatística t viesada)		
	2001	2005	2009	2005 - 2001	2009 - 2001	2009 - 2005
Pará (1)	32,0 (1,1)	34,9 (1,4)	38,2 (1,2)	3,0 (1,6)	4,6 (3,9)	1,9 (1,8)
Tocantins	22,6 (1,4)	25,9 (1,4)	34,3 (1,9)	4,2 (1,7)	5,5 (5,0)	4,4 (3,6)
Maranhão	24,8 (1,7)	29,0 (1,5)	31,9 (2,3)	3,0 (1,9)	5,9 (2,5)	1,6 (1,1)
RM de Fortaleza	34,1 (1,8)	39,1 (1,9)	41,5 (2,1)	2,5 (2,0)	3,4 (2,7)	1,0 (0,8)
Pernambuco	30,0 (0,6)	34,1 (0,8)	36,3 (1,0)	5,0 (4,0)	6,3 (5,3)	2,1 (1,7)
RM de Salvador	35,9 (2,3)	41,9 (2,8)	47,1 (3,0)	1,8 (1,7)	3,9 (3,0)	2,2 (1,3)
RM de Curitiba	28,5 (2,3)	32,5 (2,6)	39,2 (2,3)	1,4 (1,1)	3,7 (3,3)	2,6 (1,9)
Mato Grosso do Sul	24,6 (1,1)	27,8 (1,3)	28,8 (1,5)	2,3 (1,8)	3,0 (2,2)	0,9 (0,5)

Fonte: Pesquisa Nacional por Amostra de Domicílios 2001, 2005 e 2009.

(1) Exclusive população rural.

Nota: Os valores em vermelho indicam inversão na decisão ao se utilizar a estatística t viesada.

Tabela 6: Proporção de famílias com chefia feminina entre as famílias com responsável e cônjuge, erros padrão das estimativas e estatística t, considerando e desconsiderando a covariância implicada pela amostra, segundo Unidades da Federação e Regiões Metropolitanas selecionadas - 2001, 2005 e 2009

Unidades da Federação e Regiões Metropolitanas selecionadas	Proporção de Famílias com Chefia Feminina entre aquelas com Cônjuge (Erro-Padrão)			Estatística t (Estatística t viesada)		
	2001	2005	2009	2005 - 2001	2009 - 2001	2009 - 2005
Acre (1)	5,4 (1,2)	9,7 (1,5)	19,1 (3,6)	1,9 (2,2)	3,9 (3,6)	2,5 (2,4)
Amazonas (1)	5,2 (0,9)	11,1 (3,1)	20,4 (1,9)	2,0 (1,9)	7,7 (7,2)	2,2 (2,5)
RM de Salvador	6,5 (1,0)	18,1 (2,7)	25,8 (3,4)	4,2 (4,0)	5,7 (5,5)	2,5 (1,8)

Fonte: Pesquisa Nacional por Amostra de Domicílios 2001, 2005 e 2009.

(1) Exclusive população rural

Nota: Os valores em vermelho são aqueles onde haveria inversão na decisão ao se utilizar a estatística t viesada.

Para a análise das famílias com chefia feminina, destacam-se as Unidades da Federação e Regiões Metropolitanas do Nordeste, além de duas UFs da região Norte, da RM de Curitiba e do Mato Grosso do Sul. Para os níveis geográficos mencionados, as diferenças entre o princípio e o final da década são sempre significativas. Ainda, quanto às famílias com chefia feminina e com presença de cônjuge, em apenas três localidades foi verificado que a decisão relativa ao aumento na proporção do indicador seria contrária ao se desconsiderar a covariância. Na comparação entre 2005/2001 para o Acre, observou-se uma estatística t menor que a estatística t viesada. Isso ocorreu porque a covariância observada entre os dois instantes para o indicador é negativa, e por isso foi observada uma inversão na decisão, porém no sentido contrário àquele normalmente observado, ou seja, a não rejeição da igualdade na taxa de chefia feminina em famílias com cônjuge, enquanto a análise desconsiderando a covariância indicaria o aumento na taxa (rejeição da igualdade).

Dessa forma, comprova-se empiricamente que a avaliação de mudanças em fenômenos sociais através do uso de indicadores da PNAD deve levar em conta não somente a estrutura amostral complexa da pesquisa, mas também a dependência temporal implicada pelo processo de seleção da amostra. Isso está relacionado ao comportamento das covariâncias temporais entre as estimativas, o que indica o interesse da exploração sistemática dos dados levando em conta esta covariância, ademais dos indicativos de totais e distância entre as porcentagens das estimativas.

5. Conclusões e sugestões

Tendo em vista a importância da PNAD como principal fonte de acompanhamento anual das condições de vida da população, a análise das alterações nos indicadores sociais produzidos a partir desta pesquisa é de grande relevância. Em alguns casos, estas alterações estão relacionadas a mudanças sociais; em outros casos, podem ser frutos de políticas públicas. Contudo, as alterações observadas podem estar relacionadas apenas a flutuações amostrais. Desse fato vem a necessidade de se avaliar estatisticamente estas mudanças e, conforme se verificou na seção anterior, deve-se levar em conta a estrutura dependência presente na amostra da pesquisa ao longo de uma mesma década.

A não utilização da informação da covariância entre as pesquisas pode levar a conclusões incorretas. Na maior parte dos casos sublinhados, valer-se da covariância mostrou que as mudanças nas proporções são significativas, sendo a conclusão contrária no caso da utilização da estatística “viesada”. Pela observação dos indicadores de taxa de analfabetismo e chefia feminina, tem-se que a aplicação da técnica é tão mais relevante quando os contingentes são menores e as diferenças ano a ano são menos relevantes. Chega-se à mesma conclusão com a observação de outros indicadores: média de anos de estudo e proporção entre os jovens de 18 a 24 anos com 11 anos ou mais de estudo (dados completos apresentados no Anexo 2).

Este trabalho motivou a discussão quanto à forma usual de disponibilizar as informações de estrato e UPA na PNAD. Com as variáveis desidentificadas disponibilizadas com o mesmo código para anos distintos dentro da mesma década, estudos similares a este poderão ser realizados pelos usuários externos, possibilitando assim a avaliação nas mudanças socioeconômicas de forma comparativa ao longo dos anos.

Para fornecer um bom serviço aos usuários externos é essencial que o IBGE disponibilize a informação de forma completa e também organize treinamentos, com atenção às ferramentas utilizadas fora do IBGE por pesquisadores da academia e ligados a órgãos públicos. Por exemplo, é possível fazer análises levando em conta o plano amostral complexo da PNAD com o módulo *Complex Samples Procedure* do software SPSS ou com a biblioteca *Survey* do R (software livre).

O presente estudo pode ser aprofundado também no contexto de modelos de regressão, por exemplo, utilizando o ano e suas interações com outras variáveis para explicar fenômenos socioeconômicos. Nesse caso, incorporar ao modelo a informação de que as unidades elementares de análise (correspondentes a anos distintos da pesquisa) são oriundas de um mesmo estrato e conglomerado poderá alterar as estimativas de efeito das preditoras e, principalmente, as estimativas de variabilidade (e assim, de significância) desses efeitos.

6. Bibliografia

Bianchini Z.M. e Albieri S. 1999. **Uma revisão dos principais aspectos dos planos amostrais das pesquisas domiciliares realizadas pelo IBGE**. Revista Brasileira de Estatística 60(213): 7-23.

IBGE. Pesquisa Nacional por Amostra de Domicílios. **Síntese de Indicadores Sociais**. Rio de Janeiro: IBGE, 2001, 2005 e 2009. Disponível em <<http://www.ibge.br>>. Acessado em Out 2012.

Lila, M.F. e Freitas, M.P.S. 2007. **Estimação de intervalos de confiança para estimadores de diferenças temporais na Pesquisa Mensal de Emprego**. Texto para Discussão n.22. Rio de Janeiro: IBGE.

Lumley, T. 2009. **Survey: analysis of complex survey samples**. R package version 3.14. New Jersey: John Wiley & Sons.

Lumley, T. 2010. **Complex Surveys: a Guide to Analysis Using R**. Hoboken. New Jersey: John Wiley & Sons.

Pessoa, D.G.C. e Silva, P.L.N. 1998. **Análise de Dados Amostrais Complexos**, São Paulo: Associação Brasileira de Estatística.

OLIVEIRA, Sônia et al, 2002. **Dimensões Preliminares da Responsabilidade Feminina pelos Domicílios: Um Estudo do Fenômeno a partir dos Censos Demográficos de 1991 e 2000**. Rio de Janeiro, IBGE.

RIBEIRO, Rosa, SABOIA, Ana Lúcia, **Família nas Pesquisas Domiciliares: questões e propostas alternativas**, Texto para Discussão nº 30, DPE, IBGE, 2008.

SABOIA, Ana Lucia, SOARES, Cristiane, 2004. **O Conceito de Chefia nas Pesquisas Domiciliares através do Recorte por Sexo e Presença do Cônjuge – Uma Contribuição à discussão da “Feminização da Pobreza”**. In: Gênero, Revista do Núcleo Transdisciplinar de Estudos de Gênero (NUTEG) nº 2, V. 4. Niterói, Universidade Federal Fluminense (UFF), 1º semestre.

Särndal C.E., Swensson B. e Wretman J.H. 1992. **Model Assisted Survey Sampling**. Nova Iorque: Springer-Verlag.

Shah B.V., Barnwell B.G., Hunt P.N. e LaVange L.M. 1992. **SUDAAN User Manual - professional software for survey data analysis for multi-stage sample designs - release 6.0**. Research Triangle Park, NC: Research Triangle Institute.

Silva, P. L. N., D. G. C. Pessoa e M. F. Lila. 2002. **Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral**. *Ciência & Saúde Coletiva*, 7(4): 659-670.

Anexo 1. Exploração das Variáveis de Estrato e UPA da PNAD

As variáveis de estrato e UPA de cada unidade elementar de interesse nos arquivos de microdados da PNAD são as variáveis v4617 e v4618, respectivamente. Essas estão desidentificadas nos arquivos divulgados na página do IBGE.

Para a estimação das variâncias e covariâncias dos indicadores é necessário calcular quantidades linearizadas ao nível do conglomerado primário. Ao fazer um merge entre dois anos para quantidades ao nível do conglomerado primário, há o risco de se parear conglomerados diferentes, devido ao processo de desidentificação das variáveis.

