

Maria Cecilia Goi Porto Alves¹

Nilza Nunes da Silva^{II}

Métodos de estimação de variância em amostras provenientes de inquéritos domiciliares

Variance estimation methods in samples from household surveys

RESUMO

OBJETIVO: O conhecimento dos erros de amostragem é necessário à correta interpretação dos resultados de inquéritos domiciliares e à avaliação dos seus planos de amostragem. A composição das amostras de domicílios utilizadas em inquéritos caracteriza situação complexa de estimação. Nesse sentido, realizou-se estudo com o objetivo de avaliar o desempenho de estimadores de variância em inquéritos efetuados em populações urbanas brasileiras.

MÉTODOS: A população de referência do estudo constituiu-se de amostra sorteada pela Fundação Sistema Estadual de Análise de Dados Estatísticos, para a realização da Pesquisa de Emprego e Desemprego na Região Metropolitana de São Paulo. Para estimar variâncias foram utilizados: o método de linearização de Taylor e as técnicas de replicação Jackknife e BRR. Repetidas amostras foram retiradas da população de referência utilizando amostragem estratificada, por conglomerados, em dois estágios: setor censitário e domicílio. Três delineamentos foram utilizados e 2.000 amostras foram sorteadas sob cada um deles. Para um estimador razão, foi avaliada a acurácia dos estimadores de variância, por meio do erro quadrático médio, e a cobertura dos intervalos de confiança.

RESULTADOS: Os resultados relacionados ao erro quadrático médio relativo dos estimadores foram semelhantes. As razões de vício ficaram em torno de 0,10 para as menores amostras. As coberturas dos intervalos de confiança indicaram que os níveis de confiança observados foram menores que os fixados (95%), ficando em torno de 90% para as menores amostras.

CONCLUSÕES: Os estimadores de variância mostraram desempenhos semelhantes quanto à acurácia e cobertura dos intervalos de confiança. Os vícios foram irrelevantes frente às dimensões do erro-padrão. Os níveis de confiança reais foram menores que os níveis nominais da distribuição normal, mas as alterações não impedem que estimativas intervalares sejam feitas com razoável confiança.

DESCRITORES: Técnicas de estimativa. Coleta de dados. Amostragem estratificada. Amostragem por conglomerados. Linearização de Taylor. Planos complexos de amostragem.

¹ Instituto de Saúde. Secretaria de Estado da Saúde de São Paulo. São Paulo, SP, Brasil

^{II} Departamento de Epidemiologia. Faculdade de Saúde Pública. Universidade de São Paulo. São Paulo, Brasil

Correspondência | Correspondence:
Maria Cecilia Goi Porto Alves
R. Santo Antônio, 590 – 5º. Andar
01314-000 São Paulo, Brasil
E-mail: cecilia@isaude.sp.gov.br

ABSTRACT

OBJECTIVE: Knowledge of sampling errors is essential for correctly interpreting the results from household surveys and evaluating their sampling designs. The composition of household samples used in surveys gives rise to situations of complex estimation. In this light, the study was conducted with the aim of evaluating the performance of the variance estimators in surveys carried out among urban populations in Brazil.

METHODS: The reference population was the sample drawn by the Fundação Sistema Estadual de Análise de Dados Estatísticos (SEADE – State Statistical Data Analysis System Foundation) for carrying out an employment and unemployment survey in the metropolitan region of São Paulo. Three techniques were used for estimating variance: Taylor linearization and Jackknife and BRR replication. Repeated samples were selected from the reference population, using stratified cluster sampling in two stages (census tracts and households). Three different designs were used and 2,000 samples were drawn within each design. To obtain an estimator ratio, the accuracy of the variance estimators was evaluated by means of the mean square error and the confidence interval coverage.

RESULTS: According to the mean square error, the three techniques provided similar accuracy. The bias ratios were approximately 0.10, for the smaller samples. The confidence interval coverage indicated that the confidence levels observed were lower than what was set (95%), and were around 90% for the smaller samples.

CONCLUSIONS: The variance estimators showed similar performance with regard to accuracy and confidence interval coverage. The bias was irrelevant in relation to the magnitude of the standard error. Although the real confidence levels were lower than the nominal levels for normal distribution, the changes did not prevent construction of interval estimates with reasonable confidence.

KEY WORDS: Estimation techniques. Data collection. Stratified sampling. Cluster sampling. Taylor series linearization. Complex sample design.

