



**Universidade de Brasília
Departamento de Estatística**

**Estimação da Proporção de Livros com Necessidade de Restauração na
Biblioteca Central da Universidade de Brasília**

Maria Luiza Barreto Escobar

Relatório apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

**Brasília
2024**

Maria Luiza Barreto Escobar

**Estimação da Proporção de Livros com Necessidade de Restauração na
Biblioteca Central da Universidade de Brasília**

Orientador: Prof. Dr. Alan Ricardo da Silva

Relatório apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

**Brasília
2024**

Agradecimentos

Gostaria de expressar minha gratidão a todas as pessoas que contribuíram de maneira significativa para a realização deste trabalho.

Agradeço primeiramente ao meu pai, Prof. Me. Daniel Mattos Escobar, que foi meu professor em sala de aula e continua sendo na minha vida. Me trouxe muitos ensinamentos que com certeza me ajudaram a chegar onde estou e a ir muito além.

Agradeço ao meu orientador, Prof. Dr. Alan, por ser um profissional que admiro e por ter sido extremamente atencioso em suas orientações, sempre com muita dedicação e interesse em repassar seus conhecimentos.

Agradeço aos técnicos e servidores da BCE, que disponibilizaram de seu tempo para auxiliar no acesso à biblioteca durante este período e na obtenção de todas as informações necessárias para a realização deste estudo.

Agradeço aos meus colegas que me acompanharam ao longo do curso, oferecendo ajuda muito além do que eu poderia pedir. Por fim, agradeço aos meus familiares e amigos, que foram meu suporte e minha distração nos momentos de dificuldade.

Resumo

Este trabalho tem como objetivo geral desenvolver um plano de amostragem para a estimação da proporção de livros com necessidade de restauração na Biblioteca Central da Universidade de Brasília. Foram propostos dois métodos distintos: amostragem aleatória simples e amostragem por conglomerados em dois estágios, para comparação e decisão sobre o melhor método, considerando diversas características complexas do acervo. A amostragem aleatória simples foi realizada por meio do sorteio de livros do banco de dados, já a amostragem por conglomerados considerou os módulos – estrutura divisória das estantes – como unidades primárias amostrais e os livros como as unidades secundárias amostrais. Os resultados indicaram que a adoção de uma amostragem por conglomerados em dois estágios, pode diminuir consideravelmente os custos relativos ao tempo de coleta, em troca de estimativas menos precisas. Sugere-se, portanto, a utilização da amostragem por conglomerados em dois estágios com o objetivo de reduzir custos com o tempo de coleta.

Palavras-chave: Amostragem por Conglomerados em Dois Estágios, Restauração, Biblioteca Central da Universidade de Brasília.

Lista de Tabelas

2.1	Classificação da CDU por área de conhecimento	12
3.1	Resumo dos estimadores por técnica amostral	22
4.1	Variáveis utilizadas do banco de dados da BCE	26
4.2	Ordenação utilizada na BCE	27
5.1	Resumo dos resultados da estimação para a AAS	33
5.2	Ordenação ASCII utilizada no SAS	35
5.3	Tabela de correlação entre a contagem física e a estimada pelo banco de dados na classe 2	36
5.4	Módulos sorteados na AC2 6 × 7	37
5.5	Módulos sorteados na AC2 3 × 14	37
5.6	Resumo dos resultados da estimação para a AC2 6 × 7	37
5.7	Resumo dos resultados da estimação para a AC2 3 × 14	37
5.8	<i>Deffs</i> da AC2 6 × 7 e AC2 3 × 14	38
5.9	Variâncias específicas da AC2 6 × 7 e AC2 3 × 14	38

Lista de Figuras

2.1	Ficha de diagnóstico do exemplar	13
2.2	Capa solta	14
2.3	Folhas cortadas	14
2.4	Costura solta	14
2.5	Folha rasgada	14
2.6	Mofa e ferrugem	14
2.7	Capa rasgada	14
2.8	Carimbo de identificação	15
2.9	Adesivo de conscientização	15
2.10	Distribuição de classes das Coleções Especiais no primeiro andar	16
2.11	Distribuição de classes do Acervo Geral e das Referências no térreo	16
2.12	Distribuição de classes do Acervo Geral no subsolo 1	17
4.1	Ordenação do objeto de estudo	27
5.1	Frequência total de exemplares por classe no Acervo Geral	31
5.2	Frequência total de exemplares por tipo de aquisição no Acervo Geral	32
5.3	Frequência total de exemplares por década e tipo de aquisição no Acervo Geral	32
5.4	Frequência de exemplares com necessidade de restauração por década de aquisição na AAS	34
5.5	Gráfico de correlação entre a contagem física e a estimativa pelo banco de dados na classe 2	36

Sumário

1	Introdução	8
1.1	Objetivos	9
2	Biblioteca Central de Brasília	10
2.1	Introdução.	10
2.2	História	10
2.3	Acervo	11
2.3.1	Classificação dos Exemplares	11
2.4	Restauração do Acervo.	12
2.5	Organização Física	15
3	Técnicas de Amostragem	18
3.1	Introdução.	18
3.2	Amostragem Aleatória Simples	18
3.3	Amostragem Estratificada.	18
3.4	Amostragem Sistemática	19
3.5	Amostragem por Conglomerados	19
3.5.1	Custo	20
3.5.2	Peso Amostral	20
3.6	Efeito do Planejamento	21
3.7	Estimadores.	21
3.8	Tamanho Amostral	23

4	Materiais e Métodos	25
4.1	Introdução.	25
4.2	Materiais.	25
4.3	Métodos	26
5	Estudo de Caso	30
5.1	Introdução.	30
5.2	Análise Exploratória	30
5.3	Amostra Aleatória Simples	33
5.4	Amostra por Conglomerados em Dois Estágios	34
5.5	Custos Estimados	38
6	Conclusão	41
	Referências.	42
	Apêndice	45
A	Tabela da CDU para dois dígitos	45

Capítulo 1

Introdução

A Biblioteca Central (BCE) é um dos principais e mais importantes recursos da Universidade de Brasília. Amplamente utilizada pelos discentes, docentes, técnicos e pela comunidade externa, a biblioteca armazena um imenso acervo composto majoritariamente por livros, mas também por folhetos, quadrinhos e periódicos.

No entanto, esse tipo de material sofre deterioração causada por diversos fatores, tais como agentes químicos, físicos e biológicos, poluentes atmosféricos, desastres naturais e pela ação do homem (Antunes, 2010). Assim, a preservação e o monitoramento do estado físico do acervo é um dos pilares para o bom funcionamento de uma biblioteca.

Devido ao grande número de exemplares situados no espaço físico da BCE, bem como a constante movimentação de livros, seja por meio de consultas momentâneas ou de empréstimos e da entrada e remoção de itens do acervo, torna-se impraticável examinar todos os livros em um prazo de tempo razoável. Além disso, os custos da análise poderiam ser consideráveis, tendo em vista o tempo e mobilização necessária para tal atividade.

É possível simplificar a atividade através de um levantamento por amostragem, conforme sugerido por Cochran (1977). Esse tipo de levantamento tem como objetivo coletar informações específicas num menor conjunto de unidades. No caso específico retratado, ao observar a distribuição dos livros no espaço físico, a seleção amostral torna-se complicada devido à presença de diversas áreas do conhecimento, com uma quantidade diferente de livros, alocados em estantes, módulos e prateleiras.

Portanto, no cenário analisado, é necessário realizar uma amostragem envolvendo primeiramente a estratificação (ou subdivisão) das áreas do conhecimento e, então, conglomeração com seleção em dois estágios. No primeiro estágio, a seleção está relacionada aos módulos, enquanto o segundo estágio se concentra nos livros. Poder-se-ia ainda utilizar a

seleção das prateleiras, dentro dos módulos, como um estágio adicional. Esse método de amostragem, que emprega múltiplos estágios, estratificação e pesos desiguais, é conhecido como amostragem complexa (Lohr, 2010).

Considerando que a BCE da Universidade de Brasília é um espaço de uso coletivo e que a preservação do acervo é fundamental para seu adequado funcionamento, é essencial conhecer o estado de seu patrimônio. Assim, o monitoramento de métricas gerais relacionadas ao estado dos livros pode resultar em uma melhor conservação do acervo.

1.1 Objetivos

Esse trabalho tem como objetivo geral desenvolver um plano de amostragem para a estimação da proporção de livros com necessidade de restauração na Biblioteca Central da Universidade de Brasília.

Os objetivos específicos são:

- Mensurar o tipo de avaria que levará à restauração;
- Criar um instrumento estruturado para a coleta dos dados;
- Comparar o tempo e a variância dos planos amostrais por conglomerados e por amostragem aleatória simples.

Capítulo 2

Biblioteca Central de Brasília

2.1 Introdução

A Biblioteca Central da Universidade de Brasília, conhecida como BCE, desempenha um papel essencial na vida acadêmica e cultural da instituição. Neste Capítulo é fornecida uma visão geral da BCE, incluindo sua história e caracterização de seu acervo, com ênfase nos espaços destinados aos livros, que constituem o objeto de estudo deste trabalho.

2.2 História

O conceito da Biblioteca Central foi estabelecido pelo Estatuto da Universidade, em 1962, sendo esta uma unidade principal de obras gerais e de consulta, com serviços de aquisição, catalogação, documentação e intercâmbio científico e cultural (Brasil, 1962). A primeira instalação se deu no edifício do Ministério da Educação, com acervo bibliográfico de emergência – exemplares doados por entidades públicas e privadas, sem critério seletivo (Fonseca, 1973). Apenas na década de 1970 que o prédio definitivo, utilizado atualmente, foi construído e inaugurado (BCE, 2023).