Foram então utilizadas as variáveis originais de estrato e conglomerado primário, construídas a partir das informações do município, distrito, subdistrito, setor, estrato do município e pseudo-município de cada unidade elementar de interesse, presentes nos arquivos de controle da PNAD. Criou-se também uma variável auxiliar, que assume valor 2 para as áreas rurais das Unidades da Federação Rondônia, Acre, Amazonas, Roraima, Pará e Amapá (não pesquisadas até o ano de 2003), e assume o valor 1 para as demais áreas pesquisadas do país.

As variáveis de estrato e UPA foram geradas de acordo com o seguinte algoritmo:

(situação = 1) se unidade de interesse pertence à área urbana, e (situação = 2) se pertence à área rural;

(nc = 1) se unidade de interesse pertence à amostra normal, e (nc = 2) se pertence ao cadastro de novas construções;

(variável auxiliar = situação) se (uf < 17), e (variável auxiliar = 1) caso (uf >= 17)

Unidade de interesse pertence à região metropolitana (RM) ou município auto-representativo:

Se (1 <= v4107 <= 2) então faça:

Se (nc = 1) então faça:

Se pseudo-Município = 0 então estrato = (uf*100000000) + município;

Se pseudo-município ≠ 0 então estrato = (uf*100000000) + (pseudo-município * 100000);

UPA = (município*100000000) + (distrito*1000000) + (subdistrito*10000) + setor;

Finaliza;

Se (nc = 2) então faça:

Se pseudo-município = 0 então estrato = (uf * 100000000) + 60000000 + município;

Se pseudo-município ≠ 0 então estrato = (uf * 100000000) + 60000000 + (pseudo-município * 100000);

UPA = v0102 * 1000 + v0103;

Finaliza;

Finaliza;

Unidade de interesse pertence a município não auto-representativo:

Se (v4107 = 3) então faça:

Se (nc = 1) então estrato = v4602 * 100000000;

Se (nc = 2) então estrato = v4602 * 100000000 + 60000000;

UPA = município;

Finaliza;

As variáveis de estrato, UPA e a variável auxiliar derivadas a partir do algoritmo acima identificam unicamente os estratos e conglomerados primários. Para os anos de 2001, 2005 e 2009, as quantidades de estratos e unidades primárias de amostragem podem ser vistas no quadro 1.

Quadro 1: Total de Estratos e UPAs obtidos a partir do algoritmo

Ano	Estratos	UPAs
2001	545	5120
2005	945	6964
2009	1022	8543

Apesar de o número de estratos e UPAs aumentar bastante ao longo da década, a dependência não pode ser descartada, pois a maior parte dos domicílios que compõem a amostra de 2005 e 2009 pertencem a UPAs que estavam presentes no início da década. Por exemplo, 5.096 unidades primárias de amostragem da amostra de 2001 permaneceram na amostra até 2005, sendo que 137.071 dos 142.471 domicílios da amostra de 2005 pertencem a algum desses conglomerados. Observou-se ainda que das 5.120 UPAs presentes em 2001, 5.084 permaneceram na amostra até 2009, sendo que 145.700 dos 153.837 domicílios da amostra em 2009 pertencem a estas UPAs, confirmando assim a estrutura de dependência ao longo da década.

Anexo 2. Tabelas de indicadores sociais da PNAD

Tabela 7: Taxa de Analfabetismo para pessoas de 15 anos ou mais, erros padrão das estimativas e estatística t, considerando e desconsiderando a covariância implicada pela amostra, segundo Grandes Regiões, Unidades da Federação e Regiões Metropolitanas - 2001, 2005 e 2009

Grandes Regiões, Unidades da Federação e Regiões Metropolitanas	Taxa de Analfabetismo (Erro-Padrão)			Estatística t (Estatística t viesada)		
	2001	2005	2009	2005 - 2001	2009 - 2001	2009 - 2005
Brasil	12,4 (0,2)	11,0 (0,2)	9,6 (0,1)	-11,5 (-6,5)	-22,7 (-14,4)	-13,1 (-7,3)
Norte	11,2 (0,5)	9,4 (0,5)	8,8 (0,5)	-5,8 (-2,7)	-6,5 (-3,5)	-2,4 (-0,8)
Rondônia	10,2 (0,7)	8,7 (1,1)	7,8 (0,8)	-1,3 (-1,2)	-2,5 (-2,2)	-0,7 (-0,7)
Acre	16,7 (1,4)	13,8 (1,1)	12,7 (1,1)	-2,5 (-1,7)	-4,9 (-2,3)	-1,2 (-0,7)
Amazonas	7,7 (1,3)	5,2 (1,1)	6,0 (0,9)	-4,3 (-1,5)	-2,5 (-1,1)	2,5 (0,6)
Roraima	11,5 (1,8)	11,1 (1,2)	4,8 (0,7)	-0,2 (-0,2)	-3,9 (-3,6)	-5,7 (-4,6)
Pará	11,2 (0,7)	9,6 (0,9)	9,9 (0,9)	-3,2 (-1,3)	-2,1 (-1,1)	0,7 (0,2)
RM de Belém	4,2 (1,6)	4,2 (2,2)	4,2 (2,5)	0,0 (0,0)	0,0 (0,0)	0,0 (0,0)
Amapá	7,2 (0,9)	6,8 (0,9)	2,7 (0,4)	-0,2 (-0,2)	-4,4 (-4,4)	-4,5 (-4,2)
Tocantins	18,5 (1,6)	16,3 (1,0)	13,5 (0,9)	-2,2 (-1,2)	-4,8 (-2,7)	-3,0 (-2,1)
Nordeste	24,2 (0,4)	21,9 (0,5)	18,7 (0,4)	-6,9 (-3,7)	-15,6 (-9,7)	-11,1 (-5,5)
Maranhão	23,4 (1,5)	23,0 (2,4)	19,1 (1,6)	-0,2 (-0,1)	-2,8 (-2,0)	-2,8 (-1,3)
Piauí	29,4 (3,7)	27,4 (3,2)	23,4 (2,3)	-1,4 (-0,4)	-3,1 (-1,4)	-2,5 (-1,0)
Ceará	24,8 (1,1)	22,6 (0,9)	18,6 (0,8)	-3,2 (-1,5)	-9,4 (-4,6)	-7,1 (-3,3)
RM de Fortaleza	12,5 (2,6)	12,0 (2,1)	9,6 (1,9)	-0,3 (-0,2)	-1,9 (-0,9)	-1,7 (-0,9)
Rio Grande do Norte	24,2 (0,9)	21,5 (0,8)	18,1 (0,6)	-3,0 (-2,3)	-5,5 (-5,5)	-4,1 (-3,4)
Paraíba	27,2 (2,0)	25,2 (2,0)	21,6 (2,0)	-1,2 (-0,7)	-2,3 (-1,9)	-2,8 (-1,3)
Pernambuco	22,0 (0,5)	20,5 (0,6)	17,6 (0,7)	-3,6 (-1,9)	-8,2 (-5,2)	-4,7 (-3,4)
RM de Recife	11,0 (1,2)	9,6 (1,3)	8,2 (1,4)	-1,4 (-0,8)	-2,4 (-1,5)	-1,0 (-0,8)
Alagoas	30,6 (1,9)	29,3 (1,1)	24,6 (1,0)	-0,6 (-0,6)	-3,6 (-2,8)	-3,9 (-3,2)
Sergipe	21,4 (1,8)	19,7 (2,0)	16,3 (1,5)	-3,0 (-0,7)	-8,7 (-2,2)	-4,6 (-1,4)
Bahia	22,8 (0,7)	18,7 (0,7)	16,7 (0,5)	-8,7 (-4,2)	-12,2 (-6,7)	-4,5 (-2,3)
RM de Salvador	6,7 (2,9)	6,1 (2,5)	5,4 (2,0)	-0,3 (-0,2)	-0,7 (-0,4)	-0,4 (-0,2)
Sudeste	7,5 (0,2)	6,6 (0,2)	5,7 (0,1)	-6,3 (-4,4)	-12,8 (-9,1)	-6,4 (-4,6)
Minas Gerais	11,7 (0,4)	10,1 (0,4)	8,5 (0,3)	-4,2 (-2,8)	-9,4 (-5,8)	-4,4 (-2,9)
RM de Belo Horizonte	5,7 (1,7)	5,3 (1,6)	4,7 (1,3)	-0,3 (-0,2)	-0,8 (-0,5)	-0,4 (-0,3)
Espírito Santo	11,5 (0,9)	8,7 (0,7)	8,5 (0,6)	-3,8 (-2,5)	-3,5 (-2,8)	-0,3 (-0,2)
Rio de Janeiro	5,6 (0,2)	4,9 (0,2)	4,0 (0,2)	-2,9 (-2,3)	-6,7 (-5,5)	-3,9 (-3,0)
RM do Rio de Janeiro	4,9 (0,3)	3,9 (0,3)	3,2 (0,3)	-2,9 (-2,3)	-4,9 (-4,0)	-2,0 (-1,6)
São Paulo	6,0 (0,2)	5,4 (0,2)	4,8 (0,2)	-3,0 (-2,2)	-6,3 (-4,8)	-3,5 (-2,6)
RM de São Paulo	5,2 (0,4)	4,6 (0,4)	4,2 (0,4)	-1,6 (-1,1)	-2,4 (-1,8)	-0,9 (-0,7)
Sul	7,1 (0,3)	5,9 (0,3)	5,5 (0,2)	-5,6 (-3,2)	-7,3 (-4,5)	-2,3 (-1,4)
Paraná	8,7 (0,6)	7,1 (0,4)	6,7 (0,5)	-3,7 (-2,1)	-5,1 (-2,7)	-1,3 (-0,7)
RM de Curitiba	4,4 (2,0)	3,8 (1,4)	3,5 (1,5)	-0,4 (-0,2)	-0,7 (-0,4)	-0,3 (-0,2)
Santa Catarina	5,9 (0,5)	5,2 (0,7)	4,9 (0,4)	-1,7 (-1,0)	-2,3 (-1,7)	-0,5 (-0,3)
Rio Grande do Sul	6,3 (0,3)	5,2 (0,3)	4,6 (0,4)	-4,6 (-2,5)	-4,8 (-3,6)	-2,3 (-1,4)
RM de Porto Alegre	4,2 (0,8)	3,5 (0,8)	3,0 (1,0)	-1,2 (-0,7)	-1,2 (-1,0)	-0,6 (-0,4)
Centro-Oeste	10,2 (0,4)	8,9 (0,3)	8,0 (0,3)	-4,9 (-2,9)	-7,7 (-5,1)	-3,5 (-2,4)
Mato Grosso do Sul	10,3 (0,7)	9,1 (0,5)	8,7 (0,5)	-2,1 (-1,4)	-2,3 (-1,9)	-0,8 (-0,5)
Mato Grosso	11,2 (1,2)	9,8 (0,9)	10,2 (0,8)	-1,9 (-1,0)	-1,5 (-0,8)	0,7 (0,4)
Goiás	11,7 (0,5)	10,2 (0,4)	8,6 (0,4)	-3,6 (-2,2)	-6,4 (-4,9)	-3,3 (-2,7)
Distrito Federal	5,5 (0,4)	4,7 (0,3)	3,5 (0,3)	-2,1 (-1,6)	-5,1 (-4,0)	-3,6 (-2,8)

Fonte: Pesquisa Nacional por Amostra de Domicílios 2001, 2005 e 2009.