INTRODUÇÃO

O conhecimento dos erros de amostragem é necessário à correta interpretação dos resultados de inquéritos domiciliares e à avaliação dos seus planos de amostragem.¹⁶ No entanto, a composição de amostras de domicílios, pelo sorteio de conglomerados em vários estágios, caracteriza situação complexa de estimação, em que o uso dos métodos de estimação de variância com base na amostragem aleatória simples (AAS), comumente adotados, podem ser considerados inadequados.⁹

Na AAS, em que as unidades têm iguais probabilidades de seleção e são independentes, a variância do estimador da média de uma variável “y”, caso o sorteio seja

com reposição, é expressa por $\frac{s^2}{n} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n(n-1)}$, onde

n é o número de elementos da amostra, y_i são os valores desses elementos e $\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$ sua média.⁸

Em inquéritos domiciliares, a dispersão dos elementos da população em áreas geográficas extensas, torna necessário o sorteio de conjuntos de elementos para diminuir os custos da amostra. Com a adoção de conglomerados de tamanhos desiguais, o número de elementos da amostra passa a ser uma variável aleatória, o que transforma a média em uma razão de variáveis, com conseqüências na estimação do erro de amostragem. Nesse contexto, a maior parte dos estimadores de primeira ordem são não-lineares, e a inexistência de expressões exatas para calcular suas variâncias remete ao uso de aproximações ajustadas à natureza complexa do delineamento da amostra e do processo de estimação.

As abordagens atualmente mais utilizadas para estimar variância em planos complexos de amostragem são: o método de linearização de Taylor e as técnicas de replicação. O primeiro fornece uma aproximação linear para o estimador não-linear de interesse, por meio da Série de Taylor, a qual são aplicadas as fórmulas usuais

de estimadores lineares. Para que a aproximação da variância de um estimador razão pela série de Taylor seja adequada, o tamanho da amostra (denominador da razão) não pode estar sujeito a grande variação, o que ocorre quando os conglomerados diferem muito em tamanho. A dificuldade em manter sob controle a variabilidade do denominador aumenta quando a estimação é dirigida para subclasses, pois o aparecimento de unidades pertencentes a elas na amostra não pode ser controlado, em especial para características raras e/ou distribuídas desigualmente pelos conglomerados.

No caso do estimador razão, $r = \frac{y}{x}$, a expressão de variância derivada da aplicação do método de linearização de Taylor tem sido amplamente utilizada na área de amostragem, sendo da forma: $\text{var}(r) = \frac{1}{x^2} [\text{var}(y) + r^2 \text{var}(x) - 2r \text{cov}(xy)]$.⁸ Apesar de ser apresentada com frequência na literatura estatística, essa expressão, que envolve as variâncias do numerador, $\text{var}(y)$, do denominador, $\text{var}(x)$, além da covariância entre eles, $\text{cov}(xy)$, não pode ser considerada simples do ponto de vista computacional.

As técnicas de replicação surgiram como uma alternativa de simplificação do processo de estimação de variância. Consistem em se obterem subconjuntos da amostra (réplicas), repetir o processo de estimação para cada um dos subconjuntos e calcular a variância a partir dessas estimativas. Dessa forma, sendo θ o parâmetro de interesse e $\hat{\theta}$ o seu estimador, são formadas K réplicas e obtidas estimativas para cada réplica utilizando $\hat{\theta}_k$ ($k = 1, \dots, K$). O estimador da variância de θ é: $\text{var}(\hat{\theta}) = \frac{1}{c} \sum_{k=1}^K (\hat{\theta}_k - \hat{\theta})^2$, em que c é uma constante associada ao procedimento adotado.¹⁷

Essa construção segue a simplicidade da estimação da variância na AAS, que utiliza os desvios das observações em relação à média. A idéia básica foi dada por Mahalanobis, em trabalhos de 1944 e 1946, nos quais as réplicas, denominadas “*interpenetrating samples*”, foram propostas para facilitar a estimação de erros de amostragem e também para investigar erros não amostrais, tais como vícios relacionados a entrevistadores.⁷

As técnicas de replicação mais comuns são: “*Balanced Repeated Replication*” (BRR) e “*Jackknife*”. BRR surgiu como uma alternativa de estimação em delineamentos que, por considerações de eficiência, utilizam um número muito grande de estratos e, conseqüentemente, o menor número possível de unidades primárias de amostragem por estrato: duas unidades. Originou-se no esquema de pseudo-replicação denomi-

nado “*half-sample replication*”, proposto pelo *United States Bureau of the Census*, posteriormente adaptado e modificado pelos técnicos do *National Center of Health Statistics*. Em 1966, McCarthy¹² introduziu o balanceamento a que se refere o nome da técnica. As réplicas são compostas por uma das unidades de amostragem de cada estrato.

Na técnica *Jackknife*, cada réplica é formada omitindo-se sucessivamente uma unidade de amostragem em um dos estratos. Procedimentos de estimação *Jackknife* foram concebidos, originalmente, com o objetivo de reduzir o vício de um estimador de coeficiente de correlação em séries temporais, por Quenouille, em 1949. Tukey¹⁵ sugeriu que os estimadores individuais das sub-amostras criados pela técnica poderiam ser vistos como variáveis aleatórias independentes e identicamente distribuídas, levando-se a um estimador muito simples de variância. Deu-lhe o nome de *Jackknife* para fazer referência a um instrumento de múltiplos usos, o canivete. Sua utilização no contexto de populações finitas parece ter sido considerada, pela primeira vez, por Durbin, em 1959, associada ao estimador razão.¹⁸

