



PRESIDÊNCIA DA REPÚBLICA  
CONTROLADORIA-GERAL DA UNIÃO  
SECRETARIA FEDERAL DE CONTROLE INTERNO

## **MANUAL OPERACIONAL**

# **ROTINAS PARA PLANOS AMOSTRAIS**

CONTROLADORIA-GERAL DA UNIÃO  
SECRETARIA FEDERAL DE CONTROLE INTERNO

**Ministro de Estado do Controle e da Transparência**  
Jorge Hage Sobrinho

**Secretaria-Executiva**  
Luiz Augusto Fraga Navarro de Brito Filho

**Secretaria Federal de Controle Interno**  
Valdir Agapito Teixeira

**Diretoria de Planejamento e Coordenação das Ações de Controle**  
Jerri Eddie Xavier Coelho

**Coordenação – Geral de Planejamento e Avaliação**  
Márcia da Rosa Pereira

**Equipe Técnica**  
Beatriz Bernardes Ribeiro

Janeiro/2007

## SUMÁRIO

<b><i>APRESENTAÇÃO</i></b>	<b>5</b>
<b><i>I – ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA ATIVIDADES DE AUDITORIA</i></b>	<b>6</b>
<b><i>I.1 – FLUXO DA ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA ATIVIDADES DE AUDITORIA</i></b>	<b>8</b>
<b><i>II – ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA A ATIVIDADE DE ACOMPANHAMENTO SISTEMÁTICO DE PROGRAMAS/AÇÕES DE GOVERNO</i></b>	<b>9</b>
<b><i>II.1 – DO FLUXO DA ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA A ATIVIDADE DE ACOMPANHAMENTO SISTEMÁTICO DE PROGRAMAS/AÇÕES DE GOVERNO</i></b>	<b>10</b>
<b><i>III – ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA FISCALIZAÇÃO DE MUNICÍPIOS DE SORTEIO PÚBLICO</i></b>	<b>11</b>
<b><i>III.1 – DO FLUXO DA ROTINA PARA ELABORAÇÃO DE PLANOS AMOSTRAIS PARA FISCALIZAÇÃO DE MUNICÍPIOS DO SORTEIO</i></b>	<b>12</b>
<b><i>IV – ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA DEMANDAS EXTERNAS</i></b>	<b>13</b>
<b><i>IV.1 – DO FLUXO DA ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA DEMANDAS EXTERNAS</i></b>	<b>14</b>

## **APRESENTAÇÃO**

Como a Coordenação de Planejamento e Avaliação das Ações de Controle (DCPLA) tem como uma de suas atribuições a de elaboração de planos amostrais para atender às unidades da Secretaria Federal de Controle, bem como, com base nos levantamentos efetuados, elaborar as estimativas e inferências relativas a essas amostras produzidas, foram desenvolvidas algumas rotinas de modo a tornar esse atendimento o mais adequado possível.

Como existem rotinas diferentes no processo de elaboração de planos amostrais, a saber, uma para cada ação de controle: Auditorias (Acompanhamento da Gestão/Avaliação da Gestão) e Fiscalizações (Sorteio/Denúncias/Acompanhamento Sistemático), se fez necessário a descrição detalhada dessas rotinas de forma a permitir a sua melhor compreensão.

O presente Manual tem por objetivo apresentar o mapeamento de cada etapa da rotina do procedimento de elaboração de amostras probabilísticas e estabelecer os procedimentos a serem observados pelas Unidades de Controle Interno da Controladoria-Geral da União – CGU para a realização dos trabalhos no que concerne à realização de amostragem probabilística nas ações de controle efetuadas no âmbito da CGU, auxiliando assim a compreensão do fluxo operacional.

O Manual apresenta os procedimentos que extrapolam a elaboração da amostra, indo até a consolidação dos resultados obtidos nos levantamentos feitos com base nessas amostras e a elaboração de estimativas para, posteriormente, serem realizadas as devidas inferências que irão subsidiar as generalizações que são incorporadas nos relatórios finais, de forma a fornecer suporte às evidências, de acordo com as hipótese formuladas na estratégia.

Cada etapa da rotina no fluxo pressupõe uma ou mais atividades de cada unidade responsável, com as atribuições inerentes a cada uma, conforme os fluxos e a descrição de cada etapa apresentados a seguir para cada rotina.

Além disso, o Manual apresenta também um Apêndice onde consta o referencial teórico necessário para melhor compreensão dos conceitos inerentes a elaboração de planos amostrais. Adicionalmente, trata de temas relacionados e necessários à confecção de uma boa amostra (e suas estimativas) e cuja melhor compreensão auxiliará para o melhor desempenho dessa atividade, que, por envolver a unidade demandante, requer um entendimento mínimo desses temas.

## **I – ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA ATIVIDADES DE AUDITORIA**

Para a elaboração de planos amostrais para as atividades de auditoria as unidades de controle interno deverão seguir as seguintes etapas:

1ª) A Coordenação-Geral define as Unidades Jurisdicionadas (UJ), quando das auditorias, para as quais serão realizadas amostragem probabilística, e os gastos objeto de elaboração de amostra. Nessa etapa entende-se como “gasto” as unidades de observação de amostragem que serão utilizadas para o levantamento da informação individualizada, como por exemplo: convênios, contratos, processos de concessão de diárias, licitações, entre outros;

2ª) A Coordenação-Geral extrai os registros (unidades amostrais) da população referenciada para fins de elaboração de amostra dos sistemas corporativos, da internet e/ou cadastros, ou ainda outras fontes e encaminha os mesmos para a DCPLA.INFO, solicitando elaboração de amostra. A DCPLA poderá orientar as Coordenações na extração dos dados de forma padrão (nomes de variáveis, conteúdos, etc). É necessária uma reunião prévia conjunta da DCPLA, DCPLA.INFO e Coordenação demandante para estabelecer os critérios a serem considerados na amostra, exceto quando este procedimento já tenha sido feito anteriormente;

3ª) A DCPLA.INFO analisa a consistência dos dados recebidos, caso haja consistência providencia o tratamento aos dados e elabora a amostra, conforme a estratégia discutida, de vários tamanhos para níveis de confiança e erros diferentes. A Coordenação escolhe o tamanho da mesma. Caso não haja consistência nos dados, a DCPLA.INFO devolve os mesmos à Coordenação para as devidas correções;

4ª) A Coordenação-Geral, por sua vez, faz as correções e remete os dados para nova verificação da consistência por parte da DCPLA.INFO;

5ª) A Coordenação-Geral escolhe o tamanho da amostra de acordo com a estratégia, capacidade operacional e a experiência adquirida, e envia o tamanho escolhido para a DCPLA.INFO;

6ª) A DCPLA.INFO seleciona a amostra de forma aleatória conforme o método escolhido no dimensionamento da amostra e envia para a Coordenação;

7ª) A Coordenação-Geral, ao receber a amostra faz a validação da mesma, verificando se está de acordo com a estratégia discutida e os campos e conteúdos consistentes;

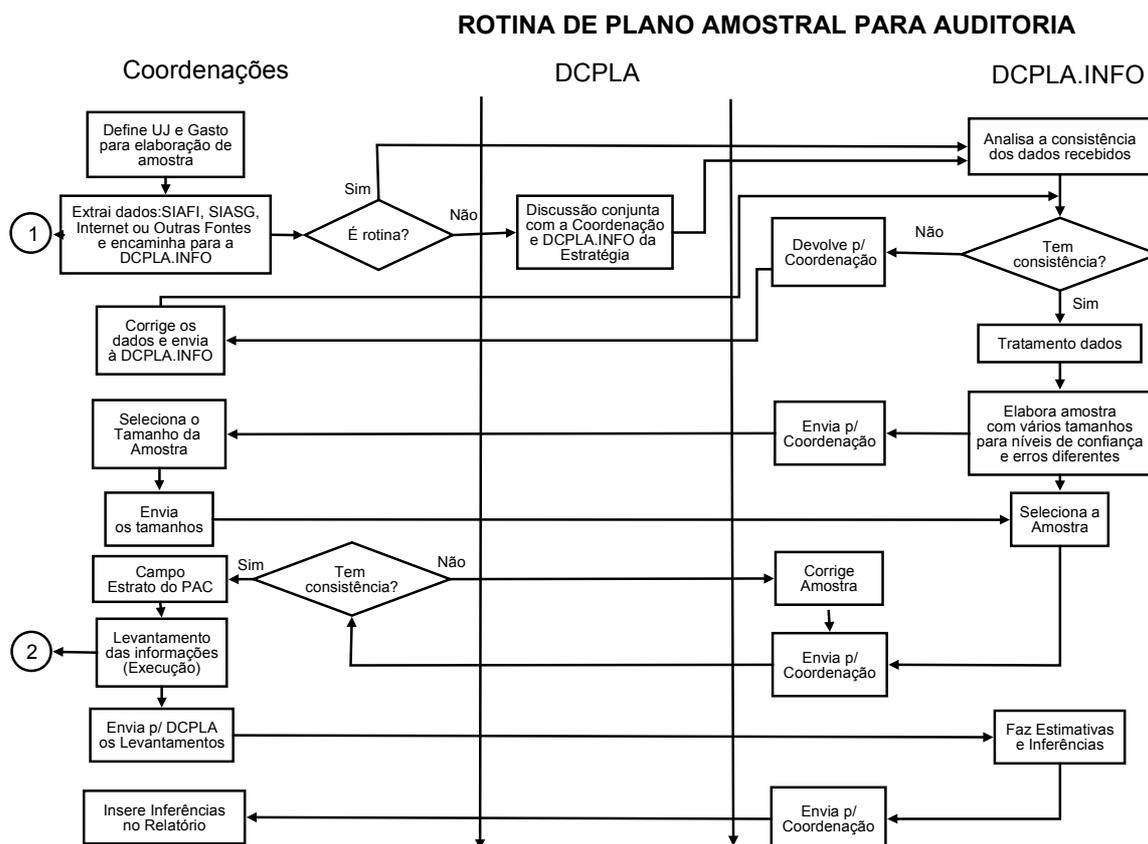
8ª) Caso esteja de acordo a Coordenação-Geral insere o PAC no sistema com o campo ESTRATO contendo a identificação da existência de uma amostra probabilística, caso contrário envia para a DCPLA.INFO para que sejam feitas as correções;

9ª) A Coordenação-Geral faz o levantamento das informações por meio de questionários ou procedimentos padronizados de forma que permita a quantificação do levantamento para possibilitar o cálculo de estimativas de proporções, médias ou totais, e envia esses levantamentos para a DCPLA.INFO, por meio de planilhas sejam elas alimentadas pelas equipes ou por via de extração do sistema ATIVA;

10ª) A DCPLA.INFO faz as estimativas de proporções, médias ou totais, baseada na solicitação da Coordenação com base nos questionários ou procedimentos e envia para a Coordenação;

11<sup>a</sup>) A Coordenação analisa os resultados dessas estimativas e insere os mesmos nos Relatórios, referenciando o tamanho do universo (população) e suas características, as fontes de informação, escopo, nível de confiança adotado na amostra, etc, ou seja, informações relevantes para a compreensão do evento/fenômeno levantado na amostra

# I.1 – FLUXO DA ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA ATIVIDADES DE AUDITORIA



1 - A DCPLA poderá orientar as Coordenações-Gerais na extração dos dados de forma padrão (nomes das variáveis, conteúdos, etc).

2 - Para a elaboração de Estimativas e Inferências faz-se necessário a existência de procedimentos e/ou questionários padronizados de forma a permitir a quantificação dos resultados leve

## **II – ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA A ATIVIDADE DE ACOMPANHAMENTO SISTEMÁTICO DE PROGRAMAS/AÇÕES DE GOVERNO**

Para a elaboração de planos amostrais para as atividades de acompanhamento sistemático de programas/ações de governo as unidades de controle interno deverão seguir as seguintes etapas:

1ª) A Coordenação-Geral, após a Priorização, momento em que define as ações em acompanhamento sistemático, define a estratégia e levanta a necessidade, ou não, de elaboração de plano amostral e solicita reunião com a DCPLA e a DCPLA.INFO para discussão da estratégia;

2ª) A DCPLA, DCPLA.INFO e Coordenação discutem a estratégia conhecendo os universos objeto de ação de controle bem como a própria ação de governo objeto de acompanhamento;

3ª) A Coordenação-Geral obtém os dados junto ao Gestor da ação de governo e envia os mesmos de forma padronizada a DCPLA.INFO;

4ª) A DCPLA.INFO analisa a consistência dos dados recebidos. Caso haja consistência providencia o tratamento aos dados e elabora a amostra, caso não haja consistência devolve os dados para a Coordenação fazer as devidas correções;

5ª) A Coordenação-Geral, por sua vez, faz as correções e remete os dados para nova verificação da consistência por parte da DCPLA.INFO;

6ª) A DCPLA.INFO seleciona a amostra de forma aleatória e de acordo com o método discutido na estratégia e envia para a Coordenação;

7ª) A Coordenação-Geral, ao receber a amostra faz a validação da mesma, verificando se está de acordo com a estratégia discutida e os campos e conteúdos consistentes;

8ª) Caso esteja de acordo a Coordenação-Geral insere o PAC no sistema com o campo ESTRATO contendo a identificação da existência de uma amostra probabilística, caso contrário envia para a DCPLA.INFO para que sejam feitas as correções;

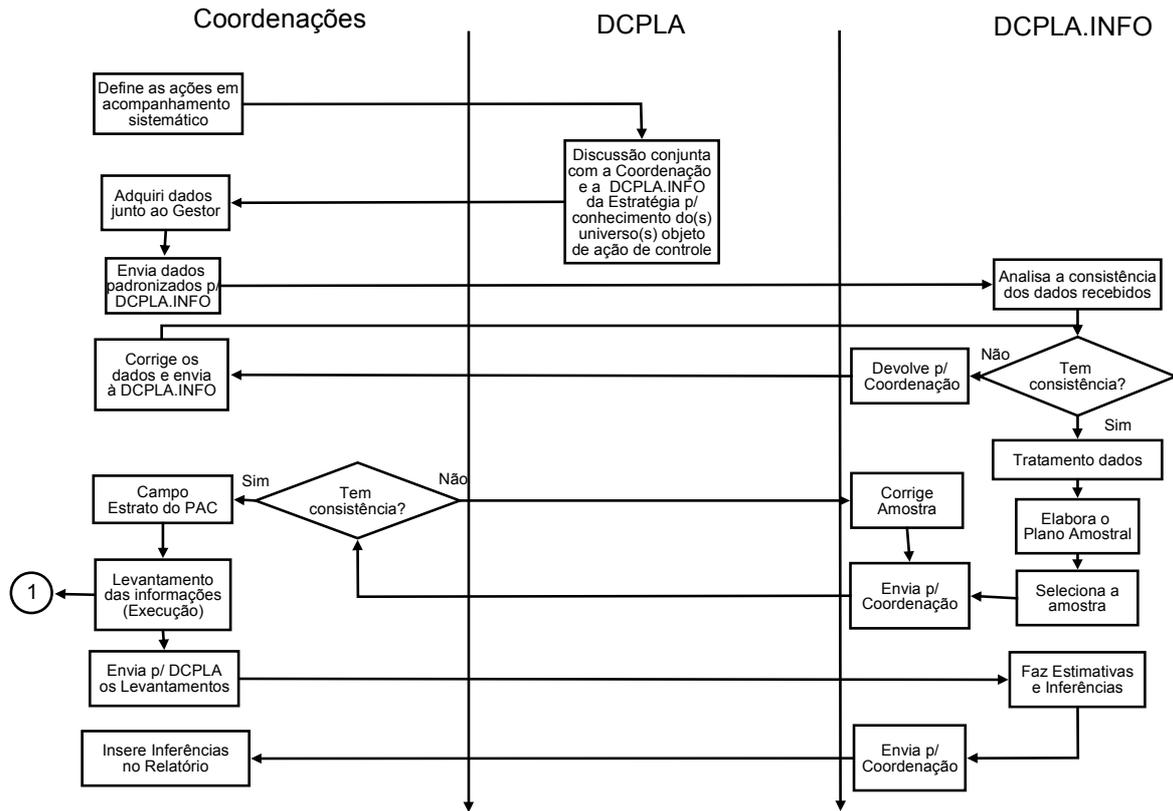
9ª) A Coordenação-Geral faz o levantamento das informações por meio de questionários ou procedimentos padronizados de forma que permita a quantificação do levantamento para possibilitar o cálculo de estimativas de proporções, médias ou totais, e envia esses levantamentos para a DCPLA.INFO, por meio de planilhas sejam elas alimentadas pelas equipes ou por via de extração do sistema ATIVA;

10ª) A DCPLA.INFO faz as estimativas de proporções, médias ou totais, baseada na solicitação da Coordenação com base nos questionários ou procedimentos e envia para a Coordenação;

11ª) A Coordenação analisa os resultados dessas estimativas e insere os mesmos nos Relatórios, referenciando o tamanho do universo (população) e suas características, as fontes de informação, escopo, nível de confiança adotado na amostra, etc, ou seja, informações relevantes para a compreensão do evento/fenômeno levantado na amostra.