Nota: (1) Exclusiva população rural de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá nos três anos de estudo. (2) Os valores em vermelho são aqueles onde haveria inversão na decisão ao se utilizar a estatística t viesada.

Tabela 8: Proporção de famílias com chefia feminina, erros padrão das estimativas e estatística t, considerando e desconsiderando a covariância implicada pela amostra, segundo Grandes Regiões, Unidades da Federação e Regiões Metropolitanas - 2001, 2005 e 2009

Grandes Regiões, Unidades da Federação e Regiões Metropolitanas	Proporção de Famílias com Chefia Feminina (Erro-Padrão)			Estatística t (Estatística t viesada)		
	2001	2005	2009	2005 - 2001	2009 - 2001	2009 - 2005
Brasil	27,3 (0,2)	30,8 (0,2)	35,2 (0,2)	15,6 (13,0)	30,7 (26,2)	17,7 (15,0)
Norte	29,1 (0,6)	33,5 (0,9)	38,7 (0,7)	6,0 (3,9)	11,1 (10,1)	4,5 (4,4)
Rondônia	24,4 (1,1)	32,5 (1,5)	37,2 (1,7)	4,7 (4,5)	6,8 (6,3)	2,3 (2,0)
Acre	32,4 (1,6)	37,4 (1,7)	39,2 (3,0)	2,1 (2,2)	2,0 (2,0)	0,5 (0,5)
Amazonas	30,0 (1,3)	33,2 (2,8)	41,4 (1,7)	1,4 (1,0)	5,2 (5,3)	2,4 (2,5)
Roraima	29,2 (2,6)	44,4 (2,3)	42,1 (2,3)	4,6 (4,3)	3,5 (3,7)	-0,8 (-0,7)
Pará	32,0 (1,1)	34,9 (1,4)	38,2 (1,2)	3,0 (1,6)	4,6 (3,9)	1,9 (1,8)
RM de Belém	40,3 (2,3)	42,1 (3,6)	44,7 (3,1)	0,7 (0,4)	1,3 (1,1)	0,6 (0,6)
Amapá	21,3 (2,1)	30,4 (2,2)	41,1 (2,3)	3,1 (2,9)	7,1 (6,4)	4,7 (3,4)
Tocantins	22,6 (1,4)	25,9 (1,4)	34,3 (1,9)	4,2 (1,7)	5,5 (5,0)	4,4 (3,6)
Nordeste	27,8 (0,3)	31,3 (0,4)	34,8 (0,5)	9,0 (7,0)	17,8 (12,4)	8,4 (6,0)
Maranhão	24,8 (1,7)	29,0 (1,5)	31,9 (2,3)	3,0 (1,9)	5,9 (2,5)	1,6 (1,1)
Piauí	27,8 (1,7)	28,4 (1,8)	30,4 (1,8)	0,4 (0,2)	1,7 (1,1)	1,3 (0,8)
Ceará	28,5 (0,8)	31,0 (0,8)	35,6 (0,9)	2,9 (2,3)	7,6 (6,1)	4,5 (3,8)
RM de Fortaleza	34,1 (1,8)	39,1 (1,9)	41,5 (2,1)	2,5 (2,0)	3,4 (2,7)	1,0 (0,8)
Rio Grande do Norte	25,2 (1,2)	29,2 (1,3)	35,9 (1,5)	2,1 (2,3)	6,3 (5,4)	4,1 (3,3)
Paraíba	27,5 (1,0)	31,5 (1,2)	32,4 (1,7)	3,5 (2,6)	3,1 (2,6)	0,7 (0,4)
Pernambuco	30,0 (0,6)	34,1 (0,8)	36,3 (1,0)	5,0 (4,0)	6,3 (5,3)	2,1 (1,7)
RM de Recife	34,9 (1,4)	42,1 (1,7)	45,2 (2,4)	4,2 (3,2)	4,5 (3,8)	1,3 (1,1)
Alagoas	25,3 (1,6)	29,7 (1,5)	32,6 (1,5)	2,6 (2,0)	4,4 (3,3)	1,8 (1,3)
Sergipe	31,6 (1,4)	34,1 (1,6)	39,7 (1,3)	1,9 (1,2)	4,4 (4,2)	2,7 (2,7)
Bahia	27,7 (0,6)	31,8 (0,7)	35,9 (0,8)	4,6 (4,3)	10,9 (8,4)	6,2 (3,9)
RM de Salvador	35,9 (2,3)	41,9 (2,8)	47,1 (3,0)	1,8 (1,7)	3,9 (3,0)	2,2 (1,3)
Sudeste	28,0 (0,3)	31,4 (0,3)	36,0 (0,4)	9,0 (8,0)	17,5 (16,2)	10,2 (9,5)
Minas Gerais	27,7 (0,5)	29,5 (0,6)	35,2 (0,6)	2,8 (2,3)	10,0 (9,5)	8,1 (7,0)
RM de Belo Horizonte	33,0 (2,1)	35,3 (2,2)	43,1 (2,3)	0,9 (0,7)	3,4 (3,2)	2,9 (2,5)
Espírito Santo	26,1 (1,2)	30,3 (1,1)	33,0 (1,5)	3,5 (2,6)	3,8 (3,6)	1,5 (1,5)
Rio de Janeiro	31,3 (0,5)	34,4 (0,6)	37,0 (0,6)	4,2 (4,0)	7,4 (6,9)	3,4 (3,1)
RM do Rio de Janeiro	32,6 (0,7)	35,7 (0,8)	39,2 (0,9)	3,1 (2,9)	6,2 (5,8)	3,3 (3,0)
São Paulo	27,0 (0,5)	31,2 (0,4)	36,2 (0,7)	7,0 (6,3)	12,4 (11,3)	6,7 (6,4)
RM de São Paulo	30,9 (1,0)	35,4 (0,9)	41,5 (1,4)	3,6 (3,2)	6,7 (6,1)	3,9 (3,7)
Sul	24,4 (0,4)	27,6 (0,5)	33,0 (0,6)	6,2 (4,9)	14,4 (11,8)	10,0 (7,3)
Paraná	24,9 (0,7)	27,5 (0,8)	31,6 (0,7)	3,0 (2,5)	7,5 (6,7)	4,9 (3,8)
RM de Curitiba	28,5 (2,3)	32,5 (2,6)	39,2 (2,3)	1,4 (1,1)	3,7 (3,3)	2,6 (1,9)
Santa Catarina	20,1 (0,9)	24,5 (1,3)	29,0 (1,4)	4,1 (2,8)	6,7 (5,4)	3,6 (2,4)
Rio Grande do Sul	26,2 (0,7)	29,2 (0,6)	36,4 (1,0)	4,1 (3,1)	10,5 (8,2)	8,3 (6,2)
RM de Porto Alegre	31,6 (2,0)	34,7 (1,6)	44,2 (2,6)	1,6 (1,3)	4,9 (3,8)	4,1 (3,1)
Centro-Oeste	26,7 (0,6)	30,3 (0,5)	36,2 (0,6)	6,0 (4,7)	13,6 (11,5)	9,3 (7,2)
Mato Grosso do Sul	24,6 (1,1)	27,8 (1,3)	28,8 (1,5)	2,3 (1,8)	3,0 (2,2)	0,9 (0,5)
Mato Grosso	19,1 (1,2)	26,0 (1,3)	36,2 (1,2)	4,6 (3,9)	10,1 (10,2)	6,6 (5,7)
Goiás	27,7 (0,9)	30,2 (0,8)	36,0 (1,1)	2,9 (2,1)	7,4 (5,9)	5,6 (4,3)
Distrito Federal	35,2 (1,2)	37,9 (1,2)	43,7 (1,0)	1,9 (1,6)	6,6 (5,5)	4,6 (3,8)

Fonte: Pesquisa Nacional por Amostra de Domicílios 2001, 2005 e 2009.

Nota: (1) Exclusiva população rural de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá nos três anos de estudo. (2) Os valores em vermelho são aqueles onde haveria inversão na decisão ao se utilizar a estatística t viesada.

