

# Notas metodológicas

---

## Fundamento legal e sigilo das informações

O Censo Demográfico 2000 segue os princípios normativos determinados na Lei nº 5.534, de 14 de novembro de 1968.

Conforme esta lei, as informações são confidenciais e obrigatórias, destinam-se exclusivamente a fins estatísticos e não podem ser objeto de certidão e nem ter eficácia jurídica como meio de prova.

Já a periodicidade dos Censos Demográficos é regulamentada pela Lei nº 8.184, de 10 de maio de 1991, que estabelece um máximo de dez anos para o intervalo intercensitário.

## O Censo Demográfico 2000 no contexto internacional

No planejamento do Censo Demográfico 2000, foram consideradas as recomendações de organismos internacionais e a experiência dos Órgãos Nacionais de Estatística de diversos países.

Desde 1997, o Brasil esteve representado pelo IBGE em seminários e fóruns internacionais de debates sobre censos.

Especialistas do IBGE realizaram visitas técnicas aos principais órgãos de estatística do mundo. Também se contou com a participação de especialistas dessas instituições no Brasil, por meio de missões de cooperação técnica que, em conjunto com as equipes locais, atuaram na formação e treinamento de equipes de análise de dados censitários e na difusão de novas metodologias de apuração.

As Nações Unidas organizaram diversas reuniões, como parte das atividades de apoio à preparação dos censos de população e domicílios na rodada de 2000. O IBGE teve uma participação efetiva na discussão dos conceitos e definições incluídos nas Recomendações Internacionais, elaboradas pelas Nações Unidas e, ainda, nos debates para a confecção dos manuais sobre melhores práticas na área de consistência e imputação de dados censitários. Esta experiência foi aproveitada no planejamento do Censo Demográfico 2000.

O Brasil participa desde 1997 do Projeto do Censo Comum do Mercosul, em conjunto com os órgãos de estatística dos demais países do bloco continental, constituído por Argentina, Brasil, Paraguai e Uruguai, aos quais se somaram Chile e Bolívia. Os institutos de estatística dos países envolvidos, neste Projeto, realizaram acordos de trabalho conjunto, no sentido da obtenção de informações homogêneas para a região, da otimização de recursos metodológicos, humanos e tecnológicos de cada país e de potencializar a troca de experiências. O objetivo do Projeto do Censo Comum do Mercosul é contribuir para a homogeneização gradativa das estatísticas econômicas e sociodemográficas entre os países do bloco e a geração de um banco de dados. Nesse sentido, os representantes dos censos da rodada de 2000 da região, entre outros trabalhos, discutiram a harmonização de um conjunto básico de características e de sua conceituação, estando em fase de definição e geração de um banco de dados único com informações dos seis países do grupo.

## **Âmbito**

O Censo Demográfico 2000 abrangeu as pessoas residentes, na data de referência, em domicílios do Território Nacional.

As embaixadas, consulados e representações do Brasil no exterior são considerados Território Nacional, porém não foram incluídos no censo. Atualmente, a maioria dos funcionários brasileiros reside em domicílios fora das representações diplomáticas.

## **Data de referência**

A investigação dos domicílios e das pessoas neles residentes teve como data de referência o dia 1º de agosto de 2000.

## **Período de coleta**

A coleta do Censo Demográfico 2000 foi realizada por cerca de 200 mil recenseadores e 30 mil supervisores, no período de 1º de agosto a 30 de novembro de 2000.

## **Base territorial**

Base territorial é a denominação dada ao sistema integrado de mapas, cadastros e bancos de dados, construído segundo metodologia própria para dar organização e sustentação espacial às atividades de planejamento operacional, coleta e apuração de dados e divulgação de resultados do Censo Demográfico 2000.

O setor censitário é a unidade territorial criada para fins de controle cadastral da coleta. Para este censo, o Território Nacional foi dividido em 215 811 áreas contíguas, respeitando-se os limites da divisão político-administrativa, do quadro urbano e rural legal e de outras estruturas territoriais de interesse, além dos parâmetros de dimensão mais adequados à operação de coleta.

O sistema cartográfico, elaborado exclusivamente para os fins estatísticos, é composto por uma série de mapas municipais e outra de mapas de localidades. Os mapas municipais, elaborados em escalas topográficas, apresentam a cobertura completa do município, sua divisão distrital e os setores rurais. Os mapas de localidades, elaborados em escalas cadastrais, apresentam a cobertura das cidades e vilas, suas divisões intra-urbanas e os setores urbanos. Os mapas de setores censitários, uma terceira série, são mapas individuais em escalas diversas, oriundos dos respectivos mapas de localidades ou municipais.

O mapeamento do Censo Demográfico 2000 foi construído utilizando-se tecnologia digital e teve como principais etapas de trabalho: a aquisição das bases cartográficas disponíveis, a atualização cartográfica; a representação da divisão político-administrativa, das divisões intra-urbanas e dos setores censitários; e a conversão para o padrão digital IBGE. Constituiu-se, portanto, no primeiro acervo digital de cobertura nacional, produzido pela instituição.

O sistema alfanumérico da base territorial é composto pelos arquivos de referência territorial: identificação (código e nome) das Unidades da Federação, dos municípios, dos distritos, dos subdistritos, dos bairros legalmente estabelecidos e de outras estruturas territoriais de interesse; descrição do perímetro dos setores censitários; e identificação de cada setor censitário e de seus atributos espaciais.

## Divisão territorial

A organização político-administrativa da República Federativa do Brasil compreende a União, o Distrito Federal, os estados e os municípios, todos autônomos nos termos da Constituição Federal de 5 de outubro de 1988.

### Distrito Federal

É a unidade autônoma onde tem sede o Governo Federal com seus poderes Executivo, Legislativo e Judiciário. Tem as mesmas competências legislativas reservadas aos estados e municípios e é regido por lei orgânica, sendo vedada sua divisão em municípios.

Brasília é a Capital Federal.

### Estados

Em número de 26, os estados constituem as unidades de maior hierarquia dentro da organização político-administrativa do País; são subdivididos em municípios e podem incorporar-se entre si, subdividir-se ou desmembrar-se para se anexarem a outros, ou formarem novos estados ou territórios federais, mediante aprovação da população diretamente interessada, através de plebiscito, e do Congresso Nacional, por lei complementar. Organizam-se e regem-se pelas constituições e leis que adotarem, observados os princípios da Constituição Federal.

A localidade que abriga a sede do governo denomina-se Capital.

## Municípios

Os municípios instalados até 1º de agosto de 2000 eram 5 507 (incluindo o Distrito Estadual de Fernando de Noronha e o Distrito Federal). Os municípios constituem as unidades autônomas de menor hierarquia dentro da organização político-administrativa do Brasil. Sua criação, incorporação, fusão ou desmembramento se faz por lei estadual, observada a continuidade territorial, a unidade histórico-cultural do ambiente urbano e os requisitos previstos em lei complementar estadual. Estas transformações dependem de consulta prévia às populações diretamente interessadas, através de plebiscito. No Censo Demográfico 2000 a população referente aos municípios criados e instalados após aquela data foi computada nos municípios que lhe deram origem.

Regem-se por leis orgânicas, observados os princípios estabelecidos na Constituição Federal e na constituição do estado onde se situam, e podem criar, organizar e suprimir distritos, observada a legislação estadual.

A localidade onde está sediada a Prefeitura Municipal tem a categoria de cidade.

## Divisão regional

As divisões regionais, em seus diferentes níveis, constituem parte da missão institucional do IBGE e têm a finalidade básica de subsidiar o levantamento e a divulgação de dados estatísticos. A primeira divisão regional, estabelecida em 1942, baseava-se nas características do quadro natural, privilegiando o nível macrorregional.

Era constituída por cinco Grandes Regiões: Norte, Nordeste, Leste, Sul e Centro-Oeste, sendo que a Região Nordeste se subdividia em Nordeste Ocidental e Nordeste Oriental e a Região Leste em Leste Setentrional e Leste Meridional. Em 1945, foram incorporados a esta estrutura níveis mais desagregados de divisão regional: 30 regiões, 83 sub-regiões e 198 zonas geográficas ou fisiográficas.