**II.1 – DO FLUXO DA ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA A ATIVIDADE DE ACOMPANHAMENTO SISTEMÁTICO DE PROGRAMAS/AÇÕES DE GOVERNO**

**ROTINA DE PLANO AMOSTRAL PARA ACOMPANHAMENTO SISTEMÁTICO DE PROGRAMAS/AÇÕES DE GOVERNO**



1 - Para a elaboração de Estimativas e Inferências faz-se necessário a existência de procedimentos e/ou questionários padronizados de forma a permitir a quantificação dos resultados lev:

### **III – ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA FISCALIZAÇÃO DE MUNICÍPIOS DE SORTEIO PÚBLICO**

Para a elaboração de planos amostrais para as atividades de fiscalização de municípios de sorteio público as unidades de controle interno deverão seguir as seguintes etapas:

1ª) A Coordenação-Geral, após a Priorização, momento em que define a estratégia e levanta a necessidade, ou não, de elaboração de plano amostral, define as ações objeto de amostragem para fiscalização de municípios do sorteio público e solicita reunião com a DCPLA e a DCPLA.INFO para discussão da estratégia;

2ª) A DCPLA, DCPLA.INFO e Coordenação discutem a estratégia conhecendo os universos objeto de ação de controle bem como a própria ação de governo objeto de fiscalização;

3ª) A Coordenação-Geral obtém os dados junto ao Gestor da ação de governo dos municípios sorteados e envia os mesmos de forma padronizada a DCPLA.INFO;

4ª) A DCPLA.INFO analisa a consistência dos dados recebidos, caso não haja consistência a DCPLA.INFO devolve os dados para a Coordenação fazer as devidas correções;

5ª) A Coordenação-Geral, por sua vez, faz as correções e remete os dados para nova verificação da consistência por parte da DCPLA.INFO;

6ª) Caso haja consistência a DCPLA.INFO se questiona sobre os dados serem de todo o universo ou somente dos municípios selecionados, caso sejam do universo ela providencia os devidos filtros;

7ª) A DCPLA.INFO elabora o plano amostral de acordo com a estratégia adotada, seleciona a amostra de forma aleatória e envia para a Coordenação;

8ª) A Coordenação-Geral, ao receber a amostra faz a validação da mesma, verificando se está de acordo com a estratégia discutida e os campos e conteúdos consistentes;

9ª) Caso esteja de acordo a Coordenação-Geral insere o PAC no sistema com o campo ESTRATO contendo a identificação da existência de uma amostra probabilística, caso contrário envia para a DCPLA.INFO para que sejam feitas as correções;

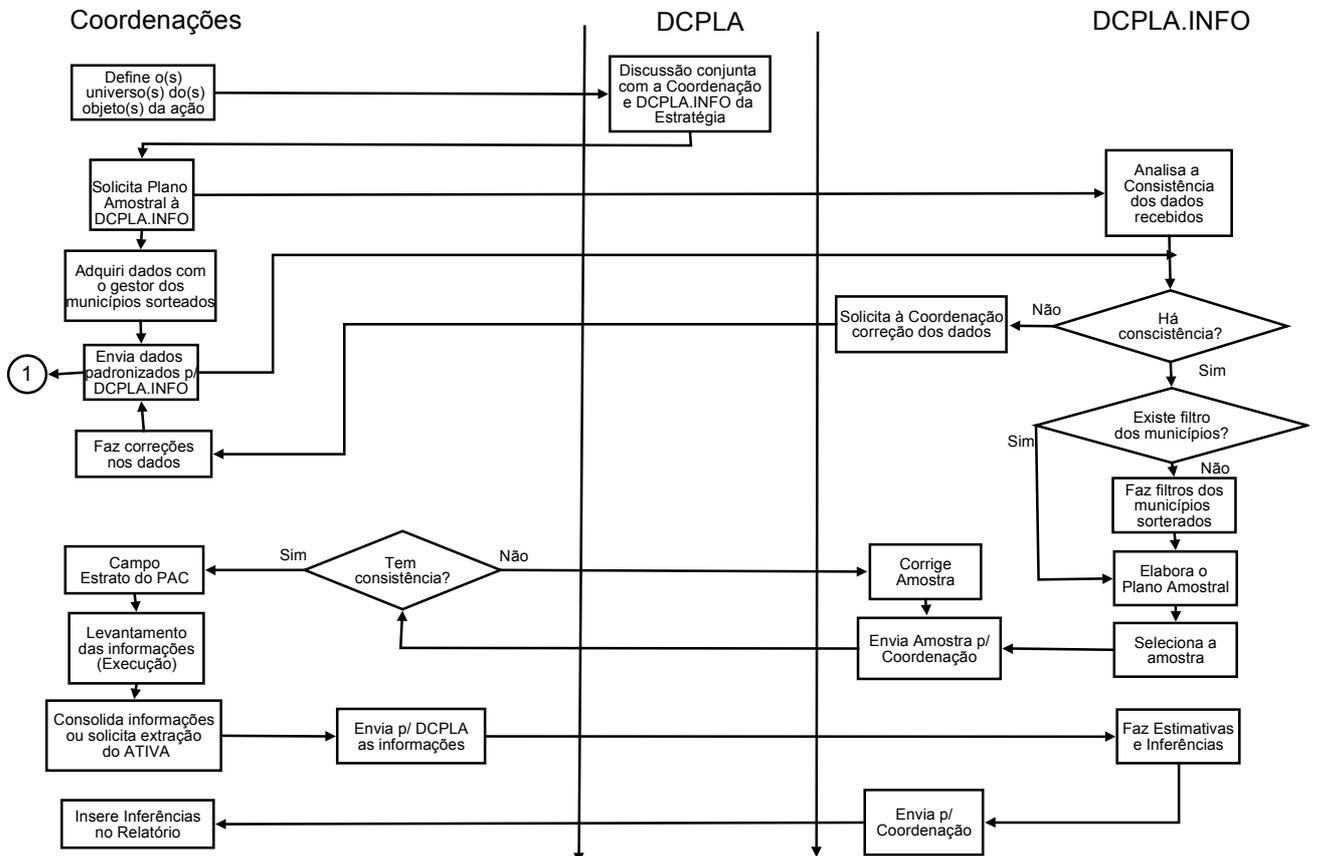
10ª) A Coordenação-Geral faz o levantamento das informações por meio de questionários ou procedimentos padronizados de forma que permita a quantificação do levantamento para possibilitar o cálculo de estimativas de proporções, médias ou totais, e envia esses levantamentos para a DCPLA.INFO, por meio de planilhas sejam elas alimentadas pelas equipes ou por via de extração do sistema ATIVA;

11ª) A DCPLA.INFO faz as estimativas de proporções, médias ou totais, baseada na solicitação da Coordenação com base nos questionários ou procedimentos e envia para a Coordenação;

12ª) A Coordenação analisa os resultados dessas estimativas e insere os mesmos nos Relatórios, referenciando o tamanho do universo (população) e suas características, as fontes de informação, escopo, nível de confiança adotado na amostra, etc, ou seja, informações relevantes para a compreensão do evento/fenômeno levantado na amostra.

### III.1 – DO FLUXO DA ROTINA PARA ELABORAÇÃO DE PLANOS AMOSTRAIS PARA FISCALIZAÇÃO DE MUNICÍPIOS DO SORTEIO

#### ROTINA DE PLANOS AMOSTRAIS PARA FISCALIZAÇÃO DE MUNICÍPIOS DO SORTEIO



1 - Os dados devem ser encaminhados conforme padrão estabelecido pela DCPLA de acordo com o que for definido na estratégia do plano amostral.

## **IV – ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA DEMANDAS EXTERNAS**

Para a elaboração de planos amostrais para as atividades de demandas externas as unidades de controle interno deverão seguir as seguintes etapas:

1ª) A Coordenação-Geral define as unidades amostrais oriundas de processos de denúncias que serão objeto de amostragem (famílias, escolas, projetos, contratos, convênios, etc);

2ª) A Coordenação-Geral informa à DCOPE da liberação do Pedido de Ação de Controle (PAC) para atendimento de demandas externas e, paralelamente, caso a DCOPE não autorize, envia a posição para a Coordenação;

3ª) A Coordenação encaminha os dados para a DCPLA.INFO se a ação já está sendo acompanhada. A DCPLA.INFO elabora o plano amostral conforme definido. Caso não seja objeto de acompanhamento a Coordenação deve procurar a DCPLA.INFO para discussão da estratégia do plano amostral;

4ª) A DCOPE analisa o processo e caso autorize a liberação de OS bem como elaboração de amostra encaminha o processo para a DCPLA;

5ª) A DCPLA.INFO elabora o plano amostral de acordo com a estratégia adotada, seleciona a amostra de forma aleatória e envia para a Coordenação;

6ª) A Coordenação-Geral, ao receber a amostra faz a validação da mesma, verificando se está de acordo com a estratégia discutida e os campos e conteúdos consistentes;

7ª) Caso esteja de acordo a Coordenação-Geral insere o PAC no sistema com o campo ESTRATO contendo a identificação da existência de uma amostra probabilística, caso contrário envia para a DCPLA.INFO para que sejam feitas as correções;

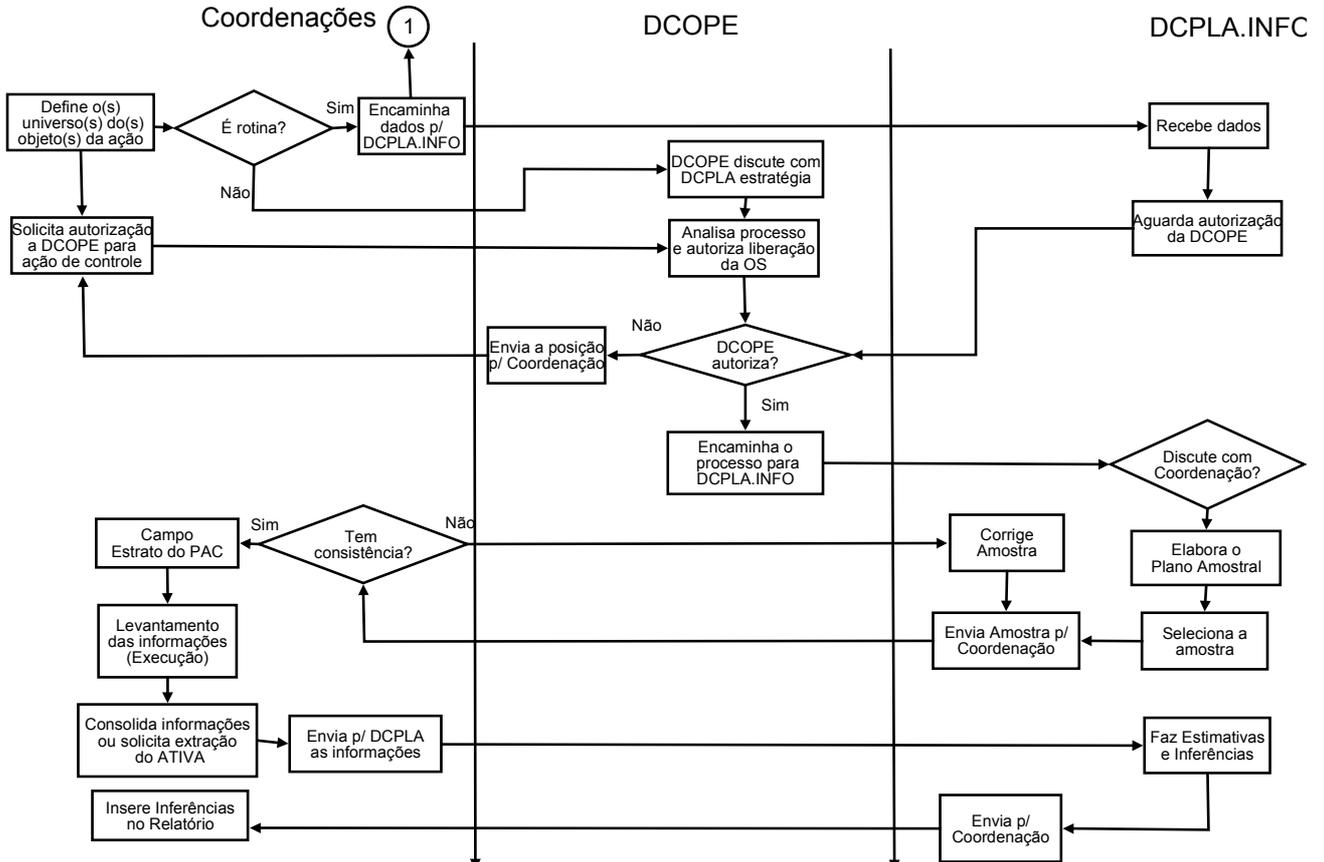
8ª) A Coordenação-Geral faz o levantamento das informações por meio de questionários ou procedimentos padronizados de forma que permita a quantificação do levantamento para possibilitar o cálculo de estimativas de proporções, médias ou totais, e envia esses levantamentos para a DCPLA.INFO, por meio de planilhas sejam elas alimentadas pelas equipes ou por via de extração do sistema ATIVA;

9ª) A DCPLA.INFO faz as estimativas de proporções, médias ou totais, baseada na solicitação da Coordenação com base nos questionários ou procedimentos e envia para a Coordenação;

10ª) A Coordenação analisa os resultados dessas estimativas e insere os mesmos nos Relatórios, referenciando o tamanho do universo (população) e suas características, as fontes de informação, escopo, nível de confiança adotado na amostra, etc, ou seja, informações relevantes para a compreensão do evento/fenômeno levantado na amostra.

**IV.1 – DO FLUXO DA ROTINA PARA ELABORAÇÃO DE PLANO AMOSTRAL PARA DEMANDAS EXTERNAS**

**ROTINA DE PLANO AMOSTRAL PARA DEMANDAS EXTERNAS**



1 - Os dados devem ser encaminhados conforme padrão estabelecido pela DCPLA de acordo com o que for definido na estratégia do plano amostral.