Tabela 9: Proporção de famílias com chefia feminina entre as famílias com responsável e cônjuge, erros padrão das estimativas e estatística t, considerando e desconsiderando a covariância implicada pela amostra, segundo Grandes Regiões, Unidades da Federação e Regiões Metropolitanas - 2001, 2005 e 2009

Grandes Regiões, Unidades da Federação e Regiões Metropolitanas	Proporção de Famílias com Chefia Feminina entre aquelas com Cônjuge (Erro-Padrão)			Estatística t (Estatística t viesada)		
	2001	2005	2009	2005 - 2001	2009 - 2001	2009 - 2005
Brasil	3,6 (0,1)	7,2 (0,2)	14,3 (0,3)	20,7 (20,6)	38,6 (38,0)	24,9 (23,6)
Norte	5,8 (0,3)	10,0 (0,9)	18,0 (0,9)	4,6 (4,4)	13,1 (12,7)	6,1 (6,3)
Rondônia	2,3 (0,8)	11,1 (1,2)	19,5 (1,8)	5,4 (6,0)	7,9 (8,8)	5,7 (3,9)
Acre	5,4 (1,2)	9,7 (1,5)	19,1 (3,6)	1,9 (2,2)	3,9 (3,6)	2,5 (2,4)
Amazonas	5,2 (0,9)	11,1 (3,1)	20,4 (1,9)	2,0 (1,9)	7,7 (7,2)	2,2 (2,5)
Roraima	8,5 (1,6)	26,1 (2,6)	22,5 (2,6)	6,4 (5,9)	4,3 (4,7)	-1,1 (-1,0)
Pará	8,0 (0,5)	9,7 (1,1)	15,6 (1,6)	1,4 (1,4)	4,6 (4,5)	3,2 (3,1)
RM de Belém	17,4 (1,3)	16,6 (3,1)	21,8 (4,4)	-0,3 (-0,3)	1,0 (1,0)	1,0 (1,0)
Amapá	5,9 (1,2)	8,1 (1,4)	23,0 (2,1)	1,1 (1,2)	7,6 (7,1)	7,6 (5,9)
Tocantins	2,1 (0,2)	5,2 (0,9)	17,9 (1,9)	3,7 (3,5)	8,2 (8,2)	5,8 (6,0)
Nordeste	2,9 (0,2)	7,0 (0,3)	11,9 (0,4)	14,2 (14,2)	23,3 (21,2)	11,5 (10,2)
Maranhão	1,5 (0,7)	4,6 (0,6)	8,9 (1,7)	3,0 (3,4)	5,6 (3,9)	2,1 (2,3)
Piauí	1,9 (0,5)	5,0 (1,0)	5,4 (0,8)	3,0 (2,8)	4,1 (3,9)	0,4 (0,3)
Ceará	3,0 (0,3)	7,9 (0,6)	12,2 (0,8)	7,7 (7,7)	11,5 (11,1)	5,8 (4,5)
RM de Fortaleza	6,9 (0,7)	15,9 (1,5)	17,2 (2,0)	5,6 (5,5)	5,2 (4,9)	0,7 (0,5)
Rio Grande do Norte	2,4 (0,6)	5,3 (0,9)	15,6 (2,0)	2,6 (2,6)	6,4 (6,4)	5,9 (4,8)
Paraíba	2,7 (0,4)	6,6 (1,0)	9,6 (1,7)	3,8 (3,7)	4,0 (3,9)	1,8 (1,6)
Pernambuco	4,2 (0,3)	7,8 (0,5)	11,5 (0,8)	6,4 (6,5)	9,2 (8,6)	4,3 (4,1)
RM de Recife	7,2 (0,8)	14,2 (1,1)	19,7 (2,1)	5,5 (5,3)	6,2 (5,6)	2,6 (2,3)
Alagoas	2,4 (0,9)	6,7 (1,1)	9,8 (1,0)	3,7 (3,0)	6,0 (5,5)	2,1 (2,1)
Sergipe	6,1 (1,2)	10,6 (1,4)	17,3 (2,2)	2,2 (2,4)	4,3 (4,4)	2,8 (2,5)
Bahia	2,5 (0,2)	7,7 (0,6)	14,1 (0,8)	8,1 (8,0)	15,2 (14,8)	8,7 (6,6)
RM de Salvador	6,5 (1,0)	18,1 (2,7)	25,8 (3,4)	4,2 (4,0)	5,7 (5,5)	2,5 (1,8)
Sudeste	3,7 (0,2)	6,9 (0,2)	14,9 (0,5)	11,2 (11,2)	21,5 (21,8)	15,8 (15,1)
Minas Gerais	2,8 (0,2)	4,1 (0,3)	12,6 (0,7)	3,4 (3,3)	12,5 (12,7)	11,1 (10,7)
RM de Belo Horizonte	6,7 (1,0)	9,4 (1,3)	21,1 (3,0)	1,7 (1,7)	4,5 (4,5)	3,8 (3,6)
Espírito Santo	1,9 (0,4)	6,2 (0,7)	12,0 (1,5)	5,7 (5,3)	6,4 (6,6)	3,6 (3,5)
Rio de Janeiro	2,3 (0,2)	7,8 (0,5)	12,9 (0,7)	10,7 (10,7)	13,9 (14,1)	6,2 (6,0)
RM do Rio de Janeiro	2,7 (0,3)	8,8 (0,6)	15,1 (1,0)	8,7 (8,7)	11,7 (11,9)	5,6 (5,4)
São Paulo	4,7 (0,3)	8,0 (0,4)	17,0 (0,8)	6,6 (6,6)	14,0 (14,2)	10,9 (10,4)
RM de São Paulo	7,8 (0,7)	11,1 (0,8)	22,0 (1,8)	3,1 (3,1)	7,3 (7,4)	5,9 (5,6)
Sul	3,4 (0,2)	6,8 (0,4)	14,5 (0,5)	8,4 (8,3)	20,6 (19,2)	12,7 (12,0)
Paraná	3,6 (0,4)	7,1 (0,7)	12,3 (0,7)	4,6 (4,5)	11,1 (11,5)	6,3 (5,6)
RM de Curitiba	7,5 (1,3)	11,4 (2,4)	20,7 (2,4)	1,5 (1,5)	4,8 (4,9)	3,3 (2,8)
Santa Catarina	2,3 (0,4)	5,0 (0,8)	12,8 (1,3)	3,0 (3,0)	8,1 (7,6)	5,4 (5,1)
Rio Grande do Sul	3,9 (0,3)	7,6 (0,4)	17,7 (0,9)	7,3 (7,1)	16,1 (14,0)	10,0 (9,9)
RM de Porto Alegre	6,4 (0,9)	12,2 (1,2)	25,2 (2,7)	4,0 (3,9)	7,5 (6,6)	4,5 (4,4)
Centro-Oeste	4,4 (0,3)	8,8 (0,5)	16,2 (0,7)	7,8 (7,5)	15,4 (15,3)	9,2 (8,6)
Mato Grosso do Sul	2,0 (0,4)	4,7 (0,6)	6,2 (1,1)	4,0 (3,9)	3,8 (3,7)	1,3 (1,2)
Mato Grosso	2,3 (0,4)	9,8 (1,6)	19,9 (1,4)	4,8 (4,6)	11,8 (11,8)	5,5 (4,7)
Goiás	4,0 (0,5)	7,6 (0,7)	16,0 (1,3)	4,6 (4,4)	8,4 (8,4)	5,8 (5,7)
Distrito Federal	11,1 (1,0)	14,5 (1,0)	22,4 (1,1)	2,4 (2,5)	7,9 (7,6)	5,4 (5,2)

Fonte: Pesquisa Nacional por Amostra de Domicílios 2001, 2005 e 2009.

Nota: (1) Exclusiva população rural de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá nos três anos de estudo. (2) Os valores em vermelho são aqueles onde haveria inversão na decisão ao se utilizar a estatística t viesada.

Tabela 10: Proporção de jovens entre 18 e 24 anos com 11 anos de estudo ou mais, erros padrão das estimativas e estatística t, considerando e desconsiderando a covariância implicada pela amostra, segundo Grandes Regiões, Unidades da Federação e Regiões Metropolitanas - 2001, 2005 e 2009

