

Caio Graco Valle  
Cobério

# Usos e experiências da amostragem documental

O crescimento da massa documental produzida contemporaneamente torna cada vez mais necessário desenvolver o conhecimento requerido na formulação de critérios norteadores para a aplicação de métodos e técnicas de amostragem, com apoio da Estatística e das tecnologias de informação.

ARQUIVO PÚBLICO MINEIRO  
Amostragem de documentos  
eliminados  
Guias para pagamento de selo para  
certidão de documento  
1918-1934

> Este trabalho não comporta a pretensão de se constituir como matéria conclusiva no que toca à amostragem documental; pelo contrário, é um texto base que, no máximo, busca contribuir para fomentar discussões na área e permitir o exercício de algumas reflexões que são pertinentes a essa prática, com base em alguns exemplos coletados por meio de vivências e relatos cotidianos e pela bibliografia pertinente, de forma sistematizada. Pode ser considerado não como um produto definitivo, mas como o lançamento inicial sobre um vasto campo de investigações que – se sujeito a novas indagações e, ainda, aberto aos desafios que se colocam para a seleção da massiva produção documental contemporânea – pode oferecer um frutífero retorno para o meio arquivístico em geral e, conseqüentemente, para todos os beneficiários da guarda documental.

Os usos e as possibilidades da amostragem documental – tema central do 2º Congresso Brasileiro de Arquivos do Poder Judiciário, que debateu as práticas arquivísticas e suas relações com as inovações recentes trazidas pelo campo da Tecnologia da Informação – podem, certamente, multiplicar-se. É sabido que aperfeiçoamentos tecnológicos só podem ocorrer sobre a base de um conhecimento construído, elaborado dentro de princípios científicos e com rigor metodológico. É um fato que não se impõe por si só – por meio eletrônico, informático ou cibernético, exclusivamente –, mas pelo esforço coletivo e pela produção sistemática dos pesquisadores e profissionais da área, empenhados numa investigação conjunta que proporcione alcançar a interface dos diversos domínios envolvidos.

Assim, cabe ainda uma necessária discussão dos fundamentos da prática da amostragem documental, isto é, a compilação dos resultados já alcançados e a sistematização de um conhecimento mais elaborado e atualizado, conforme o exame das orientações

teóricas em evidência e a observação empírica dos universos documentais que suportarão a aplicação da técnica. Esses são alguns dos propósitos a que se referem os tópicos a seguir, que tratam de pontos pertinentes às variadas experiências e situações analisadas: a Estatística e suas aplicações, casos excepcionais, aspectos quantitativos e qualitativos, representatividade, validade da amostra, entre outros.

Este texto foi produzido com base nos estudos elaborados por um grupo de trabalho resultante de convênio entre o Arquivo Público do Estado (APE), órgão central do Sistema de Arquivos do Estado de São Paulo (Saesp), por meio de seu Departamento de Gestão (DGS/AESP). Esse órgão buscava, desde agosto de 2010, a elaboração de estudos no campo da amostragem documental, desejo esse concretizado na parceria firmada, em julho de 2011, com a Fundação Sistema Estadual de Análise de Dados (Seade) – órgão da Secretaria de Planejamento e Desenvolvimento Regional do Governo do Estado de São Paulo.<sup>1</sup> Foi este texto base de apresentação realizada no 2º Congresso Brasileiro dos Arquivos do Poder Judiciário.<sup>2</sup>

Naquela ocasião, foi feita uma proposta segundo a qual se previu a formação de um Grupo de Trabalho multidisciplinar, com profissionais de Estatística, Arquivologia e História, visando à definição de critérios técnicos de amostragem para guarda permanente de documentos públicos estaduais destinados à eliminação; em contrapartida, haveria um esforço conjugado para a formulação e efetiva implantação de programa de gestão, acesso e preservação de documentos públicos produzidos e/ou acumulados pela Fundação Seade. Após algumas reuniões, pesquisa, levantamento bibliográfico etc., o Grupo de Trabalho chegou à produção de um primeiro material de consulta, pelos técnicos daquela fundação. Nessa primeira proposta de procedimentos foram explicitados

os conceitos gerais da teoria da amostragem, tais como unidade de análise, população, amostra, amostra representativa, erro amostral, precisão, tamanho da amostra e as técnicas mais apropriadas para a amostragem documental, entre as quais: amostra aleatória simples (AAS), amostra sistemática, amostra por conglomerados, amostra multietapas e amostra estratificada.

As considerações de cunho estatístico revelaram que são vários os fatores que interferem – e portanto devem ser observados – no processo de escolha, definição e aplicação das técnicas de amostragem.<sup>3</sup> Para que sejam obtidos resultados mais satisfatórios e precisos é necessário conhecer, apriorística e profundamente, os objetivos a serem alcançados, as características e o conteúdo dos documentos e, até mesmo, sua forma de armazenamento. Cuidou-se também da elaboração de um fluxograma do “processo decisório para seleção de amostra representativa de documentos”, indicando possíveis correspondências entre situações hipotéticas e opções metodológicas de amostragem; e a descrição passo a passo dos procedimentos, da definição do tamanho da amostra até a seleção dos documentos, das técnicas de amostragem aleatória sistemática e por conglomerados. Por fim, foi sublinhada a importância de documentar-se o processo de amostragem, guardando algumas informações sobre como foi aplicada a técnica, tornando possível futuras inferências para o conjunto da população.

Um segundo grupo de trabalho, instituído em setembro de 2011,<sup>4</sup> pretendeu dar continuidade a essa primeira tarefa realizada, adentrando ainda mais a Teoria da Amostragem aplicada à Arquivística, a investigação dos critérios que orientam a determinação dos valores documentais e a população a ser amostrada, propriamente dita, por meio de um exame minucioso dos tipos documentais, naquelas séries que serão destinadas à eliminação. Posteriormente,

seriam desenvolvidos modelos e procedimentos amostrais passíveis de serem aplicados aos diferentes subconjuntos de documentos, além de testes e ajustes, para, por fim, homologá-los.

### Experiências em arquivos do Judiciário

Sobre o Poder Judiciário, em termos arquivísticos, recai uma das maiores demandas de gestão documental, senão a maior, dada a natureza contínua, intensa e crescente da atividade exercida pelos órgãos dele emanados, cujo teor é assistir e amparar a sociedade nos preceitos da Justiça. Constantemente requisitados pelos mais diversos motivos, são gerados por esses órgãos milhares de documentos que, diariamente, ingressam nas diferentes instâncias jurídicas, configurando milhões de autos de processos judiciais em fase corrente, arquivados, para prescrição, em guarda permanente ou destinados à eliminação. Nesse contexto, a aplicação de técnicas amostrais para a preservação de processos para a guarda permanente é uma necessidade imperiosa, como apoio à gestão documental desses acervos.

No Brasil, destaca-se a diretriz formulada pelo Conselho Nacional de Justiça (CNJ), como parte das ações desenvolvidas no âmbito do Programa Nacional de Gestão Documental e Memória do Poder Judiciário (Proname), lançado em dezembro de 2008. Por meio da Recomendação nº 37, de 15 de agosto de 2011, o CNJ recomenda a observância das normas de funcionamento do Proname e de seus instrumentos, tal como o Plano para Amostra Estatística Representativa,<sup>5</sup> constante como anexo XI do *Manual de Gestão Documental do Poder Judiciário*. Nesse plano, que buscou sanar problemas decorrentes do grande volume de arquivos do Judiciário, respeitou-se a heterogeneidade dos acervos provenientes da variada gama de segmentos da Justiça, para se obter um “plano amostral simples, didático, que

contemple as especificidades dos tribunais”.<sup>6</sup> O plano recomenda a aplicação de uma técnica probabilística de amostragem estratificada, ou seja, em estratos homogêneos que são produtos da divisão prévia de uma população heterogênea.

O ano de distribuição dos processos é um dos critérios de estratificação, enquanto o outro pode ser indicado pelos diversos tribunais especializados de acordo com os assuntos mais comuns, predominantes ou repetitivos. Por exemplo, na Justiça Federal, as matérias referentes a direito previdenciário e execução fiscal correspondem à quase totalidade dos processos. Entre outros, podem ser citadas ações sobre FGTS, poupança, empréstimo compulsório, acordos trabalhistas etc. Apesar de fornecer essas e outras orientações, tais como as noções de cálculo do tamanho da amostra, o plano sublinha que os trabalhos devem ser acompanhados de um estatístico responsável, pertencente ao Núcleo de Estatística e Gestão Estratégica.