As grandes transformações econômico-sociais, ocorridas nas décadas de 1950 e de 1960, impuseram a necessidade de atualização do quadro de divisões regionais, utilizando como base características de homogeneidade de produção. Em 1969, foi divulgado, então, novo quadro composto por cinco Grandes Regiões: Norte, Nordeste, Sudeste, Sul e Centro-Oeste, subdivididas em 360 microrregiões homogêneas. O quadro de divisões regionais em nível intermediário só foi elaborado na década de 1970 e divulgado em 1976, totalizando 86 mesorregiões homogêneas.

O conjunto das Grandes Regiões, adotado em 1969, está vigente até hoje. Suas alterações decorrem de mudanças no quadro político-administrativo do País com a criação do Estado do Mato Grosso do Sul, incorporado à Região Centro-Oeste e do Estado do Tocantins, incorporado à Região Norte.

## Regiões Metropolitanas

A identificação de regiões metropolitanas, no Brasil, foi iniciada na década de 1960 e institucionalizada pelo Congresso Nacional nos anos de 1973 e de 1974, tendo sido definidas nove Regiões Metropolitanas: Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba e Porto Alegre.

A partir de 1988, a Constituição Federal do Brasil facultou aos estados a instituição de regiões metropolitanas, "constituídas por agrupamentos de municípios

limítrofes, com o objetivo de integrar a organização, o planejamento e a execução de funções públicas de interesse comum" (Artigo 25, Parágrafo 3º). Assim, a partir de 1998, as Unidades da Federação, buscando solucionar problemas de gestão do território estadual, definiram um total de 22 regiões metropolitanas.

## Tratamento dos dados

Todos os dados dos volumes temáticos passaram pelo processo de crítica eletrônica, cuja finalidade é eliminar inconsistências entre as informações dos diversos quesitos do questionário que podem ter sua origem na coleta dos dados ou na fase de reconhecimento de marcas e caracteres:

1. Na crítica das informações referentes à Estrutura Familiar, Gênero e Nupcialidade utilizou-se o Sistema NIM (*New Imputation Methodology*) desenvolvido pelo *Statistics Canada* que tratava os domicílios com até seis moradores e adaptado pelo IBGE para os domicílios com até oito moradores. Os registros com erro foram corrigidos através de imputação gerada pelo sistema.

Para os domicílios com mais de oito moradores, foi utilizado o Sistema IMPS (*Integrated Microcomputer Processing System*) desenvolvido pelo *Bureau of Census*, EUA. Os registros com erro foram corrigidos a partir de regras preestabelecidas, com intervenção de operadores.

2. Os demais temas do questionário foram tratados através do Sistema DIA (*Detección e Imputación Automática de Errores para Datos Cualitativos*) desenvolvido pelo Instituto Nacional de Estadísticas da Espanha - INE -, além de um conjunto de críticas determinísticas, gerado pela análise dos dados antes do processo de imputação. As variáveis de rendimento passaram por um processo inicial de crítica dentro do Sistema DIA, sendo que os registros daí provenientes com algum campo com valor correspondente a ignorado receberam imputação, conforme descrito no item a seguir.

3. As variáveis de rendimento foram imputadas com o uso da técnica de Árvores de Regressão (BREIMAN et al., 1984). O objetivo da utilização dessa técnica foi o de corrigir a não-resposta diferencial nos quesitos de rendimento. Inicialmente, os moradores foram estratificados com respeito aos seus rendimentos. Para isso, foram usadas variáveis presentes no questionário da amostra e outras delas derivadas. Em seguida, para cada morador que não declarou rendimento foi determinado o estrato onde se encontrava e seu rendimento foi imputado pelo rendimento de um morador presente no mesmo estrato selecionado aleatoriamente.

As tabelas 1.1.14, 2.1.1.14 a 2.5.1.14 e 3.1.1.14 a 3.27.1.14 do CD-ROM apresentam grupos de ocupação do trabalho principal, considerando-se a classificação adotada no Censo Demográfico 1991. As tabelas 1.1.16, 2.1.1.16 a 2.5.1.16 e 3.1.1.16 a 3.27.1.16 do CD-ROM apresentam setores de atividade do trabalho principal, considerando-se a classificação adotada no Censo Demográfico 1991. Em função de restrições operacionais advindas dos sistemas de codificação e crítica adotados no Censo Demográfico 2000, as duas variáveis do Censo Demográfico 1991 (ocupação e atividade) não passaram por procedimentos de crítica.

## **Classificações utilizadas para Trabalho e Rendimento**

Os quesitos sobre ocupação e atividade têm, no questionário, formato texto, ou seja, não são pré-codificados. Assim, para efeito de apuração, estes quesitos receberam um tratamento, antes da tabulação, que teve por objetivo a atribuição, ao texto preenchido pelos recenseadores a partir das respostas dos entrevistados, do código da categoria a que correspondiam na classificação usada. Esta etapa denomina-se codificação e foi realizada através de um sistema informatizado especialmente desenvolvido para o Censo Demográfico 2000.

Este sistema consiste, em linhas gerais, da aplicação automática de códigos aos textos que encontram um único correspondente em um banco de dados denominado banco descritor ou, no caso de multiplicidade de textos no banco descritor para o texto registrado pelo Recenseador, a aplicação de códigos recebe o nome de codificação assistida na medida em que o operador seleciona, sob a supervisão de um orientador, o texto no banco descritor que mais se adequa ao preenchido pelo recenseador.

### **Classificação de ocupações**

O Censo Demográfico 2000 utilizou, pela primeira vez, a Classificação Brasileira de Ocupações, adaptada para as pesquisas domiciliares - CBO Domiciliar. Como esta classificação representou uma grande mudança em relação àquela utilizada no censo anterior, as ocupações foram também codificadas segundo a classificação utilizada no Censo Demográfico 1991 para permitir a avaliação das diferenças existentes.

A CBO Domiciliar mantém-se idêntica à CBO no nível mais agregado - grande grupo. Algumas famílias ocupacionais, grupos e subgrupos principais foram reagrupados tendo em vista as dificuldades de serem captados com precisão nas pesquisas domiciliares. Desta forma, o Censo Demográfico 2000 adere à padronização nacional e internacional de classificação de ocupações, uma vez que a Classificação Brasileira de Ocupação - CBO - tem como referência a Classificação Internacional Uniforme de Ocupação - CIUO-88.

Figura em Anexo a descrição mais detalhada da metodologia utilizada na classificação de ocupações no Censo Demográfico 2000.

A estrutura completa da CBO adaptada para atender às pesquisas domiciliares e a relação dos códigos usada no Censo Demográfico 1991 são apresentadas como anexos no CD-ROM.

### **Classificação de atividades**

O Censo Demográfico 2000 utilizou, pela primeira vez, a Classificação Nacional de Atividades Econômicas, adaptada para as pesquisas domiciliares - CNAE Domiciliar.

Com o objetivo de garantir que a transição para uma nova base classificatória, com estrutura bastante diferenciada da anterior, pudesse ser feita possibilitando a comparação entre elas, a atividade econômica foi codificada simultaneamente em duas classificações: Classificação Nacional de Atividades Econômicas, adaptada para as pesquisas domiciliares - CNAE Domiciliar e a Classificação de Atividades do Censo Demográfico 1991.

A CNAE Domiciliar mantém-se idêntica à CNAE nos níveis mais agregados - seção e divisão, com exceção das divisões do comércio em que não distingue o varejo e atacado - reagrupa classes onde o detalhamento foi considerado inadequado para pesquisas domiciliares e desagrega algumas atividades de serviços que têm nestas pesquisas sua única fonte de cobertura. Desta forma, o Censo Demográfico 2000 adere à padronização nacional e internacional da classificação de atividades econômicas, garantindo uma maior comparabilidade das informações relativas à

atividade econômica no nível nacional (comparações com outras fontes nacionais) e com informações de outros países (comparações internacionais).

A metodologia da classificação foi discutida no âmbito do projeto Censo Comum do Mercosul, resultando na definição da Classificação de Atividades Econômicas para Encuestas Sociodemográficas del Mercosur - CAES, articulada com a Classificação Industrial Internacional Uniforme - CIU Rev 3, dentro dos princípios usados na CNAE Domiciliar.