Grandes Regiões, Unidades da Federação e Regiões Metropolitanas	Proporção de Jovens de 18 a 24 anos com 11 anos ou mais de estudo (Erro-Padrão)			Estatística t (Estatística t viesada)		
	2001	2005	2009	2005 - 2001	2009 - 2001	2009 - 2005
Brasil	33,6 (0,3)	44,7 (0,4)	53,1 (0,4)	29,3 (22,3)	49,9 (38,6)	22,5 (16,1)
Norte	24,4 (0,8)	35,2 (1,1)	46,8 (1,0)	10,6 (7,8)	20,1 (17,0)	10,6 (7,7)
Rondônia	25,2 (2,5)	39,7 (2,9)	53,4 (2,1)	4,5 (3,8)	10,2 (8,7)	5,1 (3,9)
Acre	29,2 (3,5)	40,8 (3,0)	54,0 (2,6)	4,8 (2,5)	7,3 (5,7)	4,5 (3,3)
Amazonas	28,7 (1,8)	40,0 (1,9)	49,1 (2,0)	5,4 (4,3)	8,1 (7,6)	4,4 (3,3)
Roraima	31,5 (4,6)	48,9 (3,8)	61,1 (4,8)	4,2 (2,9)	5,5 (4,5)	3,0 (2,0)
Pará	21,0 (1,3)	29,5 (2,0)	40,9 (1,8)	5,3 (3,6)	11,1 (9,0)	5,8 (4,2)
RM de Belém	28,6 (3,0)	39,2 (5,6)	45,8 (5,2)	2,3 (1,7)	3,5 (2,9)	1,2 (0,9)
Amapá	28,0 (3,3)	46,4 (3,9)	50,7 (4,0)	3,9 (3,6)	4,8 (4,4)	1,4 (0,8)
Tocantins	23,6 (1,6)	33,2 (2,0)	53,6 (2,3)	4,3 (3,7)	10,7 (10,8)	9,7 (6,7)
Nordeste	20,4 (0,6)	29,4 (0,7)	40,4 (0,8)	16,0 (9,9)	30,8 (20,9)	18,9 (10,5)
Maranhão	18,3 (2,2)	26,9 (3,5)	39,4 (3,8)	3,1 (2,1)	7,0 (4,9)	7,4 (2,4)
Piauí	18,4 (3,6)	26,0 (3,1)	38,1 (4,1)	3,6 (1,6)	7,0 (3,6)	3,6 (2,4)
Ceará	24,0 (1,2)	34,9 (1,2)	45,5 (1,6)	10,7 (6,4)	17,2 (10,7)	9,1 (5,2)
RM de Fortaleza	33,7 (2,6)	45,6 (2,6)	53,7 (3,6)	5,0 (3,2)	6,9 (4,5)	3,0 (1,8)
Rio Grande do Norte	23,5 (2,0)	33,3 (2,1)	38,6 (2,5)	4,1 (3,4)	5,3 (4,8)	2,3 (1,6)
Paraíba	17,3 (2,5)	26,7 (3,3)	37,5 (3,1)	4,1 (2,3)	8,9 (5,0)	3,7 (2,4)
Pernambuco	23,0 (1,3)	31,3 (1,1)	41,9 (1,5)	8,7 (4,9)	16,6 (9,7)	8,3 (5,8)
RM de Recife	34,1 (3,3)	43,6 (2,7)	56,4 (3,9)	3,7 (2,2)	7,7 (4,4)	3,8 (2,7)
Alagoas	13,8 (2,0)	21,0 (2,3)	33,4 (2,4)	3,4 (2,4)	8,7 (6,4)	5,5 (3,8)
Sergipe	23,5 (2,6)	32,3 (2,1)	46,0 (3,4)	4,1 (2,6)	7,1 (5,2)	4,6 (3,4)
Bahia	19,4 (0,9)	28,3 (1,2)	39,4 (1,1)	10,2 (6,1)	18,2 (13,9)	10,8 (7,0)
RM de Salvador	33,8 (3,6)	42,9 (4,8)	51,4 (4,5)	2,5 (1,5)	4,2 (3,1)	2,0 (1,3)
Sudeste	42,6 (0,6)	55,1 (0,6)	61,8 (0,6)	18,7 (15,3)	29,7 (23,5)	10,2 (8,1)
Minas Gerais	34,8 (1,0)	45,8 (1,1)	52,4 (1,0)	9,1 (7,4)	17,5 (12,7)	5,3 (4,4)
RM de Belo Horizonte	40,6 (3,6)	56,5 (4,7)	62,5 (3,9)	3,3 (2,7)	5,6 (4,2)	1,2 (1,0)
Espírito Santo	34,0 (2,2)	48,0 (2,2)	53,0 (2,4)	5,4 (4,4)	8,3 (5,8)	2,3 (1,5)
Rio de Janeiro	37,5 (1,2)	50,1 (1,2)	57,6 (1,2)	8,8 (7,5)	13,6 (11,9)	5,0 (4,4)
RM do Rio de Janeiro	40,0 (1,6)	52,7 (1,6)	59,3 (1,7)	6,5 (5,6)	9,6 (8,4)	3,3 (2,8)
São Paulo	48,8 (0,9)	62,0 (0,9)	68,7 (0,9)	13,3 (10,7)	20,4 (15,6)	7,1 (5,3)
RM de São Paulo	50,8 (1,8)	64,2 (1,7)	68,3 (1,8)	6,7 (5,3)	8,8 (6,7)	2,1 (1,6)
Sul	39,3 (0,9)	52,6 (0,9)	58,2 (0,9)	14,7 (10,3)	17,5 (14,7)	5,8 (4,4)
Paraná	41,0 (1,6)	54,4 (1,5)	58,7 (1,4)	9,1 (6,1)	10,5 (8,3)	2,6 (2,1)
RM de Curitiba	44,9 (4,9)	60,4 (4,9)	64,4 (4,4)	3,4 (2,3)	3,7 (3,0)	0,8 (0,6)
Santa Catarina	42,1 (2,4)	58,0 (2,2)	65,4 (1,7)	7,8 (4,9)	10,3 (7,9)	4,0 (2,7)
Rio Grande do Sul	36,1 (1,0)	47,8 (1,2)	53,5 (1,5)	8,8 (7,5)	9,6 (9,5)	3,7 (2,9)
RM de Porto Alegre	40,4 (2,7)	51,2 (3,1)	55,6 (4,0)	3,1 (2,6)	3,2 (3,1)	1,1 (0,9)
Centro-Oeste	31,5 (0,9)	44,4 (0,9)	53,8 (0,8)	12,3 (10,1)	20,3 (17,9)	10,7 (7,7)
Mato Grosso do Sul	35,1 (2,2)	38,3 (2,3)	50,4 (2,1)	1,2 (1,0)	6,4 (5,1)	6,4 (4,0)
Mato Grosso	27,6 (2,5)	38,2 (1,7)	46,5 (1,9)	4,5 (3,5)	7,3 (6,1)	4,7 (3,2)
Goiás	28,1 (1,3)	44,6 (1,4)	52,7 (1,3)	10,2 (8,6)	14,1 (13,4)	5,4 (4,1)
Distrito Federal	40,5 (1,9)	56,3 (1,8)	66,8 (1,6)	8,6 (6,0)	13,7 (10,5)	5,7 (4,3)

Fonte: Pesquisa Nacional por Amostra de Domicílios 2001, 2005 e 2009.

Nota: (1) Exclusive população rural de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá nos três anos de estudo. (2) Os valores em vermelho são aqueles onde haveria inversão na decisão ao se utilizar a estatística t viesada.

Tabela 11: Média de anos de estudo para pessoas de 15 anos ou mais, erros padrão das estimativas e estatística t, considerando e desconsiderando a covariância implicada pela amostra, segundo Grandes Regiões, Unidades da Federação e Regiões Metropolitanas - 2001, 2005 e 2009

Grandes Regiões, Unidades da Federação e Regiões Metropolitanas	Média de Anos de Estudo para pessoas de 15 anos ou mais (Erro-Padrão)			Estatística t (Estatística t viesada)		
	2001	2005	2009	2005 - 2001	2009 - 2001	2009 - 2005
Brasil	6,4 (0,0)	7,0 (0,0)	7,6 (0,0)	40,3 (16,1)	74,6 (31,5)	37,6 (15,4)
Norte	6,3 (0,1)	7,0 (0,1)	7,6 (0,1)	15,4 (7,0)	32,4 (12,9)	12,0 (5,4)
Rondônia	6,1 (0,2)	6,7 (0,2)	7,7 (0,2)	4,8 (2,5)	12,3 (6,3)	7,1 (3,5)
Acre	6,2 (0,3)	7,1 (0,3)	7,9 (0,2)	6,3 (2,4)	12,7 (4,9)	4,9 (2,3)
Amazonas	6,9 (0,2)	7,8 (0,2)	8,1 (0,2)	8,8 (4,0)	12,8 (5,3)	3,1 (1,4)
Roraima	6,0 (0,3)	7,3 (0,2)	8,6 (0,3)	7,2 (3,4)	11,6 (6,4)	6,8 (3,5)
Pará	6,2 (0,1)	6,8 (0,1)	7,3 (0,1)	8,2 (3,4)	19,3 (6,5)	6,0 (2,6)
RM de Belém	7,6 (0,2)	8,0 (0,3)	8,3 (0,3)	2,1 (1,0)	4,9 (1,9)	1,9 (0,8)
Amapá	7,6 (0,2)	7,9 (0,2)	8,3 (0,2)	1,7 (1,2)	3,7 (2,5)	2,0 (1,3)
Tocantins	5,4 (0,2)	6,3 (0,2)	7,3 (0,2)	6,4 (3,2)	14,6 (6,8)	7,4 (3,7)
Nordeste	4,9 (0,1)	5,6 (0,1)	6,3 (0,1)	24,8 (8,0)	43,2 (16,4)	23,3 (8,0)
Maranhão	4,7 (0,3)	5,2 (0,3)	6,2 (0,3)	4,5 (1,3)	9,6 (3,6)	6,0 (2,1)
Piauí	4,4 (0,4)	5,1 (0,4)	5,8 (0,4)	5,2 (1,2)	10,5 (2,3)	5,5 (1,2)
Ceará	5,0 (0,1)	5,8 (0,1)	6,5 (0,1)	11,3 (4,8)	19,6 (9,2)	13,1 (4,5)
RM de Fortaleza	6,7 (0,3)	7,3 (0,3)	8,0 (0,3)	3,5 (1,5)	6,9 (3,3)	5,0 (1,8)
Rio Grande do Norte	5,5 (0,2)	6,0 (0,1)	6,5 (0,1)	5,8 (2,7)	10,0 (4,9)	4,9 (2,5)
Paraíba	4,6 (0,3)	5,4 (0,3)	6,1 (0,3)	6,4 (2,0)	9,7 (3,8)	6,8 (1,6)
Pernambuco	5,4 (0,1)	6,0 (0,1)	6,6 (0,1)	12,1 (5,0)	25,9 (9,4)	10,3 (4,9)
RM de Recife	7,0 (0,2)	7,6 (0,2)	8,3 (0,2)	4,6 (2,0)	11,5 (4,1)	4,8 (2,3)
Alagoas	4,2 (0,2)	4,8 (0,2)	5,7 (0,2)	5,6 (2,5)	15,5 (5,8)	8,8 (3,6)
Sergipe	5,4 (0,2)	5,9 (0,2)	6,7 (0,3)	5,8 (1,7)	12,0 (4,1)	8,3 (2,4)
Bahia	4,9 (0,1)	5,7 (0,1)	6,4 (0,1)	16,7 (5,6)	24,3 (10,6)	12,0 (4,6)
RM de Salvador	7,5 (0,4)	8,0 (0,4)	8,5 (0,3)	2,8 (1,0)	4,1 (1,9)	2,0 (0,8)
Sudeste	7,1 (0,0)	7,7 (0,0)	8,2 (0,0)	21,5 (10,0)	42,0 (19,7)	21,0 (9,9)
Minas Gerais	6,2 (0,1)	6,8 (0,1)	7,4 (0,1)	12,0 (5,4)	25,5 (10,9)	13,2 (6,0)
RM de Belo Horizonte	7,4 (0,3)	8,0 (0,3)	8,5 (0,3)	2,9 (1,3)	5,7 (2,5)	2,7 (1,2)
Espírito Santo	6,4 (0,2)	7,3 (0,2)	7,6 (0,2)	9,3 (3,4)	10,7 (4,5)	2,8 (0,9)
Rio de Janeiro	7,4 (0,1)	7,9 (0,1)	8,4 (0,1)	9,1 (4,8)	18,3 (9,4)	9,7 (4,9)
RM do Rio de Janeiro	7,7 (0,1)	8,3 (0,1)	8,7 (0,1)	7,9 (4,1)	13,9 (7,1)	6,3 (3,1)
São Paulo	7,4 (0,1)	8,0 (0,1)	8,5 (0,1)	14,5 (6,7)	28,1 (13,3)	13,9 (6,6)
RM de São Paulo	7,8 (0,1)	8,3 (0,1)	8,8 (0,1)	6,8 (3,1)	12,7 (6,0)	6,0 (2,9)
Sul	6,8 (0,1)	7,4 (0,1)	7,9 (0,1)	18,5 (6,6)	31,4 (11,7)	15,4 (5,4)
Paraná	6,6 (0,1)	7,3 (0,1)	7,9 (0,1)	11,2 (4,2)	21,3 (7,6)	11,4 (3,8)
RM de Curitiba	7,5 (0,4)	8,3 (0,4)	8,8 (0,4)	3,9 (1,4)	6,4 (2,3)	2,9 (1,0)
Santa Catarina	6,9 (0,2)	7,6 (0,2)	8,2 (0,2)	8,5 (2,7)	13,2 (4,5)	5,5 (1,9)
Rio Grande do Sul	6,8 (0,1)	7,4 (0,1)	7,8 (0,1)	12,1 (4,8)	19,6 (8,9)	10,4 (3,8)
RM de Porto Alegre	7,7 (0,2)	8,3 (0,2)	8,6 (0,2)	4,7 (1,9)	6,3 (2,9)	2,5 (0,9)
Centro-Oeste	6,5 (0,1)	7,2 (0,1)	7,9 (0,1)	17,6 (7,4)	28,3 (14,1)	13,8 (6,6)
Mato Grosso do Sul	6,3 (0,2)	6,9 (0,2)	7,4 (0,2)	5,2 (2,3)	7,8 (4,1)	4,8 (2,1)
Mato Grosso	6,0 (0,2)	6,8 (0,2)	7,4 (0,2)	7,9 (3,3)	13,4 (6,2)	6,6 (2,8)
Goiás	6,1 (0,1)	6,9 (0,1)	7,5 (0,1)	11,9 (5,2)	19,5 (10,4)	8,6 (4,9)
Distrito Federal	8,2 (0,2)	9,0 (0,2)	9,6 (0,2)	9,4 (3,5)	14,7 (6,3)	7,1 (2,8)