Um dos atrativos das técnicas de replicação, além de suas expressões de variância serem mais simples que as do método de linearização de Taylor, é o fato de que o procedimento de análise, dado um determinado delineamento, é o mesmo para quase todo tipo de estatística, independente de sua complexidade. Outra propriedade das técnicas de replicação é criar condições para que usuários de dados secundários possam estimar erros de amostragem, sem conhecer pormenores relacionados ao delineamento, utilizando as réplicas criadas pelos pesquisadores envolvidos no inquérito e anexadas ao arquivo de dados. Isto é particularmente útil quando questões de confidencialidade envolvem os elementos da amostra, sendo necessário evitar a divulgação de qualquer informação relacionada à sua identificação.^{14,*}

As facilidades atuais de processamento eletrônico permitem a aplicação de quaisquer das técnicas referidas para estimar erros em inquéritos realizados com o uso da amostragem por conglomerados. Os estudos comparativos disponíveis fundamentam-se em resultados empíricos obtidos em diversos países, sob diferentes delineamentos.^{1,3,4,9,10,13}

O presente trabalho teve por objetivo mostrar o desempenho das técnicas de estimação de variância: *Jackknife*, BRR e linearização de Taylor, em amostras que reflitam a estrutura sociodemográfica da região metropolitana. Nesse sentido, amostras de setores censitários, que constituem as unidades geográficas de maior emprego em inquéritos domiciliares brasileiros, foram utilizadas

* Brick JM, Morganstein D, Valliant R. Analysis of complex samples using replication. Rockville: Westat; 1998 [Acesso em dia/mes/ano]. Disponível em: <http://www.westat.com/wesvar/techpapers/ACS-Replication.pdf>

para comparar a acurácia dos estimadores de variância. Com isso, pretende-se contribuir para o conhecimento e a divulgação das alternativas existentes para estimação de erros de amostragem, estimulando o uso de técnicas adequadas à realização de inferências estatísticas feitas a partir de inquéritos domiciliares.

MÉTODOS

A amostra sorteada pela Fundação Sistema Estadual de Análise de Dados Estatísticos para a realização da Pesquisa de Emprego e Desemprego na Região Metropolitana de São Paulo foi tomada como a população referência para desenvolvimento do estudo.⁶

Dessa população, foram retiradas repetidas amostras mediante sorteio estratificado, por conglomerados em dois estágios. Os setores censitários constituíram as unidades primárias de amostragem (UPA) e os domicílios, as secundárias. Em cada estrato, dois setores censitários foram sorteados com probabilidade proporcional ao número de domicílios. E em cada um desses setores, cinco domicílios foram sorteados, perfazendo o total de dez domicílios por estrato. A fração de amostragem no estrato *h* foi: $f_h = \frac{2 \cdot M_{h\alpha}}{M_h} \cdot \frac{5}{M_{h\alpha}} = \frac{10}{M_h}$, sendo M_h o número de domicílios do estrato *h* e $M_{h\alpha}$ o número de domicílios do setor censitário *α* do estrato *h*.

Para avaliar a acurácia dos estimadores de variância em função do aumento do número de unidades primárias de amostragem, foram definidos três delineamentos, mantendo-se o modelo geral anteriormente apresentado. Para o primeiro, segundo e terceiro delineamentos, a população foi organizada em, respectivamente, oito, 16 e 32 estratos. Uma vez que foram sorteados dois setores por estrato, os números de setores incluídos nas amostras foram 16, 32 e 64. Considerando, ainda, que de cada setor foram sorteados cinco domicílios, os tamanhos finais de amostra nos três delineamentos foram de 80, 160 e 240 domicílios.

Sob cada um dos delineamentos foram sorteadas 2.000 amostras.

Obteve-se, em cada uma das amostras, a renda média

$$r = \frac{\sum_h w_h \sum_{\alpha} \sum_{\beta} y_{h\alpha\beta}}{\sum_h w_h \sum_{\alpha} \sum_{\beta} x_{h\alpha\beta}}$$

onde $y_{h\alpha\beta}$ é a renda familiar do domicílio β , do setor α , do estrato h ; $x_{h\alpha\beta} = 1$, para os domicílios com informação e $x_{h\alpha\beta} = 0$ para aqueles sem resposta; $w_h = 1/f_h$, o peso de cada domicílio, dado pelo inverso da fração de amostragem do estrato a que pertence o domicílio, $f_h = 10/M_h$.

A partir da distribuição de frequências, calculou-se, em cada delineamento, a variância populacional:

$$S^2 = \frac{\sum_{i=1}^{2000} [r_i - E(r)]^2}{2000}$$

sendo $E(r)$ o valor esperado de r , correspondente à média de sua distribuição de amos-

tragem, $E(r) = \frac{\sum_{i=1}^{2000} r_i}{2000}$, e r_i a renda média calculada na *i*-ésima amostra.