Outra referência sobre técnica de amostragem nos arquivos do Poder Judiciário é o Plano para Seleção de Amostras Representativas de Ações Judiciais, adotado pelo Conselho da Justiça Federal, como Anexo III da Resolução nº 23, de 19 de setembro de 2008.<sup>7</sup> Nesse plano, o critério para a seleção de amostras baseia-se, principalmente, na capacidade de representação do universo populacional com o menor número de variáveis possível. Assim, sendo os processos da Justiça Federal constituídos, em sua maioria, por questões repetitivas, a amostra deve representar essa realidade, evitando-se valorizações de mérito ou ineditismo *a priori* de processos amostrados. Para isso, a metodologia sugerida é da amostra estratificada aleatória de até duas etapas, em que cada classe de processo condiz com um estrato e a seleção é feita por meio de um conjunto de números aleatórios, com tamanho definido em função das estimativas sobre o todo populacional. A relevância do critério qualitativo,

contudo, permanece para a seleção final, a cargo da Comissão Permanente de Avaliação.

Carlos Vieira<sup>8</sup> discute a utilização de ferramentas estatísticas para a coleta de amostra representativa de documentos, com base no método da estratificação e na aplicação da fórmula para definição do tamanho amostral. A título de ilustração, o Poder Judiciário do Estado do Rio de Janeiro caracteriza o uso da técnica de amostragem nos documentos produzidos posteriormente ao ano de 1940, visto que os anteriores a essa data são predeterminados como de guarda permanente. O mesmo procedimento é observado para aqueles documentos destinados à eliminação, desde que cumprido o prazo de guarda em arquivo intermediário. Daqueles cuja eliminação é feita imediatamente depois de findada a fase corrente, também não são retiradas amostras. Para a definição dos estratos, dentro do universo documental em que será feita a amostragem, estabeleceram-se as seguintes variáveis: 1) Para documentos judiciais: comarca, unidade organizacional, assunto/matéria, item documental/tipo de ação, período. 2) Para documentos administrativos: comarca, unidade organizacional, assunto, item documental, período.

No Tribunal Regional do Trabalho de São Paulo (TRT/SP) 2ª Região, está sendo desenvolvido um projeto para a aplicação de técnica de amostragem documental em um acervo de 400.000 caixas aproximadamente, correspondentes a 2.000.000 de “Ações Judiciais Trabalhistas”, que, após o cumprimento do prazo de guarda de cinco anos, são eliminadas. Portanto, a amostragem visa preservar para a guarda permanente, entre os lotes destinados à eliminação, uma fração representativa da totalidade.

Contando com um conjunto heterogêneo, de cinco processos, em média, por caixa, ainda que existam processos em mais de uma caixa – todos localizados

por meio de sistema eletrônico com código de barras e metadados diversos, como empregador e trabalhador –, atualmente é realizada uma pré-seleção subjetiva para salvaguardar fontes valiosas, identificadas em meio à riqueza processual. Sobre esse procedimento, almeja-se aplicar uma técnica complementar, de cunho estatístico, ou seja, uma técnica de amostragem probabilística e representativa. Tomando como base os estudos em desenvolvimento, a ideia é usar uma amostragem estratificada. Para definir o tamanho da amostra, segue-se a sugestão do CNJ de aplicar o procedimento estatístico denominado “cálculo do tamanho amostral”.

O instrumento de TI a ser utilizado baseia-se precisamente na possibilidade de utilizar os metadados já incluídos no sistema de cadastro dos autos judiciais arquivados como variáveis correspondentes aos estratos das amostras, por exemplo: data de autuação, comarca, classe processual, reclamantes e reclamados, objeto da ação e resultado.

### Teoria da amostragem

Para o chamado conhecimento científico da realidade a teoria da amostragem é um instrumento capaz de ser operacional onde outros ramos do conhecimento não se mostram suficientes. Ela estuda as relações isomórficas existentes entre uma dada população e as amostras extraídas dela, sendo útil para avaliar grandezas desconhecidas. Enquanto a amostragem é o processo de determinação da amostra a ser pesquisada, esta representa uma parcela extraída da população, sem que sejam modificadas as suas características principais. A amostragem, que faz o estudo da parte, é o contrário do censo, que faz um exame de todos os elementos do universo em questão.

A amostragem populacional consiste na seleção e observação de parte de uma população para projetar

estimativas sobre uma ou mais características da totalidade. São noções básicas da teoria da amostragem: a) população: é o grupo inteiro de objetos, unidades ou indivíduos dos quais se pretende obter alguma informação; b) unidade: é o elemento individual da população; c) amostra: parte, porção ou subconjunto da população usada para se obter informação do todo; d) variáveis: são as diferentes características da unidade, medidas dentro da amostra.<sup>9</sup>

Para se executar o processo de amostragem, existem algumas etapas necessárias, como definição dos objetivos e da população, coleta de dados segundo metodologia pré-elaborada, identificação da unidade de análise, escolha do método de amostragem, verificação preliminar e análise dos dados coletados.<sup>10</sup> Recomenda-se a prévia criação de um plano amostral.<sup>11</sup>

### Amostragem e Estatística: tipos de métodos e aplicações

As aplicações da amostragem são variadas. Frank Yates, da Subcomissão de Amostragem Estatística das Nações Unidas, em conferência pronunciada em 29 de agosto de 1951, discutia a tensão entre teoria e prática por meio de um exemplo no campo das pesquisas biológicas e agrícolas, relacionado à extração de “amostras de lotes de terra cultivável, a fim de medir características da plantação em crescimento, bem como para análises químicas, e, algumas vezes, para determinar o rendimento dos lotes”.<sup>12</sup> Em outra situação bastante distinta, ilustra-se a viabilidade das técnicas de amostragem para a execução dos procedimentos de auditoria pelos quais são verificadas, analisadas e testadas as transações ocorridas em uma empresa, além de documentos, livros, sistemas de controle interno e demonstrações contábeis.<sup>13</sup> Assim, pode-se dizer que a amostragem documental na Arquivística também se encontra dentro do grande

campo da Estatística e da teoria da amostragem, a despeito de suas peculiaridades, sobre as quais este trabalho procura lançar alguma luz. Antes, porém, devem-se identificar conhecidas técnicas de amostragem estatística que são comuns à amostragem documental, frequentemente aplicadas no contexto arquivístico. São elas:

1 – Amostra aleatória simples (AAS): não se trata de selecionar “ao acaso” (uma porcentagem qualquer, por exemplo), mas de aplicá-la quando todos os elementos da população possuam a mesma probabilidade de seleção. Dá-se por intermédio da utilização de técnicas matemáticas com números aleatórios (tábua de números aleatórios) e da existência de uma lista com todas as unidades da população, pela qual se determina um tamanho representativo da amostra, de modo a associar cada objeto dessa no conjunto total daquela.

2 – Amostra sistemática: tal como na amostra aleatória, todos os elementos da população possuem a mesma probabilidade de ser selecionados a partir da já mencionada lista de objetos. Ela é resultante da aplicação de critérios estabelecidos *a priori* baseados em características objetivas dos documentos e é selecionada por meio da escolha de objetos em intervalos de tamanhos iguais na lista (passo do intervalo, calculado pela razão entre o tamanho da população e o tamanho da amostra), partindo-se de um número aleatório dentro do primeiro intervalo.

3 – Amostra por conglomerados: inexistindo uma lista das unidades que compõem a população, ou apenas uma lista de conjuntos de unidades, a amostra deverá ser selecionada por conglomerados. Trata-se de uma técnica na qual a seleção da amostra pode ser feita em uma, duas ou mais etapas: no primeiro caso, seleciona-se uma amostra aleatória simples (ou sistemática) dos conglomerados e todos os objetos dentro dos conglomerados selecionados compõem a

amostra; no segundo caso, seleciona-se uma amostra aleatória simples (ou sistemática) de conglomerados e assim sucessivamente, conforme o número de etapas necessárias; por último, e dentro dos conglomerados selecionados, elabora-se a listagem de objetos e realiza-se uma amostra aleatória simples (ou sistemática) de objetos dentro dos conglomerados sorteados.

4 – Amostra estratificada: às vezes existe a necessidade de se obterem inferências para a população segundo grupos específicos, qualitativos. A amostragem estratificada implica que a seleção deve ser feita em cada estrato com a precisão indispensável. Para o conjunto da população, a precisão final é menor ou igual à precisão das amostras dos estratos. A forma de seleção da amostra dentro dos estratos pode ser conglomerada, aleatória simples ou sistemática.<sup>14</sup>

#### Amostragem para arquivos: métodos estatísticos e não estatísticos

Para além do enorme campo teórico da amostragem que existe em Estatística, a profusão de textos, sobretudo em língua inglesa – conhecidos como estudos sobre *Archival Sampling* – permite um debate enriquecedor sobre a “amostragem documental”, especificamente. O crescimento excessivo do volume de documentos produzidos pelas instituições contemporâneas resulta em “explosões” documentais nos arquivos atuais, tornando inviável a guarda da totalidade da documentação produzida. A eliminação de documentos públicos é uma condição necessária da gestão documental, assim como a destinação dos mesmos, mas deve ser feita dentro de critérios muito bem definidos.

