Documentação Técnica para a Amostra de Microdados do RGPH 2017 para o Uso do Público

David J. Megill
Consultor de Amostragem, INE

Carlos Creva Singano
Estatístico, Instituto Nacional de Estatística

Basílio Cubula
Estatístico, Instituto Nacional de Estatística

Janeiro de 2020

Índice

1-Introdução

Uma amostra dos dados definitivos do Recenseamento Geral da População e Habitação de 2017 (RGPH 2017) foi seleccionada para proporcionar ao público ficheiros de microdados representativos que podem ser usados para diferentes tipos de análise até ao nível de distrito. A metodologia usada para elaborar esta amostra de microdados do RGPH 2017 também protege a confidencialidade dos dados, para que agregados familiares e indivíduos não possam ser identificados, e estes microdados só podem ser usados para propósitos de análise estatística. O propósito deste documento é de descrever a metodologia usada para a selecção da amostra para as bases de microdados do RGPH 2017, e os procedimentos que devem ser usados para ponderar os dados da amostra e estimar medidas de precisão das estimativas destes dados.

2-Desenho da Amostra de Microdados do RGPH 2017

O plano de amostragem para seleccionar os dados do RGPH 2017 para a base de microdados para uso do público é baseada nos objectivos da divulgação dos dados do Censo, e para maximizar a eficiência estatística. O objectivo principal da base de microdados para o público é de disponibilizar ao público ficheiros com uma amostra representativa de microdados de agregados familiares e indivíduos dos dados definitivos do RGPH 2017 que podem ser usados para diferentes tipos de análise estatística até o nível de distrito. A taxa de amostragem vai afectar o volume de dados como também a precisão dos resultados da amostra de microdados. Outro aspecto importante é de manter a confidencialidade dos dados do RGPH 2017, então os códigos geográficos abaixo de localidade serão suprimidos nos ficheiros de microdados para o público. Para facilitar a análise introduzimos códigos novos para identificar as áreas de enumeração (AE) e agregados familiares de forma anónima.

A amostra de microdados do RGPH 2017 nos ficheiros para o público vai ser limitado a dados para agregados familiares individuais e seus membros. Não vai incluir dados da população que que vivam nos alojamentos colectivos como hospitais, prisões e dormitórios durante o período do RGPH 2017. A unidade de amostragem é o agregado familiar. Uma amostra estratificada de uma etapa foi usada para a selecção de agregados familiares para a amostra de microdados. A estratificação explícita é ao nível de distrito, estratos urbano e rural, e a estratificação implícita é ao nível de áreas de enumeração. Dentro de cada estrato, os agregados familiares foram seleccionados em base de amostragem aleatória sistemática. A amostra de indivíduos na base de microdados consiste em todos os membros dos agregados familiares seleccionados. A metodologia de amostragem é descrito nesta secção.

Depois de examinar a distribuição dos agregados familiares e população dos dados do RGPH 2017 por província, distrito, postos administrativos e localidades ou vila, foi definido que o distrito deve ser o menor domínio geográfico de análise que pode ser considerado para a base de micrdodados. Em 2017 Moçambique teve 161 distritos, que variam muito em tamanho, entre

1301 agregados familiares para o distrito de Ka Nyaka e 236,063 agregados familiares para o distrito da Cidade da Matola. A cidade de Maputo tem 7 distritos, chamados distritos municipais.

Um total de 38 distritos de Moçambique está acima de 50.000 agregados familiares, e, somente 5 distritos têm uma população acima de 100.000 agregados familiares. Por isso uma amostra de 10% deve ser efectivo para analisar os microdados para a maioria dos distritos.

O procedimento de amostragem para a base de microdados do RGPH 2017 deve ser uniforme para todas as áreas de Moçambique para simplificar a selecção da amostra e a metodologia de ponderação. A análise de microdados ao nível de distrito também deve ser prático para os usuários, dado que muitos programas do governo são administrados ao nível de distrito.