Figura em Anexo a descrição mais detalhada da metodologia utilizada na classificação de atividades no Censo Demográfico 2000.

A estrutura completa da CNAE Domiciliar (CNAE adaptada às pesquisas domiciliares) e a relação dos códigos usada no Censo Demográfico 1991 são apresentadas como anexos no CD-ROM.

## **Plano de divulgação Trabalho e rendimento do Censo Demográfico 2000**

A publicação dos resultados referentes ao Trabalho e Rendimento do Censo Demográfico 2000, captadas por meio do Questionário da Amostra (em Anexo) está estruturada por níveis geográficos e temas.

## **Aspectos de amostragem**

### **Planejamento e seleção da amostra**

Desde 1960, vem sendo utilizada a técnica de amostragem na coleta do Censo Demográfico do Brasil. O desenho amostral adotado compreende a seleção sistemática e com equi-probabilidade, dentro de cada setor censitário, de uma amostra dos domicílios particulares e das famílias ou componentes de grupos conviventes recenseados em domicílios coletivos, com fração amostral constante para setores de um mesmo município. Para a realização do Censo Demográfico 2000, da mesma forma que no Censo Demográfico 1991, foram definidas duas frações amostrais distintas: 10% para os municípios com população estimada<sup>1</sup> superior a 15 000 habitantes e 20% para os demais municípios, conforme pode ser visto em Silva e Bianchini (1990).

Na coleta das informações do Censo Demográfico 2000, foram usados dois modelos de questionário:

1. um questionário básico aplicado nas unidades não-selecionadas para a amostra e contendo perguntas referentes às características que foram investigadas para 100% da população; e
2. um segundo questionário aplicado somente nos domicílios selecionados para a amostra contendo, além das perguntas que também constam do questionário básico, outras perguntas mais detalhadas sobre características do domicílio e de seus moradores, referentes aos temas religião, cor ou raça, deficiência, migração, escolaridade, fecundidade, nupcialidade, trabalho e rendimento.

Em todo o Território Nacional foram selecionados 5 304 711 domicílios para responder ao questionário da amostra, o que significou uma fração amostral da ordem de 11,7%. Nesses domicílios, foram levantadas as informações para todos os seus moradores, totalizando 20 274 412 pessoas. Em Anexo (CD-ROM) é apresentado o quadro com a Divisão Territorial em 2000, com seus respectivos códigos e frações de amostragem efetivas.

<sup>1</sup> Estimativas de população para o ano de 2000, baseadas nas projeções realizadas pela Coordenação de População e Indicadores Sociais da Diretoria de Pesquisas do IBGE.

## **Expansão da Amostra**

### **Ponderação das unidades da amostra**

Para expansão dos dados coletados pelos questionários da amostra do Censo Demográfico 2000, foram calculados pesos para cada um dos domicílios pesquisados, sendo tais pesos atribuídos ao próprio domicílio e a cada um de seus moradores.

O método utilizado para obtenção dos pesos foi um processo de calibração em relação a um conjunto de variáveis auxiliares (restrições) para as quais se conhecem os totais populacionais, já que tais variáveis auxiliares foram levantadas pelo questionário básico. A calibração buscou ajustar os pesos iniciais (inverso da fração amostral de domicílios) de maneira que, dentro de uma determinada área geográfica, denominada área de ponderação, ao se aplicar os pesos calibrados às variáveis auxiliares fossem obtidos os totais já conhecidos para todas as unidades da população que constituem o universo da pesquisa. Dessa maneira, além de melhorar a precisão dos estimadores, obtêm-se estimativas mais consistentes para as variáveis pesquisadas somente pelo questionário da amostra.

O cálculo dos pesos calibrados foi baseado no método dos Mínimos Quadrados Generalizados — MQG —, porém com a imposição de limites nos pesos finais para evitar pesos muito pequenos ou muito grandes. Os limites utilizados foram 1, de maneira que um domicílio representasse pelo menos o próprio, e 25 ou 50, de acordo com a fração amostral correspondente à área de ponderação em questão. Sem a utilização desses limites, o método MQG pode gerar pesos negativos ou muito grandes, o que não teria sentido prático.

A metodologia para utilização do método MQG baseou-se em proposta de Bankier (1990) e, para sua implementação, um sistema em linguagem SAS foi desenvolvido por técnicos do IBGE.

O produto final da aplicação dessa metodologia é um peso ajustado para cada unidade domiciliar da amostra, ou seja, cada um dos questionários da amostra, que é repetido nos registros de cada pessoa moradora na unidade domiciliar.

### **A definição das áreas de ponderação**

Define-se área de ponderação como sendo uma unidade geográfica, formada por um agrupamento mutuamente exclusivo de setores censitários, para a aplicação dos procedimentos de calibração das estimativas com as informações conhecidas para a população como um todo.

Foram definidas, para todo o Brasil, 9 336 áreas de ponderação e, tal como nos censos anteriores, a metodologia de expansão da amostra foi aplicada independentemente para cada uma delas.

O tamanho dessas áreas, em termos de número de domicílios e de população, não pode ser muito reduzido, sob pena de perda de precisão de suas estimativas. As áreas de ponderação foram definidas considerando essa condição e, também, os níveis geográficos mais detalhados da base operacional, como forma de atender a demandas por informações em níveis geográficos menores que os municípios.

Para o Censo Demográfico 2000, foram usados métodos e sistemas automáticos de formação de áreas de ponderação que conjugam critérios, tais como: tamanho (para permitir estimativas com qualidade estatística em áreas pequenas); contigüidade (no sentido de serem constituídas por conjuntos de setores limítrofes com sentido geográfico); e homogeneidade em relação a um conjunto de características populacionais e de infra-estrutura conhecidas.

As áreas de ponderação foram criadas considerando os seguintes critérios:

1. o maior nível geográfico utilizado é o município; isto significa que uma área de

- ponderação é composta por setores censitários dentro de um único município, podendo ser o próprio município;
2. o menor tamanho de uma área de ponderação não-municipal é de 400 domicílios particulares ocupados na amostra;
  3. em alguns municípios as áreas de ponderação foram definidas considerando suas divisões administrativas, sempre respeitando o critério de tamanho mínimo; alguns municípios tiveram apenas duas áreas definidas: uma considerando todos os setores do distrito-sede e outra considerando todos os setores dos demais distritos; em outros municípios, cujos distritos possuem tamanho que ferem o critério de tamanho mínimo, também foram definidas duas áreas: uma constituída por todos os seus setores urbanos e outra por todos os seus setores rurais, mesmo que isso significasse setores não-contíguos;
  4. para um conjunto de municípios grandes em termos de população, foi feita uma consulta aos órgãos de planejamento municipal para que as áreas de ponderação fossem definidas em conjunto. Nesses municípios também foram considerados os critérios de tamanho mínimo e de contigüidade do conjunto de setores para a definição das áreas de ponderação; e
  5. os municípios que não se enquadram nas situações de 1 a 4, acima, tiveram suas áreas de ponderação definidas automaticamente, usando uma metodologia de agregação de setores implementada por meio de um sistema computacional que faz uso de informações georreferenciadas especialmente desenvolvido; essa metodologia considera os critérios de tamanho mínimo, vizinhança entre os setores e a homogeneidade dos setores em relação a um conjunto de características conhecidas para o universo no nível dos setores. Entre as 15 variáveis utilizadas constavam, por exemplo: rendimento médio dos responsáveis pelos domicílios no setor, número médio de pessoas por domicílio particular permanente, proporção de domicílios particulares permanentes ligados à rede geral de água e média de anos de estudo dos responsáveis por domicílios.

Em Anexo (CD-ROM) é apresentada a relação dos 484 municípios que tiveram mais de uma área de ponderação. Os demais 5 023 municípios tiveram apenas uma área de ponderação.

## **A definição das variáveis auxiliares**

A escolha das variáveis auxiliares cujos valores são utilizados como restrições no processo de ajustamento do qual decorrem os pesos é um aspecto importante do método aplicado. A forma ou prioridade de tratamento dessas variáveis, sobretudo quando não existe uma solução que atenda simultaneamente a todas as restrições, é outro ponto sensível do método.