As decisões para destruir documentos devem ser tomadas corretamente, baseadas na perfeita análise e no exame acurado dos mesmos

por funcionários executivos, cujo interesse se baseia no uso corrente dos documentos, e pelos arquivistas, cujo interesse reside no uso ou usos secundários dos mesmos.<sup>15</sup>

Do ponto de vista arquivístico, toda eliminação de documentos deve ser, incondicionalmente, precedida da atividade de avaliação. Essa, por sua vez, é das mais importantes atividades da gestão documental porque implica a sua destruição ou preservação, segundo os valores predeterminados. Estando avaliado, classificado e constando de uma tabela de temporalidade, o documento tem sua destinação predefinida e, assim, sua eliminação torna-se irrecusável como procedimento regular e vital.

Partindo-se do raciocínio de que toda a produção documental possui valor e que, paradoxalmente, a não eliminação de documentos dificulta a manifestação expressa desses valores, a amostragem torna-se, então, a técnica mais adequada para assegurar a preservação dos documentos e suas respectivas informações, que são considerados dignos de guarda. Em todo caso, a técnica da amostragem está estreitamente relacionada à avaliação, porque poderá ser aplicada tanto a documentos que já foram avaliados ou que já tiveram sua temporalidade e destinação definidas, quanto à massa documental ou séries documentais muito extensas que ainda estão por ser avaliadas.<sup>16</sup> No entanto, em um sentido mais geral, essa técnica também pode ser entendida como uma segunda seleção, para além daquela que, ao considerar os valores primários e secundários dos documentos, determinou, preliminarmente, quais seriam de guarda permanente e quais seriam descartados. Nesse sentido, a técnica de amostragem só poderia ser utilizada nos casos em que as séries documentais já foram avaliadas e cuja análise de valor permitira a eliminação.<sup>17</sup>

Referindo-se ao campo específico da amostragem documental, o *Dicionário Brasileiro de Terminologia*

*Arquivística* a define como “técnica de seleção em que, de um dado conjunto de documentos, elege-se um subconjunto representativo do todo”.<sup>18</sup> Já o *Dicionário de Biblioteconomia e Arquivologia* sugere a diferenciação de método entre diferentes áreas do conhecimento, caracterizando três definições possíveis e, inclusive, diferenciando a definição arquivística da definição estatística:

1. ARQ. escolha realizada no curso da triagem, segundo critérios variáveis, notadamente numéricos, alfabéticos, topográficos ou qualitativos, de uma certa proporção de documentos que, como os únicos conservados, serão considerados como representando o conjunto do qual se originam [...].
2. EST. obtenção de uma amostra da população.<sup>19</sup>

Esse verbete exprime mais correta e precisamente a relação entre os dois campos de conhecimento que se encontram na aplicação da técnica da amostragem de documentos, a Arquivística e a Estatística. De um lado, a Estatística possui um vasto campo de aplicações, dentro do qual a amostragem de documentos é uma delas, evidenciando a contribuição do método estatístico a essa área específica. De outro lado, a Arquivística necessita extrapolar a Estatística, para cobrir lacunas de sua técnica, operando, concorrentemente, um método não estatístico de amostragem documental. O que aqui se denomina como método não estatístico pode ser considerado método não probabilístico e subjetivo, ou técnica de amostragem intencional.<sup>20</sup>

Em contrapartida, o método estatístico é essencialmente quantitativo e pode ser considerado mais objetivo,<sup>21</sup> aplicável para séries extensas, homogêneas (documentos com informações similares e variabilidade pequena) e matematicamente quantificáveis. É um meio de apreender todas as características importantes das séries documentais ao

mesmo tempo que se reduz seu volume. Como técnica, tem se tornado o método mais circunspecto, de modo a alcançar um grau de precisão que conquistou o meio arquivístico, podendo atingir plenamente seus objetivos quando utilizado corretamente. Segundo Kepley

Existem dois tipos principais de amostras estatísticas: sistemática e aleatória. Depois que o tamanho da amostra estatisticamente válido foi estabelecido, na amostra sistemática, todo “n” documento é selecionado. Uma amostra aleatória difere naquilo que uma tabela de números aleatórios é utilizada para determinar as unidades a serem retidas. Amostragem aleatória tem sido descrita como estatisticamente “mais pura” porque ela reduz drasticamente a possibilidade da amostra estar sendo tendenciosa pelo próprio processo de seleção sistemática. Se, por exemplo, alguém escolhe todo décimo documento de uma série onde o décimo documento sempre se relaciona ao mesmo objeto, a amostra não seria representativa; pior, seria distorcida na direção daquele objeto, em detrimento dos outros. A maioria das autoridades concorda, entretanto, que se o processo de seleção é feito com cautela, a diferença matemática entre uma amostra sistemática e uma aleatória pode ser insignificante. Em qualquer caso, a amostragem sistemática é usualmente mais preferida porque ela é de longe mais fácil e menos custosa de ser empregada.<sup>22</sup>

Com o desenvolvimento do uso das técnicas de amostragem pelos arquivistas e observada uma maior complexidade da organização dos documentos e, ainda, dada a necessidade de se obterem inferências para o universo documental, segundo grupos especificados, o método estatístico alcançou níveis mais elevados de sofisticação. Assim acontece quando do uso da

estratificação ou da combinação entre técnicas diversas, já que as amostras dentro dos estratos podem ser ou aleatórias ou sistemáticas. Ainda de acordo com Kepley,

Estratificação é uma espécie de amostragem estatística, na qual certas partes do universo a ser amostrado são “pesadas” diferentemente de outras. Isto se diferencia de uma amostra sistemática ou aleatória, propriamente dita, nas quais, teoricamente, todas as partes do universo têm uma chance igual de ser selecionadas na amostra. Por exemplo, em uma série documental extensa que atravessa um determinado número de anos dentro dos quais há diferenças substantivas, na mesma série, é desejável que se consiga uma percentagem mais alta das porções mais valiosas dessa série, através da estratificação.<sup>23</sup>

O método não estatístico é essencialmente qualitativo e pode ser considerado mais subjetivo, pois não se almeja a retirada de amostras estatisticamente representativas, mas de documentos significativos, únicos, que são definidos por meio de critérios pessoais e não matemáticos. Como foram os arquivistas, geração após geração, diretamente responsáveis por esse tipo de amostragem – cuja relevância era atribuída de maneira subjetiva –, declarou-se que alguns documentos não tinham valor suficiente que justificasse sua guarda permanente. Essa foi a base para a amostragem subjetiva se ajustar ao nível imediatamente abaixo ao da série documental, configurando uma situação na qual o valor existia apenas em alguns documentos da respectiva série, os quais deveriam ser retidos, já que eram dotados de valor secundário, ou seja, histórico.

Criação de precedentes, interesse das mídias, convulsões político-econômicas eram alguns dos critérios utilizados para identificar os documentos que serviriam

de amostras.<sup>24</sup> O potencial de pesquisa hipotético levantado pelos investigadores para esses documentos, no campo das instituições públicas, relacionava-se, contudo, à possibilidade de utilizar as informações nele contidas para conhecer a agência produtora em si e como ela trabalhava, mas não os grupos ou indivíduos com os quais ela lidava. Esses são os valores que T. R. Schillenberg denominou de probatórios, diferindo, portanto, dos valores informativos, que dizem respeito às pessoas, entidades, coisas, problemas, condições etc., com que o órgão governamental tenha lidado. Os valores probatórios se referem à organização e ao funcionamento desse órgão, cuja importância deriva do caráter da matéria provada, como a origem ou fim de um programa substantivo, norma ou diretriz, e não da prova em si. Para identificar o valor probatório, o arquivista deve aplicar um tipo de “teste”, pelo qual conhecerá a fundo a documentação produzida e o órgão produtor.<sup>25</sup>

### Aplicações

Existem diversas maneiras de aplicar o método não estatístico de amostragem. Resumidamente, o rol a seguir apresenta algumas das mais conhecidas técnicas da amostragem subjetiva de documentos:

1 – Amostragem excepcional: técnica de seleção de subconjuntos documentais, baseada em qualidades de importância incomuns, inusuais ou excêntricas. Exemplos de critérios usados para amostragem excepcional incluem objetos controversos, assunto notório, indivíduos famosos e ilustres. Embora não seja estatisticamente válida, a amostragem excepcional, frequentemente, pode identificar materiais por meio de solicitações pessoais.<sup>26</sup>