Durante o período do regime militar em 1964, a Polícia Militar de Minas Gerais invadiu o campus e interditou a Biblioteca Central, em uma busca por materiais considerados subversivos. A reintegração das obras confiscadas ao acervo ocorreu somente em 1986, 22 anos após o ocorrido (BCE, 2023).

Em 1967, a BCE começou a funcionar aos domingos e feriados para atender às necessidades dos leitores, uma prática que continua até os dias atuais (BCE, 2023). Duas décadas mais tarde, em 1987, foi estabelecido o Setor de Conservação e Restauração de

Livros e Documentos, conhecido atualmente como RES. Em 2004, esse setor liderou o projeto de higienização e restauração do acervo, envolvendo a participação de egressos do sistema prisional (BCE, 2023).

Ao longo dos anos, a BCE também desempenhou um papel significativo em contribuições sociais, incluindo parcerias como o convênio com a APAE (Associação de Pais e Amigos dos Excepcionais) em 2006, para a formação de pessoas com deficiência mental e/ou múltipla. Em 2008, a Biblioteca Digital e Sonora foi criada, proporcionando cabines sonoras para deficientes visuais (BCE, 2023).

Assim, a BCE se adaptou e evoluiu para atender às demandas acadêmicas, promover a inclusão social e preservar o conhecimento para as gerações futuras. Sua trajetória é marcada por uma constante busca pela excelência e pelo comprometimento com a comunidade acadêmica. Não é por acaso que, na década de 1990, a BCE participou da “Exposição do Livro Ilustrado Brasileiro” em Haia, Holanda, e recebeu a classificação de nível A pelo Programa de Acesso Público à Informação pelo Conselho de Desenvolvimento Científico e Tecnológico (CNPq) (BCE, 2023).

2.3 Acervo

O acervo da BCE possui mais de 1.4 milhão de volumes distribuídos entre diversos tipos de materiais, dos quais, em 2022, 513.861 eram exemplares de livros (BCE, 2022). Estes estão espalhados em 4 coleções do acervo físico: Acervo Geral, Coleções Especiais, Obras Raras e Referência (BCE, 2022).

O Acervo Geral é o espaço que abrange a maior parte dos livros e folhetos, já as Coleções Especiais incluem livros da Editora da UnB, teses e dissertações da Universidade, publicações de organismos internacionais, acervos especiais, mapas e multimídia (BCE, 2022). A Coleção de Obras Raras, por sua vez, possui manuscritos antigos e modernos, bem como primeiras edições de autores brasileiros e portugueses, entre outros tesouros bibliográficos. Finalmente, a seção de Referência engloba recursos essenciais para a pesquisa, tais como dicionários, enciclopédias, guias, índices, catálogos e outros materiais de consulta (BCE, 2022).

2.3.1 Classificação dos Exemplares

Os livros são organizados em estantes, módulos e prateleiras por meio do código de classificação. A BCE utiliza o sistema de Classificação Decimal Universal (CDU) (BCE,

2023), que propõe 9 classes definidas de acordo com a área de conhecimento (Tabela 2.1). A classe 4 está reservada para uso posterior ou para categorias específicas não incluídas na versão principal do sistema.

Tabela 2.1: Classificação da CDU por área de conhecimento

Classe	Conhecimento
0	Generalidades, Informação, Organização
1	Filosofia, Psicologia
2	Religião, Teologia
3	Ciências Sociais, Economia, Direito, Política, Assistência Social, Educação
5	Matemática, Ciências Naturais
6	Ciências Aplicadas, Medicina, Tecnologia
7	Arte, Belas-Artes, Recreação, Diversões, Esportes
8	Linguística, Linguagem, Literatura
9	Biografia, Geografia, História

Cada grande classe é representada por um algarismo único na classificação. Essas classes podem ser subdivididas em classes mais específicas, chamadas subclasses. A extensão do número de classificação reflete o grau de detalhamento, e números com a mesma quantidade de algarismos representam níveis de generalidade similares. Por exemplo, o livro “Técnicas de Amostragem”, do autor Cochran, tem o número de chamada 311.2, que pode ser visualizado da seguinte maneira (FID, 1987):

3 Ciências Sociais, Economia, Direito, Política, Assistência Social, Educação;

311 Estatística como ciência, teoria estatística;

311.2 Técnica de pesquisa, preparação, tabulação.

Esse tipo de sistema permite a síntese, ou seja, a construção de números compostos para categorização de assuntos inter-relacionados. A descrição das classificações em dois algarismos, conforme FID (1987), está no Apêndice A .

2.4 Restauração do Acervo

A restauração é um procedimento que envolve a intervenção estrutural em um objeto, abrangendo desde a limpeza até a reconstrução de partes faltantes, procurando realizar a mínima intervenção possível (Froner e Souza, 2008). Para isso, é importante que se siga uma metodologia para determinar as intervenções necessárias por meio de estudos preliminares, que envolvem a análise da utilização do exemplar, das características morfológicas, do estado de conservação, entre outras (Froner e Souza, 2008).

A BCE conta com uma equipe de restauração no Setor de Conservação e Restauração (RES), dentro da Coordenadoria de Formação e Desenvolvimento de Acervos (FDA) (BCE, 2022). A RES possui sua própria metodologia para seleção dos exemplares a serem restaurados, passando por uma classificação de prioridades, definidas por meio da avaliação de diversos critérios. O diagnóstico físico do exemplar é registrado em uma ficha, representada na Figura 2.1.

Diagnóstico de Conservação - Selecionado	
Nº Chamada:	_____
Exemplar:	_____
Acervo:	_____
Falta folha de rosto:	Sim () Não ()
Falta capa/contracapa:	Sim () Não ()
Encadernação/capa solta:	Sim () Não ()
Costura solta:	Sim () Não ()
Página solta:	Sim () Não ()
Nº	_____
Página rasgada:	Sim () Não ()
Nº	_____
Falta página:	Sim () Não ()
Nº	_____
Pág. colada c/fita adesiva:	Sim () Não ()
Nº	_____
Apresenta mofo/insetos:	Sim () Não ()
Nº	_____
Responsável:	_____
Data:	___/___/___
Observações:	_____
Biblioteca Central - Conservação e Restauração	

Figura 2.1: Ficha de diagnóstico do exemplar
Fonte: RES (2024)

A ficha de diagnóstico traz a identificação do exemplar e as características que necessitam de intervenção, a identificação do colaborador responsável e a data da avaliação, além das observações que podem constatar outras avarias não listadas ou informações específicas pertinentes ao trabalho do restaurador. Para tornar o processo de classificação e avaliação das avarias menos subjetivo, utiliza-se referências visuais, como as encontradas em Paglione (2017). Assim, o processo de trabalho pode ser reproduzido por novos servidores/técnicos mantendo certa consistência. As Figuras de 2.2 a 2.7 ilustram as avarias dos próprios exemplares do Acervo Geral como modelos das avarias que são consideradas no processo de restauração.

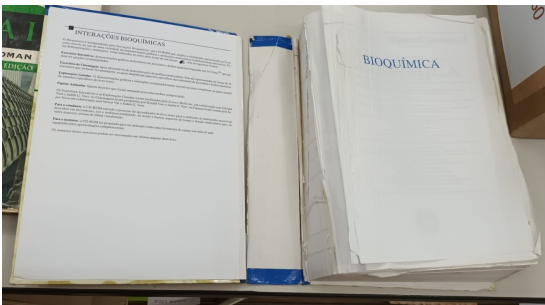


Figura 2.2: Capa solta

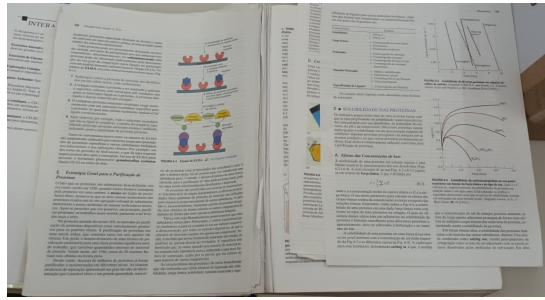


Figura 2.3: Folhas cortadas

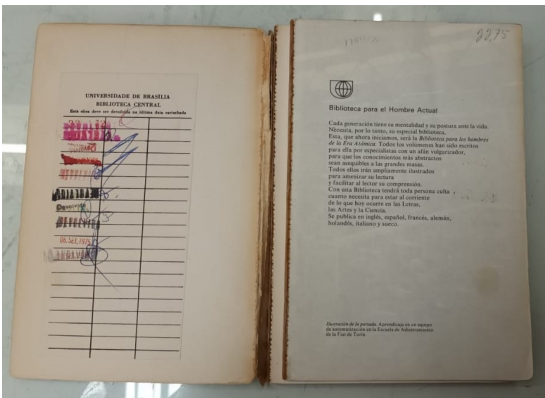


Figura 2.4: Costura solta

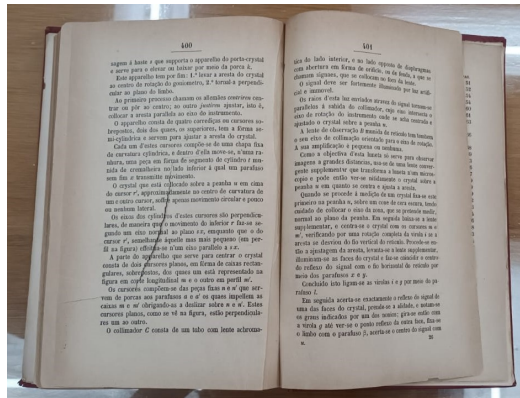


Figura 2.5: Folha rasgada

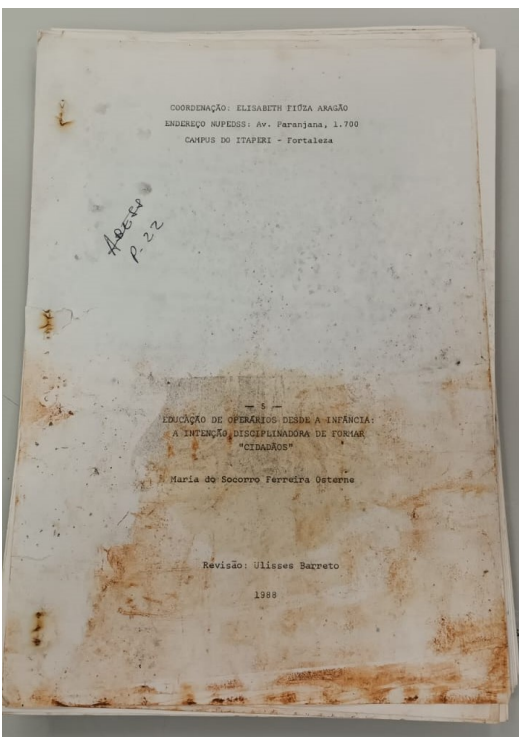


Figura 2.6: Mofo e ferrugem



Figura 2.7: Capa rasgada

Os exemplares que são restaurados recebem um carimbo de identificação, para conhecimento dos usuários e dos próprios servidores. O carimbo é encontrado ou nas primeiras ou nas últimas páginas do exemplar, como mostra a Figura 2.8. Este traz a data da restauração e a assinatura do restaurador. A Biblioteca Central também tem