As variâncias estimadas pelas técnicas *Jackknife* e BRR foram calculadas no software *WesVar*, versão 4.0,¹⁷ utilizando-se a expressão:

$$var(r) = \frac{\sum_{g=1}^G (r_{(g)} - r)^2}{c}$$

sendo r a estimativa obtida sobre o total de unidades primárias, $r_{(g)}$ a estimativa referente a *g*-ésima réplica, *G* o número de réplicas, e *c* uma constante que depende da técnica de replicação, (*c*=*G* para BRR ; *c*=1 para *Jackknife*).

As estimativas de variância pelo método de linearização de Taylor, foram obtidas pelas expressões:⁸

$$var(r) = \frac{1}{x^2} \left[\sum_h var(y_h) + r^2 \sum_h var(x_h) - 2r \sum_h cov(y_h, x_h) \right]$$

$$var(y_h) = \frac{1}{a_h - 1} \left[a_h \sum_{\alpha} (w_{h\alpha} y_{h\alpha})^2 - \left(\sum_{\alpha} w_{h\alpha} y_{h\alpha} \right)^2 \right]$$

$$var(x_h) = \frac{1}{a_h - 1} \left[a_h \sum_{\alpha} (w_{h\alpha} x_{h\alpha})^2 - \left(\sum_{\alpha} w_{h\alpha} x_{h\alpha} \right)^2 \right]$$

$$cov(y_h, x_h) = \frac{1}{a_h - 1} \left[a_h \sum_{\alpha} (w_{h\alpha} y_{h\alpha})(w_{h\alpha} x_{h\alpha}) - \left(\sum_{\alpha} w_{h\alpha} y_{h\alpha} \right) \left(\sum_{\alpha} w_{h\alpha} x_{h\alpha} \right) \right]$$

sendo $a_h = 2$ o número de unidades primárias de amostragem; w_{ha} o peso no domicílio α do estrato h ; y_{ha} o valor da variável em estudo no domicílio α do estrato h ; $x_{ha} = 1$ se domicílio tem informação e $x_{ha} = 0$, caso contrário.

Considerando que a acurácia refere-se a pequenos erros totais, incluindo vício e variabilidade amostral, a acurácia dos estimadores foi avaliada por meio do erro quadrático médio: $EQM[\text{var}(r)] = \text{Var}[\text{var}(r)] + \text{Vic}^2[\text{var}(r)]$, e

por sua medida relativa: $EQMR = \frac{EQM[\text{var}(r)]}{E[\text{var}(r)]^2}$,

A variância das distribuições de amostragem foi expressa

por: $\text{Var}[\text{var}(r)] = \frac{\sum_{i=1}^{2000} \{\text{var}(r_i) - E[\text{var}(r)]\}^2}{2000}$, e $E[\text{var}(r)]$,

o valor esperado da variância de r , foi calculado por:

$$E[\text{var}(r)] = \frac{\sum_{i=1}^{2000} \text{var}(r_i)}{2000}$$

O vício dos estimadores, medido pela distância entre os valores médios, da distribuição de amostragem e da população, foi expresso por: $\text{Vic}[\text{var}(r)] = E[\text{var}(r) - S^2]$; cujos termos já foram apresentados anteriormente. A contribuição do vício na acurácia dos estimadores foi avaliada pela razão de vício, que mede o vício em unidades de desvio-padrão: $RV[\text{var}(r)] = \frac{\text{Vic}[\text{var}(r)]}{\sqrt{\text{Var}[\text{var}(r)]}}$.

Com o objetivo de verificar o real nível de confiança dos intervalos construídos, foi estudada a distribuição da razão padronizada $\frac{r - E(r)}{\sqrt{\text{var}(r)}}$. Calculou-se a proporção de vezes que essa razão pertenceu aos limites $[-2,306; +2,306]$, $[-2,120; +2,120]$ e $[-2,064; +2,064]$, correspondentes a valores da distribuição *t-Student*, com 8, 16 e 24 graus de liberdade, respectivamente, verificando sua proximidade a 0,95. Uma vez que,

para estimadores complexos, a distribuição exata dessa razão padronizada é desconhecida, o cálculo de cobertura, verificando-se sua pertinência a intervalos pré-fixados, permite avaliar a aplicabilidade dos intervalos de confiança.¹

RESULTADOS

Foram construídas as distribuições de amostragem do estimador r para os três delineamentos definidos, considerando que a forma da distribuição de 2.000 estimativas de r é estável o suficiente para ser tomada como tal. Adotou-se a variância dessas distribuições como sendo a variância real de r para o delineamento fixado, assim como fizeram Bean & Kish em seus estudos.^{1,9}

Da mesma forma, foram construídas as distribuições de amostragem dos estimadores de variância BRR, *Jackknife* e Taylor, ilustradas nas Figuras 1 a 3, cada uma delas referindo-se a um delineamento de amostragem. As médias e os erros padrão das distribuições estão indicados em cada figura.