2 – Amostragem de exemplares: técnica de seleção de um número limitado de itens típicos para representar um grupo grande. Por exemplo, quando se tomam

todos os registros de uma seccional regional para inferir como outras seccionais regionais operam.<sup>27</sup>

3 – Amostragem ilustrativa: método de avaliação que seleciona de uma série grande uma porção de documentos para preservação, baseado no julgamento do selecionador, que deve ser formado por critérios específicos. A amostragem ilustrativa não é nem sistemática nem aleatória.<sup>28</sup>

4 – Amostragem sistemática: técnica de seleção de itens de um grupo baseada em algumas características formais, sem considerar o conteúdo desses itens. Exemplos de amostragem sistemática incluem puxar todos os documentos de um dado tamanho, ou todos os documentos nos quais o sobrenome começa com uma dada letra. Embora relativamente fácil de programar, a técnica não é estatisticamente válida.<sup>29</sup>

5 – Método do processo volumoso: técnica para a avaliação individual de documentos processuais segundo a qual os mais volumosos são, presumivelmente, os mais valiosos da série à qual pertencem. O uso do tamanho para eleger determinados documentos a serem preservados é um exemplo de amostragem sistemática. Tal método pressupõe que o maior número de documentos agregados na complexidade de ações que envolvem o processo sugere que ele foi mais importante quando estava em sua fase corrente.<sup>30</sup>

### Outras experiências e situações

Como conhecimento a ser construído, não existe nada que comprove o uso em definitivo e a supremacia de uma técnica sobre a outra. Tudo depende da complexidade do problema, do conteúdo, da forma, da organização espacial dos documentos, entre outros aspectos. Por isso, a necessidade fundamental de

bem observar as particularidades de cada acervo. Certamente, soluções singulares devem surgir para arquivos diferentes, o que não implica a exclusão mútua dos métodos empregados, ou ainda o embasamento das práticas de uma instituição a partir do que é recomendado por outra. Pelo contrário, a troca de experiências se mostra uma importante fonte de aquisição de conhecimento empírico que pode ser aplicado integralmente ou, na medida do possível, adaptado para a configuração de uma técnica adequada a um acervo determinado. Os relatos dos próximos subtópicos assinalam algumas dessas perspectivas.

#### Arquivo Nacional do Reino Unido

Esse é um caso de instituição que formulou política para a seleção e amostra de documentos conhecidos pela terminologia de *Case File*.<sup>31</sup> Entre alguns princípios predefinidos destaca-se o estado atual do conhecimento desenvolvido para áreas específicas de interesse da pesquisa, que pode estar sendo aplicado na elaboração e no uso de critérios de seleção.<sup>32</sup> No caso de haver documentos digitais natos, para os quais se busca um espaço de armazenamento mais reduzido em relação aos documentos de papel, foi proposta a guarda da totalidade dos conjuntos. Em relação àqueles documentos, foi dada preferência para sua guarda no formato original de produção. Em alguns casos, pode ser preferível armazenar apenas cópias, devido às facilidades de guarda e acesso.<sup>33</sup> Quanto às técnicas de amostragem, são feitas escolhas sobre as opções metodológicas existentes: o não aleatório e o aleatório.<sup>34</sup>

Entre os métodos não aleatórios, três opções são elencadas:

1 – Amostra por conveniência: é usada para séries homogêneas e consiste na retirada de pequenos segmentos do conjunto até que se obtenha a amostra

desejada. Considera-se de fácil utilização, porém não é estatisticamente válida.<sup>35</sup>

2 – Amostra exemplar: estabelece subdivisões em uma série em razão do tipo de informações contidas em cada documento. Por exemplo, quando em processos de “Concessões de Benefícios” se separam todos os processos que tratam dos benefícios concedidos por conta de uma dada doença. Apesar do critério fixo e predefinido, o método também não é estatisticamente válido.

3 – Amostra excepcional: seleciona documentos definidos como interessantes/excepcionais em uma série, preservando as informações definidas como significativas *per se*, dando menor importância aos elementos formais do documento e a sua dinâmica administrativa. Preserva parcelas significativas do acervo, mas omite a produção documental rotineira.

Entre os métodos aleatórios de amostragem, são elencadas seis opções:

1 – Amostra aleatória simples (AAS); e

2 – Amostra sistemática (AS): todos os objetos da população têm a mesma chance de ser selecionados. Define-se o tamanho da amostra pretendida, sendo necessária uma listagem ou organização mínima dos documentos. Enquanto na AAS se realiza um sorteio para obter cada um dos documentos, na AS há um intervalo padrão para a seleção desses. Ambos são estatisticamente válidos.

3 – Amostra aleatória estratificada: é usada em séries que comportem divisões internas significativas, cujo grau de organização possibilita a identificação desses “estratos”. Por meio da ASS ou da AS, garante-se a representatividade de cada um dos segmentos identificados.

4 – Amostra por agrupamentos: selecionam-se documentos provenientes de diferentes órgãos, mas que

guardam semelhanças marcantes entre si. Por exemplo, amostra do conteúdo integral de uma série documental de um entre dez hospitais subordinados a uma mesma estrutura administrativa. Definida por sorteio, com os cuidados necessários, é estatisticamente válida.

5 – Amostra multietapas: combinação da amostra aleatória estratificada e da amostra por agrupamentos, pela qual a técnica dos agrupamentos é usada para reduzir conjuntos muito grandes a segmentos nos quais são identificados os estratos. Objetiva representar as especificidades dentro de cada conjunto, fazendo uso da própria estrutura interna de cada série, demandando conhecimento apurado do acervo.

6 – Amostra por quotas: método subjetivo que, em parte, escapa da definição dos métodos aleatórios, pois antes da escolha aleatória são separadas amostras intencionais da série, por alguma característica específica das informações ou dos referentes, para que então possa se proceder, em cada um deles, à amostragem simples por sorteio.

#### Arquivo Histórico da Polícia Nacional da Guatemala

Outro exemplo do uso de técnicas de amostragem para o trato com acervos é o caso do Arquivo Histórico da Polícia Nacional da Guatemala. Não se trata da seleção de documentos para a preservação permanente, mas da tentativa de conhecer o conteúdo de um gigantesco acervo abandonado,<sup>36</sup> em curto período de tempo, por meio da análise de segmentos separados por técnicas estatisticamente válidas. Frente a um grande volume de documentos divididos em diversos cômodos e prédios, a decisão da equipe responsável pela análise foi a de empregar uma técnica de amostragem aleatória representativa multietapas, levada a cabo por meio de “ondas” de amostras repetitivas.

Para a definição inicial do projeto de amostragem foram estabelecidos três objetivos: “Compreender o

alcance e a heterogeneidade deste inexplorado Arquivo; Coletar amplos dados sobre os padrões macro das operações policiais, como o comando, estruturas e fluxos de comunicação; Estimar a proporção de documentos que registraram certos atos e políticas de interesse para o projeto, incluindo desaparecimentos, detenções e mortes”.<sup>37</sup> Além disso, alguns desafios tiveram de ser considerados, como a exiguidade de recursos humanos, tempo hábil, a movimentação do acervo e a realização de outros trabalhos (como conservação, restauro e organização dos documentos), o constante “crescimento” do acervo (em razão da descoberta de novos arquivos correlatos), a ausência de listagens da população, inexistindo a possibilidade de amostragem direta dos documentos. Tornou-se então imperiosa a elaboração de uma lista mestra, desde os prédios que continham documentos até as unidades de armazenamento, o que possibilitou a seleção aleatória de diversas dessas unidades em uma primeira “onda” de seleção. O tamanho da amostra a ser retirada de cada ambiente, dentro desses prédios, foi definida com base em um primeiro diagnóstico do quantitativo dos acervos. Após serem identificados, sumariamente, os documentos da primeira “onda”, tiveram lugar as “ondas” de amostras seguintes. Ao final da décima onda, foi programada uma pausa para avaliação do processo.

#### Arquivo do Estado de Nova York

No caso dos documentos do Arquivo do Estado de Nova York, a técnica de amostragem tem se revelado subjetiva. De acordo com Geoffrey A. Huth, diretor do Serviço de Documentação Governamental do Arquivo Público do Estado de Nova York,<sup>38</sup> a amostragem é realizada em conjunto com as unidades produtoras, de modo que possam ser identificados aqueles documentos que constituem os processos mais significativos. Na maior parte das vezes,

esse procedimento é aplicado para grandes séries documentais, como as mantidas na Procuradoria Geral do Estado. Nessa instituição, por exemplo, existem séries documentais descomunais, compostas de mais de 14 mil caixas, contendo as investigações da Procuradoria, muitas das quais são de rotina, enquanto em outras se revela algum tipo de especial significância.