As variáveis auxiliares constituem um subconjunto das variáveis comuns à amostra e ao universo e são referentes a características de domicílios ou de pessoas, apesar de o ajustamento ser realizado de forma a fornecer pesos para cada uma das unidades domiciliares.

A metodologia de ajuste de um modelo linear generalizado multivariado envolve cálculos com matrizes, inclusive inversão. Por essa razão, as restrições definidas, que por sua vez dão origem a uma dessas matrizes, devem satisfazer algumas condições essenciais, sendo a principal delas a de não serem linearmente dependentes (redundantes). Além disso, é também considerado o conceito de restrições quase linearmente dependentes (e, portanto, quase redundantes), que afetam a estabilidade da solução do modelo.

Outras duas condições impostas para a aplicação dessa metodologia referem-se à sua significância estatística. O tamanho da restrição, medido como o número de

domicílios aos quais a restrição se aplica a uma dada área de ponderação, não deve ser muito pequeno sob pena de tornar instável o processo de estimação. Quando uma restrição não atinge um número mínimo de unidades domiciliares, fixado em função da fração de amostragem, essa restrição é considerada rara.

Além disso, uma restrição definida pode causar a obtenção de um peso muito grande ou muito pequeno, quando comparado com o peso médio esperado em função da fração amostral adotada na área de ponderação ou até um peso negativo, constituindo-se em restrição geradora de peso extremo.

Dessa forma, o programa de ajuste do modelo incorpora procedimentos de eliminação de restrições que se enquadrem nas condições acima, observando a ordem que segue: restrições raras, restrições redundantes, restrições quase redundantes e restrições responsáveis por pesos extremos.

Convém ressaltar que a eliminação de restrições pode implicar diretamente no fato de não se ter a garantia da calibração desejada para as variáveis eliminadas para a presente área de ponderação.

As restrições inicialmente definidas para a aplicação da metodologia MQG, para cada uma das áreas de ponderação, encontram-se na relação a seguir.

**Para as unidades domiciliares**

Número total de pessoas

Número total de unidades domiciliares

Número de pessoas do sexo masculino

Número de pessoas na faixa de idade de 0 a 4 anos

Número de pessoas na faixa de idade de 5 a 9 anos

Número de pessoas na faixa de idade de 10 a 14 anos

Número de pessoas na faixa de idade de 15 a 19 anos

Número de pessoas na faixa de idade de 20 a 24 anos

Número de pessoas na faixa de idade de 25 a 29 anos

Número de pessoas na faixa de idade de 30 a 34 anos

Número de pessoas na faixa de idade de 35 a 39 anos

Número de pessoas na faixa de idade de 40 a 44 anos

Número de pessoas na faixa de idade de 45 a 49 anos

Número de pessoas na faixa de idade de 50 a 59 anos

Número de pessoas na faixa de idade de 60 a 69 anos

Número de pessoas na faixa de idade de 70 anos ou mais

Número de pessoas do sexo masculino na faixa de idade de 0 a 4 anos

Número de pessoas do sexo masculino na faixa de idade de 5 a 9 anos

Número de pessoas do sexo masculino na faixa de idade de 10 a 14 anos

Número de pessoas do sexo masculino na faixa de idade de 15 a 19 anos

Número de pessoas do sexo masculino na faixa de idade de 20 a 24 anos

Número de pessoas do sexo masculino na faixa de idade de 25 a 29 anos

Número de pessoas do sexo masculino na faixa de idade de 30 a 34 anos

Número de pessoas do sexo masculino na faixa de idade de 35 a 39 anos

Número de pessoas do sexo masculino na faixa de idade de 40 a 44 anos

Número de pessoas do sexo masculino na faixa de idade de 45 a 49 anos

Número de pessoas do sexo masculino na faixa de idade de 50 a 59 anos

Número de pessoas do sexo masculino na faixa de idade de 60 anos ou mais

Número de pessoas moradoras na situação urbana

Número de pessoas do sexo feminino moradoras na situação urbana

Número de pessoas do sexo feminino moradoras na situação rural

### **Para os domicílios particulares permanentes ocupados**

Número de pessoas do sexo masculino que são responsáveis pelo domicílio

Número total de pessoas

Número total de domicílios

Número de domicílios urbanos

Número de domicílios com 1 ou 2 moradores

Número de domicílios com 3 moradores

Número de domicílios com 4 moradores

Número de domicílios com 5 moradores

Número de domicílios com 6 ou mais moradores

### **Análise da qualidade da calibração**

As restrições acima apresentadas foram agrupadas em dez conjuntos que foram utilizados em ordem de prioridade. O primeiro conjunto foi formado por todas as restrições, como listadas, e os demais formados pela agregação de faixas etárias, agregação de faixas de moradores por domicílio ou mesmo a retirada de grupos de restrições.

No cálculo dos pesos calibrados, para cada área de ponderação, foi utilizado inicialmente o conjunto de restrições número 1. Quando não se obteve uma solução satisfatória, a área foi processada novamente utilizando o conjunto 2 e assim sucessivamente até o conjunto 10, caso anteriormente não tenha sido atingida a qualidade de ajuste adequada.

A análise da qualidade do ajuste (calibração) era feita automaticamente pelo sistema através das diferenças entre os valores populacionais conhecidos para as restrições e os valores estimados utilizando-se os pesos calculados. Para cada grupo de restrições, foram definidos limites específicos tolerados para essas diferenças.

Para as áreas de ponderação onde não ocorreu o ajuste para nenhum dos dez conjuntos de restrições, o sistema automaticamente escolheu o conjunto que proporcionou o melhor ajuste, no sentido de minimizar a soma dos quadrados das diferenças entre o valor conhecido das restrições e o valor estimado para essas mesmas restrições.

Deve-se ressaltar que para algumas áreas de ponderação onde houve um desequilíbrio forte entre a fração amostral de domicílios e a fração amostral de pessoas, ou seja, em áreas onde os números médios de pessoas por domicílio no universo e na amostra diferem muito pode ter ocorrido falta de ajuste na variável total de pessoas. Portanto, é possível a ocorrência de divergência entre o valor do número de pessoas, calculado através da expansão da amostra e o valor verificado no universo, para alguns municípios.

Para o total de 9 336 áreas de ponderação definidas para o Brasil, mais de 91% delas tiveram solução para o conjunto 1 de restrições.

No final do processo, foi garantido que pelo menos a restrição Número Total de Domicílios fosse respeitada para todas as áreas de ponderação.

### **Estimação de totais**

As estimações de totais para domínios de interesse, como, por exemplo, as células de uma tabela, devem ser feitas utilizando-se, para cada unidade (pessoa, família ou domicílio), o peso correspondente, que foi determinado para cada unidade domiciliar da amostra e atribuído a cada pessoa dessa unidade. Assim, para estimar o total de uma característica  $y$  utiliza-se o estimador,  $\hat{Y}$ , definido por:

$$\hat{Y} = \sum_{i=1}^n p_i y_i$$

onde:

$P_i$  é o peso associado à *i*-ésima unidade da amostra no domínio em questão;  
 $y_i$  é o valor de  $y$  associado à *i*-ésima unidade da amostra no domínio; e  
 $n$  é o número de unidades na amostra do domínio em questão.

Dessa forma, é possível calcular estimativas para quaisquer variáveis investigadas no censo, independente de serem de pessoas, famílias ou domicílios.

Os pesos calculados com a metodologia adotada não são necessariamente inteiros e não devem ser substituídos por pesos inteiros para não provocar a quebra na consistência das restrições efetivamente utilizadas no ajuste no modelo. O uso de pesos fracionários preserva o método de expansão da amostra e produz resultados mais precisos do ponto de vista estatístico. Assim, para o cálculo das estimativas das tabelas de divulgação do censo, foi utilizado o peso fracionário com 8 casas decimais, sendo, então, arredondadas as estimativas resultantes.