Para melhorar a eficiência estatística da amostra de microdados, a base de agregados familiares do RGPH primeiro foi ordenado por códigos geográficos (distrito, posto administrativo, localidade/vila, bairro e AE), como também por características da habitação. Usando amostragem aleatória sistemática dentro de cada estrato, este ordenamento dos agregados familiares na base proporciona um estraificação ímplícita até o nível de AE, como também pelas características da habitação. As seguintes características das habitações foram identificadas do questionário do RGPH 2017 para ordenar os agregados familiares na base de amostragem, com suas categorias respectivas:

(1) D1 – Tipo de habitação:
    11 – Casa convencional com casa de banho e cozinha dentro de casa
    12 - Casa convencional sem casa de banho ou cozinha dentro de casa
    13 – Flat/ apartamento
    14 – Palhota
    15 – Casa improvisada (barraca, lata, cartão)
    16 – Casa mista
    17 – Casa básica (casa comboio)
    18 – Parte dum edifício comercial
    19 – Outro

(2) D4 – A casa é:
    1 – Própria
    2 – Alugada
    3 – Cedida, imprestada temporariamente
    4 – Outra

(3) D6 – A casa é construída com:
    1 – Bloco de cimento
    2 – Bloco de tijolo
    3 – Madeira/ zinco
    4 – Bloco de adobe

5 – Caniço/ paus/ bambú/ palmeira
6 – Paus maticados (pau a pique)
7 – Lata/ cartão/ papel/ saco/ casca
8 – Outros

(4) D7 – A casa é coberta de:
1 – Laje de betão
2 – Telha
3 – Chapa de lusalite
4 – Chapa de zinco
5 – Capim/ colmo/ palmeira
6 – Outros

(5) D8 – O pavimento da casa é de:
1 – Madeira/ parquet
2 – Mármore/ granulito
3 – Cimento
4 – Mosaico/ tijoleira
5 – Adobe (terra batida)
6 – Sem nada
7 – Outros

(6) D10 – Principal fonte de água para beber:
1 – Água canalizada dentro de casa
2 – Água canalizada fora de casa/ quintal
3 – Água canalizada na casa do vizinho
4 – Água de fontanário / torneira pública
5 – Água de furo / poço protegido com bomba manual
6 – Água de poço protegido sem bomba manual
7 – Água de poço não protegido
8 – Água de nascente
9 – Água de superfície (rio, lago, lagoa)
10 – Água da chuva
11 – Água de tanques camiões / carregada em tambores
12 – Água mineral/ água engarrafada
13 - Outra

(7) D11 – A casa tem:
1 – Retrete com autoclismo dentro de casa
2 – Retrete sem autoclismo fora de casa
3 – Retrete sem autoclismo
4 – Latrina melhorada
5 – Latrina tradicional melhorada
6 – Latrina não melhorada

7 – Sem retrete/ latrina

(8) D13 – Principal fonte de energia que se usa para a iluminação da casa:
1 – Electricidade da rede pública
2 – Gerador/ placa solar
3 – Gás
4 – Petróleo/ parafina/ querosene
5 – Velas
6 – Baterias
7 – Lenha
8 – Pilhas
9 – Outras

Depois de ordenar a base de agregados familiares dentro do estrato urbano e rural de cada distrito, um número aleatório entre 1 e 10 foi seleccionado para identificar o primeiro agregado familiar a ser seleccionado no estrato. Começando com este arranque aleatório, cada décimo agregado familiar foi seleccionado sistemáticamente para o estrato inteiro.

3-Ponderação dos Microdados do RGPH 2017

Dado que a base de microdados é uma amostra dos dados completos do RGPH 2017, vai ser necessário aplicar um ponderador ou factor de expansão aos dados para a tabulação e análise dos dados. Isto vai ser necessário para obter estimativas para a população inteira (ao nível de distrito, província ou nacional). O ponderador para os microdados na amostra é calculado como o inverso da taxa de amostragem dentro de cada distrito. Os ponderadores estão incluídos nos ficheiros de microdados. Programas como SPSS, Stata e CSPro podem ponderar os microdados automáticamente uma vez que é especificado na aplicação.