implementado uma nova iniciativa que envolve a aplicação de adesivos nos exemplares restaurados (Figura 2.9), contendo informações de boas práticas para conscientização dos usuários com o objetivo de aumentar a durabilidade dos livros.

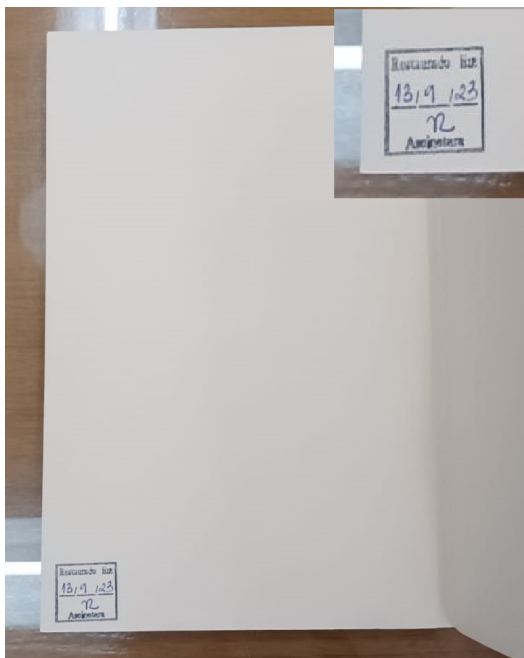


Figura 2.8: Carimbo de identificação



Figura 2.9: Adesivo de conscientização

2.5 Organização Física

A Biblioteca Central possui quatro pavimentos, sendo três de acesso público. No subsolo está localizado o Acervo Geral (Classes 0 a 3), já o térreo divide espaço para o Acervo Geral (Classes 5 a 9) e para a Referência. Por fim, no primeiro andar estão situadas as Coleções Especiais. Após a designação da coleção dos exemplares, estes são organizados de forma crescente por meio da numeração da CDU. As Figuras 2.10, 2.11 e 2.12 mostram a distribuição dos livros por classe nos pavimentos da BCE, observada em abril de 2024.

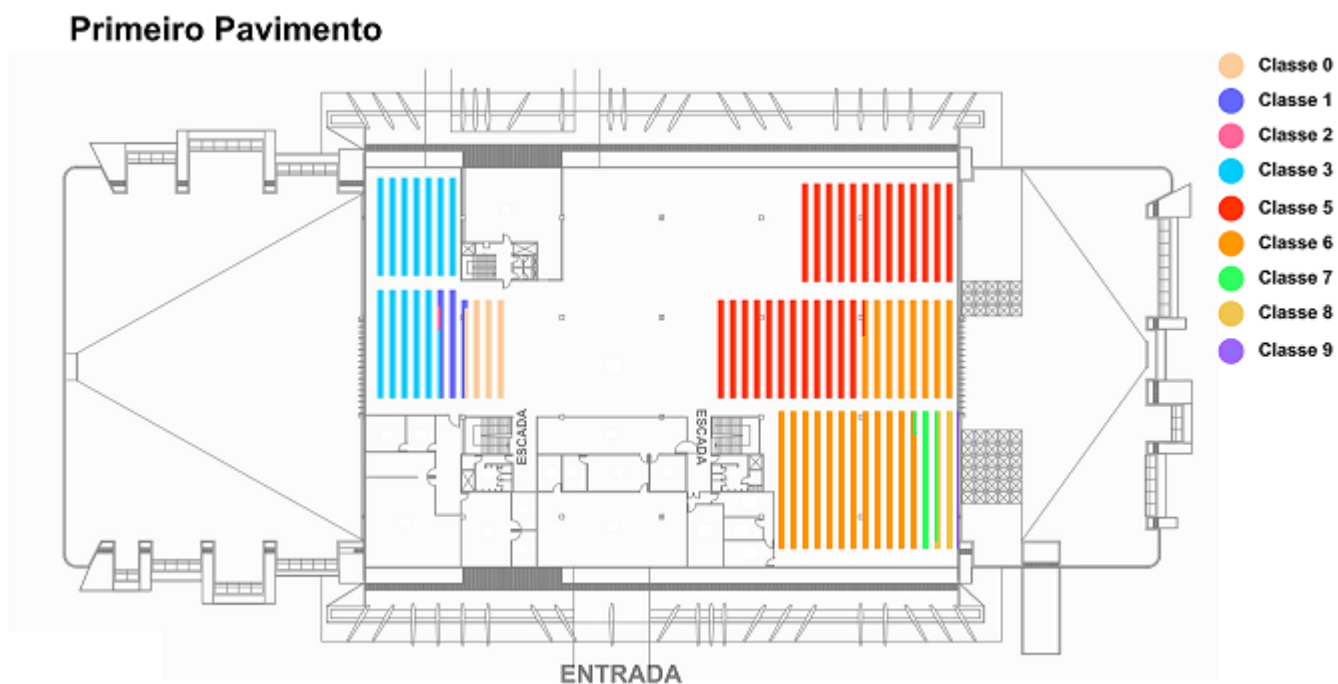


Figura 2.10: Distribuição de classes das Coleções Especiais no primeiro andar
Fonte: CEPLAN (2017) com modificações.

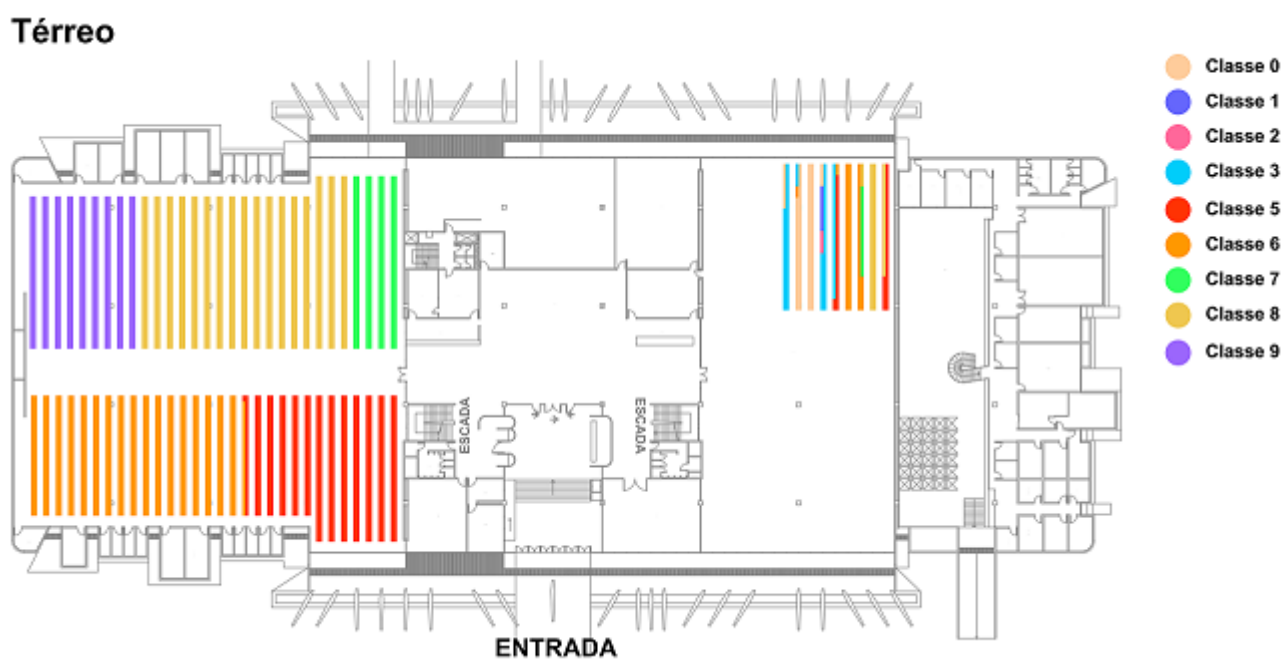


Figura 2.11: Distribuição de classes do Acervo Geral e das Referências no térreo
Fonte: CEPLAN (2017) com modificações.