Para obter consistência com as tabelas de divulgação do censo, é necessário que as estimativas sejam calculadas em cada célula básica da tabela e as linhas e colunas de totais e subtotais sejam obtidas por soma das estimativas básicas correspondentes, após terem sido arredondadas. Uma consequência desse procedimento é que os totais de uma mesma característica podem diferir ligeiramente de uma tabela para outra, ou mesmo dentro da própria tabela, em função do arredondamento das parcelas. Da mesma forma, as estimativas para o Brasil podem diferir dos valores obtidos pela soma das estimativas correspondentes publicadas por Unidade da Federação.

### **Avaliação da precisão das estimativas**

As conclusões de uma pesquisa por amostra devem ser apoiadas nas estimativas produzidas. Essas, por sua vez, embutem um erro amostral que deve situar-se dentro de um nível de confiança fixado para a tomada de decisão. Assim, a avaliação dos erros amostrais é um ponto fundamental, pois dele decorre o grau de confiança nas conclusões analíticas que subsidiam a tomada de decisão. Para cada estimativa derivada da pesquisa, é possível obter uma medida de precisão que auxilia na análise e interpretação dos dados resultantes da pesquisa.

Os erros amostrais podem ser avaliados através das estimativas dos coeficientes de variação ou dos erros padrão calculados a partir das estimativas das variâncias.

Embora seja possível estimar os erros amostrais de acordo com a metodologia usada na obtenção dos pesos, o método direto é bastante complexo (SÄRNDAL; SWENSSON; WRETMAN, 1992). Sugere-se, então, um método simples e rápido para obtenção de uma aproximação do erro padrão da estimativa, que pode ser usado para a construção de intervalos com níveis de confiança fixados. Como a amostra usada no Censo Demográfico 2000 é bastante grande e os domicílios se distribuem de forma aleatória dentro de cada setor censitário, pode-se aproximar o cálculo do erro padrão, segundo Cochran (1977), supondo que o esquema de seleção da amostra foi de amostragem aleatória simples sem reposição. Dessa maneira, um estimador do erro padrão de um estimador de total de uma característica  $y$  representado por  $\hat{Y}$ , é dado por:

$$ep(\hat{Y}) = \sqrt{\frac{(1-f)}{f} N s^2(y)}$$

onde:

$ep(\hat{Y})$  é o erro padrão do estimador de total,  $\hat{Y}$ , para o domínio em questão;

$f$  é a fração efetiva de amostragem observada no domínio em questão;

$N$  é o total de unidades da população no domínio em questão;

$s^2(y) = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$  é a variância amostral para o domínio em questão;

$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$  é a média amostral no domínio em questão;

$y_i$  é o valor da característica  $y$  na  $i$ -ésima unidade da amostra no domínio;

$n$  é o total de unidades da amostra no domínio em questão.

Em anexo (CD-ROM) são dadas as frações amostrais, em porcentagem, observadas para Brasil, Grandes Regiões, Unidades da Federação, mesorregiões, microrregiões e municípios segundo divisão territorial em 2000.

Como a maior parte das estimativas derivadas das informações coletadas na amostra do Censo Demográfico 2000 é proveniente de variáveis categóricas, para as quais  $y_i$  assume somente os valores 0 (se a unidade não pertence à categoria em questão) ou 1 (se a unidade pertence à categoria em questão), a expressão do estimador  $ep(\hat{Y})$  reduz-se a:

$$ep(\hat{Y}) = \sqrt{\frac{(1-f)\hat{Y}(N-\hat{Y})}{Nf-1}}$$

Em Anexo (CD-ROM), são apresentados quadros com valores de erro padrão calculados para alguns valores de estimativas de características de pessoas e domicílios para o Brasil, Grandes Regiões e Unidades da Federação.

O erro padrão é utilizado para construir intervalos de confiança que conterão o valor do total populacional<sup>2</sup>,  $Y$ , com uma certa probabilidade decorrente do nível de confiança desejado na tomada de decisão, ou seja,

$$P[\hat{Y} - z_{\alpha/2} ep(\hat{Y}) < Y < \hat{Y} + z_{\alpha/2} ep(\hat{Y})] = 1 - \alpha$$

onde:

$\alpha$  é o nível de significância e  $(1 - \alpha)$  é o nível de confiança;

$z_{\alpha/2}$  é a abscissa da distribuição Normal padrão no ponto  $\alpha/2$ .

Assim, para um nível de confiança de 95% tem-se  $z_{\alpha/2} = 1,96$  e o intervalo de confiança é dado por:

$$[\hat{Y} - 1,96ep(\hat{Y}); \hat{Y} + 1,96ep(\hat{Y})]$$

<sup>2</sup>O valor da população é, de um modo geral, desconhecido, exceto para as características investigadas censitariamente.

O Quadro 1, a seguir, apresenta o erro padrão para alguns tamanhos de estimativas para domicílios e pessoas, no nível Brasil. Em anexo (CD-ROM) encontram-se disponíveis quadros com informações para Brasil, Grandes Regiões e Unidades da Federação.

**Quadro 1 - Erro padrão aproximado para alguns tamanhos de estimativas para características de pessoas e domicílios - Brasil**

Características de pessoas		Características de domicílios	
Tamanho da estimativa	Erro padrão aproximado	Tamanho da estimativa	Erro padrão aproximado
100	28	100	28
500	62	500	62
1 000	87	1 000	87
2 000	123	2 000	123
5 000	195	5 000	195
10 000	275	10 000	275
20 000	389	20 000	389
50 000	615	50 000	615
100 000	870	100 000	870
150 000	1 066	150 000	1 064
200 000	1 230	200 000	1 228
500 000	1 944	250 000	1 373
1 000 000	2 745	500 000	1 936
2 000 000	3 870	1 000 000	2 722
3 000 000	4 726	2 000 000	3 807
4 000 000	5 440	3 000 000	4 608
5 000 000	6 064	4 000 000	5 258
6 000 000	6 623	5 000 000	5 808
7 000 000	7 132	6 000 000	6 283
8 000 000	7 601	7 000 000	6 700
9 000 000	8 037	8 000 000	7 069
10 000 000	8 445	9 000 000	7 397
15 000 000	10 180	10 000 000	7 690
20 000 000	11 563	15 000 000	8 730
30 000 000	13 681	20 000 000	9 217
40 000 000	15 222	25 000 000	9 240
50 000 000	16 350	30 000 000	8 802
100 000 000	17 650	35 000 000	7 826
120 000 000	16 331	40 000 000	6 057
130 000 000	15 196	45 507 516	0
140 000 000	13 645		
150 000 000	11 513		
160 000 000	8 365		
169 799 170	0		

Fonte: IBGE, Censo Demográfico 2000.

Assim, por exemplo, caso haja interesse em estimar um total de uma característica relativa às pessoas e essa estimativa para Brasil seja da ordem de 10 000 000, vê-se que seu erro padrão seria da ordem de 8 445. Portanto, de acordo com as fórmulas anteriores, um intervalo de 95% de confiança para o total da característica de interesse será dado por [9 983 448; 10 016 552]. Em termos percentuais, pode-se dizer que a estimativa da característica desejada é 10 000 000, com uma margem de erro relativo de 0,17% para cima ou para baixo.

Na prática, um intervalo de confiança de 95%, por exemplo, indica que, em cada 100 amostras selecionadas com o mesmo desenho, 95 produzirão estimativas cujo intervalo de confiança conterá o valor verdadeiro da população e em apenas 5 amostras este valor estará fora do intervalo de confiança.

Naturalmente, quanto maior o nível de confiança, maior será a amplitude do intervalo de confiança. A decisão sobre o nível de confiança decorre do grau de precisão que o usuário necessita em seu trabalho analítico.

## Conceitos e definições

Apresentam-se, a seguir, os conceitos e definições utilizados na classificação das características que são objeto desta divulgação.

### Domicílio

Domicílio é o local estruturalmente separado e independente que se destina a servir de habitação a uma ou mais pessoas, ou que esteja sendo utilizado como tal.

Os critérios essenciais desta definição são os de separação e independência.

A separação fica caracterizada quando o local de habitação é limitado por paredes, muros ou cercas, coberto por um teto, permitindo a uma ou mais pessoas, que nele habitam, isolar-se das demais, com a finalidade de dormir, preparar e/ou consumir seus alimentos e proteger-se do meio ambiente, arcando, total ou parcialmente, com suas despesas de alimentação ou moradia.