# APÊNDICE



*CONTROLADORIA GERAL DA UNIÃO – CGU*

# **TÉCNICAS DE ESTATÍSTICA E AMOSTRAGEM ADOTADAS NO CONTROLE INTERNO**

**2007**

**CONTROLADORIA GERAL DA UNIÃO – CGU**

**MINISTRO DE ESTADO DO CONTROLE E DA TRANSPARÊNCIA**

Jorge Hage Sobrinho

**SECRETÁRIO EXECUTIVO**

Luiz Augusto Fraga Navarro de Britto Filho

**SECRETÁRIO FEDERAL DE CONTROLE INTERNO**

Valdir Agapito Teixeira

**DIRETORA DE PLANEJAMENTO E COORDENAÇÃO DAS AÇÕES  
DE CONTROLE**

Jerri Eddie Xavier Coelho

**COORDENADOR-GERAL DE PLANEJAMENTO E AVALIAÇÃO**

Márcia da Rosa Pereira

**TÉCNICA**

Beatriz Bernardes Ribeiro

*Embora o homem individual seja um enigma insolúvel, o agregado humano representa uma certeza matemática. Nunca se pode prever, por exemplo, o que fará um homem, mas é possível prever as atitudes de certo número deles. Os indivíduos variam, mas as porcentagens permanecem constantes.*

Arthur Conan Doyle

(fala do personagem de Sherlock Holmes, em *O signo dos quatro*, 1999, p. 135)

## SUMÁRIO

1 - Introdução.....	4
2 – Conceitos Básicos de Estatística e Amostragem.....	6
3 - População e Amostra.....	7
4 - Tipos de Amostra.....	12
4.1) Elementos fundamentais na definição da amostra.....	14
4.2) Intervalo de Confiança e Margem de Erro.....	16
4.3) Conceitos Básicos de Média e Desvio Padrão.....	20
4.4) Métodos de Amostragem.....	25
4.5) Amostragem utilizando algoritmos do aplicativo ACL versus Distribuição Normal.....	32
5 - Cadastro e/ou Registros.....	34
5.1) Definição da Estratégia e Objetivo.....	34
5.2) Identificação dos dados.....	35
5.3) Caracterização e delimitação dos dados.....	35
5.4) Aquisição dos dados (variáveis e formato).....	36
5.5) Verificação da solicitação.....	39
6 – Questionário.....	39
6.1) Formulação de questões.....	40
6.2) Tipologia das alternativas de respostas.....	47
6.3) Questionário-Piloto.....	49
7 – Consolidação.....	51
7.1) Crítica.....	51
7.2) Tabulação.....	52
7.3) Cruzamentos relevantes.....	52
7.4) Estimativas e Inferências.....	53
7.5) Informação para relatório.....	56
8 – Bibliografia.....	60

## 1 - Introdução

A palavra estatística surgiu da expressão em Latim “*statisticum collegium*” que se refere aos assuntos do Estado, de onde surgiu a palavra em língua italiana *statista*, que significa “homem de estado”, ou político, e a palavra alemã *Statistik*, designando a análise de dados sobre o Estado. A palavra adquiriu um significado de coleta e classificação de dados, no início do século XIX.

A Estatística é um método que utiliza teorias probabilísticas para a explicação de eventos, estudos e experimentos. Tem por objetivo obter, organizar e analisar dados, determinar as correlações que apresentem, tirando delas suas consequências para descrição e explicação do que passou e previsão e organização do futuro.

A Estatística é também um método de desenvolvimento de conhecimento humano através do uso de dados empíricos. Baseia-se na teoria estatística, um ramo da matemática aplicada.

Nós descrevemos o nosso conhecimento (e ignorância) de forma matemática e tentamos aprender mais sobre aquilo que podemos observar. Na teoria estatística, a aleatoriedade e a incerteza são modeladas pela teoria da probabilidade. A estatística é a disciplina que trata das incertezas sobre as quantidades não observadas ou não observáveis pelo investigador. Algumas práticas estatísticas incluem, por exemplo, o planejamento, a sumarização e a interpretação de observações. Porque o objetivo da estatística é a produção da “melhor” informação possível a partir dos dados disponíveis, alguns autores sugerem que a estatística é um ramo da teoria da decisão.

Como o campo de atuação do controle de gastos públicos é vasto e as atividades a ele ligadas são bastante variadas, e como a observação exaustiva dos objetos de ações de controle é inviável, é imprescindível a utilização de

técnicas de amostragem que permitam se obter o máximo de informações com a capacidade operacional de que se dispõe. Mesmo porque, por meio de pesquisas feitas utilizando amostragem, além do custo ser menor fornecem informações em tempo hábil para a tomada de decisões. Sendo assim, esse material tem o objetivo de esclarecer alguns conceitos básicos relativos ao tema de estatística e amostragem de forma a auxiliar no desenvolvimento das atividades exercidas pela Secretaria Federal de Controle Interno.

A evolução dos métodos de amostragem está ligada à própria evolução da pesquisa. Conforme as necessidades de se obter, principalmente pelos Governos, um retrato da população em termos de distribuição geográfica e características populacionais, ou seja, informações sobre os fenômenos, sejam eles sociais ou não, visando o estabelecimento de mecanismos causais para auxiliar a formulação de políticas públicas, intensificaram-se os levantamentos de dados. Mas foi nos anos de 1930 que a pesquisa teve seu maior impulso, com um aumento considerável durante o período da Segunda Guerra Mundial.

A prática da pesquisa se consolidou devido ao reconhecimento por parte dos administradores do valor dos dados para a tomada de decisões. Sendo assim, à medida que se aumentou o interesse pelos resultados das pesquisas a aplicação dessas foi ampliada, exigindo, então, a elaboração de técnicas cada vez mais especializadas desde a sua concepção, incluindo nesta fase desde a elaboração do plano amostral até a coleta e análise dos dados, fase esta que também deve considerar o plano amostral adotado.

Além disso, com o aumento na demanda por estimativas regionais atualizadas, devido à crescente utilização de programas sociais para distribuição dos recursos entre as unidades geográficas, a necessidade de pesquisas por amostragem também aumenta, e não somente devido ao custo,

mas mais ainda pelo fato de as pesquisas por amostragem fornecerem informações em tempo hábil para tomada de decisões.

## 2 – Conceitos Básicos de Estatística e Amostragem

A teoria estatística se divide em duas áreas: estatística descritiva e a inferencial:

- **Estatística descritiva:** ocupa-se da descrição das informações contidas nos dados observados, provenientes de uma amostra ou de uma população. Ela se preocupa com a organização e a descrição dos dados observados. Esta descrição é feita através de: gráficos, tabelas, diagramas, distribuição de frequência, medidas de posição (média, moda, mediana), medidas de dispersão (variância, desvio padrão); e
- **Inferência estatística:** faz afirmações sobre as características de uma população com base na amostra extraída dessa população. É a generalização de resultados da parte (amostra) para o todo (população).

A estatística descritiva pode ser utilizada tanto para dados oriundos de uma amostra como do censo. Considera-se como censo a observação exaustiva de um determinado evento. Já a inferência estatística é baseada somente na observação de eventos que tiverem origem através de amostras.

A seguir tem-se algumas das razões que motivam a adoção de amostragem em detrimento ao censo:

- **Economicidade de meios:** quando os recursos humanos e materiais são escassos;
- **Dados mais atualizados:** quando as informações das quais se necessita se alteram no tempo sofrendo mudanças significativas;

- **Maior poder de criticidade:** como as informações são em volume menor que em um censo é possível dar mais atenção aos casos individuais e corrigir possíveis erros nas respostas;
- **Dimensionar erros:** mesmo com a adoção de censo seriam cometidos erros, porém não seria possível dimensioná-los, diferentemente da amostra onde os erros cometidos são quantificáveis; e
- **Maior qualidade dos dados:** com um número mais reduzido de unidades a serem observadas é possível se obter um ganho em qualidade dos trabalhos;

E a seguir tem-se as razões que motivam a adoção de censo, ou seja, o exame da totalidade da população:

- Quando a população é considerada muito pequena e sua amostra fica relativamente grande;
- Quando as características da população são de fácil mensuração, mesmo que a população não seja pequena;
- Quando há necessidade de alta precisão.

### **3 – População e Amostra**

*“Amostragem é o processo ou ato de construir (selecionar) uma amostra”* (Bolfarine e Bussab, 2005). A amostra, por sua vez, é um *“subconjunto de uma população, por meio do qual se estabelecem ou estimam as propriedades e as características dessa população”*. População entendida como um conjunto de todos os elementos sob investigação.

Na definição de Thompson (1992) a “*amostragem consiste na seleção de alguma parte da população observada da qual deseja-se estimar algo sobre toda a população*”. Segundo o autor, a obtenção de boas estimativas das observações amostradas envolve questões tais como o tamanho da amostra, a maneira de selecionar a amostra, qual o método de observação utilizado, quais medidas serão registradas, ou seja, é preciso estar atento aos aspectos relevantes a respeito dos dados sobre a opção de se utilizar informações auxiliares e a escolha do estimador.

**População ou Universo:** é um conjunto de unidades individuais, que podem ser pessoas, objetos ou resultados, com pelo menos uma característica comum, que se pretende conhecer. População entendida como um conjunto de todos os elementos sob investigação.

Segundo Cochran (1977), “*a população a ser amostrada deve coincidir com a população sobre a qual se deseja a informação (população alvo). Às vezes, por razões de praticabilidade ou conveniência, a população amostrada é mais restrita que a população alvo. Caso isso aconteça essa restrição deve ser mencionada nas conclusões dos resultados*”.

Não se deve confundir a população-alvo com a população referenciada nem com a população realmente amostrada (ou de pesquisa). Entende-se por **população-alvo** aquela para qual se gostaria de obter a informação. E por **população referenciada** aquela cuja definição foi baseada num sistema de referência (cadastro) e não necessariamente é a mesma da população-alvo. Já a **população de pesquisa** ou **amostrada** é aquela que será realmente coberta pela pesquisa, que pode conter unidades não previstas, ou ainda faltar unidades que, apesar de previstas, foram perdidas.

**Amostra:** é um subconjunto de uma população, por meio do qual se estabelecem ou estimam as propriedades e as características dessa população.

O objetivo da utilização de amostragem é obter informações sobre uma parte da população e fazer afirmações válidas a respeito de suas características. Mas para que essa validade se confirme é preciso que a amostra seja uma réplica em pequena escala de toda a população da qual se deseja obter informações e fazer generalizações. Sendo, portanto, imprescindível a definição apropriada da população alvo.

Além da premissa de similaridade suficiente entre os elementos presentes na amostra com os da população, outra premissa importante é a variação a ser levada em conta por se estar trabalhando com a amostra e não com o todo. Ou seja, é sempre bom lembrar que as estimativas possuem uma amplitude de variação quando da análise dos resultados, pois mesmo quando se fornece estimativas pontuais, deve-se lembrar sempre que essas não são exatas.

Como todo levantamento de dados é influenciado por limitações de custo e de tempo, fica claro que a utilização da amostragem se justifica, apesar do mito de que a eficiência dos resultados amostrados está, única e exclusivamente, ligada ao tamanho da amostra, ou seja, quanto maior a amostra, mantendo-se tudo o mais constante, maior a precisão. É sempre bom ressaltar que há um determinado tamanho de amostra a partir do qual aumentar o número de unidades não oferece ganhos significativos adicionais na precisão, isso sem mencionar os custos adicionais crescentes associados a esse aumento.

Depois de um certo tamanho de população, o ingresso de novos elementos na mesma não mais sensibiliza o tamanho da amostra, ou seja, há um certo patamar para o tamanho das amostras. *“A teoria estatística explica esse fenômeno mostrando que, depois de atingido tal patamar, a inclusão de novos elementos amostrais não agrega informações significativas, ou seja, a*

*partir desse ponto, nenhum elemento extra traz novidades ao processo, pois todas as nuances da população já foram suficientemente capturadas”*(Oliveira, 2004).

O gráfico a seguir ilustra bem isso, pois a curva, inicialmente, mostra o crescimento do tamanho da amostra em função do tamanho da população, que cresce rapidamente, para depois esse crescimento diminuir, até chegar ao ponto que ocorre praticamente uma estagnação da dimensão amostral.

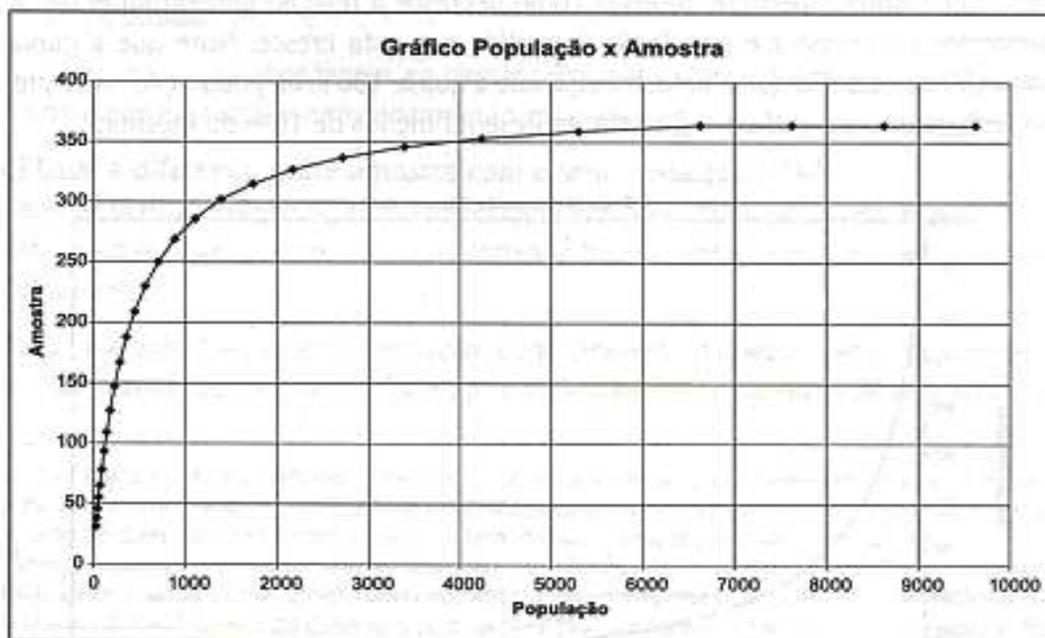


Gráfico 1 – População versus Amostra

O gráfico seguinte ilustra a disparidade entre o crescimento dos tamanhos da população comparado com os das amostras. As dimensões estão equiparadas para tamanhos menores da população e, com o crescimento da população, passam a distanciar-se progressivamente. Confirmando, assim, que o tamanho da amostra não aumenta proporcionalmente ao tamanho da população, e que isso não passa de um mito.

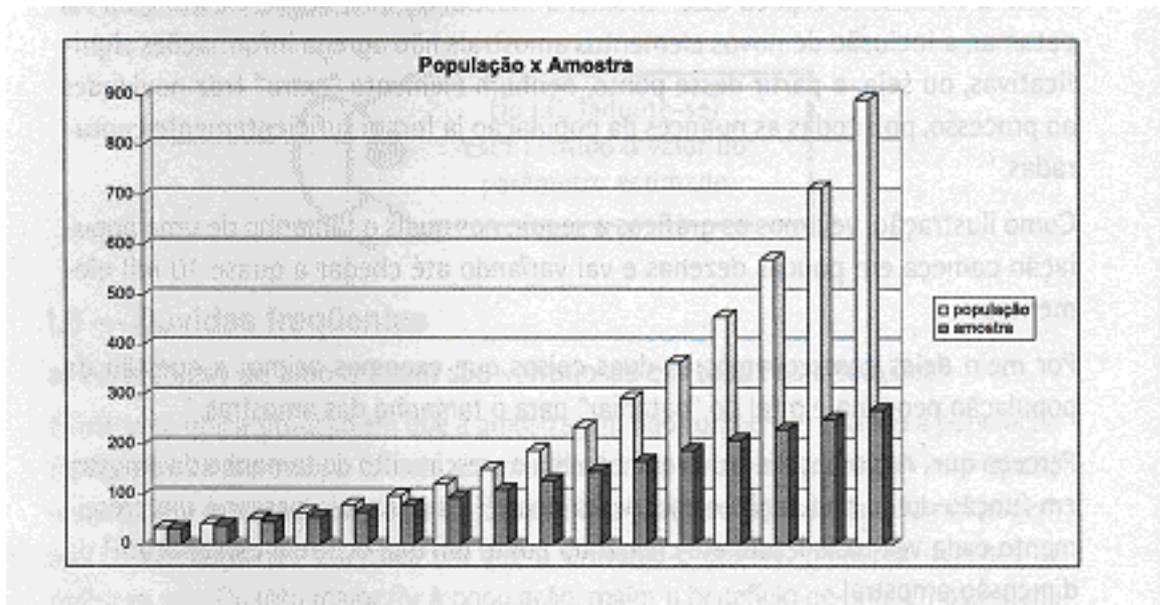


Gráfico 2 – População versus Amostra (comparação)

Outro mito que existe em torno do tema de amostragem é o de que uma boa amostra é aquela que tem 10% (dez por cento) da população. Isto não é verdade, pois o tamanho da amostra gerado por tal percentual pode ser bem menor que o necessário ou pode ser muitíssimo maior do que seria preciso (toma-se, como exemplo, as pesquisas eleitorais, imagine se fosse necessário dez por cento da população de eleitores para se obter informações a respeito das eleições presidenciais?).