3.1-Ponderador para Agregados Familiares

A base de amostragem de agregados familiares (AFs) do RGPH 2017 com códigos geográficos consistenciados que foi disponível para a selecção da amostra de AFs para os microdados foi preparado para a Amostra Mãe. Algumas áreas de enumeração (AEs) pequenas foram excluídas desta base, mas tinham menos de 1 porcento da população. Dado que o número total de agregados familiares na base completa do RGPH 2017 era um pouco maior que o número correspondente na base de amostragem da Amostra Mãe, a taxa de amostragem efectiva média por distrito foi um pouco menor de 10%, mas esta amostra de microdados é bem representativa de cada distrito. Para expandir os microdados ao nível da base do RGPH 2007 completo, o ponderador do AF para cada distrito foi calculado usando a seguinte fórmula:

$$W_{(af)D} = \frac{M_D}{m_D}$$

onde:

$W_{(af)D}$ = ponderador para os agregados familiares na amostra de microdados do RGPH 2017 no distrito *D*

M_D = número total de agregados familiares na base completa do RGPH 2017 para o distrito *D*

m_D = número de agregados familiares na amostra de microdados do RGPH 2017 para o distrito *D*

As estimativas do número total de agregados familiares com uma característica particular seria a soma dos ponderadores dos agregados familiares na amostra com esta característica. No caso de estimativas de proporções e razões, o numerador e denominador seriam totais ponderados.

3.2 Ponderador para Dados de Pessoas

A amostra de pessoas para a base de microdados corresponde a todos os membros dos agregados familiares seleccionados para a amostra de 10 porcento. Os usários dos microdados do RGPH 2017 estariam interessados em estimar o número total ou percentagem de pessoas com características diferentes. Alguns quadros podem incluir categorias por sexo e grupos de idade. Então é recomendável calcular os ponderadores para os dados de pessoas de acordo à distribuição da população de cada distrito na base completa do RGPH 2017 por sexo e grupo de idade. Isso vai assegurar consistência entre os resultados ponderados para pessoas na amostra de microdados e os resultados correspondentes dos dados de 100% do RGPH 2017. Então é recomendável calcular também os ponderadores de pessoas por sexo e grupos de idade dentro de cada distrito. Os usuários podem usar os seguintes grupos de idades com intervalos de 5 anos:

1. 0-4 anos
2. 5-9 anos
3. 10-14 anos
4. 15-19 anos
5. 20-24 anos
6. 25-29 anos
7. 30-34 anos
8. 35-39 anos
9. 40-44 anos
10. 45-49 anos
11. 50-54 anos
12. 55-59 anos
13. 60-64 anos
14. 65+ anos

A última categoria de 65 anos ou mais é consistente com os quadros do RGPH 2017. Dado estes

ponderadores por sexo e grupos de idade, temos 28 categorias para os ponderadores de pessoas em cada distrito.

O ponderador de pessoas para cada categoria de sexo e idade dentro do distrito seria calculado usando a seguinte fórmula:

$$W_{Dsg} = \frac{N_{Dsg}}{n_{Dsg}}$$

onde:

W_{Dsg} = ponderador para as pessoas de sexo *s* e grupo de idade *g* dentro do distrito *D*

N_{Dsg} = número total de pessoas de sexo *s* e grupo de idade *g* que vivem em agregados familiares, nos dados de 100% do RGPH 2017 para o distrito *D*

n_{dsg} = número de pessoas de sexo *s* e grupo de idade *g* nos dados do RGPH 2017 para os agregados familiares na amostra de microdados para o distrito *D*

Também é importante ter um número suficiente de pessoas na amostra do distrito para cada grupo de sexo e idade para estabilizar os ponderadores de pessoas. É recomendável combinar qualquer categoria com menos de 20 pessoas na amostra do distrito com o próximo grupo de idade para o mesmo sexo. Dado a população relativamente grande para a maioria dos distritos, devemos encontrar poucos casos de categorias de sexo e grupo de idade com menos de 20 pessoas na amostra.