Fonte: Pesquisa Nacional por Amostra de Domicílios 2001, 2005 e 2009.

Nota: (1) Exclusiva população rural de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá nos três anos de estudo. (2) Os valores em vermelho são aqueles onde haveria inversão na decisão ao se utilizar a estatística t viesada.

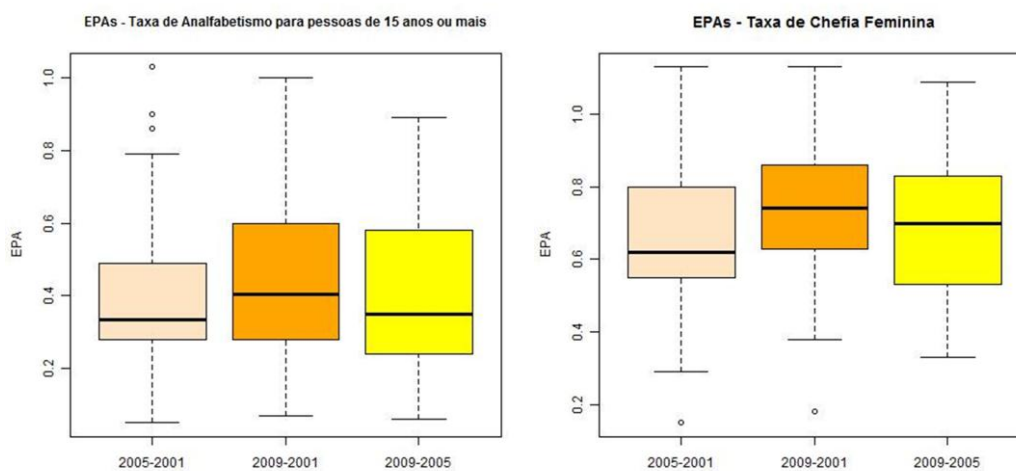
Anexo 3. Box Plots dos Efeitos de Plano Amostral Ampliados (EPAs)

A medida apresentada nos *box plots* a seguir, denominada EPA ampliado, mede o efeito de se desconsiderar a covariância implicada pela dependência presente na amostra da PNAD para uma mesma década na avaliação de diferenças temporais.

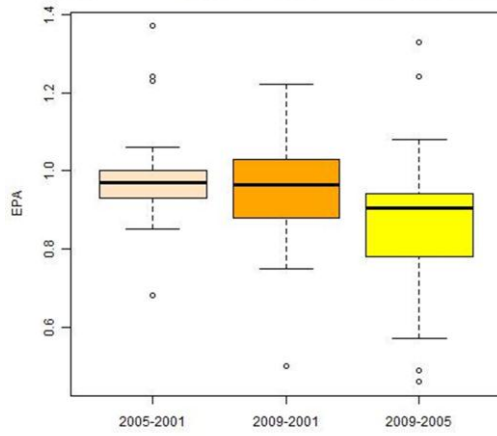
A equação do Efeito de Plano Amostral Ampliado (EPA) é dada por:

$$EPA = \left(\frac{t_v}{t_R} \right)^2$$

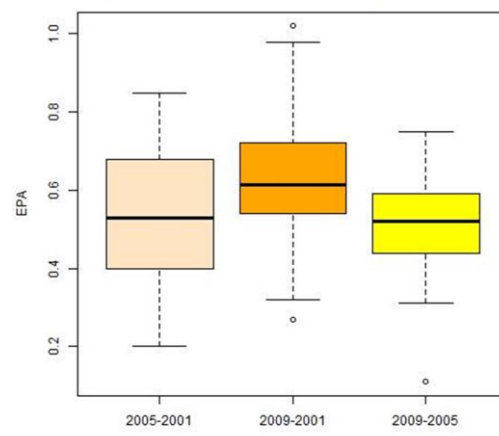
onde t_v é a estatística t desconsiderando a covariância entre dois anos de uma mesma década da PNAD, e t_R é a estatística t corretamente especificada (vide Seção 3.4). A medida foi obtida para todos os níveis geográficos, indicadores apresentados e para cada comparação entre dois instantes, no caso, 2005/2001, 2009/2001 e 2009/2005.



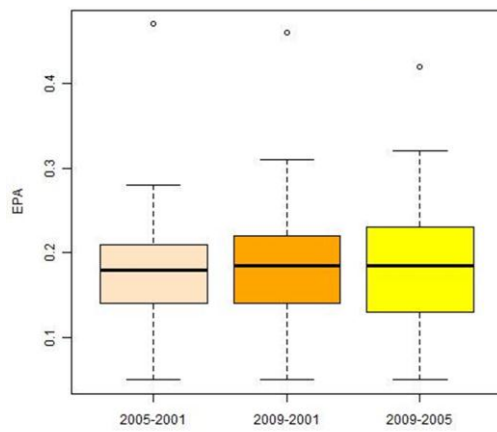
EPAs - Taxa de Chefia Feminina em Famílias com Cônjuge



EPAs - Proporção de Jovens de 18 a 24 anos com ao menos 11 anos de estudo



EPAs - Média de Anos de Estudo para pessoas de 15 anos ou mais



Textos para Discussão já publicados

Antiga série

- 📖 Pesquisas Contínuas da Indústria - Vol. 1, nº 1, janeiro 1988
- 📖 Pesquisas Agropecuárias Contínuas: Metodologia - Vol. I, nº 2, 1988
- 📖 Uma Filosofia de Trabalho: As experiências com o SNIPC e com o SINAPI - Vol. I, nº 3, março 1988
- 📖 O Sigilo das Informações Estatísticas: Idéias para reflexão - Vol. I, nº 4, abril 1988
- 📖 Projeções da População Residente e do Número de Domicílios Particulares Ocupados: 1985-2020 - Vol. I, nº 5, maio 1988
- 📖 Classificação de Atividades e Produtos, Matérias-Primas e Serviços Industriais: Indústria Extrativa Mineral e de Transformação - Vol. 1, nº 6, agosto 1988
- 📖 A Mortalidade Infantil no Brasil nos Anos 80 - Vol. I, nº 7, setembro 1988
- 📖 Principais Características das Pesquisas Econômicas, Sociais e Demográficas - Vol. I, número especial, outubro 1988
- 📖 Ensaio sobre o Produto Real da Agropecuária - Vol. I, nº 9, setembro 1988
- 📖 Novo Sistema de Contas Nacionais, Ano Base 1980 - Resultados Provisórios - Vol. I, nº 10, dezembro 1988
- 📖 Pesquisa de Orçamentos Familiares - Metodologia para Obtenção das Informações de Campo - nº 11, janeiro 1989
- 📖 De Camponesa a Bóia-fria: Transformações do trabalho feminino - nº 12, fevereiro 1989
- 📖 Pesquisas Especiais do Departamento de Agropecuária - Metodologia e Resultados - nº 13, fevereiro 1989
- 📖 Brasil - Matriz de Insumo-Produto - 1980 - nº 14, maio 1989
- 📖 As Informações sobre Fecundidade, Mortalidade e Anticoncepção nas PNADs - nº 15, maio 1989
- 📖 As Estatísticas Agropecuárias e a III Conferência Nacional de Estatística - nº 16, junho 1989
- 📖 Brasil - Sistema de Contas Nacionais Consolidadas - nº 17, agosto 1989
- 📖 Brasil - Produto Interno Bruto Real Trimestral - Metodologia - nº 18, agosto 1989
- 📖 Estatísticas e Indicadores Sociais para a Década de 90 - nº 19, setembro 1989
- 📖 Uma Análise do Cotidiano da Pesquisa no DEREN (As Estatísticas do Trabalho) - nº 20, outubro 1989
- 📖 Coordenação Estatística Nacional - Reflexões sobre o caso Brasileiro - nº 21, novembro 1989
- 📖 Pesquisa Industrial Anual 1982/84 - Análise dos Resultados - nº 22, novembro 1989
- 📖 O Departamento de Comércio e Serviços e a III Conferência Nacional de Estatística - nº 23, dezembro 1989
- 📖 Um projeto de Integração para as Estatísticas Industriais - nº 24, dezembro 1989
- 📖 Cadastro de Informantes de Pesquisas Econômicas - nº 25, janeiro 1990
- 📖 Ensaio sobre a Produção de Estatística - nº 26, janeiro 1990
- 📖 O Espaço das Pequenas Unidades Produtivas: Uma tentativa de delimitação - nº 27, fevereiro 1990
- 📖 Uma Nova Metodologia para Correção Automática no Censo Demográfico Brasileiro: Experimentação e primeiros resultados - nº 28, fevereiro 1990
- 📖 Notas Técnicas sobre o Planejamento de Testes e Pesquisas Experimentais - nº 29, março 1990