Observa-se, pela sobreposição das curvas, que os resultados relacionados a precisão e vício dos estimadores foram muito próximos. Esses resultados constam na Tabela 1.

A comparação das técnicas pelo erro quadrático médio relativo indicou que BRR, *Jackknife* e Taylor têm acurácia semelhante, e as diferenças entre as medidas situaram-se na terceira casa decimal, à exceção de *Jackknife*, no delineamento que definiu amostras de 16 unidades primárias de amostragem.

As razões de vício ficaram em torno de 0,10, sob o primeiro delineamento, de 16 setores censitários, indicando que o vício correspondeu a 10% do erro padrão da variância. Diminuíram para valores próximos a 0,04 e 0,06, sob os delineamentos com 32 e 48 setores, respectivamente.

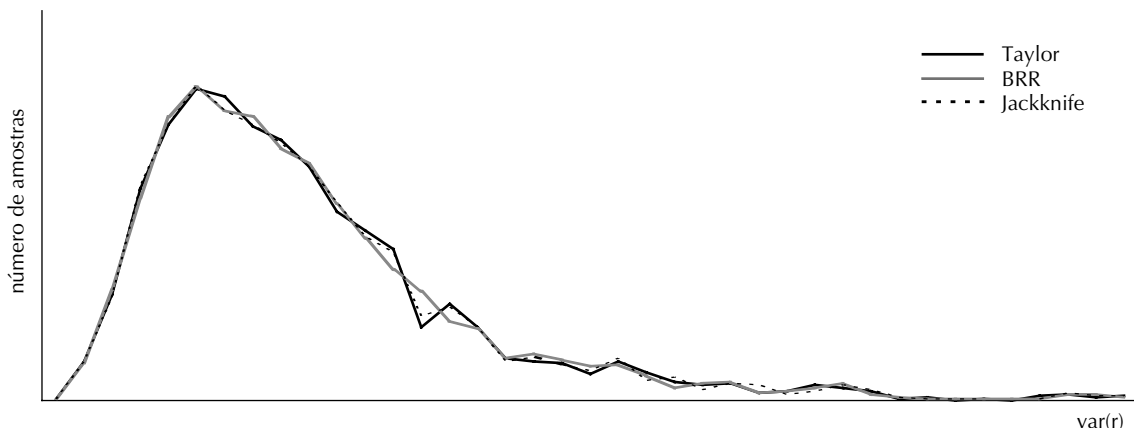


Figura 1. Distribuição de amostragem da variância do estimador razão (média 1353,955 e erro padrão 256,750), sob delineamento com 8 estratos e 16 unidades primárias de amostragem, segundo técnica de estimação de variância.

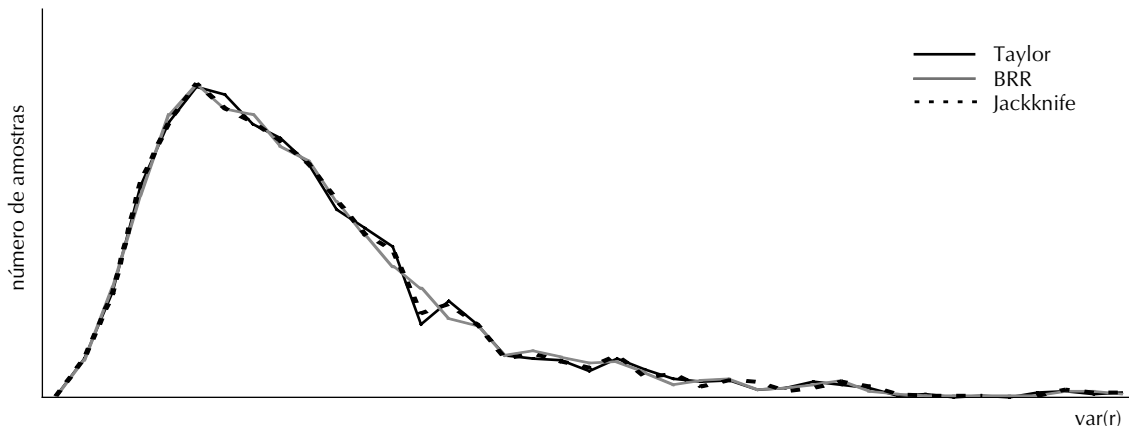


Figura 2. Distribuição de amostragem da variância do estimador razão (média 1353,767 e erro padrão 171,616), sob delineamento com 8 estratos e 16 unidades primárias de amostragem, segundo técnica de estimação de variância.