Um outro mito é a idéia de que uma amostra de 30 (trinta) elementos é sempre suficiente. Em alguns casos pode até ser, mas não sempre. A origem dessa idéia vem da interpretação equivocada decorrente do Teorema do Limite Central, que preconiza que, para se realizar inferências estatísticas utilizando parâmetros da distribuição Normal, a amostra deve ser suficientemente grande, e que isso ocorre quando essa compreende, no mínimo, trinta ou mais elementos. Caso a amostra tenha menos de trinta elementos a distribuição a

ser adotada deve ser a  $t$  de student. Sendo assim, o número trinta não é um número mágico que serve para dimensionar qualquer amostra.

A definição do tamanho de amostra adequado depende da homogeneidade da população e do grau de precisão desejado.

A incerteza das estimativas está ligada à estratégia amostral; quanto mais apropriada for a estratégia, mais precisas serão as estimativas. Quando se fala em estratégia ou desenho amostral adequado significa dizer que o método escolhido deverá ir de encontro com a maneira que a população de interesse se apresenta e ainda de acordo com a maneira como se deseja obter os resultados. As características da população observada devem ser consideradas no momento de escolha da estratégia.

Na elaboração da amostra deve-se refletir sobre as variações do fenômeno associadas às diferenças regionais, tais como as diferenças entre as unidades da federação e os municípios. Mais, ainda, pode-se pensar sobre os municípios e seus respectivos portes, um município de grande porte tem diferenças estruturais em relação aos municípios de médio e pequeno porte, que podem influenciar no desempenho do programa de governo de forma crucial.

Sendo assim, no momento de elaborar um plano amostral tais diferenças devem ser consideradas. E com relação aos resultados, ao se fazer uma amostra vislumbrando estimativas em nível estadual, a amostra deve ser suficientemente representativa em termos de quantidade de municípios para garantir a possibilidade de se efetuar estimativas estaduais.

#### **4 – Tipos de Amostra**

Existem dois tipos de amostras: as probabilísticas e as não-probabilísticas.

- As **probabilísticas**, como o nome já sugere, são aquelas cujas unidades da população, sem exceção, têm alguma probabilidade de seleção na amostra, ou seja, têm a probabilidade de seleção conhecida e diferente de zero (Silva, 1998). Esse método tem como característica fundamental o fato de permitir que os resultados coletados possam ser submetidos ao tratamento estatístico, podendo ser generalizados para a população em questão. As probabilidades de seleção são produzidas por mecanismos aleatórios de sorteio, como as tabelas de números aleatórios ou algoritmos computacionais para geração de números pseudo-aleatórios.
- As **não-probabilísticas** possuem unidades da população que não têm chance de serem selecionadas e por isso este tipo de amostra não permite que se faça inferências com relação ao universo. Como exemplo clássico tem-se a amostragem por julgamento ou intencionais, sendo impossível a mensuração da probabilidade de seleção, por ser um critério puramente subjetivo (Fauze, 2001). Como esse método se baseia, principalmente, na experiência do servidor, não é possível a aplicação de tratamento estatístico aos seus resultados. A limitação se deve à impossibilidade de demonstração segundo as normas de cálculo hoje existentes de extrapolar os resultados. Contudo, ele é útil em determinados contextos, como a busca exploratória de informações ou sondagem, ou ainda quando se deseja obter informações detalhadas sobre questões particulares, durante um espaço pequeno de tempo.

#### **4.1) Elementos fundamentais na definição da amostra**

Para se obter uma amostra se faz necessário obter alguns elementos fundamentais, sem os quais fica impraticável o uso da técnica:

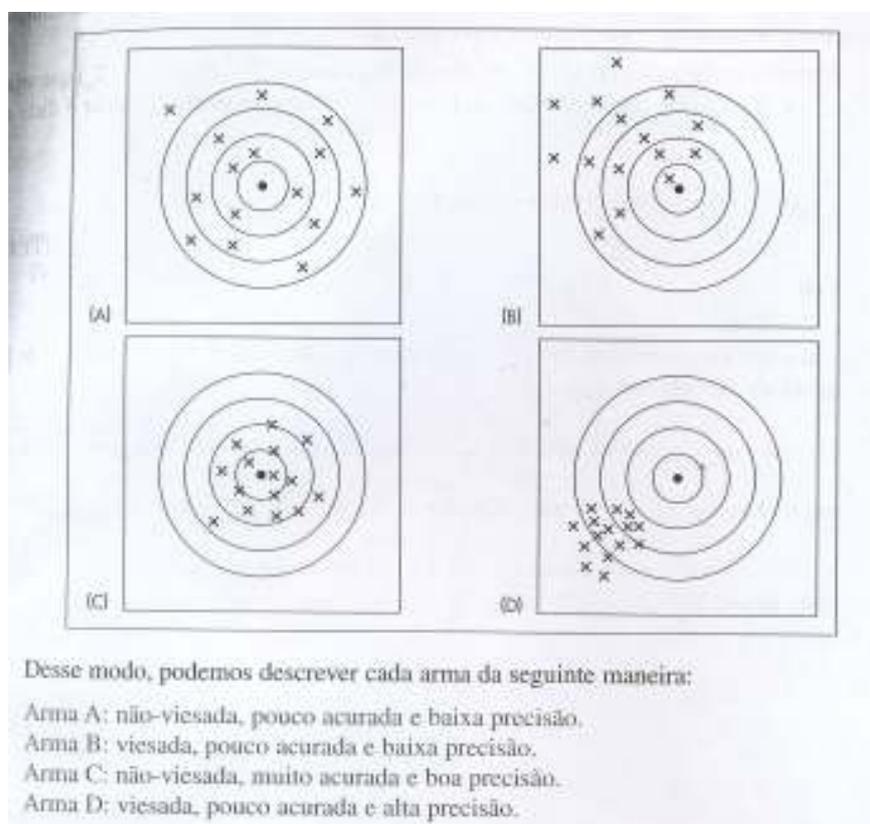
- **Definição do(s) objetivo(s):** para se elaborar um plano amostral é preciso saber exatamente qual o objetivo da ação de controle, ou seja, o que se deseja obter com os resultados e de posse deles a quais conclusões deseja-se chegar. Um exemplo é obter uma estimativa do percentual de repetições no cadastro do Programa Bolsa Família, para tanto é necessário se fazer uma amostra de famílias cadastradas e verificar por meio de entrevistas se a família existe mesmo, ou se foi erro de cadastramento, ou ainda, fraude;
- **Definição do universo ou população de interesse:** quando os objetivos da ação estão bem definidos fica claro o estabelecimento de qual é a população alvo a ser considerada para a elaboração da amostra, bem como a definição das unidades de amostragem (elementos). A definição da população é necessária, pois de acordo como essa se apresenta, será possível a escolha do método de amostragem a ser adotado, bem como todas as suas considerações *a priori*, como por exemplo, quando se deseja fazer uma amostra com base no cruzamento de bases do PRONAF e do PROGER, a base encaminhada para o PROGER, ou seja, o universo, continha somente as coincidências com o PRONAF. Portanto, futuramente, será inviável inferir sobre o PROGER como um todo, bem como sobre o PRONAF. Sendo assim, as afirmações e conclusões somente poderão dizer respeito ao universo de coincidência;
- **Definição do elemento (unidade amostral):** a unidade amostral é aquela que sofrerá a ação de controle diretamente, ou seja, que será objeto de algum questionamento por parte do controle via . Por exemplo, no caso do acompanhamento do Programa Bolsa Família, a

unidade de observação para saber sobre as questões relativas aos critérios de elegibilidade e ao cumprimento das condicionalidades, é a família beneficiária;

- **Estabelecimento da precisão:** a precisão nada mais é que o máximo que se deseja errar ao se obter os resultados. A precisão influenciará diretamente no tamanho da amostra, quanto mais precisos forem os resultados que se deseja obter, maior será o dimensionamento da amostra, e vice-versa;
- **Aleatoriedade:** a seleção dos elementos da amostra devem ser rigorosamente aleatória. A aleatoriedade é um pressuposto fundamental para se evitar inclusão de tendenciosidade nos resultados, ou seja, que não se está influenciando na escolha dos elementos da amostra com seu juízo pessoal, que além de não ser padronizado, coloca em risco os resultados da amostra, tornando-a enviesada e portanto resultando em informações equivocadas a respeito da população alvo da amostragem. Sendo assim, uma amostra sem aleatoriedade é tendenciosa, levando o servidor a conclusões errôneas;
- **Representatividade:** esse é outro fator fundamental quando se fala em amostragem. Para que a amostra seja representativa ela deve ser, de fato, um retrato da população que ela objetiva representar só que em uma dimensão menor, ou seja, a amostra deve conter todos os tipos de elementos da população. Isso significa que todas as características que a população possui devem ser preservadas na amostra, por exemplo, se uma população possui elementos de várias classes e em proporções diferentes, a amostra deve retratar essas diferenças.

## 4.2) Intervalo de Confiança e Margem de Erro

Ao se falar em variação é preciso abordar algumas características necessárias a uma amostra consistente, tais como: precisão, eficiência e ausência de viés (Almeida, 2003). A precisão está relacionada com a exatidão que a estimativa possui, ou seja, é a medida do erro amostral. A eficiência está ligada à comparação entre diversos planos de amostragem, um plano é considerado mais eficiente quando produz estimativas mais confiáveis ou precisas que outro, mantendo-se os custos, ou ainda com custos menores. A ausência de viés se refere à ausência de erros sistemáticos, ou não amostrais, que são erros provenientes de falhas na coleta, erros de digitação, erros de medida, ou de interpretação no momento da codificação das respostas subjetivas. Tais erros podem influenciar os resultados em determinadas direções. A figura 3 a seguir ilustra um pouco os conceitos de precisão e viés.



### Gráfico 3 – Precisão e Viés

Por isso, sempre que se fala em estimativas deve-se fazer referência tanto às estimativas pontuais quanto às intervalares, porque quando se fala somente em estimativas pontuais não é possível julgar qual a magnitude do erro que estamos cometendo. Daí surge a idéia de construir os intervalos de confiança, é o procedimento que permite julgar qual é a possível magnitude do erro que estamos cometendo no momento de utilizar uma amostra para falarmos sobre uma população.

Os intervalos de confiança são baseados na distribuição amostral do estimador pontual. Ele pode ou não conter o parâmetro que se deseja conhecer, no caso, uma média, um total, ou uma proporção. Então, em todas as possibilidades de amostras de uma determinada população, quando se estabelece um nível de confiança de 95%, significa dizer que em 95% dessas possíveis amostras, quando estimamos o parâmetro de interesse, o intervalo estimado tem o parâmetro, que é desconhecido, dentro do intervalo e em 5% das possíveis amostras o parâmetro de interesse não se encontra dentro do intervalo estimado. A figura a seguir ilustra bem esse conceito.

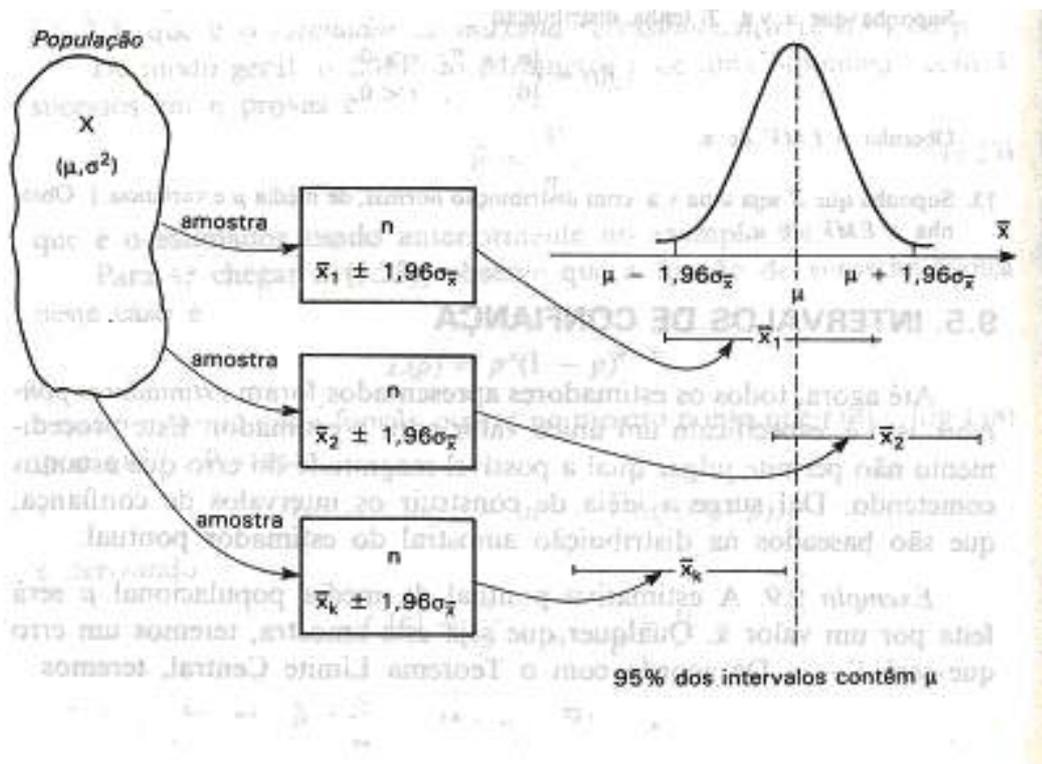


Gráfico 4 – Intervalo de Confiança

Conforme pode-se observar na figura 3 ao escolher uma amostra sua média e desvio padrão estimados permite construir o intervalo a seguir:

$$\left( \bar{x}_0 - 1,96\sigma_{\bar{x}_0}, \bar{x}_0 + 1,96\sigma_{\bar{x}_0} \right)$$

Esse intervalo pode ou não conter o parâmetro  $\mu$ , mas em 95% de confiança de que contenha.

Para ilustrar melhor esse conceito, considere o exemplo seguinte de experimento de simulação. Foram geradas 20 amostras de tamanho  $n = 25$  de uma distribuição normal de média  $\mu = 5$  e desvio padrão  $\alpha = 3$ . Para cada amostra que foi contruída o intervalo de confiança para  $\mu$ , com o coeficiente

de confiança  $\gamma = 0,95$ , na figura 4, tem-se os intervalos representados e nota-se que em três deles (amostra 5, 14 e 15) não contêm a média  $\mu = 5$ .