Subsolo 1

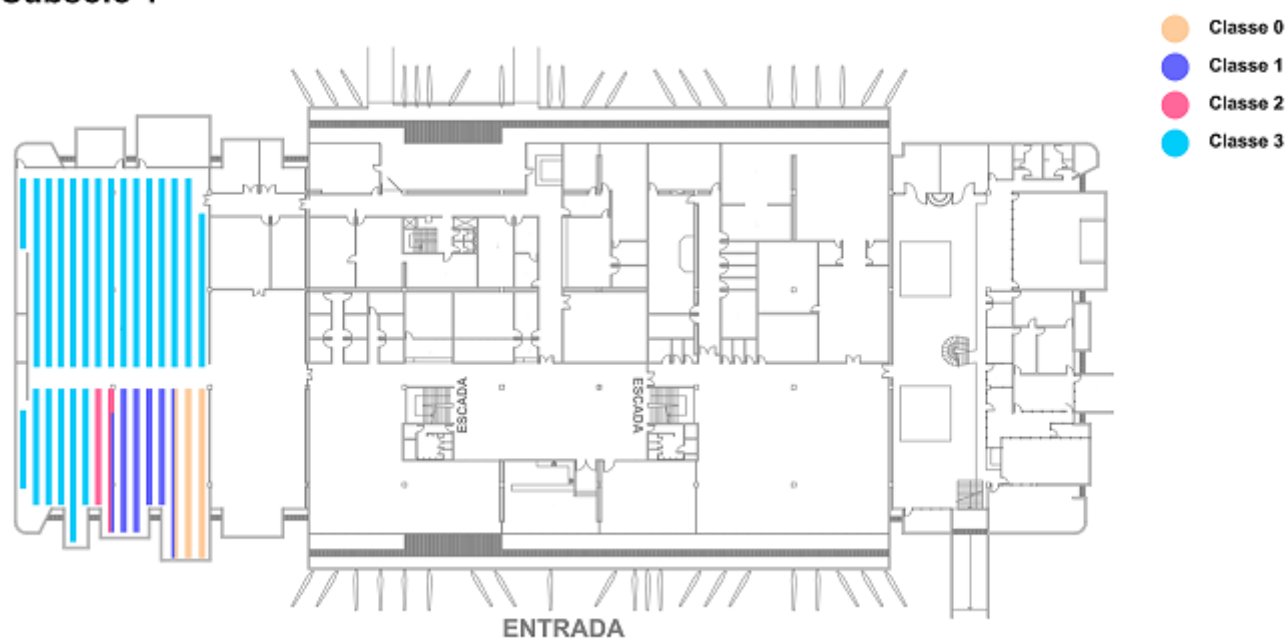


Figura 2.12: Distribuição de classes do Acervo Geral no subsolo 1

Fonte: CEPLAN (2017) com modificações.

Capítulo 3

Técnicas de Amostragem

3.1 Introdução

Amostragem é uma estratégia de pesquisa que visa adquirir medidas suficientemente precisas da população através da escolha de uma amostra, que consiste em uma parcela menor dos elementos da população. A população, por sua vez, engloba a totalidade dos elementos dos quais se pretende obter informações detalhadas (Cochran, 1977).

A amostra pode ser escolhida com ou sem reposição, sendo a escolha entre essas abordagens determinada pelas particularidades do problema e pelas suposições feitas (Cochran, 1977). A seguir, serão apresentadas algumas técnicas de amostragem e conceitos que serão utilizados no trabalho, considerando amostras sem reposição.

3.2 Amostragem Aleatória Simples

A Amostragem Aleatória Simples (AAS) é um procedimento no qual um subgrupo de tamanho n é escolhido aleatoriamente de uma população de tamanho N . Nesse método, cada unidade da população e cada amostra possível tem mesma probabilidade de ser selecionada (Bolfarine e Bussab, 2005).

3.3 Amostragem Estratificada

A Amostragem Estratificada (AE) considera a divisão da população em estratos, ou seja, grupos onde as unidades possuem características semelhantes. Após isso, uma AAS é escolhida para a seleção de uma amostra para cada estrato selecionado (Bolfarine e Bussab, 2005). Costuma ser utilizada quando existem grupos na população com tamanhos

muito diferentes, ou quando há o interesse de obter uma precisão conhecida da estimativa de cada subgrupo. Além disso, se aplicada corretamente, trará estimativas mais precisas da população como um todo (Lohr, 2010).

3.4 Amostragem Sistemática

A Amostragem Sistemática (AS) é aplicada quando há uma lista ordenada por um fator aleatório das unidades populacionais disponíveis ou quando não existe uma lista com as unidades populacionais. A primeira unidade é escolhida por meio de um número aleatório entre 1 e k , sendo $k = \frac{N}{n}$, e as subsequentes são selecionadas em intervalos regulares de k em k , a partir deste número inicial (Lohr, 2010). Muitas vezes a amostragem sistemática é utilizada como uma alternativa conveniente em situações onde a ordenação dos itens se dá por um fator aleatório, ou seja, não relacionado à variável de estudo. Nesses casos ela é equivalente à amostragem aleatória simples (Cochran, 1977).

3.5 Amostragem por Conglomerados

A Amostragem por Conglomerados (AC), diferentemente da AAS, AE e AS que selecionam diretamente os elementos a serem analisados da população, considera que as unidades disponíveis para a seleção são grupos de elementos, os chamados conglomerados. A seleção dos agrupamentos e de seus elementos, seja por questão de custo ou por questões administrativas, pode ser feita de duas formas:

Amostragem por conglomerados de um estágio (AC1): Nesse método, a população é subdividida em conjuntos disjuntos chamados conglomerados, e esses conglomerados são selecionados através de uma AAS para formar a amostra, englobando todas as unidades dentro dos conglomerados escolhidos (Bolfarine e Bussab, 2005). Essas unidades são chamadas de Unidades Primárias de Amostragem (UPA).

Amostragem por conglomerados de dois estágios (AC2): Ao contrário da AC, as unidades dentro dos conglomerados selecionados são também escolhidas utilizando o mesmo ou outro plano amostral dos conglomerados (Lohr, 2010). As unidades dentro dos conglomerados são chamadas de Unidades Secundárias de Amostragem (USA).

Embora a eficiência desses métodos amostrais tenda a ser inferior aos métodos mencionados anteriormente, a vantagem está na redução de custos, especialmente quando a população pode ser dividida em diferentes conglomerados (Bolfarine e Bussab, 2005).

3.5.1 Custo

O custo mencionado é relacionado as despesas vinculadas a coleta da amostra. Pode envolver gastos com mão de obra, tempo para coleta, deslocamento e outros. Na amostragem por conglomerados, o custo é comumente utilizado para cálculo do tamanho da amostra. Uma função de custo mais simples para conglomerado em dois estágios foi descrita por Lohr (2010) como:

$$C = c_1a + c_2am \quad (3.5.1)$$

na qual a é a quantidade de unidades no primeiro estágio, c_1 é o custo de coleta por unidade primária, m é a quantidade de unidades no segundo estágio e c_2 é o custo por unidade secundária.

3.5.2 Peso Amostral

O peso amostral é um fator utilizado nos cálculos amostrais para ponderar a amostra em relação à população total. O peso amostral pode ser calculado como:

$$w_i = \frac{1}{p_i}, \quad (3.5.2)$$

em que p_i é a probabilidade de selecionar a i -ésima unidade. Essa probabilidade é geralmente calculada de forma proporcional ao tamanho da unidade, comumente chamada de “PPT - Probabilidade Proporcional ao Tamanho”. A PPT para a amostragem por conglomerados no primeiro estágio se dá por:

$$p_i = \frac{aM_i}{M_0}, \quad M_0 = \sum_{i=1}^N M_i. \quad (3.5.3)$$

No segundo estágio, considerando um número fixo m de subunidades a serem selecionadas, é dada por:

$$p_i = \frac{m}{M_i}. \quad (3.5.4)$$

Assim, considerando ambos os estágios realizados por PPT, a probabilidade irá resultar em:

$$p_i = \frac{aM_i}{M_0} \frac{m}{M_i} = \frac{am}{M_0}, \quad (3.5.5)$$

gerando o peso amostral de (Cochran, 1977):

$$w_i = \frac{M_0}{am}, \quad (3.5.6)$$

em que M_i é o tamanho das unidades secundárias no i -ésimo conglomerado, n é o tamanho amostral das UPA's e m é o tamanho amostral das USA's (Lohr, 2010).

Planos amostrais podem ser considerados autoponderados (ou do inglês *self-weighting*) quando são um múltiplo do total de todas as subunidades da amostra (Cochran, 1977). Como M_0 é o total de unidades possível na amostra e $n = am$, a Equação (3.5.6) se torna:

$$w_i = \frac{N}{n}, \quad (3.5.7)$$

o que é equivalente ao peso de uma AAS.

3.6 Efeito do Planejamento

O efeito do planejamento (do inglês *Deff - Design Effect*) é um critério de comparação das variâncias de diferentes planos amostrais, comumente em relação à um plano considerado padrão (Bolfarine e Bussab, 2005). Além disso, também pode ser utilizado para o cálculo do tamanho da amostra (Cochran, 1977), sendo calculado como:

$$Deff = \frac{\text{Variância do plano amostral}}{\text{Variância da AAS}}. \quad (3.6.1)$$

A relação do *Deff* com o tamanho amostral será definida na Seção 3.8.

3.7 Estimadores

Uma estatística é qualquer característica numérica observada por meio de uma amostra da população, já uma estimativa é valor numérico obtido através da associação de uma estatística com uma expressão que irá “estimar” um parâmetro populacional (Bolfarine e Bussab, 2005).