A independência fica caracterizada quando o local de habitação tem acesso direto, permitindo a seus moradores entrar e sair sem necessidade de passar por locais de moradia de outras pessoas.

### Espécie do domicílio

Quanto à espécie, classificou-se o domicílio como:

- **Domicílio particular** - quando o relacionamento entre seus ocupantes era ditado por laços de parentesco, de dependência doméstica ou por normas de convivência. Os domicílios particulares desagregam-se em:

- **permanente** - quando construído para servir, exclusivamente, à habitação e, na data de referência, tinha a finalidade de servir de moradia a uma ou mais pessoas; ou
- **improvisado** - quando localizado em unidade não-residencial (loja, fábrica, etc.) que não tinha dependências destinadas, exclusivamente, à moradia, mas que, na data de referência, estava ocupado por morador. Os prédios em construção, vagões de trem, carroças, tendas, barracas, grutas, etc., que estavam servindo de moradia na data de referência, também foram considerados como domicílios particulares improvisados.

Entendeu-se como dependência doméstica a situação de subordinação dos empregados domésticos e agregados em relação à pessoa responsável pelo domicílio, e entendeu-se por normas de convivência as regras estabelecidas para convivência de pessoas que residiam no mesmo domicílio e não estavam ligadas por laços de parentesco nem de dependência doméstica.

- **Domicílio coletivo** - quando a relação entre as pessoas que nele habitavam era restrita a normas de subordinação administrativa, como em hotéis, pensões, presídios, cadeias, penitenciárias, quartéis, postos militares, asilos, orfanatos, conventos, hospitais e clínicas (com internação), alojamento de trabalhadores, motéis, *camping*, etc.

## População residente

A população residente constituiu-se pelos moradores em domicílios na data de referência.

Considerou-se como moradora a pessoa que tinha o domicílio como local de residência habitual e que, na data de referência, estava presente ou ausente por período que não tenha sido superior a 12 meses em relação àquela data, por um dos seguintes motivos:

- viagens: a passeio, a serviço, a negócio, de estudos, etc.;
- internação em estabelecimento de ensino ou hospedagem em outro domicílio, visando a facilitar a freqüência à escola durante o ano letivo;
- detenção sem sentença definitiva declarada;
- internação temporária em hospital ou estabelecimento similar; e
- embarque a serviço (marítimos).

## Situação do domicílio

Segundo a localização do domicílio, a situação é urbana ou rural, conforme definida por lei municipal em vigor em 1º de agosto de 2000. Em situação urbana consideram-se as áreas urbanizadas ou não, correspondentes às cidades (sedes municipais), às vilas (sedes distritais) ou às áreas urbanas isoladas. A situação rural abrange toda a área situada fora desses limites, inclusive os aglomerados rurais de extensão urbana, os povoados e os núcleos. Este critério também é utilizado na classificação da população urbana e rural.

## Características das pessoas

### Idade

A investigação da idade da pessoa foi feita por meio da pesquisa do mês e ano do nascimento e, também, da declaração da idade que tinha em 31 de julho de 2000. Para a pessoa que não sabia o mês e o ano do nascimento, foi investigada a idade presumida.

A idade em anos ou meses completos foi calculada em relação ao dia 31 de julho de 2000.

## Anos de estudo

A classificação de anos de estudo foi estabelecida com objetivo de compatibilizar os sistemas de ensino anteriores e atual.

Essa classificação foi obtida em função da última série concluída com aprovação no nível ou grau mais elevado que a pessoa de 5 anos ou mais de idade estava freqüentando ou havia freqüentado, sendo a correspondência feita do seguinte modo:

- **Sem instrução e menos de 1 ano de estudo** - para a pessoa que nunca freqüentou escola ou, embora tenha freqüentado, não concluiu pelo menos a 1ª série do ensino fundamental, 1º grau ou elementar;
- **1 ano de estudo** - para a pessoa que concluiu: curso de alfabetização de adultos; ou a 1ª série do ensino fundamental, 1º grau ou elementar;
- **2 anos de estudo** - para a pessoa que concluiu a 2ª série do ensino fundamental, 1º grau ou elementar;
- **3 anos de estudo** - para a pessoa que concluiu a 3ª série do ensino fundamental, 1º grau ou elementar;
- **4 anos de estudo** - para a pessoa que concluiu: a 4ª série do ensino fundamental ou 1º grau; ou, no mínimo, a 4ª série e, no máximo, a 6ª série do elementar;
- **5 anos de estudo** - para a pessoa que concluiu: a 5ª série do ensino fundamental ou 1º grau; ou a 1ª série do médio 1º ciclo;
- **6 anos de estudo** - para a pessoa que concluiu: a 6ª série do ensino fundamental ou 1º grau; ou a 2ª série do médio 1º ciclo;
- **7 anos de estudo** - para a pessoa que concluiu: a 7ª série do ensino fundamental ou 1º grau; ou a 3ª série do médio 1º ciclo;
- **8 anos de estudo** - para a pessoa que concluiu: a 8ª série do ensino fundamental ou 1º grau; ou, no mínimo, a 4ª série e, no máximo, a 5ª série do médio 1º ciclo;
- **9 anos de estudo** - para a pessoa que concluiu a 1ª série do ensino médio, 2º grau ou médio 2º ciclo;
- **10 anos de estudo** - para a pessoa que concluiu a 2ª série do ensino médio, 2º grau ou médio 2º ciclo;
- **11 anos de estudo** - para a pessoa que concluiu, no mínimo, a 3ª série e, no máximo, a 4ª série do ensino médio, 2º grau ou médio 2º ciclo;
- **12 anos de estudo** - para a pessoa que concluiu a 1ª série do superior;
- **13 anos de estudo** - para a pessoa que concluiu a 2ª série do superior;
- **14 anos de estudo** - para a pessoa que concluiu a 3ª série do superior;
- **15 anos de estudo** - para a pessoa que concluiu a 4ª série do superior;
- **16 anos de estudo** - para a pessoa que concluiu a 5ª série do superior;
- **17 anos de estudo ou mais** - para a pessoa que concluiu a 6ª série do superior ou mestrado ou doutorado; ou
- **Não-determinados** - para a pessoa que freqüentava ensino fundamental ou 1º grau não-seriado.

## Trabalho e rendimento

A investigação de trabalho e rendimento abrangeu as pessoas de 10 anos ou mais de idade.

- **Períodos de referência** - Os períodos de referência utilizados na investigação de trabalho e rendimento foram:

- **Semana de referência** - semana de 23 a 29 de julho de 2000;
- **Período de referência de 30 dias** - período de 30 de junho a 29 de julho de 2000; e
- **Mês de referência** - mês de julho de 2000.

- **Trabalho** - considerou-se como trabalho em atividade econômica o exercício de:

- Ocupação remunerada em dinheiro, produtos, mercadorias ou somente em benefícios (moradia, alimentação, roupas, etc.) na produção de bens ou serviços;
- Ocupação remunerada em dinheiro ou benefícios (moradia, alimentação, roupas, etc.) no serviço doméstico;
- Ocupação sem remuneração na produção de bens e serviços, desenvolvida durante pelo menos uma hora na semana:
  - a) em ajuda a membro do domicílio que tinha trabalho como conta própria ou empregador;
  - b) em ajuda a membro do domicílio que tinha trabalho como empregado na produção de bens primários; ou
  - c) como aprendiz ou estagiário.
- Ocupação desenvolvida, durante pelo menos uma hora na semana, na produção de bens do ramo que compreende as atividades de agricultura, silvicultura, pecuária, extração vegetal, pesca e piscicultura, destinados à alimentação de, pelo menos, um morador do domicílio.

- **Procura de trabalho** - definiu-se como procura de trabalho a tomada de alguma providência para conseguir trabalho, tais como: consultar a empregador; participar ou inscrever-se em concurso; consultar a agência de emprego ou sindicato; colocar ou responder anúncio; consultar a parente, amigo ou colega; tomar providência para iniciar negócio próprio; ou outra providência qualquer que efetivamente tivesse como objetivo conseguir trabalho.