Dado que a équipe técnica encontrou Localidades com menos de 10 óbitos nos últimos 12 meses à data do Censo e, por ser este o nível mínimo de anonimização dos dados para o público, decidiu-se que, esta base de dados fosse separada da geral e todos os óbitos seleccionados para a amostra ao nível de cada unidade mínima e ao nível nacional. Assim, para o caso de mortalidade, o ponderador é 1 (um) já que todos os casos foram incluídos na amostra.

4-Exactidão das Estimativas da Amostra de Microdados do RGPH 2017

As estimativas ponderadas produzidas da amostra de microdados são sujeitas a dois tipos de erro: erro de amostragem e erro não de amostragem. Os erros de amostragem são devidos ao facto que as estimativas são baseadas numa amostra dos dados do RGPH 2017, não nos dados de 100%. Uma amostra diferente iria proporcionar resultados um pouco diferentes. O erro de amostragem é uma medida da variabilidade entre as estimativas de todas as amostras possíveis, então é um indicador da precisão da estimativa. Os erros não de amostragem incluem todos os outros tipos de erro nos dados do RGPH 2017, incluíndo erros de resposta, codificação e processamento; estes erros também afectam os dados de 100% do RGPH 2017. O controlo de qualidade nas operações do Censo foram elaborados para reduzir o mais possível os erros não

de amostragem.

O erro de amostragem de uma estimativa da amostra de microdados é medido estatísticamente pelo erro padrão, ou raíz quadrado da variância. O erro padrão decresce em relação ao raíz quedrado do tamanho da amostra, e depende do tamanho da população e a taxa de amostragem. A maioria das estimativas de interesse elaborados pelos usuários da amostra de microdados do RGPH 2017 serão em forma do número total ponderado de agregados familiares ou pessoas com certas características.

Dado que a amostra de microdados é seleccionada em uma etapa de amostragem, é possível estimar o erro padrão das estimativas dos dados usando as fórmulas para uma amostra aleatória simples sem reposição. Os quadros apresentados no Anexo A podem ser usados pelos usuários como uma referência simples para obter valores aproximados para os erros de amostragem de estimativas dos microdados, em base de uma amostra aleatória simples sem reposição. As fórmulas usadas para produzir os quadros de erros padrão aproximados apresentados no Anexo A estão especificadas neste documento.

Dado que o número total de agregados familiares na base de dados de 100% do RPGH 2017 para cada domínio geográfico (por exemplo, distrito ou província) é conhecido, a estimativa ponderada do número total de agregados familiares com uma característica particular seria equivalente à:

$$\hat{Y}_{Da} = \sum_{i=1}^{n_d} W_D \times y_{Di} = \frac{N_D}{n_D} \times \sum_{i=1}^{n_d} y_{Di} = N_D \times \frac{a_D}{n_D} = N_D \times \hat{p}_{Da}$$

onde:

$\hat{Y}_{Da}$ = número total ponderado de agregados familiares no domínio geográfico *D* com a característica *a*, estimado dos microdados de 10% do RGPH 2017

W_D = 10 = ponderador para os agregados familiares na amostra de 10% para o domínio *D*

y_{Di} = 1 se o i-éssimo agregado familiar na amostra do domínio *D* tem a característica *a*; ou 0 se não tem

$$a_D = \sum_{i=1}^{n_d} y_{Di}$$

= número de agregados familiares na amostra do domínio *D* com a característica *a*

N_D = número total de agregados familiares nos dados de 100% do RGPH 2017 para o domínio *D*

n_D = número de agregados familiares na amostra de microdados de 10% para o domínio *D* (10% de N_D)

$\hat{p}_{Da}$ = estimativa da proporção de agregados familiares com a característica *a* dentro do domínio *D*, da amostra de microdados de 10% do RGPH 2017

Para este tipo de estimativa de totais da amostra de microdados, a fórmula para o erro padrão de uma amostra aleatória simples sem reposição seria a seguinte:

$$se(\hat{Y}_{Da}) = \sqrt{{N_D}^2 \times (1 - f_D) \times \frac{\hat{p}_{Da} \times (1 - \hat{p}_{Da})}{n_D}}$$

onde:

f_D = 0.1 = taxa de amostragem para os agregados familiares na amostra de microdados do RGPH 2017 no domínio *D*

O componente (1- f_D) é o factor de correção para população finita. Este factor é relacionado ao tipo de amostragem sem reposição, e reduz o erro padrão quando a taxa de amostragem aumenta. No caso de uma taxa de amostragem de 10%, este factor vai reduzir a variância (quadrado do erro padrão) por um factor de 0.9.