A Tabela 3.1 traz os principais estimadores de algumas das técnicas amostrais apresentadas, apresentadas por Kish (1965):

Tabela 3.1: Resumo dos estimadores por técnica amostral

Técnica	Parâmetro	Estimador	Variância Estimada
AAS	Média	$\bar{y} = \frac{1}{n} \sum_{j=1}^n y_j$	$\left(1 - \frac{n}{N}\right) \frac{s^2}{n}$
	Proporção	$\hat{p} = \frac{b}{n}$	$\left(1 - \frac{n}{N}\right) \frac{\hat{p}(1-\hat{p})}{n-1}$
	Total	$\hat{t} = N\bar{y}$	$N^2 \left(1 - \frac{n}{N}\right) \frac{s^2}{n}$
AC1	Média	$\bar{y}_{ac} = \frac{1}{a} \sum_{i=1}^a \bar{y}_i$	$\frac{(1-f)}{nM} s^2 (1 + (M-1)\rho)$
	Total	$\hat{t} = \frac{A}{a} \sum_{i=1}^a t_i$	$A^2 \left(1 - \frac{a}{A}\right) \frac{s_1^2}{a}$
AC2	Média	$\bar{y}_{ac2} = \sum_{i=1}^a \frac{\bar{y}_i}{a}$	$\frac{A-a}{A} \frac{s_1^2}{a} + \frac{a}{A} \frac{M-m}{M} \frac{s_2^2}{am}$
	Total	$\hat{t}_i = \frac{A}{a} \sum_{i=1}^a M_i \bar{y}_i$	$A^2 \left(1 - \frac{a}{A}\right) \frac{s_1^2}{a} + \frac{A}{a} \sum_{i=1}^A \left(1 - \frac{m_i}{M_i}\right) M_i^2 \frac{s_2^2}{m_i}$

n é o número de observações na amostra;

y_j é a observação j da amostra;

N é o número de observações na população total;

s^2 é o estimador não viesado da variância de y em uma população finita, ou seja,

$$s^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1};$$

b é o número de observações, dentro da amostra, com a característica de interesse;

\bar{y}_i é a média do i -ésimo conglomerado no primeiro estágio;

A é a quantidade de conglomerados;

a é a quantidade de conglomerados na amostra;

f é a fração amostral, ou seja, $\frac{n}{N}$;

M é a quantidade de observações por conglomerado, no caso de conglomerados de tamanhos iguais;

m é a quantidade de observações na amostra do conglomerado, no caso de conglomerados de tamanhos iguais;

M_i é a quantidade de observações dentro do i -ésimo conglomerado;

m_i é a quantidade de observações na amostra do i -ésimo conglomerado;

S_1^2 é a variância populacional total dos conglomerados, definida como

$$\frac{\sum_{i=1}^N M_i^2 (\bar{Y}_i - \bar{\bar{Y}})^2}{\bar{M}^2 (N-1)}, \quad (3.7.1)$$

em que \bar{M} é a média de M_i ;

S_2^2 é a variância populacional entre os elementos no i -ésimo conglomerado, definida como

$$\sum_{i=1}^N \frac{M_i}{M_0} S_{2i}^2, \quad S_{2i}^2 = \frac{\sum_{j=1}^{M_i} (y_{ij} - \bar{Y}_i)^2}{M_i - 1}, \quad M_0 = \sum_{i=1}^N M_i \quad \text{e} \quad (3.7.2)$$

ρ é a correlação dos elementos dentro do conglomerado, definida como

$$\rho = \frac{2 \sum_{i=1}^a \sum_{j < k} (y_{ij} - \bar{Y}_i)(y_{ik} - \bar{Y}_i)}{(M-1)(NM-1)S_2^2}. \quad (3.7.3)$$

Além disso, considerando que os dados seguem uma distribuição Normal por se tratar de amostras suficientemente grandes (Cochran, 1977), o intervalo de confiança é dado por

$$IC(\hat{\theta}) = \hat{\theta} \pm t_{gl, \gamma/2} \sqrt{\widehat{Var}(\hat{\theta})}, \quad (3.7.4)$$

em que t é o quantil da distribuição t -Student, gl são os graus de liberdade e γ é o nível de confiança. Para a AAS, os graus de liberdade são dados por $gl = n - 1$, enquanto para a AC, $gl = a - 1$.

3.8 Tamanho Amostral

A determinação do tamanho amostral constitui uma etapa primordial no processo de condução de uma pesquisa. Esse procedimento é realizado ao definir a precisão requerida das estimativas, mediante a especificação do nível de confiança γ e da margem de erro ϵ . O nível de confiança γ é a probabilidade de que o intervalo de confiança contenha a verdadeira estimativa, já a margem de erro é relacionada a variabilidade da estimativa (Cochran, 1977). O cálculo se dá por meio da margem de erro obtida pelo intervalo de confiança do estimador, expresso em (Bussab e Morettin, 2017), para uma Amostra Aleatória Simples, como:

$$n' = \frac{z_{\gamma/2}^2 S^2}{\epsilon^2}, \quad (3.8.1)$$

em que $z_{\gamma/2}$ é o quantil de uma distribuição normal para um nível de confiança γ e S^2 é a variância populacional. No caso de uma proporção, pode-se utilizar a variância máxima

$S^2 = 0,25$, eliminando um parâmetro dessa equação. Destaca-se que, quando o tamanho da população N é conhecido, a correção é aplicada através do FPC (*Finite Population Correction*), gerando a expressão dada por

$$n = \frac{n'}{1 + \frac{n'}{N}}, \quad (3.8.2)$$

na qual n' é calculado por meio da Equação (3.8.1) (Cochran, 1977). Caso N seja desconhecido, ou muito grande em relação à n' , então $n = n'$. Para uma amostra por conglomerados, o custo determina o tamanho da amostra, além da quantidade de itens por conglomerado. Assim, para esse caso, calcula-se m e a como (Lohr, 2010):

$$m = \sqrt{\frac{S_2^2 c_1}{S_1^2 - \frac{S_2^2}{M} c_2}} \quad (3.8.3)$$

e

$$a = \frac{S_1^2 - \frac{S_2^2}{M} + \frac{S_2^2}{m}}{Var(\bar{y}) + \frac{S_1^2}{A}}, \quad (3.8.4)$$

em que $Var(\bar{y})$ pode ser definida como (Kish, 1965):

$$Var(\bar{y}) = (CV(\bar{y})\bar{y})^2, \quad (3.8.5)$$

na qual o $CV(\cdot)$ é o coeficiente de variação. Considerando que, para a finalidade de cálculo amostral, a margem de erro pode ser escrita como

$$\epsilon = z_{\gamma/2} \sqrt{Var(\hat{\theta})}, \quad (3.8.6)$$

é possível reescrever a Equação (3.8.5) como

$$Var(\bar{y}) = \left(\frac{\epsilon}{z}\right)^2, \quad (3.8.7)$$

caso se deseje usar a margem erro ϵ e o nível de confiança γ da mesma forma que a Equação (3.8.1). Com esses valores é possível obter o tamanho amostral n , comparável à AAS, pela relação $n = am$. Utilizando o *Deff*, o valor de m também pode ser determinado por:

$$m = \sqrt{\frac{1 - \rho c_1}{\rho c_2}}, \quad \rho = \frac{Deff - 1}{M - 1}. \quad (3.8.8)$$

Capítulo 4

Materiais e Métodos

4.1 Introdução

Este Capítulo tem como objetivo principal detalhar os esquemas amostrais adotados, os materiais disponibilizados e o desenvolvimento do plano de amostragem empregados na pesquisa. Inclui uma descrição do banco de dados e a nomenclatura a serem empregados no estudo, além do detalhamento e estruturação dos planos amostrais propostos.

4.2 Materiais

O banco de dados com todos os livros foi disponibilizado pela BCE primeiramente em 28 de setembro de 2023 e atualizado em 23 de abril de 2024. O *status* “normal” disponibilizado exclui todas as demais situações possíveis em que um exemplar poderia se encontrar, tais como: excluído, aguardando baixa, baixa ocioso, baixa antieconômico, baixa irrecuperável, retido, em restauração, e provisório, indicando que a obra não apresenta nenhuma pendência administrativa interna, possibilitando, assim, que os livros estejam disponíveis para consulta e empréstimo nas estantes, ou ainda em posse de alunos.

Os exemplares apresentados no banco de dados referem-se exclusivamente aos tipos de material “livros”, “folhetos”, “teses” e “dissertações”. Além disso, esses itens são pertencentes apenas à biblioteca “Biblioteca Central” e estão fisicamente localizados no “Acervo Geral”.

Diversas características estão listadas no banco de dados, porém, para esse estudo, as seguintes variáveis foram selecionadas:

Tabela 4.1: Variáveis utilizadas do banco de dados da BCE

Variável	Descrição
Código do exemplar	Código único de referência do livro, composto por 8 dígitos numéricos
Nome modo de aquisição	Tipo de aquisição: Assinatura, Compra, Convênio, Doação, Reposição
Data de aquisição do exemplar	Data em que o exemplar foi adquirido
Título	Título do exemplar
Página	Número de páginas do exemplar
Classificação	Classificação referente ao CDU
Ano	Ano de publicação do livro
Nome área de conhecimento	Nome da área referente à classificação
Código área de conhecimento	Classificação da área referente ao CDU

4.3 Métodos

Primeiramente, é crucial estabelecer uma padronização dos conceitos a fim de prevenir erros e inconsistências durante a coleta de dados. Nesse contexto, serão adotadas as seguintes orientações:

Estante: é uma estrutura composta por módulos alinhados horizontalmente da esquerda para a direita, onde apenas um dos lados está acessível para armazenamento ou retirada de exemplares.

Módulo: é um bloco de prateleiras com delimitações laterais, onde os livros são armazenados e ordenados pela classificação de forma crescente, da esquerda para a direita e de cima para baixo.

Prateleira: é uma fileira horizontal delimitada pelo módulo onde residem os livros.

Área Temática: são as “subclasses”, temas definidos nas descrições de cada classe.