- **Pessoas ocupadas** - foram classificadas como ocupadas na semana de referência as pessoas que tinham trabalho durante todo ou parte desse período. Incluíram-se, ainda, como ocupadas as pessoas que não exerceram o trabalho remunerado que tinham na semana de referência por motivo de férias, licença, greve, etc.

- **Pessoas desocupadas** - foram classificadas como desocupadas na semana de referência as pessoas sem trabalho nessa semana que tomaram alguma providência efetiva na procura de trabalho no período de referência de 30 dias.

- **Condição de atividade** - as pessoas foram classificadas, quanto à condição de atividade na semana de referência, em economicamente ativas e não-economicamente ativas.

- **Pessoas economicamente ativas** - as pessoas economicamente ativas na semana de referência compuseram-se das pessoas ocupadas e desocupadas nessa semana.

- **Pessoas não-economicamente ativas** - foram definidas como não-economicamente ativas na semana de referência as pessoas que não foram classificadas como ocupadas nem desocupadas nessa semana.

- **Empreendimento** - definiu-se como empreendimento a empresa, a instituição, a entidade, a firma, o negócio, etc., ou, ainda, o trabalho sem estabelecimento, desenvolvido individualmente ou com ajuda de outras pessoas (empregados, sócios ou trabalhadores não-remunerados).

#### - **Número de trabalhos**

Pesquisou-se o número de trabalhos, ou seja, em quantos empreendimentos a pessoa teve trabalho na semana de referência.

O trabalho na produção para o próprio consumo somente foi contado para a pessoa que não houvesse tido qualquer outro trabalho remunerado ou sem remuneração na semana de referência.

Por convenção, para a contagem do número de trabalhos, o exercício do serviço doméstico remunerado, independentemente do número de unidades domiciliares em que este serviço era prestado, foi contado como se fosse um único trabalho.

#### - **Trabalho principal**

Considerou-se como principal o único trabalho que a pessoa tinha na semana de referência. Para a pessoa que tinha mais de um trabalho na semana de referência, ou seja, para a pessoa ocupada em mais de um empreendimento nessa semana, adotaram-se os seguintes critérios, na ordem enumerada, para definir o principal:

- 1º) o principal era aquele ao qual a pessoa habitualmente dedicava maior número de horas por semana, independentemente de ser remunerado ou não;
- 2º) no caso de igualdade no número de horas trabalhadas, o principal era aquele que a pessoa tinha mais tempo de exercício, independentemente de ser remunerado ou não; e
- 3º) no caso de igualdade também no tempo de exercício dos trabalhos, o principal era aquele que proporcionava o maior rendimento.

- **Nível da ocupação** - é a percentagem das pessoas ocupadas (total ou de um grupo) em relação às pessoas de 10 anos ou mais de idade (total ou do mesmo grupo).

#### - **Ocupação**

Investigou-se a ocupação no trabalho principal.

Foi considerada como ocupação a função, cargo, profissão ou ofício desempenhado por uma pessoa numa atividade econômica.

- **Grandes grupos de ocupação** - referem-se ao nível mais agregado (grandes grupos) da Classificação Brasileira de Ocupação para Pesquisas Domiciliares - CBO-Domiciliar (CBO adaptada para as pesquisas domiciliares), cuja estrutura completa é apresentada em Anexo (CD-ROM).

#### - **Atividade**

Investigou-se a atividade do trabalho principal.

Foi considerada a atividade principal do empreendimento em que a pessoa trabalhava, ou seja, a principal finalidade ou o ramo do negócio, da firma, da instituição, da empresa ou da entidade a que ela prestava serviços ou, para a pessoa que trabalhava por conta própria, a natureza da atividade exercida.

- Para a pessoa que trabalhava ligada às atividades de agricultura, silvicultura, pecuária, extração vegetal ou mineral, pesca e piscicultura, considerou-se a atividade principal do estabelecimento.
  - Para a pessoa que trabalhava em empreendimento agroindustrial foi considerada a atividade principal, quando se tratava de pessoa que explorava o empreendimento, ou a atividade na qual exercia sua ocupação, quando se tratava de qualquer outro trabalhador.
  - Para a pessoa que trabalhava em empreendimento que alocava seu pessoal para prestar serviços em outros empreendimentos foi considerada a atividade daquele com o qual possuía vínculo de trabalho.
  - Para a pessoa que trabalhava na fabricação de determinados itens (rótulos, embalagens, vidros, etc.), unicamente para adicioná-los ao seu produto final, foi considerada a atividade referente à fabricação deste produto final.
- **Seção de atividade** - refere-se ao nível mais agregado (seção) da Classificação Nacional de Atividades Econômicas para Pesquisas Domiciliares - CNAE-Domiciliar (CNAE adaptada para as pesquisas domiciliares), cuja estrutura completa é apresentada em Anexo (CD-ROM).

#### - **Posição na ocupação**

Considerou-se como posição na ocupação a relação de trabalho existente entre a pessoa e o empreendimento em que trabalhava.

Foram definidas cinco posições na ocupação no trabalho principal:

- **Empregado** - para a pessoa que trabalhava para um empregador (pessoa física ou jurídica), geralmente obrigando-se ao cumprimento de uma jornada de trabalho e recebendo, em contrapartida, uma remuneração em dinheiro, mercadoria, produtos ou benefícios, como moradia, alimentação, roupa, etc.

Incluiu-se nesta posição na ocupação:

- a pessoa que prestava o serviço militar obrigatório;
- o sacerdote, ministro de igreja, pastor, rabino, frade, freira e outros religiosos; e
- o aprendiz ou estagiário sem remuneração (definido como a pessoa que exerceu uma ocupação sem remuneração, durante pelo menos uma hora na semana de referência, na qualidade de aprendiz de uma profissão ou ofício ou em cumprimento de estágio obrigatório ou voluntário ligado a uma profissão).

Os empregados foram desagregados em dois subgrupos:

- a) Trabalhador doméstico - para a pessoa que prestava serviço doméstico remunerado em um ou mais domicílios; e
  - b) Demais empregados - para os demais empregados;
- **Empregador** - para a pessoa que trabalhava explorando o seu próprio empreendimento com pelo menos um empregado;
  - **Conta própria** - para a pessoa que trabalhava explorando o seu próprio empreendimento, sozinha ou com sócio, sem ter empregado, ainda que contando com ajuda de trabalhador não-remunerado;

- **Não-remunerado em ajuda a membro do domicílio** - para pessoa que trabalhou sem remuneração, durante pelo menos uma hora na semana de referência, em ajuda a membro do domicílio que era empregado na produção de bens primários (que compreende as atividades de agricultura, pecuária, silvicultura, extração vegetal ou mineral, caça, pesca ou piscicultura) ou em ajuda no trabalho de membro do domicílio que era conta própria ou empregador em qualquer atividade; e
- **Trabalhador na produção para o próprio consumo** - para pessoa que trabalhou, durante pelo menos uma hora na semana de referência, na produção de bens do ramo que compreende as atividades de agricultura, silvicultura, pecuária, extração vegetal, pesca ou piscicultura, destinados exclusivamente à alimentação de pelo menos um morador do domicílio.

#### - Categoria do emprego

Os empregados, quanto à categoria do emprego no trabalho principal, foram classificados em: com carteira de trabalho assinada; militares e funcionários públicos estatutários; e outros sem carteira de trabalho assinada.

A categoria dos militares e funcionários públicos estatutários foi constituída pelos militares das Forças Armadas (Exército, Marinha e Aeronáutica) ou das Forças Auxiliares (Polícia Militar, Corpo de Bombeiros Militar, etc.) e pelos empregados de instituição, fundação, autarquia, etc., do Poder Público (Executivo, Legislativo ou Judiciário) regidos pelo Regime Jurídico dos Funcionários Públicos de qualquer instância (federal, estadual e municipal).

#### - Contribuição para instituto de previdência oficial

Foi pesquisado se a pessoa contribuía para um dos seguintes institutos de previdência oficial no trabalho principal:

- INSS (Instituto Nacional de Seguridade Social);
- Plano de Seguro Social da União; ou
- Instituto de Previdência Social Estadual ou Municipal.