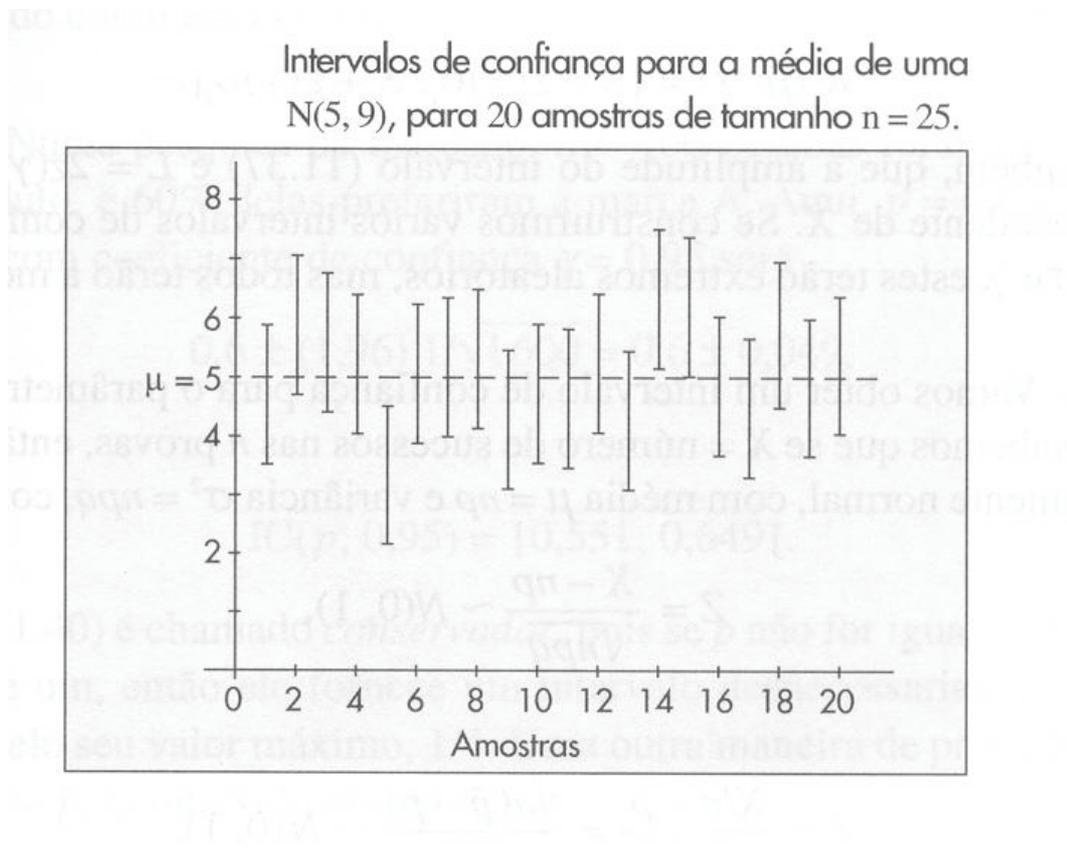


Figura 5 – Intervalos de Confiança para a média.

Portanto, não é correto afirmar que quando uma estimativa qualquer é calculada pode-se falar em 95% de certeza de que ela é coerente com o parâmetro, pois nível de confiança não é certeza, principalmente quando se está trabalhando com uma amostra que não é representativa da população em questão. Por isso, é tão importante que toda amostra seja representativa da população em estudo, para garantir que suas estimativas sejam as mais próximas possíveis dos parâmetros de interesse.

### 4.3) Conceitos Básicos de Média e Desvio Padrão

- **Média Aritmética:** é a soma das observações dividida pelo número delas, como mostra a fórmula a seguir.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

O exemplo a seguir ilustra o conceito de média.

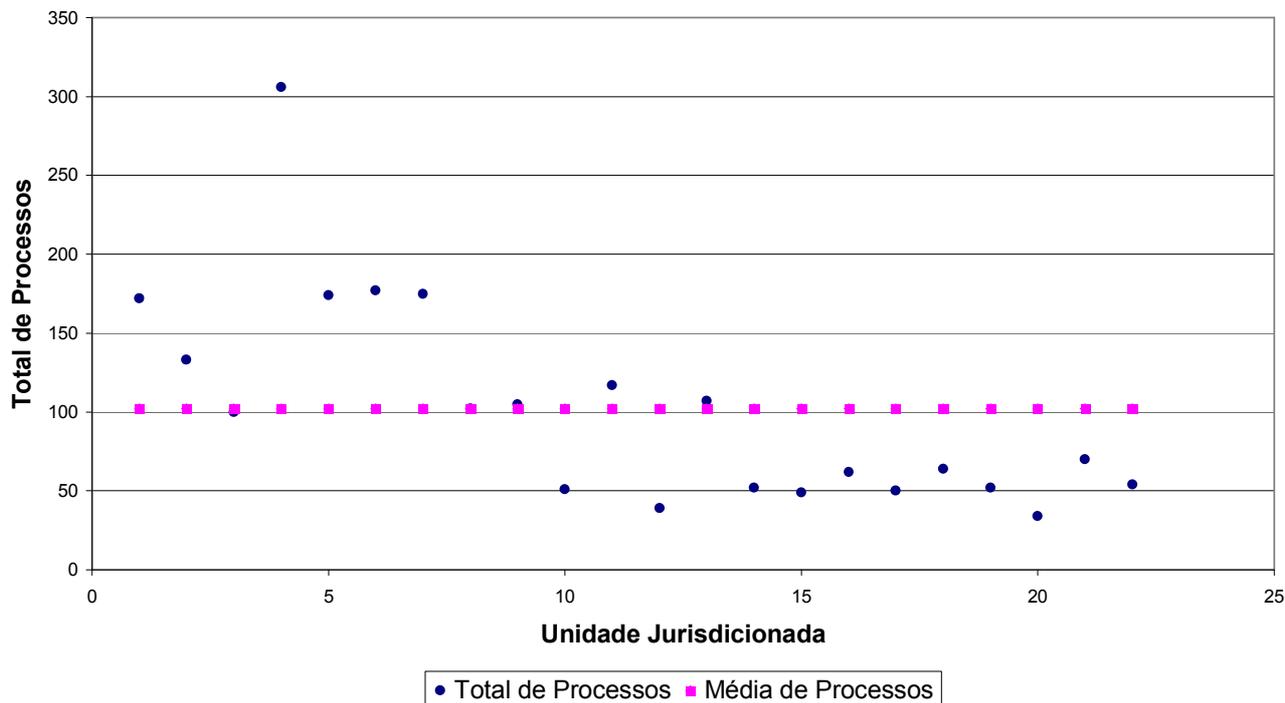
**Tabela 2 – Quantidade de Processos de Compras em algumas Unidades Jurisdicionadas do MF**

Observação	Unidade Jurisdicionada	Total de Processos
1	170016	172
2	170038	133
3	170075	100
4	170085	306
5	170114	174
6	170131	177
7	170175	175
8	170207	102
9	170214	105
10	170025	51
11	170045	117
12	170050	39
13	170055	107
14	170064	52
15	170100	49
16	170106	62
17	170166	50
18	170190	64
19	170195	52
20	170345	34
21	170346	70
22	170347	54
	<b>Total</b>	<b>2.245</b>

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{2.245}{22} = 102$$

**Gráfico 4 - Média de Processos de Compras nas UJ do MF comparada com as observações de cada UJ**

### Quantidade de Processos da UJ



- **Desvio Padrão:** é o erro médio que se está cometendo ao tentar substituir cada observação pela medida resumo do conjunto de dados, no caso, a média. O desvio é uma medida consequente da variância, que mede a distância entre média das observações e cada observação, ou seja, a raiz da variância, porque como a variância tem essa diferença de cada observação e a média elevada ao quadrado, para tornar a diferença sempre positiva, o resultado dela é muito elevado, necessitando tirar a raiz para tornar a medida da dispersão compatível com a média. A seguir tem-se as fórmulas da variância e do desvio padrão.

$$V(x) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

$$s = \sqrt{V(x)}$$

Conforme o exemplo a cima, tem-se:

$$V(x) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} = 4.378$$

$$s = \sqrt{V(x)} = 66,17$$

Diante dos resultados do exemplo, a interpretação é que a média de processos nas UJ é de 102, com um desvio de 66 processos, para mais ou para menos. O que pode se perceber, observando o gráfico 4, é que existe uma observação relativa a UJ de número 170085, cuja quantidade de processos é bastante elevada, de 306, ou seja, ao observar o conjunto dos dados essa unidade é um ponto discrepante, destoando das demais unidades, influenciando a estimativa da média. Nesse caso, quando existem dados extremos, tanto para mais, quanto para menos, o ideal é utilizar outra medida de posição, como por exemplo a mediana.

- **Mediana:** é a medida que ocupa a posição central da série que divide as observações ordenadas segundo suas grandezas (crescente ou decrescente) ao meio, tendo 50% abaixo dela e 50% a cima. Como por exemplo se em cinco observações apresentarem os valores 3, 4, 7, 8 e 8, a

mediana é 7, correspondendo à terceira observação. Quando o número de observações for par, usa-se a média aritmética das duas observações centrais. Assim como (3, 4, 7, 8, 8 e 9) a mediana é igual a 7,5.

$$Md = \frac{7+8}{2} = 7,5$$

No exemplo, a mediana das observações é igual a 85 processos. O único problema é que não é possível se medir matematicamente a dispersão da mediana.

Uma outra solução para esse problema seria retirar do conjunto dos dados a observação discrepante e recalcular as estimativas, como a seguir.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{1.939}{21} \cong 92$$

$$V(x) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \cong 2.419$$

$$s = \sqrt{V(x)} \cong 49$$

A estimativa da média agora está mais precisa e mais coerente com o conjunto dos dados, a média sem a observação discrepante é de 92 processos e com um desvio padrão de 49, para mais ou para menos. E a mediana passou a ser de 70 processos, isso significa dizer que 50% das unidades tem menos de 70 processos e 50% das unidades tem mais de 70 processos.

#### 4.4) Métodos de Amostragem

O desenho amostral consiste na escolha do método de amostragem. Os métodos freqüentemente utilizados são os seguintes:

- **Amostragem Aleatória Simples – AAS:**

A amostra aleatória simples (AAS) é o método de amostragem mais simples, como o próprio nome diz, desde que se tenha identificado as unidades de observação de uma população, caso contrário outros métodos são mais recomendáveis.

A AAS pode ser com reposição ou sem reposição. Na amostra com reposição é permitido que a unidade possa ser sorteada mais de uma vez e na sem reposição a unidade selecionada é removida da população (Bussab e Morettin, 2003).

Na definição de Cochran (1977) a “*AAS é um método de seleção de  $n$  unidades das combinações de  $N$ , tal que em cada uma das combinações tomadas  $N$  a  $n$ , distintas de amostras tem chance igual de ser selecionada*”. Thompson (1992) em sua definição de AAS complementa a anterior acrescentando que na obtenção das  $n$  unidades selecionadas “*em cada passo toda unidade da população não selecionada tem igual chance de seleção...a probabilidade de inclusão da  $i$ -ésima unidade da população na amostra é  $\pi_i = n/N$ ... mas somente na AAS em cada possível amostra de tamanho  $n$  que se tem a mesma probabilidade de seleção*”.

Quando se observa a definição de AAS de Sheaffer *et all* (1996), que afirma que se “*uma amostra de tamanho  $n$  é desenhada a partir de uma população de tamanho  $N$ , de tal forma que toda amostra possível de tamanho  $n$  tem a mesma chance de ser selecionada*”. Ele alega ainda que “*uma*

*conseqüência direta dessa definição é que todos elementos na população têm a mesma chance de serem selecionados; porém esta declaração não pode ser tomada como uma definição de AAS porque não implica que todas amostras de tamanho  $n$  tem a mesma chance de seleção”.*

Ao se falar em AAS e que os elementos selecionados devem ter a mesma chance de seleção, remete-se à possibilidade de se fazer uma AAS com probabilidades de seleção de elementos diferenciadas, que é possível, porém, quando se utiliza pesos diferenciados na seleção das unidades amostrais, deve-se considerá-los no momento de cálculo das estimativas, essa amostra passa a ser considerada uma amostra complexa, apesar do método escolhido ser AAS.

Na prática, a AAS é pouco utilizada, pois tem pouco potencial de redução de variação e, conseqüentemente, de redução do tamanho da amostra. Mas é muito útil quando não se tem muitas informações adicionais a respeito do fenômeno em estudo.

É importante lembrar que, para cada tipo de estimativa (média, total, proporção, razão, etc), existe uma fórmula correspondente de cálculo, tanto do tamanho da amostra quanto para seus respectivos estimadores. Ou seja, para cada método de amostragem adotado, existe uma fórmula apropriada àquele método para se calcular as estimativas.

## **MÉDIA POPULACIONAL**

$$\mu = \frac{\sum_{i=1}^N X_i}{N}$$

## MÉDIA AMOSTRAL – AMOSTRA ALEATÓRIA SIMPLES

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

## MÉDIA AMOSTRAL – AMOSTRA ESTRATIFICADA

$$\hat{\mu}_{st} = \bar{x}_{st} = \frac{1}{N} [N_1 \bar{x}_1 + N_2 \bar{x}_2 + \dots + N_L \bar{x}_L] = \frac{1}{N} \sum_{i=1}^L N_i \bar{x}_i$$

onde L é o número de estratos.

- **Amostragem Sistemática – AS:**

A amostra sistemática é uma variante da AAS, é, na verdade, uma regra de seleção dos elementos da amostra baseada na posição onde estes se encontram, que pode variar conforme a população se apresenta. Ela é eficiente desde que a forma como a população se apresenta seja misturada do ponto de vista da variável em estudo, ou seja, o intervalo de seleção não esteja correlacionado com alguma característica das unidades da população. Por exemplo, quando não se possui em meio digital uma lista de beneficiários do PRONAF em vários municípios objeto de denúncia e essa lista se encontra somente no município a ser fiscalizado, ou seja, em campo, mas se sabe quantos beneficiários existem por município, calcula-se a amostra, porém a forma de selecionar será a seleção sistemática dos beneficiários segundo uma regra, como: de uma lista de 100, cuja a amostra foi de 20, divide-se o total de 100 pela amostra de 20, daí tem-se 5 que será selecionado na lista de cada 5, seleciona-se um beneficiário, até o final da lista, que totalizará os 20;

- **Amostragem Estratificada – AE:**

A amostra estratificada consiste em separar os elementos da população em grupos distintos, chamados de estratos, onde são selecionadas amostras simples dentro destes estratos (Cohran, 1977). Segundo Sheaffer *et al* (1996) “o princípio da estratificação é dividir a população de tal forma que as unidades dentro do estrato são as mais semelhantes possíveis”. O objetivo é evidenciar as possíveis diferenças existentes nos grupos distintos da população, como também reduzir a variação total, pois, em sendo os elementos mais homogêneos dentro do estrato, a variância dentro do estrato é menor, e quando as variâncias dos estratos são somadas essa é menor que a variação total quando os estratos são desconsiderados.

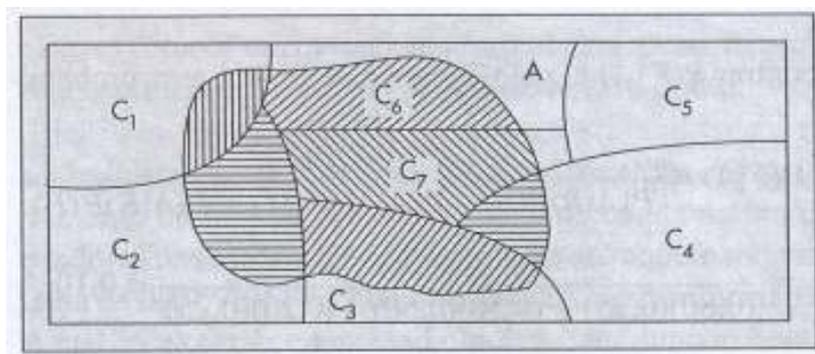


Figura 6 – Espaço Amostral Estratificado.

Sendo assim, a população é dividida em subpopulações (estratos) tal que:

$$N_1 + N_2 + \dots + N_L = N$$

E o tamanho da amostra será dada por:

$$n_1 + n_2 + \dots + n_L = n$$

A vantagem desse método é que é possível se obter estimativas em cada estrato (Sheaffer *et al*,1996), possibilitando desta forma comparações entre os mesmos. Porém, é preciso que se tenha informações teóricas sobre quais variáveis são relevantes para explicar a divisão da população de um jeito e não de outro.

Conforme o método da AAS, o método de amostragem estratificada (AE) possui uma fórmula para cada tipo de estimador que se deseja estimar (média, total, proporção, razão, etc).

Um exemplo desse tipo de amostra é a amostra de contratos de licitações para a auditoria de gestão, onde os elementos são estratificados por modalidade de licitação (concorrência, convite, pregão, etc), imagina-se que dentro do estrato esses se comportam de forma homogênea, pois tem o mesmo mecanismo legal, os mesmos procedimentos e/ou ritos, etc, enquanto se juntássemos no mesmo bolo as discrepâncias seriam bem maiores do que quando agrupados dentro das suas classificações.