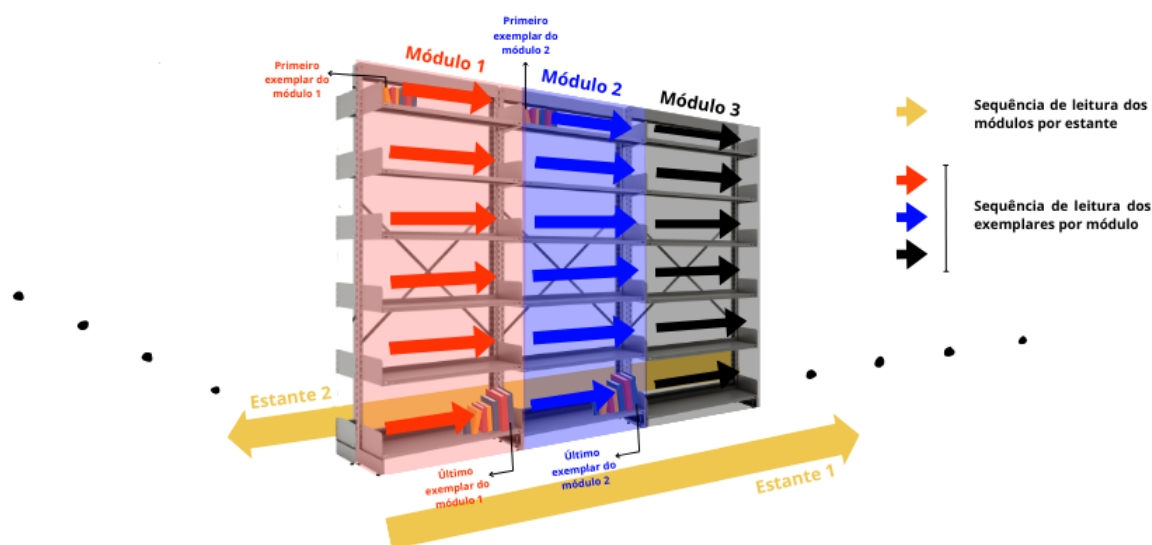


Figura 4.1: Ordenação do objeto de estudo

As estantes organizam os módulos, os quais, por sua vez, ordenam os exemplares em ordem crescente de classificação. A ordenação ocorre da esquerda para a direita e de cima para baixo dentro de cada módulo, seguindo então para o próximo módulo à direita, conforme esquema apresentado na Figura 4.1. A sequência de numeração dos livros da BCE (Tabela 4.2) possui algumas peculiaridades e deve-se tomar cuidado ao utilizar a ordenação padrão das linguagens de programação, que costumam seguir o sistema ASCII (*American Standard Code for Information Interchange*).

Tabela 4.2: Ordenação utilizada na BCE

espaço : () , - . / 0 1 2 3 4 5 6 7 8 9 =																									
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Ao ter acesso ao banco de dados, é possível implementar dois tipos de planos amostrais. Dessa forma, considerando a natureza do levantamento e da distribuição do objeto de estudo, a metodologia levará em conta as classes dos exemplares como subpopulações, o que significa que cada plano proposto será realizado separadamente para cada classe. A seguir, serão apresentados diferentes métodos concorrentes para a definição da proposta mais adequada.

Plano Amostral 1 - Amostragem Aleatória Simples

Neste plano, a amostragem é realizada por meio de uma amostra aleatória simples dos livros, utilizando o banco de dados disponível. Antes da coleta, a amostra será organizada por classe. Assim, os passos para esse método são:

1. Escolher a classe a ser trabalhada;
2. Realizar o levantamento dos dados da classe, por meio do banco de dados atualizado, disponibilizado pela BCE (deve ser requisitado periodicamente, para diminuir possíveis erros de contagem);
3. Estabelecer a margem de erro ϵ e o nível de confiança γ , além dos custos da mão de obra c_1 e c_2 , caso pertinente;
4. Determinar, se possível, os estratos da classe, por meio da separação dos temas que estão descritos na Tabela 2.1. Se não for possível, ignorar este passo;
5. Calcular o tamanho do estrato (caso haja);
6. Sortear a amostra por meio do banco de dados;
7. Recolher a amostra, anotando as observações das avarias descritas, conforme Figura 2.1;
8. Calcular a proporção de livros com necessidade de restauração por meio das equações apresentadas na Tabela 3.1 e calcular o intervalo de confiança.

Plano Amostral 2 - Conglomerado

O método de amostragem adotado é o conglomerado em dois estágios. No primeiro estágio, os módulos são selecionados com uma probabilidade proporcional ao número de módulos da classe, por meio da Equação (3.5.6). No segundo estágio, os livros são selecionados por meio de uma amostra sistemática ou aleatória simples dentro de cada módulo sorteado. Dessa forma, deve-se seguir os seguintes passos:

1. Seguir os passos 1 a 4 do Plano Amostral 1;
5. Calcular o tamanho do estrato (caso haja);
6. Contar a quantidade de módulos e registrar o código do primeiro e do último livro de cada módulo;
7. Calcular a quantidade a de módulos a serem amostrados e a quantidade de livros m a serem verificados por módulo;
8. Sortear a amostra de módulos;
9. Realizar a contagem de livros dos módulos selecionados na amostra ($M_i, i = 1, \dots, a$);

10. Calcular o valor de $k_i = M_i/m$ (intervalo entre os livros selecionados de cada módulo);
11. Sortear um valor aleatório (*seed*) entre 0 e k_i , podendo ser um valor contínuo;
12. Calcular as posições dos livros (Posic) por meio da equação $Posic_j = INT(seed + (j - 1)k_i)$, $j = 1, \dots, m$, onde $INT()$ é a parte inteira do argumento, de forma que:
$$Posic_1 = INT(seed)$$
$$Posic_2 = INT(seed + k_i)$$
$$\vdots$$
$$Posic_m = INT(seed + (m - 1)k_i)$$
Se $0 \leq Posic_i \leq 1$, então faça $Posic_i = 1$.
13. Recolher a amostra, anotando os dados do exemplar e as observações das avarias descritas, conforme Figura 2.1;
14. Calcular a proporção de livros com necessidade de restauração por meio das equações apresentadas na Tabela 3.1 e calcular o intervalo de confiança.

As características coletadas serão as avarias descritas na Seção 2.4, além do indicador de livro restaurado. Adicionalmente, será mensurado o tempo de coleta da amostra para comparação entre os planos. Será considerado um livro com necessidade de restauração aquele que possuir pelo menos uma das avarias avaliadas.

Capítulo 5

Estudo de Caso

5.1 Introdução

O objetivo deste Capítulo é aplicar o método descrito no Capítulo 4, com o intuito de realizar comparações entre os resultados alcançados, a fim de auxiliar na determinação do melhor método para esta pesquisa. Inicialmente, será realizada a análise exploratória dos dados para maior compreensão dos mesmos e, em seguida, a aplicação de cada um dos planos amostrais. Para as análises e implementações computacionais, será utilizado o *software* SAS 9.4.

5.2 Análise Exploratória

A análise exploratória de dados é fundamental para a elaboração de um plano amostral, pois permite visualizar a distribuição e a variabilidade dos dados. Essa compreensão é essencial para a seleção de amostras representativas, garantindo que as inferências realizadas sejam válidas e confiáveis, gerando estimativas precisas. Para esse fim, foram analisadas características relevantes das classes que poderiam influenciar significativamente a escolha do método de amostragem, como a distribuição da frequência por classe.

Como mostra a Figura 5.1, há uma variação considerável na quantidade de livros entre as classes. A classe 3 detém o maior número de exemplares, ao passo que a classe 2 apresenta o menor acervo. Devido à variação na frequência observada e à diversidade no perfil dos usuários de cada classe, cada classe foi considerada como uma subpopulação. Embora a estratificação fosse uma opção viável, decidiu-se tratar cada classe como uma subpopulação com o objetivo de obter informações mais precisas para cada uma.

Devido ao tamanho exacerbado da classe 3 e sua extensão temática, foi levantada

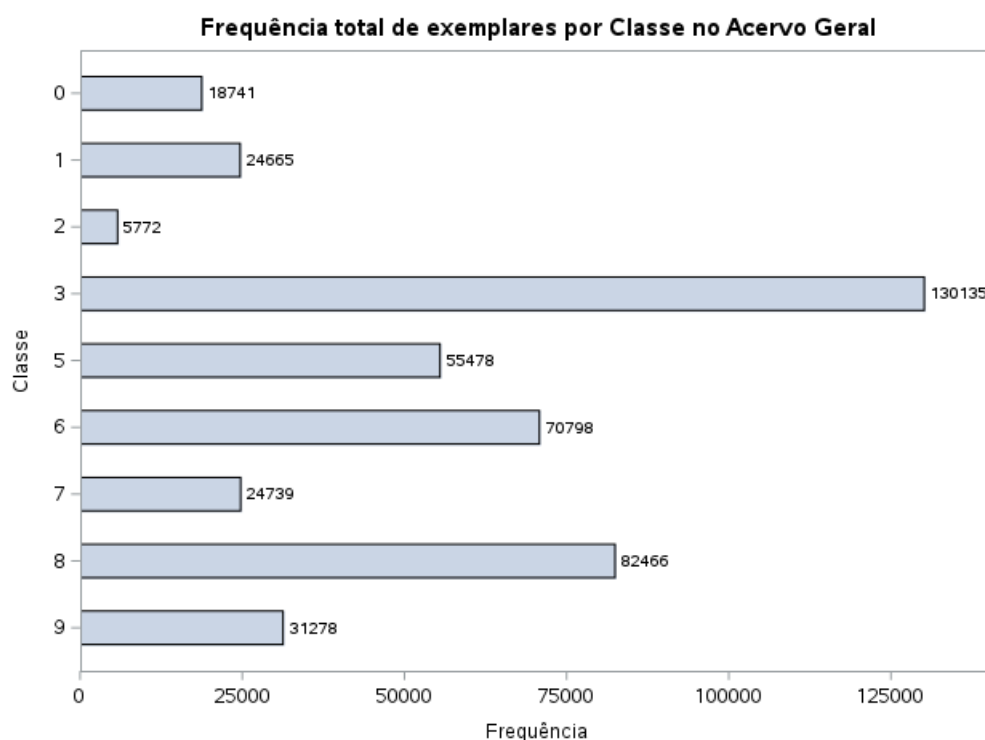


Figura 5.1: Frequência total de exemplares por classe no Acervo Geral

a hipótese de separar internamente as classes em estratos, utilizando a subclassificação de dois dígitos da CDU (ver Apêndice A). Dessa forma, o título da classe seria decomposto em seus temas correspondentes, cada um dos quais originando um estrato. Contudo, ao examinar as subclassificações, torna-se aparente a ocorrência de sobreposições temáticas, bem como uma descontinuidade na ordenação. Isso impede a detecção precisa da localização física do estrato. Como exemplo, toma-se a classe 1: Filosofia e Psicologia. Ao selecionar os dois temas, observa-se que o código 15, correspondente à psicologia, está intercalado entre os códigos de filosofia. Além disso, o código 13 (Filosofia da mente e do espírito) representa uma intersecção entre os dois tópicos.