Três quadros de erros padrão estão apresentados no Anexo A para estimativas baseadas na amostra de 10% dos agregados familiares do RGPH 2017. Estes erros padrão foram calculados usando esta fórmula para uma amostra aleatória simples sem reposição.

O Quadro 1 no Anexo A apresenta os erros padrão para estimativas de totais dos microdados na amostra de 10%. Para usar este quadro, é necessário saber o valor da estimativa do total para uma característica particular, e o número total de agregados familiares neste domínio (por exemplo, distrito) dos dados definitivos de 100% do RGPH 2017 (isto é, dos quadros publicados). Por exemplo, se a estimativa do número total de agregados familiares com casa construída de *caniço/paus/bambú/palmeira* para um distrito particular é 75.000, e o número total de agregados familiares no distrito, de acordo ao RGPH 2017, é 250.000, primeiro buscamos 75.000 nas categorias das filas, e cruzamos com a coluna correspondente ao total de 250.000 agregados familiares no Quadro 1, e encontramos o valor de 687, que é o erro padrão aproximado para esta estimativa de 75.000. Logo, podemos calcular o intervalo de confiança de 95% usando a seguinte fórmula:

$$\hat{Y}_D \pm 1.96 \times se\left(\hat{Y}_D\right) = 75.000 \pm 1.96 \times 687 = 75.000 \pm 1.347$$

Neste caso o limite inferior do intervalo de confiança de 95% seria 73.653, e o limite superior

seria 76.347. Isto significa que temos uma probabilidade de 95% que o número verdadeiro de casas construídas de *caniço/paus/bambú/palmeira* no distrito é entre 73,653 and 76,347.

Quando o valor de uma estimativa do número total de agregados familiares com certa característica dentro de um domínio é entre as categorias das filas ou colunas do Quadro 1, os valores do erro padrão neste quadro podem ser interpolados para obter o valor aproximado do erro padrão.

Alguns usuários podem preferir calcular o coeficiente de variação (CV) ou erro padrão relativo para a estimativa de um total, que é definido como o erro padrão dividido pelo valor da estimativa. Quadro 2 no Anexo A apresenta os valores dos CVs aproximados para estimativas de totais de uma amostra de 10%. Este quadro de CVs pode ser usado da mesma maneira que o Quadro 1, com o valor da estimativa nas categorias das filas e o número total de agregados familiares no domínio geográfico indicado nas categorias das colunas.

Os usuários também podem estar interessados em obter estimativas da proporção (ou percentagem) de agregados familiares com uma característica particular. No caso da estimativa de uma proporção, o erro padrão para uma amostra aleatória simples sem reposição pode ser calculada usando a seguinte fórmula:

$$se(\hat{p}_{Da}) = \sqrt{(1 - f_D) \times \frac{\hat{p}_{Da} \times (1 - \hat{p}_{Da})}{n_D}}$$

onde:

$se(\hat{p}_{Da})$ = erro padrão para estimativa da proporção de agregados familiares com característica *a* no domínio *D*

Quadro 3 no Anexo A apresenta os erros padrão aproximados para estimativas de percentagens em base de uma amostra de 10% dos agregados familiares do RGPH 2017, calculados usando a fórmula acima. Nas categorias das filas a percentagem e seu complemento estão agrupados juntos dado que têm o mesmo valor para o erro padrão. As categorias das colunas referem-se ao número total de agregados familiares no domínio geográfico, similar aos Quadros 1 e 2.