- **Horas trabalhadas por semana** - pesquisou-se o número de horas habitualmente trabalhadas por semana no trabalho principal e nos demais trabalhos que a pessoa tinha na semana de referência.

- **Pessoas aposentadas** - foram classificadas como aposentadas as pessoas que, no mês de referência, eram aposentadas, jubiladas ou reformadas por instituto de previdência oficial (Instituto Nacional do Seguro Social - INSS, Plano de Seguro Social da União ou instituto de previdência social estadual ou municipal).

#### - Rendimento nominal mensal

Considerou-se como rendimento nominal mensal da pessoa de 10 anos ou mais de idade a soma do rendimento nominal mensal de trabalho com o proveniente de outras fontes.

a) **Rendimento nominal mensal de trabalho** - pesquisou-se o rendimento nominal mensal do trabalho principal e dos demais trabalhos que a pessoa tinha na semana de referência.

- Para a pessoa que recebia rendimento fixo do trabalho, considerou-se a remuneração bruta (do empregado ou trabalhador doméstico) ou a retirada (do trabalhador por conta própria ou empregador) do mês de referência.

- Para a pessoa que recebia rendimento variável do trabalho, considerou-se a média da remuneração bruta ou da retirada do mês de referência.
- Para a pessoa licenciada por instituto de previdência oficial pelo trabalho, considerou-se o rendimento bruto do mês de referência, recebido como benefício (auxílio-doença, auxílio por acidente de trabalho, etc.).
- Entendeu-se por remuneração bruta o pagamento do empregado ou do trabalhador doméstico, sem excluir os descontos correspondentes ao instituto de previdência oficial, imposto de renda, faltas, etc., e sem incluir o 13º salário, a gratificação de férias e a participação de lucros pagos por empregador.
- Entendeu-se por retirada o ganho do trabalhador por conta própria ou do empregador com o seu empreendimento, ou seja, o rendimento bruto menos os gastos (pagamento de empregados, compra de equipamentos, matéria-prima, energia elétrica, telefone, etc.) do empreendimento.
- Para a pessoa que recebia em produtos ou mercadorias de atividade do ramo que compreende a agricultura, silvicultura, pecuária, extração vegetal ou mineral, pesca e piscicultura, considerou-se o valor, real ou estimado, relativo ao mês de referência, recebido normalmente ou em média mensal, no caso de produção sazonal.
- Não foi computado como rendimento do trabalho o valor da produção para o próprio consumo nem a parcela do pagamento efetuada em benefícios, tais como: moradia, alimentação (refeição, cesta de alimentos, vale ou tíquete-alimentação), vale ou tíquete-transporte, roupas, etc.

b) **Rendimento nominal mensal de outras fontes** - pesquisou-se o rendimento nominal mensal de aposentadoria, pensão, aluguel, pensão alimentícia, mesada ou doação recebida de não-morador, renda mínima, bolsa-escola ou seguro-desemprego, relativo ao mês de referência e, também, investigou-se o rendimento habitualmente recebido a título de abono de permanência em serviço, pensão paga integralmente por seguradora ou entidade de previdência privada aberta, juros de aplicações financeiras, dividendos, etc., relativo ao mês de referência.

Entendeu-se por rendimento de:

- **Aposentadoria** - aquele recebido de instituto de previdência oficial a título de aposentadoria, jubilação ou reforma, incluindo, se for o caso, a parcela paga por entidade de previdência privada fechada como complementação de aposentadoria;
- **Pensão** - aquele recebido de instituto de previdência oficial, deixado por alguém da qual a pessoa era beneficiária;
- **Aluguel** - aquele recebido a título de locação, sublocação ou arrendamento de móveis, imóveis, máquinas, equipamentos, animais, etc.;
- **Pensão alimentícia** - aquele recebido a título de pensão alimentícia, paga espontânea ou judicialmente;
- **Mesada** - aquele recebido, em caráter regular, sem contrapartida de serviços prestados, proveniente de pessoa não-moradora do domicílio;

- **Renda mínima** - aquele recebido de órgão governamental, com vistas em garantir um rendimento mínimo necessário à sobrevivência de uma família;
  - **Bolsa-escola** - aquele recebido de órgão governamental, com vistas em garantir a permanência na escola das crianças de uma família que estejam em idade de frequência obrigatória à escola;
  - **Seguro-desemprego** - aquele recebido por dispensa de emprego com carteira de trabalho assinada e atendendo aos requisitos necessários para o recebimento deste benefício; e
  - **Abono de permanência em serviço** - aquele recebido como benefício que era concedido a quem permanecia trabalhando, embora tivesse tempo de serviço suficiente para se aposentar.
- **Rendimento nominal mensal domiciliar** - considerou-se como rendimento nominal mensal domiciliar a soma dos rendimentos nominais mensais dos moradores do domicílio, exclusive os das pessoas cuja condição no domicílio fosse pensionista, empregado(a) doméstico(a) ou parente do(a) empregado(a) doméstico(a).
- **Rendimento nominal mensal familiar** - considerou-se como rendimento nominal mensal familiar a soma dos rendimentos nominais mensais dos componentes da família, exclusive os das pessoas cuja condição na família fosse pensionista, empregado(a) doméstico(a) ou parente do(a) empregado(a) doméstico(a).
- **Salário mínimo** - para a apuração dos rendimentos, segundo as classes de salário mínimo, considerou-se o valor que vigorava no mês de referência. O salário mínimo era de R\$ 151,00 (cento e cinquenta e um reais).

## Índice de Gini

O Índice de Gini é uma medida do grau de concentração de uma distribuição, cujo valor varia de 0 (zero) (a perfeita igualdade) até 1 (um) (a desigualdade máxima).

O Índice de Gini da distribuição do rendimento é calculado através da seguinte expressão:

$$G = \frac{\sum_{i=1}^N (2i-1)y_{(i)}}{N \sum_{i=1}^N y_i} - 1$$

onde  $y_i$  é o valor da variável  $y$  para a  $i$ -ésima unidade da população, e  $y_{(i)}$  é o valor da  $i$ -ésima estatística de ordem da variável  $y$  na população.

Um estimador consistente para  $G$  é dado por Sandström, Wretman e Waldén (1988).

$$\hat{G}_w = \frac{\sum_{i=1}^n \left[ 2\hat{F}(y_{(i)}) + w_{(i)} \right] w_{(i)} y_{(i)}}{N \sum_{i=1}^n w_i y_i} - 1$$

onde

$w_i$  é o peso da  $i$ -ésima unidade amostral, satisfazendo a seguinte propriedade

$$E \left[ \hat{N} \right] = E \left[ \sum_{i=1}^n w_i \right] \cong N \quad (\text{pesos devem permitir obter estimadores aproxima-}$$

madamente não viciados de total);

$$\hat{F}(t) = \sum_{j=1}^n w_j I(y_j < t) \quad \text{para qualquer } t \in \mathfrak{R};$$

$I(y_j < t)$  é a função indicadora do evento  $y_j < t$ , isto é, a função que toma o valor 1 quando o evento ocorre e valor zero quando não ocorre;

$y_{(i)}$  é o valor da  $i$ -ésima estatística de ordem da variável  $y$  na amostra; e

$w_{(i)}$  é o peso correspondente à  $i$ -ésima estatística de ordem  $y_{(i)}$ .

Os resultados apresentados para o valor do rendimento nominal médio mensal e valor do rendimento nominal mediano mensal referem-se às pessoas de 10 anos ou mais de idade, total e economicamente ativas, com rendimento, às pessoas de 10 anos ou mais de idade ocupadas na semana de referência com rendimento do trabalho principal e de todos os trabalhos, às famílias residentes em domicílios particulares, total e com rendimento familiar.

Os resultados apresentados para o índice de Gini referem-se às pessoas de 10 anos ou mais de idade, total, economicamente ativas e não-economicamente ativas, com rendimento, às pessoas ocupadas com rendimento de trabalho e às famílias residentes em domicílios particulares, com rendimento familiar. O índice de Gini não foi calculado para os municípios.