O objetivo é identificar aqueles processos que são historicamente relevantes e salvá-los para a guarda permanente. Alguns desses processos também são bastante grandes, chegando apenas um único deles a preencher 400 caixas. Esses são também os mais prováveis de servir a algum uso no futuro, como amostras. Nas amostras escolhidas, sempre que possível, são destruídas as duplicatas, *clippings* de jornais, assim como outros papéis considerados sem valor documental. Em todos esses casos, procura-se definir quais são os critérios que identificam aqueles processos destinados à guarda permanente, dentro daquelas séries documentais, ou seja, quais são os critérios qualitativos de amostragem. Entre os critérios que têm sido definidos pelo Arquivo do Estado de Nova York, destacam-se três situações, referentes a: “Documentos de Reexame da Qualidade Ambiental”; “Processos Arquivados do Departamento Jurídico/ Administração de Documentos e Arquivos do Estado”; “Documentos sobre Edificações e Construções”.

Na documentação sobre a qualidade ambiental, são considerados amostras aqueles documentos que obedecem, no mínimo, a um dos seguintes critérios: sejam altamente controversos, tenham impacto ambiental negativo abrangente, tenham vasto significado histórico (como os que se referem ao “11 de Setembro”) e/ou resultem de decisão histórica, regulamento ou comando superior. Nos processos jurídicos arquivados, são considerados amostras aqueles documentos que atendam aos seguintes

critérios: registros pertencentes aos processos que denotem decisão histórica, regulamento ou comando superior e que, posteriormente, forneçam evidência de precedente; registros relacionados a organizações ou figuras notórias que são objetos prováveis de pesquisas em andamento; registros pertencentes aos processos que envolvam indivíduos proeminentes que serão prováveis objetos de pesquisas futuras; registros que definam uma considerável mudança ou movimento, ou sejam ilustrativos de iniciativas ou políticas importantes do Departamento Jurídico. Além desses critérios específicos, podem ser amostrados quaisquer processos que tenham sido basicamente enviados para a Corte de Apelações (incluir-se-iam alguns processos envolvendo constitucionalidade de estatutos, entre outros ordenamentos legais); que tenham sido encerrados até 1950 (quando da emissão das primeiras Autorizações para a Disposição de Documentos – RDAs, pela Divisão de Orçamento); ou tenham sido alegados, por funcionário público ou empregado do Estado de Nova York, como dotados de objeto fraudulento.

Enfim, no terceiro caso mencionado estão entre os critérios que caracterizam relevância histórica em projetos de edificações, estruturas e construções: a fama do arquiteto ou da construtora; a autoria do projeto ser de um arquiteto local (no caso, do Estado de Nova York); o caráter original, inovador ou de vanguarda do projeto e das características da construção; alguma controvérsia considerável, associada à construção e/ou a pós-construção (por exemplo, a explosão do transformador do edifício da Secretaria de Estado de Binghamton, liberando Bifenilpoliclorado – PCBs); o grande alcance e impacto de um projeto de construção (expansão de um *campus*, alteração de uma paisagem, eliminação de uma vizinhança inteira); as características distintivas de um tipo, período, estilo regional ou método de construção; associação com eventos, pessoas ou usos significativos; seus arrolamentos em

registros nacional/estadual, individualmente ou como parte de uma região histórica.

Uma vez que tenham sido identificados edifícios ou estruturas historicamente significativos, os tipos documentais precisam ser valorizados. Os pontos a seguir são baseados nas observações feitas pelo Arquivo e Administração Nacional de Documentos (NARA), em uma avaliação de documentos arquitetônicos conduzida em Michigan<sup>39</sup> e na revisão feita pelas “agências de redesenho” do Estado de Nova York. Tais pontos indicam como tipos documentais relevantes para a preservação permanente: desenhos preliminares de *design*; esboços de apresentação (para vender a ideia ao cliente); desenhos de versão final e desenhos de como construir; desenhos e demais documentos relacionados à negociação de contratos ou licitações que mostrem a estrutura física do edifício e que não sejam suplantados pelos respectivos desenhos de versão final, ou os de como construir (excluindo-se propostas não contempladas); concorrência, contrato e projeto manual; desenhos detalhados mostrando estrutura ou aparência; desenhos artísticos mostrando traços decorativos tais como escultura, terracota, serralheria, mobiliário etc.; desenhos de fabricação preparados pelos empreiteiros, usualmente para elementos mecânicos, marcenaria, ou algum outro produto construído como elemento da estrutura; especificações relacionadas à estrutura ou aparência; correspondência com empreiteiros; relatórios de engenharia e consultoria; brochuras suplementares e amostras dos produtos usados na construção; aprovação das fases dos projetos; relatórios de teste do terreno; estimativas orçamentárias; minutas de reuniões; fotografias, negativos e outras imagens do exterior ou interior da edificação.

Quanto aos tipos documentais que, na perspectiva dos critérios adotados, não são destinados à guarda permanente, estão: esboços anteriores à versão final

que possam ser suplantados pelos desenhos de versão final ou os de como construir; desenhos de fabricação detalhados ou especificações para sistemas de eletricidade, encanamento, aquecimento e de ar-condicionado; documentos e desenhos de licitações não contempladas; documentos relativos ao pagamento de empreiteiras; ordens de alterações nas construções; lista de inspeção de falhas; manual de instruções e garantias de sistemas.

#### Validade da amostragem: questionamentos e proposições

Como já foi mencionado, não é possível admitir que seja eleita uma única proposta de amostragem documental, pois o uso dessa técnica dependerá de um conjunto de fatores, objetivos e subjetivos, próprios dos documentos, dos arquivos e dos arquivistas. Alguns aspectos relevantes sobre amostragem documental, em meio ao debate existente, podem dispensar a validação universal da técnica. Em primeiro lugar, a amostragem se baseia na premissa de que, se uma porção da totalidade de documentos é escolhida corretamente, qualquer indivíduo pode generalizar seguramente sobre as características da totalidade, utilizando-se apenas de um fragmento dela; ou seja, a amostra deve possuir representatividade em relação ao universo de onde foi extraída. Em segundo lugar, para que sejam retiradas amostras de uma série é preciso que ela apresente um valor mínimo, histórico, informativo ou probatório, de natureza substancial, que justifique sua preservação. Em terceiro lugar, é necessário que o arquivista responsável pela aplicação da técnica de amostragem conheça bem os documentos que estão sob sua responsabilidade, assim como os órgãos que os produziram, para ter certeza da adequação da técnica que será aplicada, dos recursos humanos e materiais que serão envolvidos na ação e da própria utilidade da amostra produzida ao final. Em quarto lugar, a

aplicação da técnica de amostragem deve ser evitada sempre que for necessária uma avaliação individual de cada documento ou em universos documentais muito restritos, com pequeno número de documentos, pois a amostragem é uma solução para tratar do grande acúmulo de documentos resultantes das atividades das burocracias estatais modernas.

Embora esses aspectos sejam consensuais na comunidade arquivística, existem outros que, por seu turno, expressam pontos divergentes, ambiguidades, mas que fazem parte do desenvolvimento das investigações no campo. Um desses aspectos diz respeito à própria indefinição conceitual da noção de “amostragem”, voltada para documentos arquivísticos. Isso se deve à relação intrínseca entre a técnica amostral, a Estatística e a representatividade. Muito dos pontos de vista divergentes decorrem dessa condição essencial da amostragem, que revelam uma tensão entre a teoria e a prática, pois os fatores particulares são preponderantes na determinação de quando, como e porque por que fazer a amostragem.

Alguns autores consideram que amostragem é somente aquele procedimento que possui uma representatividade estatisticamente válida, isto é, que pode ser verificado matematicamente dentro de uma margem de erro aceitável. Nesse caso, ele corresponde ao que os estatísticos denominam de “amostragem probabilística”. Dessa forma, a escolha de itens individuais de uma série, para obter-se um reflexo qualitativo de alguma importante característica predeterminada da totalidade de documentos, não seria amostra, mas seleção. É nesse sentido que T. Cook considera útil especificar e diferenciar amostra de seleção, e, essas duas, de exemplo.<sup>40</sup> Nesses casos, a dúvida é de designação, pois os dois últimos termos – seleção e exemplo – correspondem, respectivamente, ao que foi anteriormente descrito como “amostragem excepcional” e “amostragem exemplar”. Também

não há unanimidade sobre o fato de os estatísticos considerarem apenas um único tipo de amostragem correto, pois a questão vai mais além.

