

ATA “RESULTADOS DE ESTUDOS AMOSTRAIS”

Elaborada por Marcelo de Moraes Duriez

A técnica da Coordenação de Métodos e Qualidade (COMEQ), Sra. Samela Batista Arantes, apresentou os resultados dos estudos efetuados pela referida Coordenação, com o intuito de estabelecer o plano amostral para a PNAG, uma das pesquisas constituintes do SNPA.

Iniciando sua explanação, a Sra. Samela explicou que o plano amostral planejado tratará o conjunto dos estabelecimentos agropecuários de maneira distinta. Tendo como base o Censo Agropecuário 2006, serão construídos dois subconjuntos. Um subconjunto constituirá o cadastro de lista, e o outro subconjunto constituirá o cadastro de área. Cada subconjunto sofrerá um processo de amostragem diferente. O primeiro, o cadastro de lista contendo cerca de 102.000 estabelecimentos, incluirá os estabelecimentos que atendam a critérios pré-estabelecidos, de maneira geral estar acima de limites mínimos de efetivo de animais, ou de produção para determinadas lavouras. O segundo subconjunto será constituído por aqueles estabelecimentos que não atendam aos critérios para compor o grupo anterior. Para o cadastro de lista pretende-se realizar uma amostra estratificada, com estrato certo, tipo de amostra muito utilizada nas pesquisas econômicas do IBGE. Já o segundo grupo, o cadastro de área contendo cerca de 4.800.000 estabelecimentos, adotará um plano amostral mais complexo, amostra conglomerada em dois estágios. Em um primeiro estágio serão selecionados setores e, em um segundo estágio, estabelecimentos agropecuários. Aqui cabe uma rápida explicação sobre o conceito de setor. Para o IBGE, a menor porção de área do território considerada é o setor censitário. Na verdade, ele tem como princípio definidor a capacidade de trabalho de um recenseador em cobrir determinada área, durante o período de coleta das operações censitárias. Tendo este princípio em mente, são estabelecidos critérios para determinar o tamanho do setor (quantidade máxima de domicílios ou de estabelecimentos agropecuários, por setor). Os setores são classificados em rurais ou urbanos, dependendo se sua localização é na área urbana ou na área rural. O conjunto dos setores é denominado malha setorial e esta sofre alterações de uma operação censitária para outra, em função do “crescimento” dos setores. Isto é, com a adição de novas construções em uma área há o aumento do número de domicílios ou de estabelecimentos agropecuários, sendo necessária a divisão do setor, criando-se a partir do antigo, novos setores. Portanto, pode-se dizer que o cadastro de áreas é um cadastro de setores. Nos estudos iniciais sobre o plano amostral será utilizada a mesma malha setorial utilizada no Censo Agropecuário 2006, embora esta malha já tenha sido alterada, dando origem a uma nova malha (a de 2010). Posteriormente, quando acontecer a seleção da amostra definitiva, será feita uma compatibilização

entre as malhas, uma vez que, para a amostra final, será utilizada a malha mais atualizada. Dando prosseguimento à apresentação, a Sra. Samela explicou mais detalhadamente como serão construídas as UPAs (Unidades Primárias de Amostragem). A princípio um setor é uma UPA, entretanto, existem setores que contém poucos estabelecimentos e não seria possível processar uma rotação entre os estabelecimentos do setor, para a composição da amostra. Por isso, foi necessário agregar setores vizinhos, de forma que este agregado possua um número de estabelecimentos em que seja possível realizar a rotação. Os setores ou o agregado de setores é que serão selecionados no primeiro estágio da amostra. Depois, em cada uma das UPAs selecionadas, são escolhidos estabelecimentos, o que consiste no segundo estágio da amostra. A técnica da COMEQ esclareceu que, para a seleção das UPAs, utilizou-se a probabilidade proporcional ao tamanho, sendo a variável de tamanho o valor da produção. Ou seja, aquelas UPAs que tiveram um maior valor de produção, terão uma maior chance de serem selecionadas para a amostra. O valor de produção da UPA é obtido a partir da soma dos valores de produção dos estabelecimentos constituintes da mesma. Para a seleção dos estabelecimentos dentro da UPA, se adotará uma amostra aleatória simples de, no mínimo, 10 estabelecimentos. A Sra. Samela salientou que este número pode mudar de acordo com a variabilidade. Os estudos revelaram haver uma grande heterogeneidade entre as Unidades da Federação (quanto maior a variabilidade, maior o número de estabelecimentos a serem selecionados em cada UPA).

Após sua apresentação, a Sra. Samela colocou-se à disposição para esclarecimentos. O Sr. Carlos Antônio Almeida Barradas (COAGRO/IBGE) questionou se havia alguma estimativa da participação relativa, em termos de valor da produção, dos estabelecimentos que compunham o cadastro de lista. Em resposta, a técnica da COMEQ disse que não poderia fornecer a informação no momento, mas que há como fazer esta estimativa. Contudo, ela salientou que esta participação não deve ser pequena, pois estes estabelecimentos são os “grandes” em alguma característica, por exemplo: área colhida de soja, área colhida de laranja, número de bovinos, número de suínos etc. Em seguida, o Sr. Aristides P. Lima Green (COAGRO/IBGE) perguntou se haviam considerado outra variável além do valor da produção, como variável de tamanho para dimensionar a amostra, usando o modelo PPT (probabilidade proporcional ao tamanho). A técnica da COMEQ respondeu que sim, pois haviam dimensionado a amostra utilizando como variável de tamanho, o número de estabelecimentos por setor. Contudo, a amostra havia ficado três vezes maior do que quando se utilizou o valor da produção. Ela acrescentou que, quando se utiliza como variável de tamanho uma variável que tenha alguma correlação com as variáveis de interesse, os tamanhos amostrais tendem a ser menores. Dando continuidade, a Sra. Gabriela Fernandez (UERJ) quis saber qual o tamanho da amostra. Ao respondê-la, a Sra. Samela justificou não ter apresentado o tamanho, pois os trabalhos ainda estão em fase inicial. Entretanto, informou que os estudos efetuados apontavam para