**Tabela 4.1 - Amostra Estratificada por “Modalidade”, ponderada pelo valor, da Unidade Jurisdicionada 170016 do Ministério da Fazenda, para Auditoria de Acompanhamento de 2006**

<b>UASG/UG</b>	<b>Modalidade</b>	<b>Universo Processo</b>	<b>Amostra Processo</b>	<b>Universo Valor Empenho</b>	<b>Amostra Valor Empenho</b>
170016	Dispensa de licitação	65	3	9.617.919,44	7.971.091,65
170016	Inexigibilidade	10	2	3.095.221,11	1.351.258,46
170016	Pregão	76	4	43.658.267,12	30.852.208,07
170016	Tomada de preço	5	2	1.134.339,67	743.378,92
170016	concorrência	11	2	5.623.443,90	1.706.096,18
170016	convite	5	2	111.933,78	0,00
<b>Total</b>		<b>172</b>	<b>15</b>	<b>63.241.125,02</b>	<b>42.624.033,28</b>

Nota: Nível de Confiança 99 e erro 6

- **Amostragem por Conglomerados – AC:**

Quando não é possível se obter um cadastro contendo todas as unidades elementares que serão objeto de amostragem, ou esse não é adequado, ou o custo de obtê-lo é elevado, ou ainda quando o deslocamento para identificar as unidades em campo é caro ou consome muito tempo, recomenda-se a adoção do método de amostragem por conglomerados ( Bolfarine e Bussab, 2005).

A amostra por conglomerados é constituída por unidades populacionais reunidas em grupos denominados conglomerados, eles se dividem na população em unidades primárias, e essas unidades primárias serão formadas por unidades secundárias (Sheaffer *et all*, 1996). O resultado disso é a diminuição do espaço amostral, e conseqüentemente, há um aumento do espalhamento e da variação, porque os conglomerados, na prática, tem tamanhos diferentes, diminuindo assim a precisão, porém a representatividade é preservada e os custos são menores.

Ela pode ser em um ou mais estágios: em um estágio é quando selecionados os grupos, todos seus elementos farão parte da amostra; e em mais estágios quando os grupos são arrançados segundo uma hierarquia até chegar às unidades elementares. Após a seleção dos grupos, no primeiro estágio, se estabelece os grupos existentes no segundo estágio, que estão dentro dos grupos do primeiro e se procede um outro sorteio para determinar a

amostra dos grupos no segundo estágio, depois faz-se a seleção de uma amostra de elementos desses grupos. A figura a seguir apresenta um exemplo de conglomerado em múltiplos estágios.

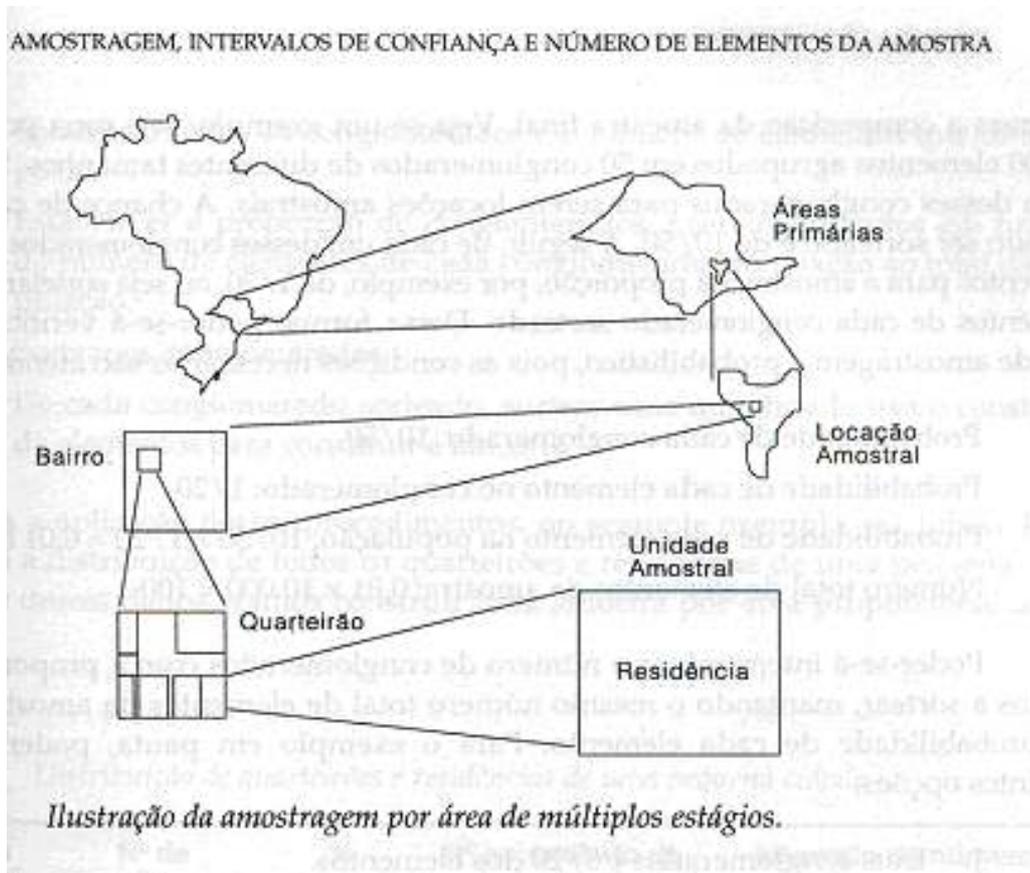


Figura 7 – Amostragem por área em múltiplos estágios.

Conforme a figura ilustra, é possível ver que a cada estágio devem ser produzidas estimativas que irão compor a estimativa nos estágios superiores. Como por exemplo, quando se deseja estimar a ocorrência de um evento qualquer em nível nacional, como as pesquisas oficiais do Instituto Brasileiro de Geografia e Estatística (IBGE) tais como a Pesquisa Nacional por Amostragem de Domicílio (PNAD) ou a Pesquisa de Orçamento Familiar (POF), desde que a amostra tenha sido elaborada conforme o esquema apresentado na figura acima, deve-se estimar a proporção em nível de

quarteirão, para posteriormente estimar em nível de bairro, depois município, unidade da federação para, finalmente, chegar na estimativa do Brasil.

Os métodos para fazer as amostras em cada um desses estágios podem variar, sendo possível se fazer inúmeras combinações de métodos, podendo ser uma amostra estratificada no primeiro estágio e uma amostra aleatória simples no segundo estágio, etc. A escolha irá depender da necessidade e/ou conveniência do pesquisador, e de como a população se apresenta.

#### **4.5) Amostragem utilizando algoritmos do aplicativo ACL versus Distribuição Normal**

O quadro 1 a seguir ilustra a comparação entre o método de amostras elaboradas utilizando algoritmos do aplicativo ACL e o método da Distribuição Normal, que é a adotada pela área de Estatística da SFC. O objetivo dessa comparação é fazer um alerta mostrando em que casos é permitido a adoção de um método e do outro, bem como os resultados que se pode obter de ambos os casos.

#### **Comparação entre Amostras feitas através de algoritmos do ACL e os algoritmos que usem a Distribuição Normal**

<b>ACL</b>	<b>Distribuição Normal</b>
Efetua cálculos visando apenas os limites superiores dos intervalos de confiança.	Podem fornecer ambos os limites: superior e inferior.
Os cálculos de tamanho de amostra se abstraem da dimensão da população, pois é suposto que as populações são sempre “grandes”, algo que impacta fortemente o tamanho das amostras... a tal ponto que, em populações pequenas, o tamanho da amostra oriunda da população pequena pode superar o tamanho das amostras de populações grandes.	Valem-se do tamanho da população para dimensionar proporcionalmente a amostra, o que evita que esta exceda o universo.
Não serve para cálculos de dimensão amostral em estimação de mais de uma proporção independente.	Podem efetuar todos os cálculos de dimensão amostral em estimação de mais de uma proporção independente.
Só se presta a estimacões relacionadas a incidências de falhas.	Podem ser aplicados em todas as estimacões, inclusive as não relacionadas com questões de inconformidade como, por exemplo, os casos de estimacões de médias e totais, o quais dizem respeito a diversas outras variáveis (idade, peso, salário, número de filhos, etc).
<p>Não considera a (estimacão da) variabilidade populacional, ou seja, independentemente do universo ser homogêneo ou heterogêneo, é determinado um mesmo valor para o dimensionamento da amostra.</p> <p>Isso significa uma potencial superestimacão do tamanho da amostra em populações muito homogêneas, o que se traduz em desperdício de tempo e dinheiro.</p>	São sensíveis à variância (estimada) da população. Assim, não existe a possibilidade de ser requerida uma amostra grande para um universo homogêneo (como pode acontecer no ACL).
Não trabalha com todas as especificacões de erro.	Podem admitir, irrestritamente, qualquer quantificacão de erro.

## 5 - Cadastro e/ou Registros:

## **5.1) Definição da Estratégia e Objetivo**

Como já foi dito anteriormente, a definição clara dos objetivos, seja na estratégia contida na abordagem do acompanhamento das ações de governo, ou nas auditorias, ou nas apurações de denúncias, é fundamental para a elaboração de um plano amostral consistente e coerente com os resultados que se deseja obter. Ou seja, é preciso se ter em mente o que se pretende responder com a ação de controle e o que se deseja com a amostra para que o levantamento das informações, bem como sua consolidação, auxilie de fato na construção das conclusões, de forma a fornecer subsídios para compor as evidências necessárias ao julgamento da ação governamental ou da gestão.

Para isso, é preciso levantar o máximo de informações possíveis a respeito do objeto da ação de controle que se está planejando. Nesse momento é que se questiona como se efetiva a ação, quais são as unidades de observação que serão examinadas para possibilitar a comprovação das hipóteses formuladas na estratégia de ação de controle. Como, por exemplo, na estratégia do Programa Bolsa Família as unidades observadas são: Prefeitura, Agente Operador (CEF), Conselho, Escolas e Famílias. Para identificá-las é preciso que estas pertençam a um cadastro.

É no cadastro/registros que se encontram as informações sobre as unidades a serem observadas. É importante que a Coordenação conheça os cadastros existentes relacionados à ação governamental de interesse, como esses foram construídos, suas características (dimensões, formatos, manutenção, atualização, acessos, restrições), ou seja, tudo que diz respeito ao mesmo. Deve-se conhecer com profundidade os cadastros, pois são através deles que é possível a identificação individual do beneficiário, se for o caso, e dos recursos a eles repassados.

## **5.2) Identificação dos dados**

Além dessas características dos cadastros/registros, deve-se estar muito atento ao conteúdo dos mesmos. Ou seja, deve-se conhecer os dados que compõem o cadastro, quais e quantas são as variáveis e seus formatos, quais as unidades de observação, como por exemplo no Bolsa Família, cada registro se refere a uma criança pertencente a família, repetindo-se então tantos registros de responsável pela família quantas crianças tenham em dada família.

Deve-se conhecer as relações existentes no cadastro, pois uma variável pode ser o resultado de outras, citando novamente o exemplo do Bolsa Família, cada família recebe o benefício básico e o variável que atinge o limite de três crianças, desde que estas estejam freqüentando a escola, então a variável TOTAL é o somatório dos benefícios básico mais o benefício variável.

## **5.3) Caracterização e delimitação dos dados**

Outro fator importante nos dados é a identificação da localização espacial do benefício, para possibilitar a fiscalização física do objeto em questão, como, por exemplo, o endereço ou o nome do estabelecimento, seja uma escola, um banco, etc. Muitos cadastros tem a identificação, mas como a Coordenação solicita ao Gestor os dados, muitas vezes alguns dados são omitidos, sendo imprescindível a solicitação expressa dos campos com suas caracterizações.

Quando da solicitação dos dados para o Gestor da ação governamental é importante a delimitação desses por período, seja por exercício, ou por período de disponibilização do recurso, folha de pagamento

mensal, por atualização cadastral, etc, pois é nesse momento que se define a população de referência objeto de alguma ação de controle.

Além da delimitação temporal, há também a delimitação espacial ou geográfica, pois existem várias ações que são regionalizadas e isso deve estar bem claro no momento da descrição da população que será objeto de amostragem e posteriormente na consolidação dos resultados que comporão os relatórios.

Outra delimitação possível é a de volume de recursos, dependendo do objetivo da ação, pode-se, por exemplo, separar para censo aqueles valores considerados muito elevados, ou aquelas unidades que já se tenha evidências de serem problemáticas, de forma que essas não influenciem tendenciosamente os resultados e seja garantida a sua observação na ação de controle.

#### **5.4) Aquisição dos dados (variáveis e formato)**

No momento da solicitação do cadastro deve-se ter bem claro quais serão os campos necessários para realização do trabalho, não somente em termos de elaboração da amostra mas também de fornecer informações suficientes para os fiscais localizarem o objeto da ação de controle.

Outra informação importante é o formato da variável, muitas vezes os campos chegam num formato não apropriado, como, por exemplo, um campo numérico que está com o formato de texto, ou vice-versa. A importância nesse caso é que campos numéricos que serão somados, caso estejam em formato texto, sejam convertidos, caso contrário, dificultará no momento da soma.

Além desses fatores, existem outros que dizem respeito ao campo, que é o seu formato e o seu tamanho, pois alguns campos são identificados no

cadastro como valores em Reais, e outros não; uns com separação contendo barra, traço, espaço ou ponto, e outros não; quando o campo é em formato texto uns estão com letras maiúsculas outros com letras minúsculas; uns com acento e cedilha, outros não. Todos esses formatos são para fins de informática identificados como diferentes. Ou seja, o seu agrupamento, quando necessário, não será possível, pois não é possível somar campos diferentes nem contabilizar coisas diferentes. A seguir, tem-se alguns exemplos do que pode acontecer em termos das variáveis e formatos dos cadastros:

**1)** Os cadastros, na sua maioria, não contém a localização da unidade de observação, objeto de amostragem. Quando da solicitação das bases aos Gestores, exigir que as variáveis de localização constem nas referidas bases, por exemplo, a identificação do bairro, ou endereço completo da unidade de observação, seja ela um beneficiário, uma agência, etc. O objetivo disso é, ao se elaborar o Plano Amostral, levar em conta a racionalização da amostra, para assim, diminuir o número de deslocamentos dos fiscais no mesmo município e possibilitar a realização da amostra por completo;

**2)** Os cadastros contém, na sua maioria, erros de digitação, como, por exemplo, no Bolsa Escola, uma única escola foi digitada 09 (nove) vezes com formatos diferentes e, sendo assim, o aplicativo de informática, seja ele qual for, identifica como sendo 09 (nove) escolas diferentes, quando na verdade não são, inflacionando as bases;

**3)** Outro tipo de erro é a troca de nome de variáveis com conteúdos adversos aos que deveriam, e ainda a existência de repetições de beneficiários, ou seja, o mesmo recebendo várias vezes os recursos, não se sabe se são homônimos, ou fraudes;

4) Ao se fazer a crítica nos cadastros ficar atento a campos em branco, preenchimento de números em campos cujo conteúdo são palavras, e vice-versa;

**Recomendações:**

a) Todas as colunas devem conter apenas as informações correspondentes ao título da variável.

**Sugestão:** Utilizar no aplicativo EXCEL, a opção, **DADOS → FILTRAR → AUTOFILTRO;**

b) Formatar as células considerando a distinção entre valores numéricos e caracteres, ou seja, manter em uma mesma coluna apenas uma opção (número ou texto).