O processo de amostragem probabilística permite a inferência de dados da amostra para a população, segundo a teoria de probabilidade e inferência estatística com mensuração da precisão dos indicadores. As amostras não probabilísticas não apresentam essa possibilidade. Desse modo, caso seja importante a realização de inferências, recomenda-se a utilização de amostragens probabilísticas para a seleção de documentos que devem ser separados para guarda permanente antes do descarte. Caso contrário, o uso de amostra não probabilística pode ser mais prática e de rápida implementação.<sup>41</sup>

Com base nessa orientação é possível admitir, portanto, a existência de dois tipos principais de métodos de amostragem documental, probabilísticos e não probabilísticos, sob os quais se poderia fazer uma correspondência, lexical e metodológica, de acordo com o proposto no Quadro 1.

De ambas as vertentes, entretanto, surgem questões que merecem ser analisadas. Basicamente, o uso do método estatístico (em suas diferentes variantes) deve ser avaliado para que seja percebido se não está provocando distorção da amostra, de acordo com a natureza da série. O método não estatístico, por sua vez, afasta a representatividade do conjunto da série documental.

Outra observação importante, anteriormente mencionada, diz respeito à relação entre avaliação e amostragem. A avaliação é uma atividade crucial para a gestão arquivística de documentos e está diretamente relacionada à técnica da amostragem. Esse fato leva alguns autores a considerar a amostragem como

apenas uma parte da avaliação. No entanto, também é preciso reconhecer visões distintas sobre essa relação, justificando diferentes acepções. Existem três situações bem delineadas: o universo documental não é conhecido; a tabela de temporalidade está obsoleta ou superada; o universo documental é destinado à eliminação.

No primeiro caso, a amostra serve para o conhecimento da série, implicando a discussão sobre a amostragem documental como um todo. Após o conhecimento da série, reutiliza-se a técnica para extrair a nova amostra que será preservada. Portanto, a amostragem, primeiramente, é um método de avaliação e, em segundo lugar, uma forma de preservar documentos.<sup>42</sup> No segundo caso, a tabela de temporalidade já foi utilizada, como instrumento principal da avaliação de documentos, mas, por algum motivo, não racionaliza a gestão arquivística, sofrendo algum tipo de obsolescência. Pode, por exemplo, justificar a guarda permanente de séries documentais muito extensas ou que contenham documentos de pouco valor. Na realidade, a avaliação já foi feita por um instrumento convencional. Contudo, em função de um conjunto de fatores novos, resolve-se pela aplicação de uma técnica de amostragem sobre a série já avaliada a fim de preservar nela um conjunto menor de documentos ou os mais significativos.<sup>43</sup> No terceiro caso, a avaliação documental também já foi realizada, mas, ao contrário do caso anterior, a destinação prevista para a série é a eliminação. Nessa condição, a técnica de amostragem deve ser aplicada com o intuito de preservar, pressupondo algum valor secundário, intrínseco ou extrínseco, que reste no universo documental, por meio da definição de critérios que especifiquem quais são as técnicas mais adequadas.<sup>44</sup>

### Considerações finais

As opções pela amostragem documental existem dentro de um repertório variado. Ao compreender e legitimar

os avanços trazidos pelos métodos estatísticos, não se abandonam por completo algumas das antigas práticas já consolidadas no meio arquivístico, mas se as reformulam para atender a necessidades agudas dos arquivos no tempo presente. A interação metodológica dos procedimentos que definem a técnica a ser aplicada para obter a amostragem, segundo as singularidades dos documentos, apresenta-se como uma solução viável frente ao volume e à diversidade da produção documental contemporânea.

A tensão entre a teoria que foi construída para a avaliação documental e a prática cotidiana junto aos acervos demonstra uma contradição em termos, cuja síntese final tem se realizado nos métodos de amostragem, como um de seus aspectos. Comentando a experiência desenvolvida por outros arquivistas, Jane Turner mostra que a articulação de princípios arquivísticos – como os de imparcialidade e proveniência – tem sido fundamental para resolver os problemas enfrentados nos arquivos, encontrando diferentes soluções metodológicas.

O princípio da imparcialidade reconhece que a autenticidade, credibilidade e a natureza imparcial dos arquivos, a qual é derivada do processo de criação e acumulação natural deles, deve ser preservada. A utilidade futura dos documentos como evidência, seja administrativa ou histórica, a qual é essencial para todos os usuários, é diretamente dependente da preservação da imparcialidade dos documentos. Isto é realizado através da aplicação do princípio do “respeito ao fundo”, o qual respeita e preserva a integridade contextual de cada fundo, se ela é física, para documentos de papel, ou contextual, para documentos eletrônicos. Se os arquivistas falham em executar esta tarefa fundamental, os documentos não serão prestáveis como evidências e, em lugar disso, serão reduzidos a discretos artefatos históricos.<sup>45</sup>



Seguindo esse raciocínio, a autora demonstra, pela análise de estudos de caso, que o método estatístico, aplicado por meio de uma amostra aleatória sobre uma série documental organizada cronologicamente, ao reter o padrão original e limitando futuras distorções, atendeu ao princípio da imparcialidade. De outro lado, o método não estatístico, aplicado por meio de uma amostra sistemática de processos volumosos, subsequente ao entendimento aprofundado da série documental, ao postular uma metodologia claramente definida para a particularidade daquela série, atendeu ao princípio da proveniência.

A relação entre as técnicas de amostragem e as inovações trazidas pelo campo da Tecnologia da Informação acompanha e se embasa nessa discussão precedente. Os diversos tipos de amostras, o conhecimento das séries, a atribuição de valor aos documentos são terrenos férteis nos quais se introduzem os instrumentos de TI. Entre alguns instrumentos conhecidos, podem ser úteis o uso de planilhas eletrônicas, bancos de dados ou outros aplicativos, como a “Calculadora de Tamanho Amostral”,<sup>46</sup> pelos quais se determina o nível de precisão em uma dada amostra. Esse aplicativo trabalha com dois conceitos fundamentais: margem de erro e nível de confiança. Quanto maior a margem de erro, maior o nível de confiança; quanto menor a margem de erro, menor o nível de confiança.

A determinação da correlação entre esses dois fatores, por sua vez, influenciará o tamanho da amostra: quanto menor a margem de erro, maior deverá ser o tamanho da amostra para um dado nível de confiança; quanto menor o nível de confiança, maior deverá ser o tamanho da amostra para uma dada margem de erro. A calculadora é útil porque, dependendo das variáveis dadas, ela indica o tamanho da amostra ou a margem de erro. É possível encontrar, em publicações sobre amostragem, as fórmulas para o cálculo do tamanho amostral,<sup>47</sup> que varia conforme o tipo de técnica escolhida: amostra

aleatória simples, conglomerada etc. Outro aplicativo é a ferramenta *Ex-Sample*, contida no *software Methodologist's Toolchest*, desenvolvido pela The Idea Works, que também calcula o tamanho da amostra em situações de maior complexidade, operacionalizando uma grande variedade de recursos. O programa inclui uma série de outros aplicativos relacionados aos aspectos metodológicos de pesquisa, que deve ser avaliada.<sup>48</sup>

Concluindo, entende-se que, no estágio contemporâneo da produção de documentos, a amostragem documental permite múltiplas opções para aplicação da técnica, com ou sem o auxílio de instrumentos da Tecnologia da Informação. Releva-se que “a teoria arquivística deve ser flexível o suficiente para responder a essa realidade [o desordenado mundo da atividade humana], precisamente, identificando a essência da atividade documentada, de modo que a teoria possa ser largamente aplicada no desordenado mundo real dos documentos”,<sup>49</sup> ordenando-o e atribuindo-lhe o sentido original.

---

**RESUMO** | A amostragem é um método com numerosas possibilidades de uso, nos mais diferentes campos, da agricultura à análise clínica, fundamentado em procedimentos estatísticos ou não estatísticos. Aplicada ao campo arquivístico, pode atender a diferentes necessidades da gestão documental, entre as quais a determinação de amostras para a guarda permanente de documentos previamente destinados à eliminação. Este artigo pretende compreender, de forma ampliada, o prospecto de abordagens sobre técnicas de amostragem e sua aplicação em arquivos, para levantar questões e reflexões pertinentes ao conhecimento requerido na formulação de critérios norteadores para tal aplicação de cunho prático.

---

**ABSTRACT** | Sampling offers a number of possible uses in the most diverse fields, from agriculture to clinical analysis, based on statistical or non-statistical procedures. For the archivist, it can meet diverse document management needs, including the determination of samples for the permanent storage of documents previously intended to be discarded. This article seeks to understand in a broad fashion various approaches to sampling technique and their use in archives in order to raise pertinent questions and engender thought on the knowledge required in the formulation of guiding criteria for their practical application.