Outra característica observada foi a origem do exemplar, ou seja, o tipo de aquisição. A Figura 5.2 mostra que a maior parcela dos livros foi adquirida por meio da compra e da doação. Esse alto volume de aquisição por doação levanta a hipótese de existirem muitos livros com necessidade de restauração caso estes não tenham passado por uma rigorosa avaliação antes de serem aceitos.

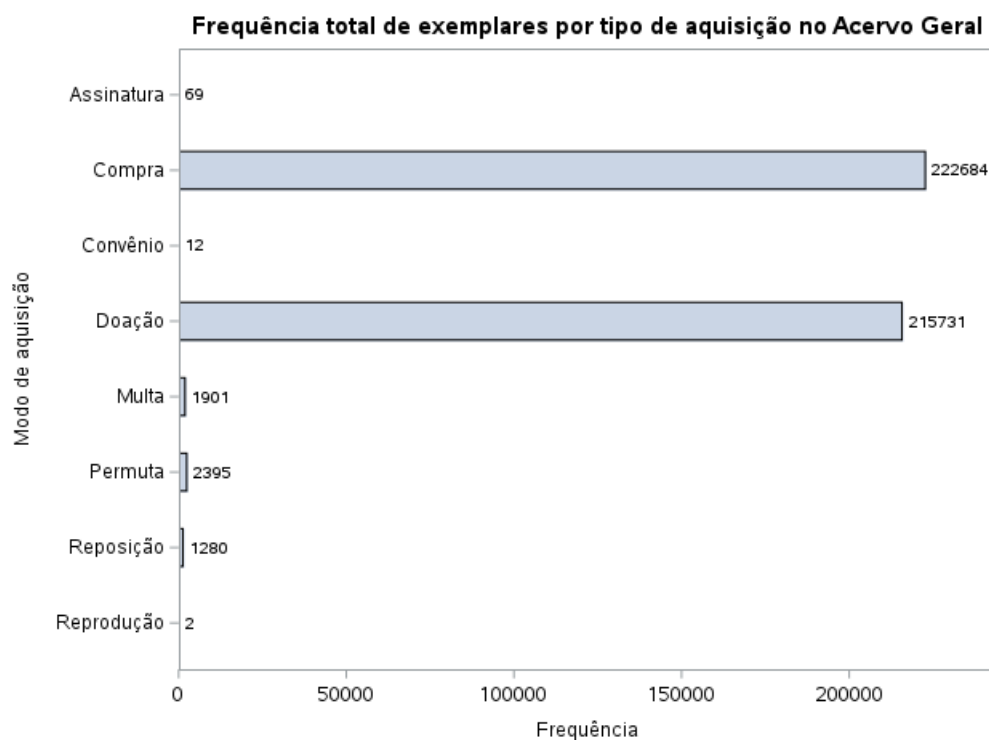


Figura 5.2: Frequência total de exemplares por tipo de aquisição no Acervo Geral

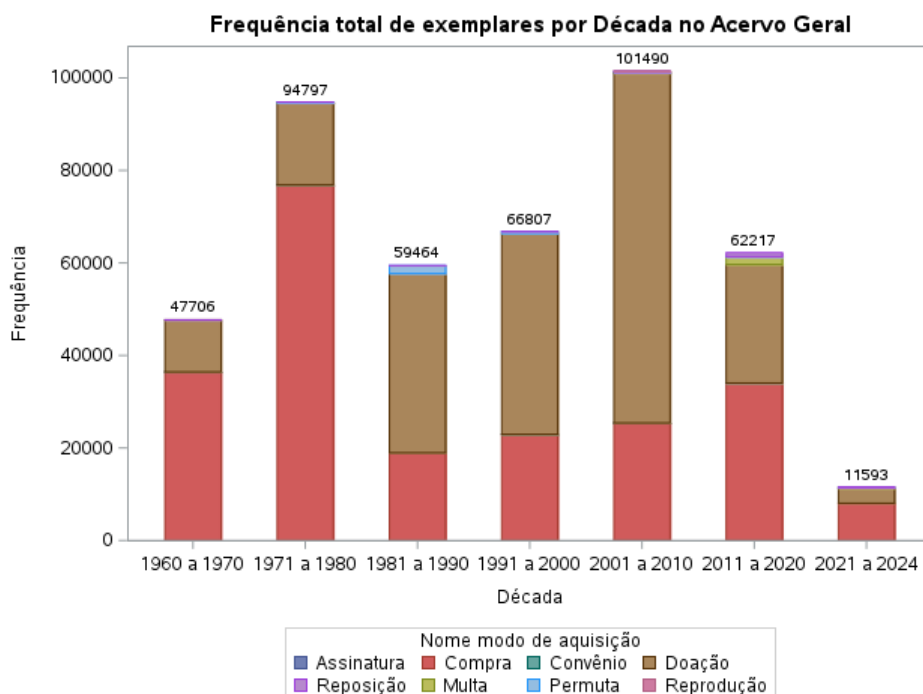


Figura 5.3: Frequência total de exemplares por década e tipo de aquisição no Acervo Geral

A distribuição dos exemplares em relação ao tipo de aquisição pode estar associada ao período de obtenção. Por meio da Figura 5.3, dois picos de aquisição são perceptíveis: o período de 1971-1980 e de 2001-2010. Apesar dos volumes semelhantes,

suas origens são opostas. Na década de 70, os livros foram em sua maioria adquiridos por compra para expandir o acervo, que era novo. Já nos anos 2000, a maioria veio de doações que foram motivadas por campanhas de arrecadação de livros. Também é interessante notar que, apesar da diminuição de aquisições por compra a partir de 1991, estas voltam a aumentar e até ultrapassam a quantidade de doações nas décadas seguintes.

5.3 Amostra Aleatória Simples

A classe 2 foi selecionada para a implementação do método proposto, em virtude de possuir o menor número de exemplares, o que facilita a realização de estudos comparativos com a população total. Conforme o método descrito no Capítulo anterior, após a seleção da classe, e da não possibilidade de uso de estratos, o próximo passo é o cálculo do tamanho da amostra.

O tamanho da amostra foi determinado usando a Equação (3.8.1). Por meio de outros levantamentos realizados anteriormente, foi possível obter valores iniciais para a proporção de livros avariados, sendo de 55% para a classe 2. Devido a restrições de tempo e acesso limitado ao acervo devido à greve dos técnicos, a análise não tem a pretensão de ser precisa, mas sim de ilustrar o método. Assim, com um nível de confiança de $\gamma = 0,95$ e uma margem de erro de 15%, o tamanho da amostra calculado foi de 43 livros, conforme abaixo:

$$n = \frac{1,96^2 \times 0,55 \times (1 - 0,55)}{0,15^2} = 42,2576 \approx 43$$

A amostra sorteada do banco de dados foi obtida por meio da *PROC SURVEY-SELECT* do *software* SAS 9.4. Durante a coleta de dados, um dos livros selecionados não pôde ser encontrado, deixando a amostra final com 42 livros. Isso se deve ao fato de que o acervo tem uma constante movimentação de livros, mas, devido à greve, estava consideravelmente estável. Foram observados 13 livros com pelo menos uma avaria, obtendo os resultados descritos na Tabela 5.1, com a estimativa pontual de $\hat{p} = 0,3095$. Ou seja, a AAS estimou a proporção em 30,95% de livros com necessidade de restauração na classe 2, com uma margem de erro de $(0,4553 - 0,1637)/2 = 0,1458$ (14,58%), abaixo dos 15% pré-determinado.

Tabela 5.1: Resumo dos resultados da estimação para a AAS

Estimador	Estimativa	Erro Padrão	CV	Limite inferior	Limite superior
\hat{p}	0,3095	0,0722	0,2333	0,1637	0,4553

Ao sortear o exemplar por meio do banco de dados, a amostra permite verificar

suas características por meio do cruzamento de dados. A Figura 5.4 exibe a relação entre os livros que necessitam de restauração na AAS e a década de aquisição, em que a variável “restauração” recebe 0 quando um livro não tem necessidade de restauração e 1 quando o livro necessita de restauração.