No caso da estimativa de pessoas, os ponderadores estão calibrados ao número total de pessoas por sexo e grupos de idade ao nível de distrito. Por isso podemos usar as mesmas fórmulas para o cálculo de erros padrão para estimativas do total ou percentagem de pessoas com certas características, só que neste caso N_D e n_D correspondem ao número de pessoas na base de 100% do RGPH 2017 e na amostra de 10%, respectivamente, em vez do número de agregados familiares. Então também podem usar os quadros no Anexo A para estimar os erros padrão, com referência à população.

Dado que usamos amostragem aleatória sistemática para seleccionar a amostra de 10% de

agregados familiares do RPGH 2017 para a base de microdados, isto resultou em uma estratificação implícita ao nível da área de enumeração (AE). Este plano de amostragem deve ter uma melhor eficiência estatística do que uma amostra aleatória simples. Por isso os quadros no Anexo A podem sobre-estimar um pouco os erros padrão e CVs. Ao mesmo tempo, a amostra de pessoas é baseada numa amostra estratificada de conglomerados de uma etapa, dado que a amostra consiste de todas as pessoas que são membros dos agregados familiares seleccionados para a amostra de 10%. Isto é, cada agregado familiar na amostra é considerado um conglomerado de pessoas.

No caso de uma amostra estratificada de conglomerados de uma etapa, o erro padrão em base de uma amostra aleatória simples sem reposição seria multiplicado pelo efeito do desenho (DEFF). O DEFF é um factor que toma em conta a estratificação e conglomeração da amostra. É definido com a variância de uma estimativa em base da amostra actual dividida pela variância correspondente de uma amostra aleatória simples do memso tamanho. É uma medida relativa da eficiência estatística da amostra. Dado que cada agregado familiar na amostra é um conglomerado de pessoas, o efeito do desenho para as estimativas de algumas características de pessoas pode ser um pouco maior que 1 devido aos efeitos de conglomeração. Ao mesmo tempo, as estimativas de características de agregados familiares pode ter efeitos de desenho um pouco menor que 1, dado a estratificação implícita por AE. Neste caso cada AE pode ser considerada um estrato.

Alguns usuários podem preferir calcular os erros padrão e outras medidas de precisão para as estimativas da amostra de microdados usando um software estatístico como SPSS ou Stata. Estes pacotes usam um estimador de variância linearizado por série de Taylor, que toma em conta a estratificação e conglomeração na amostra de microdados. Para cada tipo de software vai ser necessário especificar as variáveis para o estrato, a unidade primária de amostragem (UPA), e o ponderador. Neste caso o estrato seria a AE, e a UPA seria o agregado familiar. Por isso os ficheiros de microdados incluem variáveis de estrato e UPA como também os ponderadores.

Quadro 1. Erros padrão para estimativas de totais da amostra de microdados de 10% do RGPH 2017

Valor da estimativa (total)	Número total de agregados familiares ou população no domínio (dos dados do RGPH 2017)											
	500	**1,000**	**2,500**	**5,000**	**10,000**	**25,000**	**50,000**	**100,000**	**250,000**	**500,000**	**1,000,000**	**5,000,000**
50	20	21	21	21	21	21	21	21	21	21	21	21
100	27	28	29	30	30	30	30	30	30	30	30	30
250	34	41	45	46	47	47	47	47	47	47	47	47
500	-	47	60	64	65	66	67	67	67	67	67	67
1,000	-	-	73	85	90	93	94	94	95	95	95	95
2,500	-	-	-	106	130	142	146	148	149	150	150	150
5,000	-	-	-	-	150	190	201	207	210	211	212	212
10,000	-	-	-	-	-	232	268	285	294	297	298	300
15,000	-	-	-	-	-	232	307	339	356	362	365	367
25,000	-	-	-	-	-	-	335	411	450	462	468	473
75,000	-	-	-	-	-	-	-	411	687	757	790	815
100,000	-	-	-	-	-	-	-	-	735	849	900	939
250,000	-	-	-	-	-	-	-	-	-	1,061	1,299	1,462
500,000	-	-	-	-	-	-	-	-	-	-	1,500	2,012