**Sugestão:** para observações só numéricas, utilizar, no aplicativo EXCEL, a opção, **FORMATAR CÉLULAS → NÚMERO → NÚMERO (ATENÇÃO: NÃO “USAR SEPARADOR DE 1000 (.)”**). E ainda, **COPIAR → COLAR ESPECIAL → VALORES**. Nos casos em que a variável endereço, ou semelhantes, contem só número, **utilizar o apóstrofe antes do número para transformá-la em texto;**

c) Para os casos que necessitem fazer corte de alguma observação, verificar se existe a variável separadamente e, em caso contrário, criá-la. **Ex.:** Situações em que o corte será por **ano** e a variável disponível contem a data completa (dia, mês e ano), é necessária a criação de uma nova variável contendo apenas a observação “**ano**”;

d) Para **variável valor ou correspondentes**, que será utilizada como critério de materialidade, retirar a separação por pontos e/ou vírgulas; retirar o R\$, verificar se contém apenas uma moeda, e configurar conforme item b);

## **5.5) Verificação da solicitação**

Quando é solicitado um cadastro ao Gestor é preciso verificar se a solicitação foi devidamente atendida. É de suma importância ver se todas as variáveis solicitadas estão no cadastro, nos devidos formatos, se não há inconsistências tais como ausência de preenchimento de alguns campos ou repetição de registros, etc.

Caso haja alguma inconsistência deve-se solicitar a correção do cadastro por parte do Gestor que é o responsável prioritário pela informação. Essa preocupação com a fidedignidade do cadastro se deve ao fato de que todas as afirmações feitas posteriormente serão com base nesses registros. Por isso é tão importante conhecê-los com profundidade, para saber solicitar as variáveis no formato adequado para atender às necessidades do controle interno.

## **6 - Questionário**

Existem várias técnicas de coleta de dados, sejam elas diretas ou indiretas. Algumas das técnicas indiretas são: pesquisa documental, pesquisa bibliográfica. As diretas podem ser: pesquisa de campo, pesquisa de laboratório, observação, entrevista, questionário, formulário, etc.

O questionário é uma dessas técnicas, e bastante utilizada no âmbito da CGU, pois é por meio dele que são coletados os dados que fornecerão subsídios aos trabalhos, principalmente de fiscalização.

A elaboração do questionário é uma etapa das mais importantes no processo de levantamento de informações. Essa elaboração é um processo longo e complexo, porém é o questionário que irá possibilitar a coleta dados

de forma uniformizada tanto no que diz respeito aos questionamentos quanto às possibilidades de resposta. Para que o questionário seja bem elaborado é necessário que o formulador das questões conheça bem o assunto que está sendo objeto de coleta.

Segundo Marconi e Lakatos (1985), *“questionário é um instrumento de coleta de dados, constituído por uma série ordenada de perguntas, que devem ser respondidas por escrito e sem a presença do entrevistador. Em geral, o pesquisador envia o questionário ao informante, depois de preenchido, o pesquisado devolve o mesmo”*. Além de enviar o questionário esse deve ser acompanhado das instruções de preenchimento do mesmo, onde constarão as explicações sobre a natureza do levantamento, sua importância e a necessidade de obter as respostas, para que o informante tome ciência do que se deseja com ele.

### **6.1) Formulação de questões**

Antes de elaborar as perguntas do questionário é necessário tomar algumas decisões quanto ao seu conteúdo, tais como:

- 1) Esta pergunta é necessária?
- 2) Qual sua utilidade para os objetivos da ação de controle?
- 3) Há necessidade de várias perguntas sobre o assunto desta pergunta?
- 4) O assunto não é muito complexo para constar de uma só pergunta?
- 5) Será que a sua subdivisão não pode gerar maior número de perguntas mais simples de forma a facilitar o entendimento e obtenção das respostas?
- 6) A pergunta abrange adequadamente o tópico?
- 7) Será que todos os aspectos importantes para a pesquisa sobre este tópico serão obtidos da forma como a pergunta foi elaborada?

- 8) Em perguntas sobre opiniões, basta saber se a opinião é a favor ou contra, ou será importante medir os graus de favorabilidade/desfavorabilidade?
- 9) As pessoas têm a informação necessária para responder à pergunta?
- 10) É um assunto que pode descrever adequadamente?
- 11) A pergunta exige respostas que a pessoa não tem condições de responder, ou de responder de forma precisa?
- 12) O tópico perguntado faz parte das experiências e (ou) do conhecimento da maioria do público alvo da pesquisa?
- 13) Quando se trata de pergunta sobre o comportamento passado, será que as pessoas têm ainda viva na memória recordação suficiente para relatá-las com a precisão necessária para a pesquisa?
- 14) Há outro grupo de pessoas onde possam ser obtidos os mesmos dados de forma mais adequada e precisa do que junto ao público alvo inicial da pesquisa?
- 15) Será necessário ter diferentes versões da pergunta adaptadas aos diferentes públicos alvo da pesquisa?
- 16) Deve a pergunta ser mais direta, objetiva e ligada à experiência da pessoa, ou indireta, ampla e mais geral?
- 17) Está apresentada de forma excessivamente geral?
- 18) Pode a informação ser mais facilmente obtida se for mais direta, objetiva e relacionada com o comportamento da pessoa?
- 19) Tende o conteúdo da pergunta a captar dados sobre situações momentâneas e específicas ao invés de captar o que é típico no respondente?
- 20) O conteúdo da pergunta não estará viesado ou carregado em determinada direção, de forma a causar vieses na obtenção das respostas?

- 21) A pergunta pode ser considerada neutra?
- 22) Da forma como está elaborada será considerada neutra por pessoas com opiniões opostas sobre o assunto?
- 23) A pergunta contém opiniões ou pré-julgamentos relacionados ao assunto abordado?
- 24) As pessoas estarão predispostas a fornecer o dado solicitado?
- 25) O tema abordado é muito íntimo, perturbador ou expõe socialmente as pessoas de forma a provocar resistências, evasões, mentiras e insinceridades?
- 26) Quais as objeções que uma pessoa poderia ter para responder ao conteúdo desta pergunta?
- 27) Será que a pergunta não lhe dará a impressão de estar sendo exposta e examinada?
- 28) É possível obter o dado desejado de forma não ofensiva ou será melhor omiti-lo?
- 29) Imaginando a situação ambiental típica em que o instrumento será aplicado, será que o conteúdo da pergunta não receberá objeções e resistências para ser respondida de forma sincera?
- 30) O conteúdo da pergunta tende a provocar desejos e reações emocionais que provocarão uma resposta não verdadeira?

Além dessas decisões existem outras que dizem respeito à redação das questões:

- 1) A pergunta pode ser mal compreendida?
- 2) Contém termos difíceis ou com significado dúbio ou obscuro?
- 3) As palavras utilizadas são de domínio do público alvo?
- 4) Estão sendo utilizados termos técnicos de uso exclusivo de atividades profissionais específicas?
- 5) A redação da pergunta é simples e curta?

- 6) Existe palavreado desnecessário?
- 7) Existe outro sentido que a pergunta pode ter para quem responde?
- 8) A pergunta exprime, e de forma balanceada, todas as alternativas?
- 9) Será que não estamos apresentando apenas um dos lados do tópico em estudo?
- 10) A pergunta, ou o seu preâmbulo, deixa claro e uniforme o quadro de referência em que a resposta deve ser dada?
- 11) A pergunta apresenta a base para a resposta?
- 12) A pergunta consegue deixar explícito que a resposta a ser dada refere-se ao que a pessoa pensa ser a verdade e não o que ela deseja que seja a verdade?
- 13) A redação da pergunta está viesada?
- 14) Está emocionalmente carregada ou deformada para dirigir as respostas em determinado sentido?
- 15) Emprega termos do tipo “A opinião geral das pessoas é que...” para em seguida perguntar a opinião do respondente?
- 16) Emprega a opinião de pessoas famosas para em seguida perguntar a opinião do respondente?
- 17) Emprega termos superlativos que sugestionam a resposta para determinado sentido?
- 18) A pergunta seria considerada neutra para pessoas com opiniões opostas a respeito do mesmo assunto?
- 19) A redação da pergunta tende a despertar objeções, de qualquer tipo, do respondente?
- 20) Uma redação solicitando diretamente a renda do respondente tende a ter mais objeções do que outra que apenas solicite vários indicadores para partir deles deduzir a classe de renda do respondente.

- 21) Uma redação mais pessoal ou menos pessoal apresentaria melhores resultados?
- 22) Uma redação mais direta ou menos direta apresentaria melhores resultados?

Existem perguntas de vários tipos:

**- Abertas**

- Qual a sua opinião sobre o Programa Bolsa Família?

**- Fechadas e Dicotômicas**

- O roteiro da viagem confere com a finalidade declarada?
- a) Sim (S)
- b) Não (N)

**- Múltipla escolha:** com mostruário e estimação ou avaliação

- Qual o local do recebimento do benefício?
- 1 – Caixa
- 2 – Lotérica
- 3 – Prefeitura
- 4 - Outros
- 5 - Não se Aplica
- 6 - Impossível Avaliar

Elas variam conforme o objetivo:

- **Fato:** dizem respeito a questões concretas, tangíveis, fáceis de precisar; portanto, referem-se a dados objetivos (idade, sexo, profissão, domicílio, estado civil, religião);

- **Ação:** referem-se a atitudes ou decisões tomadas pelo indivíduo. São objetivas, às vezes diretas demais, podendo, em alguns casos, despertar

certa desconfiança por parte do informante, influenciando no seu grau de sinceridade (uso de drogas ou preservativos);

- **Intenção:** tentam averiguar o procedimento do indivíduo em determinadas circunstâncias. Não se pode confiar na sinceridade da resposta; entretanto, os resultados podem ser considerados aproximativos (pesquisas eleitorais);

- **Opinião:** representam a parte básica das pesquisas de opinião, ou seja, quando o foco é levantar a opinião do entrevistado, mais utilizadas na área de *marketing*;

- **Índice ou teste:** são utilizadas para levantar informações que suscitam medo quando formuladas diretamente, fazem parte daquelas consideradas socialmente inaceitáveis. Mediante este tipo de perguntas, procura-se estudar um fenômeno por meio de um sistema ou índice revelador do mesmo. É utilizada no caso em que a pergunta direta é considerada imprópria ou indiscreta. Normalmente falar sobre a renda é um tabu, então são elaboradas, com base em estudos, categorias sócio-econômicas de forma a possibilitar que o entrevistado responda indiretamente a renda.

As questões devem ser formuladas de forma direta e objetiva, ou seja, de forma a não confundir o fiscal no momento da aplicação. É muito comum o formulador do questionário, ao tentar fazer um questionário menor incluir mais de um questionamento na mesma questão. Isso é permitido, desde que as alternativas de respostas permitam responder a ambas perguntas ao mesmo tempo, sem perder informação, mas o mais indicado é que cada questionamento se atente a uma questão apenas. Por exemplo:

1) O número e a data do documento estão de acordo com o previsto na portaria publicada?

A) Sim

B) Não

A pergunta se refere a duas coisas diferentes (número e data) que podem estar de acordo ou não, conjuntamente ou separadamente. O ideal é que as alternativas de repostas contemplassem todas as possibilidades de resposta.

2) 1) O número e a data do documento estão de acordo com o previsto na portaria publicada?

A) Sim, somente o número

B) Sim, somente a data

C) Sim, ambos

D) Não

Existem algumas deformações provocadas por perguntas que podem influenciar o resultado, tais como:

**- Conformismo ou Deformação Conservadora:**

- tendência a responder afirmativamente;

- medo de mudança da situação vigente ou conformismo.

A) Você acha que o servidor deveria ter aceito o convite?

B) Você acha que o servidor teve razão em recusar o convite?

**- Efeitos de Palavras e Estereótipos:** o medo de determinadas palavras, em virtude de sua carga emocional, pode levar a distorções das respostas (nazista, comunista, capitalista). O exemplo a seguir sobre pesquisa realizada no Estados Unidos, em 1939, mostra bem isso:

A) Você acha que se deveria fazer uma interpretação legal da Constituição, de maneira a impedir que o Presidente possa solicitar um terceiro mandato?

B) Você estaria a favor de modificar a Constituição dos Estados Unidos para impedir que o presidente possa solicitar um terceiro mandato?

	Forma A	Forma B
Sim	36%	26%
Não	50%	65%

Sem opinião 14%                      9%

**- Influência das Personalidades:**

- O **prestígio positivo**: aceitação de opiniões ou atitudes somente por serem atribuídas a alguma personalidade conhecida e respeitada;

- O **prestígio negativo**: o informante rejeita as opiniões e atitudes precedentes de uma personalidade que se tenha desprestigiado.

- Influência da Simpatia ou da Antipatia: a mesma pergunta obterá resultados diferentes, de acordo com o aspecto afetivo

## **6.2) Tipologia das alternativas de respostas**

Outro aspecto importante da elaboração de questões é ficar atento às alternativas de respostas. Essas devem cobrir todas as possibilidades existentes, ou pelo menos ter o máximo de cobertura possível, pois caso não se tenha previsto alguma resposta, as demais não poderão assumir 100% das possibilidades e o levantamento fica comprometido, como por exemplo:

1) Houve alguma alteração do contrato?

A) Sim, houve alteração de contrato.

B) Sim, houve alteração de localidade.

C) Não houve alteração.

Nesse exemplo faltou uma alternativa não prevista, de ambas alterações de contrato e localidade, que quando o questionário foi aplicado mostrou-se necessária, porém não existia a alternativa, não permitido responder duas alternativas ao mesmo tempo.

Para facilitar a consolidação dos questionamentos deve-se codificar as respostas para que, no momento, tanto da coleta quanto da alimentação dos dados coletados em meio digital, essas sejam feitas de forma prática com o

intuito de diminuir ao máximo a incidência de erros, conforme exemplo da questão acima, as alternativas estão codificadas em A, B ou C.

Existem algumas regras que visam facilitar e evitar que sejam feitas confusões nas fases seguintes de digitação e processamento. Essas regras são as seguintes:

- as opções de respostas precisam ser **mutuamente exclusivas** e **coletivamente exaustivas**. Mutuamente exclusivas significa dizer que uma resposta somente pode ser assinalada uma única vez e o respondente não pode ter nenhuma dúvida quanto a resposta a assinalar. Coletivamente exaustiva significa dizer que as opções de respostas devem cobrir todas as opções possíveis, não permitindo que nenhuma resposta fique sem opção de registro;

- quando não é possível ser coletivamente exaustiva a questão deve possuir alternativas complementares tais como: “outras”, “nenhuma das anteriores” e essas seguidas de esclarecimentos dessas outras para possibilitar futuras catalogações e incorporações dessa alternativa na pergunta;

- preferir sempre coletar os dados em categorias mais desagregadas, pois no momento de tabulação sempre será possível agregar categorias, mas será impossível desagregá-las;

- obedecer às convenções quando o processamento for eletrônico:

- usar somente códigos numéricos;
- não utilizar nenhum caractere especial;
- não deixar nenhuma resposta em branco;
- utilizar cada coluna do *layout* da planilha eletrônica para cada variável;
- nunca utilizar o zero como opção de resposta, isto pode gerar muita confusão no processamento;
- estabelecer e utilizar um código diferente e padronizado para diferentes tipos de não respostas, como por exemplo: “Não sabe” – 90, “Não

respondeu/recusou responder” – 91, “Não se aplica” – 92, “Dado excluído pela verificação/edição” – 93, “Não resposta devido a salto da questão” – 94, etc;

- em perguntas fechadas com respostas múltiplas (em que várias opções de respostas podem ser escolhidas simultaneamente), utilizar, para codificação, cada opção de resposta como uma pergunta simples (ou como se fosse uma variável) cujas respostas possíveis fossem “sim” ou “não”.

Deve-se evitar ao máximo a inclusão de questões subjetivas ou abertas no questionário, pois essas questões não são passíveis de catalogação imediata e, devem ser lidas para posteriormente serem categorizadas. Além disso, sugere-se que: ou seja a mesma pessoa que faça essa leitura e categorização, ou que a categorização seja feita anteriormente e todos que irão participar do processo de catalogação das questões tenham essas informações de forma uniforme.