---

#### Notas |

1. Publicado em SÃO PAULO. *Diário Oficial do Estado*. Poder Executivo. Seção I. Casa Civil. Extrato de Convênio. São Paulo: Imprensa Oficial, v. 121, n. 141, 28 jul. 2011.
2. Cf. 2º CONGRESSO BRASILEIRO DOS ARQUIVOS DO PODER JUDICIÁRIO. *Arquivos do Congresso*. Brasília, DF: Tribunal Superior

Eleitoral, 28 a 30 de março de 2012. Disponível em: <<http://www.tse.jus.br/hotSites/2CongressoNacional/arquivos-congresso.html>>. Acesso em: 9 set. 2013.

3. Esse aspecto se mostra deveras importante na elaboração de quaisquer planos de amostragem, pois o processo de amostragem documental também é um procedimento de avaliação e seleção de documentos.

4. Publicado em SÃO PAULO. *Diário Oficial do Estado*. Poder Executivo. Seção I. Casa Civil. Resoluções de 26-9-2011. São Paulo: Imprensa Oficial, v. 121, n. 183, 27 set. 2011.

5. Cf. alínea g, item IV. In: CONSELHO NACIONAL DE JUSTIÇA. Recomendação nº 37, de 15 de agosto de 2011. Publicada no DJ-e nº 152/2011, em 17/08/2011. Disponível em: <<http://www.cnj.jus.br/atos-administrativos/atos-da-presidencia/322-recomendacoes-do-conselho/15447-recomendacao-n-37-de-15-de-agosto-de-2011>>. Acesso em: 1º mar. 2012.

6. Cf. 9.11 Anexo XI – Plano para Amostra Estatística Representativa. In: CONSELHO NACIONAL DE JUSTIÇA. *Manual de Gestão Documental do Poder Judiciário*. Brasília: PRONAME/CNJ, out. 2011. Disponível em: <[http://www.cnj.jus.br/images/programas/gestao-documental/manual\\_gestao\\_documental\\_poder%20judiciario.pdf](http://www.cnj.jus.br/images/programas/gestao-documental/manual_gestao_documental_poder%20judiciario.pdf)>. Acesso em: 1º mar. 2012.

7. CONSELHO DA JUSTIÇA FEDERAL. Resolução nº 23, de 19 de setembro de 2008. Plano para Seleção de Amostras Representativas de Ações Judiciais. Anexo III. Publicada em 24/09/2008. Disponível em: <<http://www2.cjf.jus.br/jspui/bitstream/handle/1234/5373/Res%20023%20de%202008%20-%20ANEXO%20III.pdf?sequence=4>>. Acesso em: 1º mar. 2012.

8. VIEIRA, Carlos Alberto de Souza. A coleta de amostras representativas de um acervo documental: O caso do Poder Judiciário do Estado do Rio de Janeiro. *Arquivística.net*, Brasília, v. 1, n. 2, 6 mar. 2006. Disponível em: <<http://www.arquivistica.net/ojs/viewarticle.php?id=35>>. Acesso em: 2 mar. 2012.

9. MEASURE Evaluation. *Teoria da amostragem*. US: USAID (Materials for Trainers). P. 1-3. Disponível em: <<http://www.cpc.unc.edu/measure/training/materials/data-quality-portuguese/Amostragem.pdf/view?searchterm=Amostragem>>. Acesso em: 11 jul. 2011.

10. GONÇALVES, Cristina Faria Fidelis. *Estatística*. Londrina: Edel, 2002. p. 5-6.

11. V. BOLFARINE, Heleno; BUSSAB, Wilton O. *Elementos de amostragem*. 2. reimp. São Paulo: Blucher, 2011.

12. YATES, Frank. Métodos de amostragem em censos e levantamentos. *Revista Brasileira de Estatística*. Sociedade Brasileira de Estatística, ano XII, n. 47, p. 279, jul.-set. 1951.

13. CUNHA, Paulo Roberto da; BEUREN, Ilse Maria. Técnicas de amostragem utilizadas nas empresas de auditoria independente estabelecidas em Santa Catarina. *Revista Contabilidade & Finanças*, São Paulo, Departamento de Contabilidade e Atuária, Fea/USP, v. 1, n. 1, p. 67-86, out. 1989.

14. FUNDAÇÃO SEADE. Proposta de Procedimento de Seleção de Documentos para Guarda Permanente: Cooperação Técnica entre a Fundação SEADE e o Arquivo Público Estadual. São Paulo, SP, 2011. (Material impresso não publicado.). p. 4-5.

15. SCHELLENBERG, Theodore R. *Arquivos modernos*: princípios e técnicas. 2. ed. Rio de Janeiro: FGV, 2002. p. 143-144.

16. No primeiro caso, a amostragem arquivística é um meio de preservação de documentos a serem eliminados e, no segundo caso, é um

método de avaliação. V. BRADSHER, James Gregory; AMBACHER, Bruce I. *Archival Sampling: A Method of Appraisal and a Means of Retention*. New York: Mid-Atlantic Regional Archives Conference/MARAC, 1992 (Technical Leaflet Series, n. 8).

17. ESCUELA IBEROAMERICANA DE ARCHIVOS. *Metodología para la identificación y valoración de fondos documentales*. Madrid: Ministerio de Educación, Cultura y Deporte, 2000. (Experiencia y Materiales). p. 46-47.

18. ARQUIVO NACIONAL. *Dicionário Brasileiro de Terminologia Arquivística*. Rio de Janeiro: Arquivo Nacional, 2005. (Publicações Técnicas, 51). p. 22.

19. CUNHA, Murilo Bastos da; CAVALCANTI, Cordélia Robaldinho de Oliveira. *Dicionário de Biblioteconomia e Arquivologia*. Brasília: Briquet de Lemos/Livros, 2008. p. 13.

20. Cf. BOLFARINE; BUSSAB. *Elementos de amostragem*, p. 15.

21. Ainda que caibam discussões a esse respeito, envolvendo a categoria de objetividade.

22. KEPLEY, DAVID R. Sampling in Archives: a Review. *American Archivist*. Chicago: Society of American Archivists/SAA, v. 47, n. 3, p. 239-240, Summer 1984.

23. KEPLEY. Sampling in Archives, p. 241.

24. KEPLEY. Sampling in Archives, p. 238. Desde então, aconselhar uma lista de critérios que possa, ao mesmo tempo, preservar os documentos em destaque e administrar a aplicação da técnica tem sido um problema importante.

25. Cf. SCHELLENBERG. *Arquivos modernos*, p. 181-199.

26. PEARCE-MOSES, Richard. *A Glossary of Archival and Records Terminology*. Disponível em: <[http://www.archivists.org/glossary/term\\_details.asp?DefinitionKey=3065](http://www.archivists.org/glossary/term_details.asp?DefinitionKey=3065)>. Acesso em: 7 out. 2011.

27. PEARCE-MOSES, Richard. *A Glossary of Archival and Records Terminology*. Disponível em: <[http://www.archivists.org/glossary/term\\_details.asp?DefinitionKey=3064](http://www.archivists.org/glossary/term_details.asp?DefinitionKey=3064)>. Acesso em: 7 out. 2011.

28. PEARCE-MOSES, Richard. *A Glossary of Archival and Records Terminology*. Disponível em: <[http://www.archivists.org/glossary/term\\_details.asp?DefinitionKey=227](http://www.archivists.org/glossary/term_details.asp?DefinitionKey=227)>. Acesso em: 7 out. 2011.

29. PEARCE-MOSES, Richard. *A Glossary of Archival and Records Terminology*. Disponível em: <[http://www.archivists.org/glossary/term\\_details.asp?DefinitionKey=3063](http://www.archivists.org/glossary/term_details.asp?DefinitionKey=3063)>. Acessado em: 7 out. 2011.

30. PEARCE-MOSES, Richard. *A Glossary of Archival and Records Terminology*. Disponível em: <[http://www.archivists.org/glossary/term\\_details.asp?DefinitionKey=3066](http://www.archivists.org/glossary/term_details.asp?DefinitionKey=3066)>. Acesso em: 7 out. 2011.