- 📖 Estatísticas, Estudos e Análises Demográficas - Uma visão do Departamento de População - **nº 30**, abril 1990
- 📖 Crítica de Equações de Fechamento de Empresas no Censo Econômico de 1985 - **nº 31**, maio 1990
- 📖 Efeito de Conglomeração da Malha Setorial do Censo Demográfico de 1980 - **nº 32**, maio 1990
- 📖 A Redução da Amostra e a Utilização de Duas Frações Amostrais no Censo Demográfico de 1990 - **nº 33**, junho 1990
- 📖 Estudos e Pesquisas de Avaliação de Censos Demográficos - 1970 a 1990 - **nº 34**, julho 1990
- 📖 A Influência da Migração no Mercado de Trabalho das Capitais do Centro-Oeste - 1980 - **nº 35**, agosto 1990
- 📖 Pesquisas de Conjuntura: Discussão sobre Variáveis a Investigar - **nº 36**, setembro 1990
- 📖 Um Modelo para Estimar o Nível e o Padrão da Fecundidade por Idade com Base em Parturições Observadas - **nº 37**, outubro 1990
- 📖 A Estrutura Operacional de Uma Pesquisa por Amostra - **nº 38**, novembro 1990
- 📖 Produção Agrícola, Agroindustrial e de Máquinas e Insumos Agrícolas no Anos 80: Novas Evidências Estatísticas - **nº 39**, dezembro 1990
- 📖 A Inflação Medida pelo Índice de Preços ao Consumidor - **nº 40**, janeiro 1991
- 📖 A Participação Política Eleitoral no Brasil - 1988, Análise Preliminar - **nº 41**, fevereiro 1991
- 📖 Ensaio sobre Estatísticas do Setor Produtivo - **nº 42**, março 1991
- 📖 A Produção Integrada de Estatística e as Contas Nacionais: Agenda para Formulação de um Novo Plano Geral de Informações Estatísticas e Geográficas - **nº 43**, março 1991
- 📖 Matriz de Fluxos Migratórios Intermunicipais - Brasil - 1980 - **nº 44**, abril 1991
- 📖 Fluxos Migratórios Intrametropolitanos - Brasil - 1970-1980 - **nº 45**, abril 1991
- 📖 A Revisão da PNAD, A Questão Conceitual e Relatório das Contribuições - **nº 46**, maio 1991
- 📖 A Dimensão Ambiental no Sistema de Contas Nacionais - **nº 47**, maio 1991
- 📖 Estrutura das Contas Nacionais Brasileiras - **nº 48**, junho 1991
- 📖 Mercado do Couro e Resultados da Pesquisa Anual do Couro - **nº 49**, junho 1991
- 📖 As Estatísticas e o Meio Ambiente - **nº 50**, julho 1991
- 📖 Novo Sistema de Contas Nacionais Séries Correntes: 1981-85 Metodologia, Resultados Provisórios e Avaliação do Projeto - **nº 51**, julho 1991 (2 Volumes: Volume 1 - Metodologia, Resultados Provisórios e Avaliação do Projeto; Volume 2-Tabelas)
- 📖 O Censo Industrial de 1985 - Balanço da Experiência - **nº 52**, agosto 1991
- 📖 Análise da Inflação Medida Pelo INPC 1989 - **nº 53**, agosto 1991
- 📖 Revisão da PNAD: A Questão Amostral: Módulo II do Anteprojeto - **nº 54**, setembro 1991
- 📖 A Força de Trabalho e os Setores de Atividade - Uma Análise da Região Metropolitana de São Paulo - 1986-1990 - **nº 55**, outubro 1991
- 📖 Revisão da PNAD: Apuração das Informações: Módulo III do Anteprojeto - **nº 56**, novembro 1991
- 📖 Novos Usos para Pesquisa Industrial Mensal: A Evolução dos Salários Industriais, O Desempenho da Pecuária - **nº 57**, novembro 1991
- 📖 Revisão da PNAD: A Disseminação das Informações Módulo IV do Anteprojeto - **nº 58**, dezembro 1991
- 📖 Estatísticas Agropecuárias : Sugestões para o Novo Plano Geral de Informações - **nº 59**, dezembro 1991
- 📖 Análise Conjuntural e Pesquisa Industrial - **nº 60**, janeiro 1992
- 📖 Exploração dos Dados da Pesquisa Industrial Mensal de Dados Gerais - **nº 61**, fevereiro 1992

- 📖 Uma Proposta de Metodologia para a Expansão da Amostra do Censo Demográfico de 1991 - **nº 62**, outubro 1993
- 📖 Expansão da Fronteira e Progresso Técnico no Crescimento Agrícola Recente - **nº 63**, novembro 1993
- 📖 Avaliação das Condições de Habitação com Base nos Dados da PNAD - **nº 64**, setembro 1993
- 📖 Análise da Taxa de Desemprego Feminino no Brasil – **nº 65**, dezembro 1993
- 📖 Aspectos da Metropolização Brasileira: Comentários sobre os Resultados Preliminares do Censo Demográfico de 1991- **nº 66**, janeiro 1994
- 📖 Estimativas Preliminares de Fecundidade Considerando os Censos Demográficos, Pesquisas por amostragem e o Registro Civil - **nº 67**, janeiro 1994
- 📖 Apuração de Dados no IBGE: Problemas e Perspectivas - **nº 68**, fevereiro 1994
- 📖 Limeira - SP: Estimativas de Fecundidade e Mortalidade 1980/1988 - **nº 69**, março 1994
- 📖 Desemprego - Uma Abordagem Conceitual - **nº 70**, abril 1994
- 📖 Apuração dos Dados Investigados no Questionário Básico (CD 1.01) do Censo Demográfico de 1991 - **nº 71**, outubro de 1994
- 📖 Deslocamento Populacional e Segregação Sócio-Espacial – Migrantes Originários do Rio de Janeiro - **nº 72**, novembro de 1994
- 📖 Projeção Preliminar da População do Brasil para o Período 1980-2020 - **nº 73**, dezembro de 1994
- 📖 Considerações Preliminares Sobre a Migração Internacional no Brasil - **nº 74**, janeiro de 1995
- 📖 Estatísticas Agropecuárias Censitárias no Âmbito do Mercosul - Brasil, Argentina e Uruguai - **nº 75**, julho de 1995
- 📖 Projeções Preliminares das Populações das Grandes Regiões para o Período 1991-2010 - **nº 76**, agosto de 1995
- 📖 Dinâmica da Estrutura Familiar no Sudeste Metropolitano, Chefia Feminina e Indicadores Sócio-Demográficos: Um exercício exploratório utilizando modelo da regressão múltipla - **nº 77**, setembro de 1995
- 📖 O Uso das Matrizes de Insumo-Produto e Matrizes de Inovação para Medir Mudanças Técnicas - **nº 78**, outubro de 1995
- 📖 Estimativas dos Fatores de Correção para o Registro de Nascimentos Utilizando Registros tardios a nível de Brasil, Grandes Regiões, Unidades da Federação e Regiões Metropolitanas 1974/1994 - **nº 79**, abril de 1996
- 📖 Aspectos de Amostragem Relativos ao Censo Cadastro de 1995 - **nº 80**, junho de 1996
- 📖 Tendências Populacionais no Brasil e Pressão Sobre o Mercado de Trabalho Futuro - **nº 81**, setembro de 1996
- 📖 Transformações Estruturais e Sistemas Estatísticos Nacionais - **nº 82**, setembro de 1996
- 📖 Metodologias para o Cálculo de Coeficientes Técnicos Diretos em um Modelo de Insumo-Produto - **nº 83**, outubro de 1996
- 📖 Avaliação da Cobertura da Coleta do Censo Demográfico de 1991 - **nº 84**, outubro de 1996
- 📖 Componentes da Dinâmica Demográfica Brasileira: Textos Selecionados - **nº 85**, novembro de 1996
- 📖 Apuração dos Dados Investigados pelo Questionário da Amostra - CD 1.02 do Censo Demográfico de 1991 - **nº 86**, dezembro de 1996
- 📖 Estudo Preliminar da Evolução dos Nascimentos, Casamentos e Óbitos 1974-1990 - **nº 87**, janeiro de 1997

- 📖 Sistema de Contas Nacionais - Tabelas de Recursos e Usos - Metodologia - nº 88, dezembro de 1997
- 📖 Aspectos de Amostragem da Pesquisa de Economia Informal Urbana 97 - nº 89, junho de 1998
- 📖 Comparações da Renda Investigada nos Questionários do Censo Demográfico de 1991 - nº 90, julho de 1998
- 📖 Uma Revisão dos Principais Aspectos dos Planos Amostrais das Pesquisas Domiciliares Realizadas pelo IBGE - nº 91, setembro de 1998
- 📖 Planejamento Amostral para as Pesquisas Anuais da Indústria e do Comércio - nº 92, outubro de 1998
- 📖 Aspectos de Amostragem da Pesquisa de Orçamentos Familiares 1995-1996 - nº 93, dezembro de 1998
- 📖 Reflexões sobre um Programa de Estatísticas Ambientais - nº 94, abril de 1999
- 📖 O Comportamento das Importações e Exportações Brasileiras com Base no Sistema de Contas Nacionais 1980 - 1997 (versão preliminar) - nº 95, maio de 1999
- 📖 Meio Ambiente: sua integração nos sistemas de informações estatísticas - nº 96, maio de 1999
- 📖 Conta da Terra: considerações sobre sua realização no Brasil - nº 97, dezembro de 1999