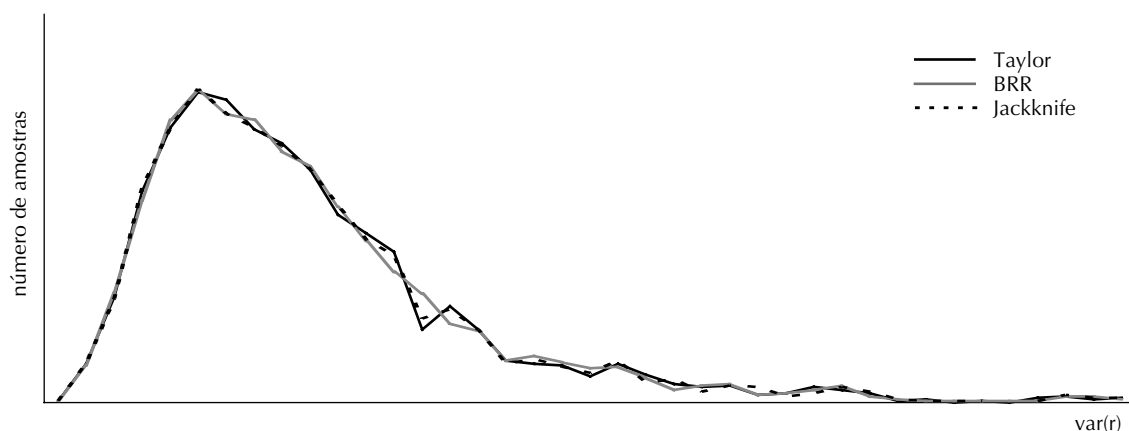


Figura 3. Distribuição de amostragem da variância do estimador razão (média 1353,107 e erro padrão 137,564), sob delineamento com 16 estratos e 32 unidades primárias de amostragem, segundo técnica de estimação de variância.

Em relação à cobertura dos parâmetros populacionais pelos intervalos de confiança, os resultados estão apresentados na Tabela 2. Os níveis de confiança observados foram menores que os fixados, ficando em torno de 90% para amostras de 16 unidades primárias e em torno de 94% para as amostras maiores. Os valores referentes às distintas técnicas foram muito próximos.

DISCUSSÃO

Com base nas distribuições de amostragem dos estimadores de variância, verificou-se que em relação ao erro-padrão e vício, indicadores da precisão e confiabilidade dos resultados obtidos e de sua validade, houve o seguinte padrão: diferenças pequenas entre os estimadores, resultados semelhantes para BRR e Taylor e desempenho ligeiramente inferior de *Jackknife*. As diferenças foram mais acentuadas para os resultados obtidos com menores graus de liberdade, deixando de existir ou tornando-se muito menores com o aumento

do número de unidades primárias de amostragem.

No que respeita a acurácia, as diferenças foram tão pequenas que não apareceriam caso o erro quadrático médio relativo fosse calculado com somente duas casas decimais, sendo difícil falar em acurácia maior para um ou outro estimador. Dados obtidos em outros estudos confirmam a proximidade dos resultados referentes aos diversos estimadores.

Bean¹ conduziu investigação empírica sobre o comportamento dos métodos de estimação de variância, linearização de Taylor e replicação (dois estimadores BRR), utilizando variáveis medidas no estudo *U.S. Health Interview Survey*. O autor observou alternância de melhores resultados entre BRR e Taylor, concluindo que este último apresentou erro quadrático médio ligeiramente menor. Kish & Frankell⁹ estudaram estimadores BRR, *Jackknife* e Taylor, utilizando dados do inquérito *Current Population Survey* do *U. S. Bureau of the Census*. Os autores concluíram que a variabilidade

Tabela 1. Esperança, erro padrão, vício e erro quadrático médio de var(r).

Estimador	Estimador de variância	Delineamento de amostragem		
		8 estratos 16 setores	16 estratos 32 setores	24 estratos 48 setores
Esperança	Taylor	$7,242 \times 10^4$	$3,198 \times 10^4$	$2,100 \times 10^4$
E[var(r)]	BRR	$7,293 \times 10^4$	$3,198 \times 10^4$	$2,105 \times 10^4$
	<i>Jackknife</i>	$7,345 \times 10^4$	$3,207 \times 10^4$	$2,107 \times 10^4$
Erro-padrão	Taylor	$7,260 \times 10^4$	$2,222 \times 10^4$	$1,211 \times 10^4$
SE[var(r)]	BRR	$7,303 \times 10^4$	$2,219 \times 10^4$	$1,214 \times 10^4$
	<i>Jackknife</i>	$7,398 \times 10^4$	$2,233 \times 10^4$	$1,217 \times 10^4$
Vício absoluto	Taylor	$6,493 \times 10^3$	$2,533 \times 10^3$	$2,079 \times 10^3$
B[var(r)]	BRR	$7,003 \times 10^3$	$2,525 \times 10^3$	$2,129 \times 10^3$
	<i>Jackknife</i>	$7,522 \times 10^3$	$2,622 \times 10^3$	$2,147 \times 10^3$
Vício relativo	Taylor	0,089	0,036	0,054
BR[var(r)]	BRR	0,096	0,036	0,055
	<i>Jackknife</i>	0,102	0,037	0,056
Erro quadrático	Taylor	$5,313 \times 10^9$	$4,942 \times 10^8$	$1,471 \times 10^8$
Médio	BRR	$5,383 \times 10^9$	$4,929 \times 10^8$	$1,478 \times 10^8$
MSE[var(r)]	<i>Jackknife</i>	$5,530 \times 10^9$	$4,992 \times 10^8$	$1,486 \times 10^8$
Erro quadrático	Taylor	1,013	0,483	0,333
Médio relativo	BRR	1,012	0,482	0,333
RMSE[var(r)]	<i>Jackknife</i>	1,025	0,485	0,335