31. Apesar de o *Dicionário Brasileiro de Terminologia Arquivística* traduzir *Case Files* como *Dossiês*, optou-se no entanto, por não utilizar essa tradução no presente texto. Dadas as diferenças entre as tradições administrativas britânicas e brasileiras, os *Case Files* possuem características que aqui podem ser associadas a *Processos*. ARQUIVO NACIONAL. *Dicionário Brasileiro de Terminologia Arquivística*, p. 210. V. THE NATIONAL ARCHIVE (UK). *Operational Selection Policy OSPO – The Selection of Case Files: Sampling Techniques*. 2005. Disponível em: <<http://collections.europarchive.org/tna/20081023125241/http://www.nationalarchives.gov.uk/recordsmanagement/selection/ospintro.htm>>. Acesso em: 3 abr. 2012. THE NATIONAL ARCHIVE (UK). *Operational Selection Policy OSP48 – Case Files*. 2009. Disponível em: <<http://collections.europarchive.org/tna/20081023125241/http://www.natio>

nalarchives.gov.uk/recordsmanagement/selection/ospintro.htm>. Acesso em: 3 abr. 2012.

32. THE NATIONAL ARCHIVE (UK). *Operational Selection Policy OSP48*, p. 4.

33. Deixa-se claro, no entanto, que essas cópias devem obedecer às prescrições técnicas estabelecidas pelo Arquivo Nacional Britânico.

34. Assim, tanto a OSPO (2005) quanto a OSP48 (2009) indicam aspectos complementares da metodologia de amostragem e seleção de documentos. A primeira dessas normas está em desuso e as demais estão sujeitas a mudanças.

35. Para os próximos parágrafos, seguimos os pontos 6.4 e 6.5 de THE NATIONAL ARCHIVE (UK). *Operational Selection Policy OSPO*, p. 6-12.

36. Trata-se de um imenso acervo descoberto acidentalmente em 2005 e que pertencia à Polícia Nacional da Guatemala. Esse órgão existiu por quase todo o século XX, sendo especialmente ativo como mecanismo de repressão durante o conflito armado no país (1960-1996). Foi extinto em 1996 como parte do pactuado nos acordos de paz que puseram fim ao conflito. Para mais detalhes, ver GUZMAN, Daniel *et al.* Studying Millions of Rescued Documents: Sample Plan at the Guatemalan National Police Archive. In: JOINT STATISTICAL MEETINGS, 2009, Washington, DC. *The Guatemalan Police Archive Project: Sample Design, Weighting, and Analysis - Topic Contributed – Papers*. Disponível em: <[http://www.amstat.org/meetings/jsm/2009/onlineprogram/index.cfm?fuseaction=activity\\_details&activityid=584&sessionid=204017](http://www.amstat.org/meetings/jsm/2009/onlineprogram/index.cfm?fuseaction=activity_details&activityid=584&sessionid=204017)>. Acesso em: 3 abr. 2012.

37. GUZMAN *et al.* Studying Millions of Rescued Documents, p. 2.

38. HUTH, Geoffrey A. *Re: Sao Paulo State Archives – Brazil* [mensagem pessoal]. Mensagem recebida por <cgdsaesp@sp.gov.br> em 2 dez. 2011. V. NEW YORK State Education Department. Cultural Education Center. Disponível em: <[www.archives.nysed.gov](http://www.archives.nysed.gov)>. Acesso em: 31 jan. 2012.

39. NELB, Tawny Ryan. Architectural Records Appraisal: Discussion of Problems and Strategies for the Documenting Michigan Architecture Project. *American Archivist*. Chicago: Society of American Archivists/SAA, n. 59, Spring 1996.

40. COOK, T. Many Are Called, but Few Are Chosen: Appraisal Guidelines for Sampling and Selecting Case Files. *Archivaria*. América do Norte, n. 32, p. 27-28, jan. 1991. Disponível em: <<http://journals.sfu.ca/archivar/index.php/archivaria/article/view/11759/12709>>. Acesso em: 3 fev. 2012.

41. FUNDAÇÃO SEADE. *Proposta de Procedimento de Seleção de Documentos para Guarda Permanente*: Cooperação Técnica entre a Fundação SEADE e o Arquivo Público Estadual. São Paulo, 2011 (material impresso não publicado). p. 2.

42. V. BRADSHER, James Gregory; AMBACHER, Bruce I. *Archival Sampling: A Method of Appraisal and a Means of Retention*. New York: Mid-Atlantic Regional Archives Conference/MARAC, 1992 (Technical Leaflet Series, n. 8).

43. Situações ilustrativas são descritas em KOLISH, Evelyn. Sampling Methodology and its Application: An Illustration of the Tension Between Theory and Practice? *Archivaria*, América do Norte, n. 38, jan. 1994. Disponível em: <<http://journals.sfu.ca/archivar/index.php/archivaria/article/view/12024/12992>>. Acesso em: 3 fev. 2012; e em SCHEINBERG, Ellen. Case File Theory: Does It Work in Practice? *Archivaria*, América do Norte, n. 38, jan. 1994. Disponível em: <<http://journals.sfu.ca/archivar/index.php/archivaria/article/view/12023/12990>>. Acesso em: 3 fev. 2012.

44. Este é o teor do convênio entre o Arquivo Público do Estado de São Paulo e a Fundação Sistema Estadual de Análise de Dados.

45. TURNER, Jane. Theoretical Dialectics: A Commentary on Sampling Methodology and its Application. *Archivaria*, América do Norte, n. 38, p. 76, jan. 1994. Disponível em: <<http://journals.sfu.ca/archivar/index.php/archivaria/article/view/12025/12994>>. Acesso em: 4 fev. 2012.

46. O uso da calculadora é livre e pode ser feito *online*, no sítio eletrônico da empresa. Para o recurso e demais informações, v. SAMPLE SIZE CALCULATOR. [Calculadora de Tamanho Amostral]. Califórnia, Estados Unidos: Creative Research Systems. Disponível em: <<http://www.surveysystem.com/sscalc.htm>>. Acesso em: 8 fev. 2012.

47. FUNDAÇÃO SEADE. *Proposta de Procedimento de Seleção de Documentos para Guarda Permanente*: Cooperação Técnica entre a Fundação Seade e o Arquivo Público Estadual. São Paulo, SP, 2011 (material impresso não publicado). p. 3-4.

48. A versão 3.0 é paga e pode ser conhecida e adquirida no sítio eletrônico de divulgação do produto. V. EX-SAMPLE. *Methodologist Toolchest*. Version 3.0. [s.l.]: The Idea Works, 2012. Disponível em: <<http://www.ideaworks.com/mt/index.html>>. Acesso em: 8 fev. 2012.

49. TURNER, Jane. Theoretical Dialectics: A Commentary on Sampling Methodology and its Application. *Archivaria*, América do Norte, n. 38, p. 77, jan. 1994. Disponível em: <<http://journals.sfu.ca/archivar/index.php/archivaria/article/view/12025/12994>>. Acesso em: 4 fev. 2012.

**Caio Graco Valle Cobério** foi servidor do Departamento de Gestão do Sistema de Arquivos do Estado de São Paulo (DGSASP) e participou do grupo de trabalho que estudou a definição de critérios norteadores para amostragem de documentos destinados à eliminação, por meio de cooperação técnica estabelecida entre o Arquivo Público do Estado de São Paulo e a Fundação Sistema Estadual de Análise de Dados (Seade). Agradecimento particular ao colega Luiz Felipe Loureiro Foresti pelas colaborações. E-mail: prof.caigraco@yahoo.com.br

## Quadro 1 | Amostragem documental

<b>Método estatístico:</b> com representatividade.	<b>Método não estatístico:</b> sem representatividade.
<b>Probabilístico:</b> é possível mensurar a precisão dos indicadores.	<b>Não probabilístico:</b> não é possível mensurar a precisão dos indicadores.
<b>Quantitativo:</b> preserva as características da série no conjunto originalmente constituído (parâmetros).	<b>Qualitativo:</b> preserva as características mais relevantes individualmente.
<b>Objetivo:</b> parte da definição e homogeneidade do universo documental.	<b>Subjetivo:</b> parte da experiência individual e interesse do arquivista. Universo documental heterogêneo.
<b>Nível de confiança:</b> quantificação da probabilidade de o estimador de uma determinada amostra selecionada ter um erro menor ou igual ao desejável.	<b>Importância:</b> o valor atribuído aos documentos da amostra é o fator preponderante para torná-lo digno de preservação.*
<b>Erro amostral:</b> diferença entre o valor observado na amostra e o parâmetro de interesse na população. O erro aceitável corresponde à quantidade considerada razoável de que a medida inferida (estimativa) pela amostra difira da quantidade populacional (parâmetro).	<b>Lista de critérios:</b> determinados atributos dos documentos são sublinhados, indicando aqueles que possuem maior importância.

\* No método não estatístico, significância remete ao documento importante, valioso, digno. Diferentemente, quando usada no sentido estatístico, a significância corresponde à proximidade com uma relação correta, isto é, provavelmente verdadeira. Adotamos “importância” como equivalente a “significância”, na primeira das acepções acima.