Textos para discussão - nova série

- 📖 **Número 1** - Sistema integrado de contas econômico-ambientais - SICEA : síntese e reflexões / Sandra De Carlo. - Rio de Janeiro : IBGE, Departamento de Contas Nacionais, 2000.
- 📖 **Número 2** - Aspectos da produção de informação estatística oficial no contexto da sociedade atual : algumas questões teórico-metodológicas / Rosa Maria Porcaro - Rio de Janeiro : IBGE, Departamento de Metodologia, 2000
- 📖 **Número 3** - A Cor denominada : um estudo do suplemento da Pesquisa Mensal de Emprego de julho/98 / José Luis Petruccelli. - Rio de Janeiro : IBGE, Departamento de População e Indicadores Sociais, 2000.
- 📖 **Número 4** - Indicadores para a agropecuária - Rio de Janeiro : IBGE, Departamento de Agropecuária, 2001.
- 📖 **Número 5** - Estudos para definição da amostra da Pesquisa Industrial Mensal de Emprego e Salário / Ana Maria Lima de Farias. - Rio de Janeiro : IBGE, Departamento de Indústria, 2001.
- 📖 **Número 6** - A declaração de cor/raça no censo 2000: um estudo comparativo / José Luis Petruccelli. - Rio de Janeiro : IBGE, Departamento de População e Indicadores Sociais, 2002..
- 📖 **Número 7** - Dimensões preliminares da responsabilidade feminina pelos domicílios: um estudo do fenômeno a partir dos censos demográficos 1991 e 2000 / Sonia Oliveira, Ana Lucia Sabóia, Bárbara Cobo - Rio de Janeiro : IBGE, Departamento de População e Indicadores Sociais, 2002.
- 📖 **Número 8** - Principais Aspectos de Amostragem das Pesquisas Domiciliares do IBGE - revisão 2002 / Zélia Magalhães Bianchini e Sônia Albieri - Rio de Janeiro : IBGE, Departamento de Metodologia, 2003.
- 📖 **Número 9** - Censo Demográfico 2000 - Resultados da Pesquisa de Avaliação da Cobertura da Coleta / Luís Carlos de Souza Oliveira, Marcos Paulo Soares de Freitas, Márcia Regina Martins Lima Dias, Cláudia Maria Ferreira Nascimento, Edie da Silva Mattos e João José Amado Ramalho Júnior - Rio de Janeiro : IBGE, Coordenação Técnica do Censo Demográfico, 2003.
- 📖 **Número 10** - Sistema de informação estatística e a sociedade da informação / Rosa Maria Porcaro - Rio de Janeiro : IBGE, Departamento de Metodologia, 2003.

- 📖 **Número 11** - Indicadores para a agropecuária - 1996 a 2001 /Julio César Perruso, Marcelo de Moraes, Duriez, Roberto Augusto Soares P. Duarte e Carlos Alfredo Barreto Guedes - Rio de Janeiro : IBGE, Coordenação de Agropecuária, 2003.
- 📖 **Número 12** - A Unidade de Metodologia e a Evolução do Uso de Amostragem no IBGE, 2003 / Sônia Albieri - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, 2003.
- 📖 **Número 13** - Estimando a Precisão das Estimativas das Taxas de Mortalidade Obtidas a Partir da PNAD / Pedro Luis do Nascimento Silva e Djalma Galvão Carneiro Pessoa. - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, 2004.
- 📖 **Número 14** - A Qualidade na Produção de Estatísticas no IBGE / Zélia Magalhães Bianchini. - Rio de Janeiro : IBGE, Diretoria de Pesquisas, 2004
- 📖 **Número 15** - Calibration Estimation: When and Why, How Much and How / Pedro Luis do Nascimento Silva . - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, 2004
- 📖 **Número 16** - Um panorama recente da desigualdade no Brasil a partir dos dados da PNAD 2002 / Ana Lucia Saboia e Barbara Cobo. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2004
- 📖 **Número 17** – Processamento das Áreas de Expansão e Disseminação da Amostra no Censo Demográfico 2000 / Ari Nascimento Silva, Luiz Alberto Matzenbacher e Bruno Freitas Cortez. - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, 2004
- 📖 **Número 18** – Fatores de correção para o registro de nascimentos utilizando registros tardios segundo os grupos de idades das mulheres - Brasil e Unidades da Federação - 1984-2001 / Fernando Roberto Pires de Carvalho e Albuquerque e Selma Regina dos Santos. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2004
- 📖 **Número 19** – O processo de Imputação dos quesitos de migração no Censo Demográfico 2000 / Fernando Roberto P. de C. e Albuquerque, Janaína Reis Xavier Senna e Antonio Roberto Pereira Garcez - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2004
- 📖 **Número 20** – Tábuas de Mortalidade por sexo e grupos de idade - Grandes Regiões e Unidades da Federação - 1980, 1991 e 2000 / Fernando Roberto P. de C. e Albuquerque e Janaína Reis Xavier Senna - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2005
- 📖 **Número 21** – Tempo, trabalho e afazeres domésticos: um estudo com base nos dados da Pesquisa Nacional por Amostra de Domicílios - 2001 e 2005/ Cristiane Soares e Ana Lucia Saboia - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2007
- 📖 **Número 22** – Estimação de Intervalos de Confiança para Estimadores de Diferenças Temporais na Pesquisa Mensal de Emprego / Mauricio Franca Lila e Marcos Paulo soares de Freitas - Rio de Janeiro: IBGE, Coordenação de Trabalho e Rendimento e Coordenação de Métodos e Qualidade, 2007
- 📖 **Número 23** – Amostra Mestra para o Sistema Integrado de Pesquisas Domiciliares / Marcos Paulo Soares de Freitas, Mauricio Franca Lila, Rosemary Vallejo de Azevedo e Giuseppe de Abreu Antonaci - Rio de Janeiro: IBGE, Coordenação de Métodos e Qualidade, 2007
- 📖 **Número 24** – Sistema Integrado de Pesquisas Domiciliares - SIPD / Coordenação de Trabalho e Rendimento - Rio de Janeiro: IBGE, 2007
- 📖 **Número 25** – Pesquisas Agropecuárias por Amostragem Probabilística no IBGE: Histórico e Perspectivas Futuras / Coordenação de Agropecuária - Rio de Janeiro: IBGE, 2007

- 📖 **Número 26** – Migração Pendular Intrametropolitana no Rio de Janeiro: Reflexões sobre o seu estudo, a partir dos Censos Demográficos de 1980 e 2000 / Antonio de Ponte Jardim e Leila Ervatti - Rio de Janeiro: IBGE, Coordenação de População e Indicadores Sociais, 2007
- 📖 **Número 27** – Características da fecundidade e da mortalidade segundo a condição migratória das mulheres, com base no quesito de "data fixa" / Fernando Roberto Pires de Carvalho e Albuquerque, Isabel Cristina Maria da Costa e Antonio Roberto Pereira Garcez - Rio de Janeiro: IBGE, Coordenação de População e Indicadores Sociais, 2007
- 📖 **Número 28** – Utilização de Modelos para Estimar a Mortalidade Brasileira nas Idades Avançadas / Jorcely Victório Franco, Juarez de Castro Oliveira e Fernando Roberto Pires de C. e Albuquerque - Rio de Janeiro: IBGE, Coordenação de População e Indicadores Sociais, 2007
- 📖 **Número 29** – Influência da mortalidade nos níveis de fecundidade da população brasileira e o intervalo médio entre duas gerações sucessivas - 1980, 1991, 2000 e 2005/ Fernando Roberto Pires de C. e Albuquerque e Maria Lúcia Pereira do Nascimento - Rio de Janeiro: IBGE, Coordenação de População e Indicadores Sociais, 2008
- 📖 **Número 30** - Família nas pesquisas domiciliares : questões e propostas alternativas / Rosa Ribeiro, Ana Lúcia Sabóia - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2008
- 📖 **Número 31** – Setor e Emprego Informal no Brasil - Análise dos resultados da nova série do Sistema de Contas Nacionais / João Hallak Neto, Katia Namir, Luciene Kozovitz, Sandra Rosa Pereira - Rio de Janeiro : IBGE, Coordenação de Contas Nacionais, 2008
- 📖 **Número 32** - Diferenciais de idade entre os casais nas famílias brasileiras / Cristiane Soares. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, 2008
- 📖 **Número 33** – Estudos de modalidades alternativas de censos demográficos : aspectos de amostragem / IBGE, Diretoria de Pesquisas, Grupo de Trabalho de Amostragem, Estimção e Acumulação de Informações. - Rio de Janeiro : IBGE, 2009.
- 📖 **Número 34** – O Acompanhamento Estatístico da Fabricação de Medicamentos na Indústria Farmacêutica Brasileira/ Marcus José de Oliveira Campos e Luiz Antônio Casemiro dos Santos. - Rio de Janeiro : IBGE, Diretoria de Pesquisas, 2009.
- 📖 **Número 35** – Áreas mínimas de Comparação / Weuber da Silva Carvalho, Gilson Flaeschen. - Rio de Janeiro : IBGE, Diretoria de Pesquisas, 2010.
- 📖 **Número 36** – Contabilizando a Sustentabilidade: principais abordagens / Frederico Barcellos, Paulo Gonzaga M. de Carvalho e Sandra De Carlo. - Rio de Janeiro : IBGE, Diretoria de Pesquisas, 2010.
- 📖 **Número 37** – Indicadores sobre Trabalho Decente: Uma contribuição para o debate da desigualdade de gênero / Cíntia Simões Agostinho e Ana Lucia Saboia. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, Diretoria de Pesquisas, 2011.
- 📖 **Número 38** – Reflexões sobre pesquisas longitudinais: uma contribuição à implementação do Sistema Integrado de Pesquisas Domiciliares / Leonardo Athias. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, Diretoria de Pesquisas, 2011.
- 📖 **Número 39** – Desafios e possibilidades sobre os novos arranjos familiares e a metodologia para identificação de família no Censo / Ana Lucia Saboia, Bárbara Cobo e Gilson Gonçalves Matos. - Rio de Janeiro : IBGE, Coordenação de População e Indicadores Sociais, Diretoria de Pesquisas, 2012.

- 📖 **Número 40** – Metodologia Estatística da Pesca: Pesca embarcada / Aristides Pereira Lima Green e Guilherme Guimarães Moreira. - Rio de Janeiro : IBGE, Coordenação de Agropecuária e Coordenação de Métodos e Qualidade, Diretoria de Pesquisas, 2012.
- 📖 **Número 41** – Pareamento Automático na Pesquisa de Avaliação da Cobertura da Coleta do Censo Demográfico / Djalma Galvão Carneiro Pessoa, Fábio Figueiredo Farias e Vinícius Layter Xavier. - Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, Diretoria de Pesquisas, 2012.
- 📖 **Número 42** – Seminários IBGE – 15 anos disseminando conhecimento / Sonia Albieri. – Rio de Janeiro : IBGE, Coordenação de Métodos e Qualidade, Diretoria de Pesquisas, 2012.