### **6.3) Questionário-Piloto**

O questionário-piloto é o teste do mesmo. O objetivo é testar se o questionário está adequado ao contexto das informações que se pretende levantar. Apesar de se ter consciência de que o piloto demanda um tempo razoável de planejamento, ele é de suma importância, pois quando bem feito, elimina problemas de etapas posteriores que podem demandar muito mais tempo na correção que o piloto demandaria. Um aspecto importante a respeito do piloto é que existem várias fases do mesmo:

- uma em que quem o elaborou deve aplicá-lo para perceber se ele tem condições de levantar as informações necessárias para se evidenciar as

hipóteses elaboradas inicialmente e, se preciso, alterar o questionário de acordo com a realidade encontrada no momento da sua aplicação;

- uma outra fase é aquela em que o piloto é aplicado por uma outra pessoa que não participou do processo de elaboração do questionário, porém na companhia de quem elaborou, mas essa deve acompanhar somente como observador e não deve fazer nenhum comentário ou explicação adicional a respeito do questionário, para quem está aplicando, devendo simplesmente anotar as dificuldades de quem aplicou e reformular o mesmo ou fornecer mais explicações a respeito do questionário no próprio corpo do mesmo, pois o questionário deve falar por si mesmo;
- a última fase do piloto é a aplicação do questionário sem a presença de quem o elaborou, devendo as dificuldades encontradas serem apresentadas ao formulador para revisão. Posteriormente, o formulador do questionário deve analisar os resultados e, se achar necessário, reformular o mesmo.

Caso as condições de tempo e recursos permitam todas as fases devem ser cumpridas. Entretanto, limitadas as condições, o indispensável é a aplicação da última fase, que é a aplicação do piloto por terceiros, sem a presença do formulador do questionário, pois o questionário deve falar por si, ou seja, ser auto-suficiente no que diz respeito a informações necessárias à sua aplicação.

O questionário-piloto deve ser testado *a priori* para que se possa verificar a aplicabilidade do mesmo e ainda incrementá-lo com questões anteriormente não previstas, bem como adicionar novas alternativas de respostas às perguntas já formuladas porém não consideradas.

Outra finalidade do teste é verificar se a linguagem é acessível, se existem perguntas supérfluas ou causam embaraço ao informante, se as questões obedecem a determinada ordem ou se são muito numerosas.

O questionário-piloto se presta também para testar outros instrumentos que envolvem o levantamento: o tempo necessário à sua aplicação, os custos, os instrumentos auxiliares à consolidação (planilhas, inserção no sistema), treinamentos, manuais, etc.

Verificadas as falhas, deve-se reformular o questionário, seja conservando, modificando, ou eliminando itens, explicitando melhor alguns ou melhorando a redação de outros.

## **7 - Consolidação**

Após o levantamento feito é preciso transformar os dados coletados em informação gerencial para evidenciar as hipóteses elaboradas no início do trabalho e tirar as conclusões a respeito

### **7.1) Crítica**

Assim que os dados estiverem digitados e armazenados no banco de dados, e antes de iniciar a tabulação/processamento, são necessárias verificações e depurações dos dados de possíveis erros e inconscistências que não tenham sido percebidos nas verificações/edições no campo ou escritório. Existem três tipos de verificação eletrônica:

- verificação de valores estranhos: valores que não foram previsto nas alternativas de respostas a perguntas fechadas;

- verificação da consistência: coerência interna, que pode ser vista observando respostas a perguntas chaves que implicam na ocorrência de outras;

- verificação de valores extremados: valores que fogem completamente ao campo de variação da variável.

A crítica é uma forma de verificação do preenchimento adequado dos campos, pois quando uma questão depende de uma antecedente afirmativa, caso esteja preenchida de forma negativa, a sua dependente não poderia ter sido preenchida.

## **7.2) Tabulação**

A tabulação é definida por Abramo (1979) como sendo “*a arrumação dos dados em tabelas, de maneira e permitir a verificação das relações que eles guardam entre si*”. Os dados obtidos através do processamento são transferidos para as tabelas para serem observados e submetidos à análise.

## **7.3) Cruzamentos relevantes**

Os cruzamentos funcionam tanto como crítica aos dados coletados, como análise dos mesmos. Quando o cruzamento é utilizado na análise tem a função de filtrar questões que foram concebidas com esse propósito, ou seja, de permitir explicação de ocorrência de eventos correlatos. Por exemplo, no Programa Bolsa Família existe uma questão relativa ao pagamento da Bolsa a menor, ou a maior, ou não ter havido o pagamento, ou esse estar correto, e outra questão onde se questiona onde esse pagamento está sendo

feito, se pela Caixa, ou pela Prefeitura, ou outros lugares, então seria interessante cruzar as respostas de pagamento a maior com o lugar onde estes estão sendo feitos. Além dessa tem outra questão levantando se houve condições impostas ao pagamento da Bolsa, que poderia ser cruzada também com a questão onde foi efetuado o pagamento.

#### **7.4) Estimativas e Inferências**

Inferência estatística é o processo pelo qual estatísticos tiram conclusões acerca da população usando informação de uma amostra. E as estimativas de uma amostra são o nosso "melhor chute" para os verdadeiros valores populacionais.

A seguir será descrito um exemplo prático de uma amostra elaborada para fiscalização do programa de Merenda Escolar, no estado do Espírito Santo.

Foi elaborada uma amostra de conglomerado em dois estágios, onde o primeiro estágio era amostra de municípios e no segundo estágio uma amostra de escolas nesses municípios. Na seleção de municípios foi feita uma ponderação de municípios pela quantidade de escolas e na seleção de escolas por quantidade de alunos nas escolas. Vale a pena lembrar que dar pesos diferenciados no momento da seleção não significa que somente serão selecionadas unidades amostrais de grande porte.

O universo de municípios beneficiados pelo programa e objeto de fiscalização é de 77 municípios e o universo de escolas no estado é de 2.767.

No primeiro estágio a amostra de municípios foi de 6 municípios, que totalizavam 322 escolas. No segundo estágio, foi elaborada uma amostra de 40

escolas do universo de 322. Sendo assim, as 40 escolas são representativas das 322, que por sua vez são representativas das 2.767, com os respectivos níveis de confiança e margem de erro estabelecidos *a priori*.

### **Merenda Escolar no Estado do Espírito Santo**

Amostra de Escolas Municipais, por Conglomerado, em Dois Estágios:

#### **Primeiro Estágio - Amostra de Municípios**

( Estratificada pela qtde de escolas por municípios )

Tamanho do Universo de Municípios: 77

Tamanho do Universo de Escolas: 2.767

Tamanho da Amostra de Municípios: 6

Tamanho do Universo de Escolas no 1º Estágio: 322

#### **Segundo Estágio - Amostra de Escolas**

Tamanho do Universo de Municípios no 1º Estágio: 6

Tamanho do Universo de Escolas no 1º Estágio: 322

Tamanho da Amostra de Municípios no 2º Estágio: 6

Tamanho do Amostra de Escolas no 2º Estágio: 40

A seguir tem-se o exemplo das estimativas e inferências relativas a uma pergunta do questionário aplicado nas fiscalizações das escolas. A margem de erro de tais estimativas e inferências variam, pois cada alternativa de resposta

corresponde a uma estimativa da proporção (presença ou ausência do atributo) e, respectivamente, à sua variação. Então para cada alternativa existe uma margem de erro.

A inferência se refere à generalização dos resultados dessas proporções para a população referenciada de escolas no estado do Espírito Santo. Ou seja, o quanto esse percentual de resposta, em números absolutos de escola significa.

### CONSTATAÇÃO - OBJETO (01.01.00)

<b>(01.00.00) - Houve merenda todos os dias?</b>	<b>Proporção</b>	<b>Qtde Escolas</b>
(01.01.01) - Sim, não houve falta de merenda.	57,69% ( 47,40% ;67,97%)	1.596 ( 1.312 ; 1.881 )
(01.01.02) - Faltou por até 10 dias letivos.	6,24% ( 3,25% ; 9,21% )	173 ( 90 ; 255 )
(01.03.00) - Faltou por até 20 dias letivos.	12,44% ( 9,16% ;15,75% )	344 ( 253 ; 436 )
(01.04.00) - Faltou por mais de 20 dias letivos.	6,25% ( 3,25% ; 9,21% )	173 ( 90 ; 255 )
(01.05.00) - Objeto inexistente.	17,38% ( 13,21% ;21,54%)	481 ( 366 ; 596 )
<b>Total</b>	<b>100%</b>	<b>2.767</b>

Diante desse resultado pode-se afirmar que em aproximadamente 173 escolas faltou merenda por mais de 20 dias letivos e que essa quantidade de escolas pode variar entre 90, no melhor caso, até 255 escolas, no pior caso.

Outra afirmação é de que das 2.767 escolas do estado em apenas 1.596 não houve falta de merenda.

### **7.5) Informação para relatório**

O objetivo da utilização de métodos estatísticos nos trabalhos de auditoria e fiscalização é proporcionar informações consistentes a respeito das constatações levantadas em campo, ou seja, fornecer elementos que contribuam na sustentação das evidências encontradas e, conseqüentemente, darão suporte na tomada de decisão, seja a respeito da regularidade de aplicação dos recursos públicos, seja na compreensão de como se dá a política pública nos vários níveis de execução.

É com base nos dados levantados em campo que o analista poderá tirar suas conclusões, então é de vital importância que as informações sejam descritas de maneira que se possa conhecer as dimensões do fenômeno em investigação. Sendo assim, o relatório deve conter a descrição da população objeto de amostragem, com suas características, como ela foi obtida, ou seja, sua origem (extração Siafi, cadastros obtidos nos ministérios, dados da internet, etc), os montantes que essa população envolve, tanto em termos de recursos quanto de quantidades.

Além das informações a respeito da população o relatório deve descrever o método de amostragem adotado, o nível de confiança adotado, bem como caso a amostra não tenha sido coletada por completo, o quanto ela representa da amostra original, com as explicações cabíveis sobre a impossibilidade de obtenção da amostra inicialmente elaborada.

A seguir tem-se dois exemplos de inclusão das informações no Relatório:

- “Especificamente no âmbito do MTur, as despesas relativas a diárias, no decorrer de 2005, totalizaram R\$ 1.211.460,40, correspondendo a 1861 OB, sendo que R\$ 814.210,74 se destinaram a diárias no país e R\$ 397.249,66, no exterior. Do volume de recursos despendido cerca de 94%, equivalente a 1759 OB, destinaram-se ao pagamento de diárias nacionais, enquanto as internacionais corresponderam a 6% do total, o que equivale à emissão de 103 OB. Esse quantitativo envolveu 370 beneficiários distintos, incluindo servidores e colaboradores eventuais. Considerando-se a realização de mais de uma viagem por beneficiário, esse quantitativo se eleva a ocorrência de 414 CPF. Das 1758 ordens bancárias relativas à concessão de diárias nacionais, 391 foram destinadas a colaboradores eventuais, ou seja, cerca de 22%, não se identificando para este segmento nenhum pagamento de diárias internacionais.

Foi definida uma amostra probabilística em dois estágios, com um nível de confiança de 80%, dos atos a serem auditados. Considerou-se, inicialmente, os CPF beneficiados com a emissão de Ordens Bancárias-OB relativas ao pagamento de diárias, e posteriormente, uma amostra das OB emitidas em favor daqueles CPF já selecionados no primeiro estágio.

A distribuição por programa/ação deu-se da seguinte forma: cerca de 70%, encontra-se no âmbito da execução do **Programa 1166 – Turismo no Brasil: Uma Viagem Para Todos**, o que equivale a 1304 OB (1234 nacionais e 70 internacionais), envolvendo o volume de recursos da ordem R\$ 794.920,12. Em seguida vem o **Programa 1001 – Gestão de Política de Turismo** com aproximadamente 28% das OB (521, sendo 491 nacionais e 30 internacionais), totalizando a aplicação de R\$ 390.290,00. O **Programa 1163 – Brasil Turístico Internacional**, foi o que teve menor participação nas despesas com diárias, ou seja, apenas cerca de 2% das OB, o que equivale a 36 (33 nacionais e 3 internacionais). Aproximadamente 78% das despesas,

*que totalizaram R\$ 26.250,28, foram com diárias internacionais, o que corresponde a R\$ 20.629,73.”; e*

- *No âmbito do Ministério do Meio Ambiente, as despesas com diárias, no exercício de 2005, totalizaram R\$ 1.064.226,65, valor somado em 1.571 OB's emitidas. Desse montante, o valor de R\$ 547.133,43 destinou-se a diárias nacionais, o que representa 51,41% do total, pagas em 1.428 OB's, e R\$ 517.093,22 foram gastos com diárias internacionais, o que equivale a 48,58% do total, pagas em 143 OB's. Esse quantitativo envolveu 346 beneficiários (CPF) distintos, até mesmo servidores e colaboradores eventuais.*

*A amostra probabilística aleatória estratificada por tipo de diária (nacional e internacional) foi definida com um nível de confiança de 80%. Desta forma, foram selecionados 73 processos para conferência e análise, que resultou nas seguintes constatações:*

- *as autorizações para concessões de diárias foram realizadas de acordo com as diretrizes normativas do órgão em 100% dos casos analisados;*
- *o afastamento do servidor estava vinculado à atividade desenvolvida ou à atribuição do cargo, com a respectiva comprovação, em todos os processos analisados;*
- *o roteiro da viagem conferia com a finalidade declarada em 93,13% dos casos. Nos outros casos observamos que as divergências estavam relacionadas à conveniência e oportunidade da atuação do gestor público;*
- *os cálculos do pagamento de diárias estavam corretos em 94,83% nas diárias nacionais, podendo variar entre 91,12% e 98,54%, e em 94,72% dos processos de diárias internacionais, podendo variar entre 87,44% e 100%. A diferença constatada já foi objeto de Nota de Auditoria n.º 175786/01, de 8/6/2006, conforme ponto específico deste Relatório;*

- *as datas dos cartões de embarque conferiram com o período da viagem em 76,98% das diárias nacionais, podendo variar entre 70,09% e 83,87%, e em 55,04% das diárias internacionais, podendo variar entre 33,41% e 76,68%. Nos processos que apresentaram divergências de datas de embarque, a impropriedade se explica por motivos alheios às atitudes do gestor;*
- *as diárias previstas para os cargos em comissão ou função de confiança foram concedidas somente aos que estavam efetivamente em exercício em 100% dos casos;*
- *as diárias não utilizadas foram devolvidas em 100 % dos casos. Ressalta-se, porém, que em alguns casos, as devoluções ocorreram intempestivamente, fato tratado na Nota de Auditoria 175786/01.*

#### **BIBLIOGRAFIA**

Almeida, A. C.(2003). *Como são feitas as pesquisas eleitorais*. Editora FGV/RJ, 2ª edição.

Bolfarine, H., Bussab, W. O. (2005). *Elementos de Amostragem*. Ed. Edgard Blücher. São Paulo.

Cochran, W. G.(1977). *Sampling Techniques*. John Wiley & Sons. 3ª edição.

Fauze, N. Mattar (2001) *Pesquisa de Marketing: edição compacta*. Ed. Atlas AS, 3ª edição.

Oliveira, P. H. F. C. (2004). *Amostragem Básica, Aplicação em Auditoria*. Ciência Moderna.

Silva, P.L. do N ,Pessoa, D.G.C. (1998). *Análise de dados amostrais complexos.13º Simpósio Nacional de Probabilidade e Estatística*.

Thompson, S. K. (1992). *Sampling*. John Wiley & Sons. 2ª edição.



**PORTARIA N.º 635, DE 28 DE MAIO DE 2007.**

Aprova o manual de Rotinas para Elaboração de Planos Amostrais nas diversas ações de controle

**O SECRETÁRIO FEDERAL DE CONTROLE INTERNO**, no uso da competência que lhe foi atribuída pelo inciso XV do artigo 21 do capítulo IV do anexo VII do Regimento Interno, aprovado pela Portaria CGU n.º 289, de 20 de dezembro de 2002, e considerando a necessidade de estabelecer rotinas padrão para elaboração de planos amostrais,

RESOLVE:

Art. 1º Aprovar, nos termos do anexo a esta Portaria, o Manual de Rotinas para Elaboração de Planos Amostrais nas diversas ações de controle, desde a fase de planejamento até a fase de inserção dos resultados nos relatórios.

Art. 2º Esta portaria entra em vigor na data e de sua publicação.

**VALDIR AGAPITO TEIXEIRA**