um tamanho de cerca de 11.000 UPAs, mas que esse número poderia mudar com a continuidade dos trabalhos. O Sr. Aristides Green acrescentou que temia que, ao se utilizar o valor da produção como variável de tamanho, não estaríamos privilegiando os estabelecimentos grandes, dando a eles uma maior probabilidade para que eles componham a amostra. A técnica da COMEQ esclareceu que o valor total da produção é o da UPA, ou seja, o valor total de produção leva em conta todos os estabelecimentos constituintes dela e, portanto, se em determinada UPA existirem muitos estabelecimentos pequenos, o valor total pode ser maior do que o de uma outra onde existam menos estabelecimentos, mesmo sendo estes de maior porte. Acrescentando, ela ressaltou que fez-se a seleção da UPA com probabilidade proporcional ao tamanho (PPT), mas dentro de cada UPA se procedeu a seleção dos estabelecimentos (Unidade Secundária de Amostragem), e esta será por amostra aleatória simples, ou seja, não se leva em conta se o estabelecimento é grande ou pequeno, tendo todos a mesma chance de serem selecionados. A seguir, o Sr. Antônio Carlos S. Florido (IBGE/GTA) fez um comentário, salientando que mesmo sem o tamanho definitivo podemos esperar um número maior que 110.000 estabelecimentos, pois além destes que virão das UPAs, teremos estabelecimentos vindos do cadastro de lista. Ele comentou que, baseado nesta cifra, a operação de coleta para a Pesquisa será de grande magnitude. Em seguida o Sr. Flávio P. Bolliger (IBGE/COAGRO) pediu a palavra e comentou que, em um evento interno anterior (ENAGRO), o Sr. Marcos Paulo S. de Freitas (IBGE/COMEQ) havia dito que um tamanho de amostra menor (menos UPAs) havia sido obtido, quando se utilizou outra variável para dimensionar a amostra. Entretanto, esta variável não apresentou boa correlação com as variáveis de interesse da Pesquisa, o que fez com que a precisão das estimativas para estas variáveis também não houvesse sido boa. Concluindo sua intervenção, o Sr. Flávio acrescentou que aquilo que se busca com os estudos para delinear o plano amostral, é encontrar uma variável que forneça um tamanho de amostra confortável, operacionalmente falando, e que possibilite produzir estimativas com baixo coeficiente de variação para as variáveis de interesse da Pesquisa. A seguir, a Sra. Sônia Albieri (IBGE/COMEQ) esclareceu aos presentes alguns pontos relativos a amostragem. Primeiramente, ela abordou a diferença existente entre variável de dimensionamento e a variável de tamanho. No primeiro caso, é estipulada a precisão desejada para uma variável, e a partir daí dimensiona-se um tamanho de amostra que seja suficiente para garantir esta precisão, para a variável em questão. Já a variável de tamanho está relacionada com o tipo da amostra. Se você utiliza amostra PPT (Probabilidade Proporcional ao Tamanho), como foi o caso da PNAG, é necessário escolher uma variável de tamanho de modo que aqueles indivíduos que tenham um maior valor para a variável, tenham maior probabilidade de serem selecionados para compor a amostra. No caso da PNAG, foram testadas duas variáveis: número de estabelecimentos na UPA e valor de produção na UPA. É desejável que a variável de tamanho seja correlacionada com as variáveis de interesse. O valor da

produção está mais correlacionado com as variáveis de interesse, por isso os resultados foram melhores (amostra menor) que com o número de estabelecimentos. Outro ponto abordado por ela, foi a existência de dois cadastros diferentes. Isto confere uma liberdade muito grande, já que pode-se escolher planos amostrais totalmente diferentes para cada cadastro, e depois juntar as amostras para o cálculo das estimativas e de suas precisões. Ela ressaltou ainda, que o número de estabelecimentos selecionados por UPA vai depender da variabilidade das variáveis de interesse da Pesquisa, em cada UPA. Se houver grande variabilidade, o número de 10 estabelecimentos/UPA inicialmente pensado pode ser maior. Segundo ela, mais estudos posteriores serão feitos para determinar este número. Então, o Sr. Antônio Florido (IBGE/GTA) fez uma intervenção e reafirmou sua preocupação com o tamanho da amostra. Conforme ele, em uma reunião anterior haviam dito que seria uma amostra de 6.000 setores, e que agora foi colocado que a amostra seria de 11.000 UPAs, sendo que há casos nos quais há mais de um setor compondo a UPA. Além do mais, há a parte da amostra advinda do cadastro de lista. Assim, o Sr. Florido afirmou que uma vez estabelecido que a primeira coleta da PNAG será realizada junto com a operação censitária, sua preocupação reside em realizar a coleta de uma Pesquisa com amostra tão grande, junto com o Censo Agropecuário. O Sr. Flávio Bolliger (IBGE/COAGRO) interveio, dizendo que o tamanho apresentado ainda é muito preliminar, e acrescentou que o número de 6.000 setores havia sido arbitrado por ele próprio, para os cálculos orçamentários preliminares. Ou seja, não houve naquela ocasião qualquer preocupação metodológica. Concluindo, o Sr. Flávio disse acreditar que o número de UPAs será menor que 11.000. Finalizando, a Sra. Samela afirmou que os estudos prosseguirão, e provavelmente no próximo Fórum, o tamanho da amostra será mais definitivo do que o apresentado agora.