BRR: *Balanced Repeated Replication*

de Taylor foi menor, embora as diferenças em relação aos outros estimadores tivessem sido pequenas, em especial quando foram utilizados estimadores razão. Em termos de precisão, foi seguido do estimador *Jackknife*. Kovar et al¹⁰ compararam estimadores Taylor, BRR, *Jackknife* e *bootstrap* em estudo de simulação, baseado em populações hipotéticas, construídas para se assemelharem à população do estudo *National Assessment of Educational Progress*. Os autores concluíram, em relação à precisão das estimativas de variância do estimador razão, que Taylor e *Jackknife* tiveram os melhores desempenhos, equivalentes entre si.

Por meio da razão de vício, observou-se que os vícios foram irrelevantes frente às dimensões do erro-padrão das estimativas de variância, para qualquer das técnicas. Esse fato leva à conclusão de que, nas condições em que foi feita a presente avaliação, problemas de acurácia dos estimadores estão associados a problemas de precisão, mais intensamente do que a vícios.

Esses resultados coincidem com os observados por Kish & Frankel⁹ e Bean,¹ que concluíram pela não existência de padrão consistente de menor vício para algum dos estimadores avaliados, tendo sido pequeno e tolerável para todos. À mesma conclusão chegaram Kovar et al,¹⁰ que avaliaram o vício em situações distintas. Para o sorteio de duas unidades primárias de amostragem por estrato, consideraram coeficientes de variação da

Tabela 2. Cobertura dos intervalos de confiança com coeficientes da distribuição *t-Student* e Normal, calculada com 2.000 amostras, segundo estimadores de variância e delineamento de amostragem.

Estimador	Delineamento de amostragem		
	8 estratos 16 setores	16 estratos 32 setores	24 estratos 48 setores
Taylor	90,20	93,90	94,10
BRR	90,10	93,85	93,95
<i>Jackknife</i>	90,15	93,85	94,10

média do denominador da razão r menores ou iguais e maiores que 10%. Esses autores¹⁰ verificaram que, para coeficientes de variação baixos, o vício dos estimadores de variância foram insignificantes. Porém, quando o coeficiente de variação da média do denominador aumentou, tornando-se maior que 10%, BRR apresentou vício positivo substancial, e *Jackknife* e Taylor tenderam a subestimar levemente a verdadeira variância.

Em relação à cobertura dos intervalos de confiança, os resultados indicaram, como nas avaliações anteriores, desempenhos semelhantes entre as três técnicas avaliadas. Embora os níveis de confiança observados tivessem sido menores que os fixados, foram sempre superiores a 90%, mesmo para o delineamento que considerou somente 16 setores censitários. Essas alterações no nível

de confiança podem ser consideradas toleráveis^{1,9} sendo importante que, ao fazerem inferências, o pesquisador esteja consciente de sua existência.

Estudos empíricos que avaliaram a aplicabilidade de intervalos de confiança em populações diversas, utilizando delineamentos e estimadores também diversos, verificaram igualmente que os níveis de confiança reais foram menores que os níveis nominais da distribuição normal. Seus autores consideraram que as alterações não impediam que estimativas intervalares fossem feitas com razoável confiança. Bean¹ observou em seu estudo coberturas superiores a 90% para os estimadores BRR e Taylor, sendo mais próximas de 95% para o primeiro. Também Kish & Frankell⁹ encontraram, para o estimador razão, coberturas mais adequadas para BRR (entre 90,4 e 94,4%), depois para *Jackknife* (entre 89,4 e 94,3) e, por último, para Taylor (entre 88,8 e 94,0). Kovar et al¹⁰ consideraram BRR e *Jackknife* equivalentes.

Os resultados do presente estudo mostraram que as inferências foram válidas até para o delineamento sob o qual foram sorteados oito setores censitários. Mesmo em inquéritos conduzidos com muitas unidades primárias de amostragem, pode haver interesse em estudar subgrupos populacionais restritos a algumas dessas unidades. Isso ocorre com frequência em inquéritos de saúde, levando à necessidade de se obterem estimativas intervalares com um número muito menor de unidades primárias do que aquelas com que o delineamento foi concebido. Burke & Rust,³ em estudo sobre o desempenho do estimador *Jackknife* para amostras sistemáticas com dois a 30 UPA, haviam mostrado que inferências válidas poderiam ser feitas com amostras de pelo menos seis unidades primárias.