1,000,000	-	-	-	-	-	-	-	-	-	-	-	2,683

Quadro 2. Coeficientes de variação para estimativas de totais da amostra de microdados de 10% do RGPH 2017

Valor da estimativa (total)	Número total de agregados familiares ou população no domínio (dos dados do RGPH 2017)											
	500	1,000	2,500	5,000	10,000	25,000	50,000	100,000	250,000	500,000	1,000,000	5,000,000
50	40.2%	41.4%	42.0%	42.2%	42.3%	42.4%	42.4%	42.4%	42.4%	42.4%	42.4%	42.4%
100	26.8%	28.5%	29.4%	29.7%	29.8%	29.9%	30.0%	30.0%	30.0%	30.0%	30.0%	30.0%
250	13.4%	16.4%	18.0%	18.5%	18.7%	18.9%	18.9%	18.9%	19.0%	19.0%	19.0%	19.0%
500	-	9.5%	12.0%	12.7%	13.1%	13.3%	13.3%	13.4%	13.4%	13.4%	13.4%	13.4%
1,000	-	-	7.3%	8.5%	9.0%	9.3%	9.4%	9.4%	9.5%	9.5%	9.5%	9.5%
2,500	-	-	-	4.2%	5.2%	5.7%	5.8%	5.9%	6.0%	6.0%	6.0%	6.0%
5,000	-	-	-	-	3.0%	3.8%	4.0%	4.1%	4.2%	4.2%	4.2%	4.2%
10,000	-	-	-	-	-	2.3%	2.7%	2.8%	2.9%	3.0%	3.0%	3.0%
15,000	-	-	-	-	-	1.5%	2.0%	2.3%	2.4%	2.4%	2.4%	2.4%
25,000	-	-	-	-	-	-	1.3%	1.6%	1.8%	1.8%	1.9%	1.9%
75,000	-	-	-	-	-	-	-	0.5%	0.9%	1.0%	1.1%	1.1%
100,000	-	-	-	-	-	-	-	-	0.7%	0.8%	0.9%	0.9%
250,000	-	-	-	-	-	-	-	-	-	0.4%	0.5%	0.6%
500,000	-	-	-	-	-	-	-	-	-	-	0.3%	0.4%
1,000,000	-	-	-	-	-	-	-	-	-	-	-	0.3%

Quadro 3. Erros padrão para estimativas de percentagens da amostra de microdados de 10% do RGPH 2017

Estimativa da percentagem	Número total de agregados familiares ou população no domínio (dos dados do RGPH 2017)												
	500	750	1,000	1,500	2,500	5,000	7,500	10,000	25,000	50,000	100,000	250,000	500,000
2 or 98%	1.9	1.5	1.3	1.1	0.8	0.6	0.5	0.4	0.3	0.2	0.1	0.1	0.1
5 or 95%	2.9	2.4	2.1	1.7	1.3	0.9	0.8	0.7	0.4	0.3	0.2	0.1	0.1
10 or 90%	4.0	3.3	2.8	2.3	1.8	1.3	1.0	0.9	0.6	0.4	0.3	0.2	0.1
15 or 85%	4.8	3.9	3.4	2.8	2.1	1.5	1.2	1.1	0.7	0.5	0.3	0.2	0.2
20 or 80%	5.4	4.4	3.8	3.1	2.4	1.7	1.4	1.2	0.8	0.5	0.4	0.2	0.2
25 or 75%	5.8	4.7	4.1	3.4	2.6	1.8	1.5	1.3	0.8	0.6	0.4	0.3	0.2
30 or 70%	6.1	5.0	4.3	3.5	2.7	1.9	1.6	1.4	0.9	0.6	0.4	0.3	0.2
35 or 65%	6.4	5.2	4.5	3.7	2.9	2.0	1.7	1.4	0.9	0.6	0.5	0.3	0.2
40 or 60%	6.6	5.4	4.6	3.8	2.9	2.1	1.7	1.5	0.9	0.7	0.5	0.3	0.2
50%	6.7	5.5	4.7	3.9	3.0	2.1	1.7	1.5	0.9	0.7	0.5	0.3	0.2