Considerando que os métodos avaliados exibiram resultados equivalentes em relação à precisão e ao

vício, questões relacionadas à operacionalização passam a ter peso preponderante na decisão sobre a aplicação de algum dos métodos nos inquéritos de saúde realizados. Nesse sentido, a disponibilidade de *softwares* que produzem estimativas de variância sob delineamentos complexos torna-se critério relevante para essa escolha.

Várias alternativas surgiram nas últimas décadas. Foram criados *softwares* específicos para estimação de variância, usando um dos métodos ou ambos.^{5,11} Além disso, *softwares* de análise de dados de ampla utilização, tais como SAS, Stata e SPSS, passaram a considerar a estimação de variância sob delineamentos complexos, ampliando o leque de alternativas disponíveis para a análise de dados provenientes de inquéritos domiciliares.²

Esforços precisam ser empreendidos pelos pesquisadores responsáveis pela realização desses inquéritos para que as informações referentes aos delineamentos de amostragem sejam sempre incluídas nos arquivos de dados. É preciso conhecer, basicamente, as unidades primárias de amostragem e os estratos a que pertencem os elementos da amostra e seus pesos, caso existam.

AGRADECIMENTOS

A Tarcisio Neves da Cunha, da Organização Pan-Americana de Saúde, pelo suporte nas atividades de computação; à Fundação Sistema Estadual de Análise de Dados, por ter cedido a base de dados da Pesquisa de Emprego e Desemprego; aos professores: Chester Luiz Galvão Cesar, da Faculdade de Saúde Pública da USP; Rosana Fiorini Puccini, do Departamento de Pediatria da Unifesp; Jair Lício Ferreira Santos, da Faculdade de Medicina de Ribeirão Preto da USP; e Benedito Galvão Benze, do Departamento de Estatística da UFSCar, pelas críticas e sugestões.

REFERÊNCIAS

1. Bean JA. Distribution and properties of variance estimators for complex multistage probability samples. An empirical distribution. *Vital Health Stat 2*. 1975;(65): i-iv,1-46.
2. Brogan D. Sampling error estimation for survey data. In: Household sample surveys in developing and transition countries. New York: United Nations Publication; 2005. p.447-90.
3. Burke J, Rust K. On the performance of Jackknife variance estimation for systematic samples with small numbers of primary sampling units. In: Proceedings of the Survey Research Methods Section, American Statistical Association. Alexandria (VA): ASA; 1995. p. 321-7.
4. Campbell C, Meyer M. Some properties of T Confidence Intervals for Survey Data. In: Proceedings of the American Statistical Association, Survey Research Methods Section; 1978. p. 437-42.
5. Carlson BL. Software for statistical analysis of sample survey data. In: Armitage P, Colton T, editores. Encyclopedia of biostatistics. Cambridge: Harvard University; 1998 [Acesso em 15/10/2007]. Disponível em: <http://www.fas.harvard.edu/~stats/survey-soft/blveob.html>
6. Fundação Sistema Estadual de Análise de Dados. Pesquisa de Emprego e Desemprego – Conceitos, Metodologia e Operacionalização. São Paulo; 1995.
7. Kalton G. Practical methods for estimating surveys sampling errors. *Bull Int Statist Inst*. 1977;47(3):495-514.
8. Kish, L. Survey sampling. New York: John Wiley & Sons; 1965.
9. Kish L, Frankel MR. Inference from complex samples. *J R Stat Soc Ser B Methodol*. 1974;36(1):1-37.
10. Kovar JG, Rao JNK, Wu CFJ. Bootstrap and other methods to measure errors in survey estimates. *Canad J Statist*. 1998;16(Supl):25-45.
11. Lepkowski J, Bowles J. Sampling error software for personal computers. *Surv Stat* [periódico na internet].1996[Acesso em 15/10/2007](35):10-7. Disponível em: <<http://www.fas.harvard.edu/~stats/survey-soft/iass.html>>.[informar link correto]
12. McCarthy PJ. Replication: an approach to the analysis of data from complex surveys. *Vital Health Stat 2*. 1966;(14):1-38.
13. Mulry MH, Wolter KM. The effect of Fisher's z-transformation on confidence intervals for the correlation coefficient. In: Proceedings of the Survey Research Methods Section of the American Statistical Association; 1981. p. 601-6.
14. Skinner CJ, Holt D, Smith TMF. Analysis of complex surveys. Chichester: John Wiley & Sons; 1989.
15. Tukey JW. Bias and confidence in not-quite large samples. Abstract. *Ann Math Statist* 1958;29:614.
16. United Nations. Sampling errors in household surveys. New York; 1993.
17. Westat. WesVar™ 4.0 User's guide. Rockville: Westat; 2000.
18. Wolter KM. Introduction to variance estimation. New York: Springer-Velag; 1985.

Trabalho baseado na tese de doutorado de MCGP Alves, apresentada ao Departamento de Epidemiologia da Faculdade de Saúde Pública da USP, em 2002.

Apresentado no VI Congresso Brasileiro de Epidemiologia da Associação Brasileira de pós-Graduação em Saúde Coletiva, de 19 a 23 de junho de 2004, em Recife.