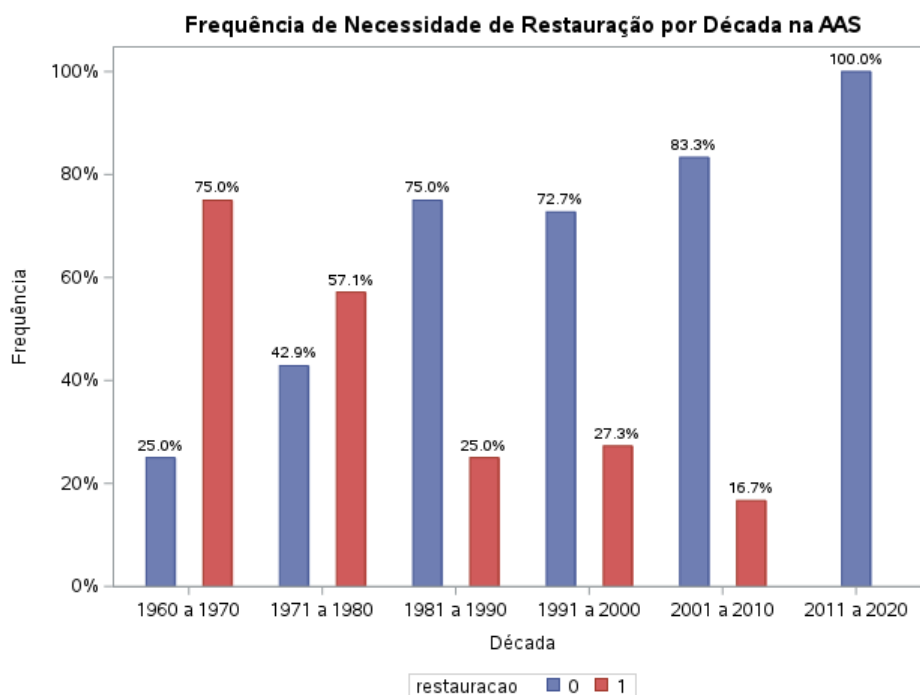


Figura 5.4: Frequência de exemplares com necessidade de restauração por década de aquisição na AAS

A amostra claramente demonstra que os livros com mais tempo na coleção são mais propensos a necessitar de restauração do que os livros mais novos, o que é logicamente plausível. O tempo de coleta para este plano amostral foi de 1 hora e 11 minutos, considerando o deslocamento, localização do exemplar e a verificação e anotação das avarias. Ou seja, em média, gastou-se 1 minuto e 42 segundos por livro na classe 3 para este plano amostral, conforme o cálculo abaixo:

$$\text{Tempo por livro AAS} = \frac{71}{42} = 1,6905 \approx 1\text{min}42\text{seg.}$$

5.4 Amostra por Conglomerados em Dois Estágios

Devido aos problemas de localização e demora no procedimento da AAS, decidiu-se avaliar a amostragem por conglomerados em dois estágios (AC2), com seleção sistemática. O objetivo da amostragem por conglomerados é diminuir os custos, que, neste caso, é o tempo. Aliada a seleção sistemática, acredita-se que o tempo de localização

dos exemplares será reduzido, além da possibilidade de conseguir analisar todos os livros selecionados na amostra. Em virtude da ausência de informação sobre a quantidade de livros por módulo no banco de dados, necessária para a amostragem por conglomerados, foi feito um levantamento completo da classe 2. Alternativamente, tentou-se também estimar a quantidade de livros em cada módulo utilizando o banco de dados, registrando as classificações inicial (do primeiro livro) e final (do último livro) de cada módulo. Em seguida, os livros de cada módulo foram contados fisicamente e comparados com os livros ordenados no banco de dados entre as observações com classificações inicial e final coletadas.

Como mencionado na Seção 4.3, a ordenação dos livros na BCE não segue a ordem comumente utilizada nas funções de ordenação dos *softwares* e deve-se tomar cuidado ao tentar ordenar os itens automaticamente. Com isso, foi necessário realizar modificações na ordenação considerada nos procedimentos do SAS 9.4. Estas modificações foram feitas por meio da *PROC TRANTAB*, procedimento para criar, editar, ou visualizar tabelas de tradução (SAS, 2024). Assim, foi possível adaptar a Tabela 5.2, referente ao ordenamento utilizado no SAS 9.4, para obter a ordenação descrita na Tabela 4.2.

Tabela 5.2: Ordenação ASCII utilizada no SAS

blank ! " # \$ % & ' () * + , - . / 0 1 2 3 4 5 6 7 8 9 : ; < = > ? @
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z [\] ^ _
a b c d e f g h i j k l m n o p q r s t u v w x y z { } ~

Fonte: SAS (2024).

Foram observados 6.151 livros na classe 2 distribuídos em 36 módulos, com média de 170,86 livros em cada módulo. A comparação entre a contagem física e a quantidade estimada pelo banco de dados revelou uma correlação linear alta (Figura 5.5), de mais de 75%, e excluindo dois valores fora da reta, de mais de 90% (Tabela 5.3).

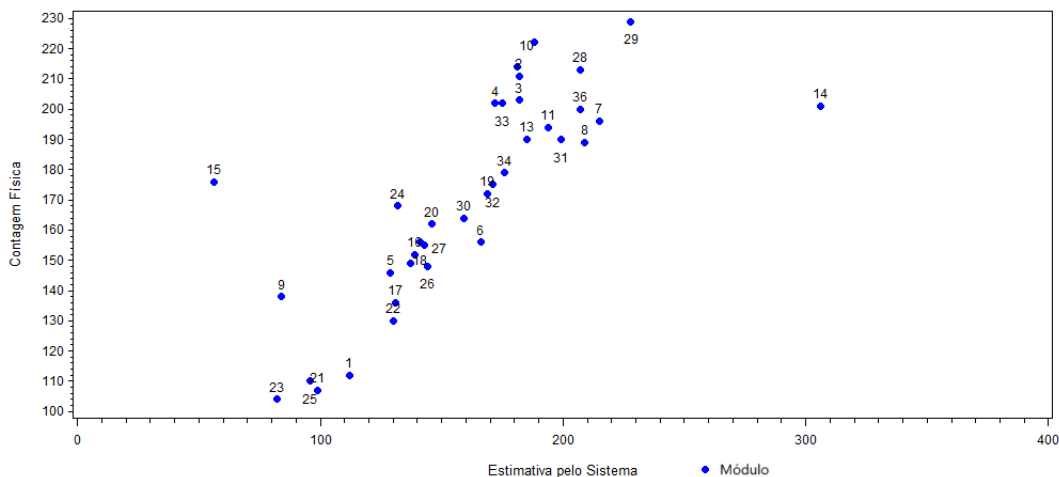


Figura 5.5: Gráfico de correlação entre a contagem física e a estimativa pelo banco de dados na classe 2

Isso significa que o banco de dados pode ser utilizado para estimar com certa precisão o número de exemplares por módulo, podendo ser realizada apenas a coleta da classificação inicial e final e, assim, diminuir os custos por tempo. No entanto, o maior problema é que essa contagem deve ser censitária, ou seja, deve-se tentar estimar a quantidade de livros em todos os módulos para que seja feito o sorteio dos módulos por Probabilidade Proporcional ao Tamanho (PPT).

Tabela 5.3: Tabela de correlação entre a contagem física e a estimada pelo banco de dados na classe 2

Variáveis	Correlação Pearson
Contagem Física × Estimativa	0,76227
Contagem Física × Estimativa (excluindo módulos 14 e 15)	0,90675

Poder-se-ia calcular o tamanho amostral por meio das Equações (3.8.3) e (3.8.4). Porém, por causa da comparabilidade entre os resultados da AAS e da AC2, foi utilizado o tamanho amostral da AAS de 42 livros. Considerando a equivalência $n = am$ mencionada na Seção 3.8, optou-se por realizar as duas possíveis variações de uma amostra com $n = 42$ exemplares: uma amostra com 42 livros distribuídos em 6 módulos, resultando em 7 livros por módulo (AC2|6 × 7) e outra de 3 módulos com 14 livros em cada (AC2|3 × 14).

As Tabelas 5.4 e 5.5 mostram os módulos sorteados de cada amostra, o valor de k calculado (segundo Seção 3.4), a ser utilizado na seleção sistemática, e a quantidade de livros avariados sorteados por módulo. Já as Tabelas 5.6 e 5.7 representam os resultados obtidos de cada amostra. Vale ressaltar que as estimativas dentro dos conglomerados foram calculadas como se fossem uma AAS, uma vez que pode-se assumir que os livros estão aleatoriamente distribuídos nos módulos e nas prateleiras, não existindo correlação com as avarias investigadas, tornando o plano sistemático equivalente ao aleatório simples.

Tabela 5.4: Módulos sorteados na AC2|6 × 7

Módulo	Contagem física	k_i	Livros avariados
1	112	16	2
3	203	29	2
14	201	28,7143	0
22	130	18,5714	3
28	207	29,5714	0
32	175	25	1

Tabela 5.5: Módulos sorteados na AC2|3 × 14

Módulo	Contagem física	k_i	Livros avariados
10	222	15,8571	4
14	201	14,3571	3
28	207	14,7857	5

A AC2|6 × 7 estimou a proporção em 19,05% de livros com necessidade de restauração na classe 2, com uma margem de erro de $(0,3720 - 0,0089)/2 = 0,1816$ (18,16%), enquanto que a AC2|3 × 14 estimou a proporção em 28,57% de livros com necessidade de restauração, com uma margem de erro de $(0,4631 - 0,1083)/2 = 0,1774$ (17,74%).

Tabela 5.6: Resumo dos resultados da estimação para a AC2|6 × 7

Estimador	Estimativa	Erro Padrão	CV	Limite Inferior	Limite Superior
\hat{p}	0,1905	0,070630	0,3708	0,0089	0,3720

Tabela 5.7: Resumo dos resultados da estimação para a AC2|3 × 14

Estimador	Estimativa	Erro Padrão	CV	Limite Inferior	Limite Superior
\hat{p}	0,2857	0,0412	0,1443	0,1083	0,4631

Como as estimativas são diferentes, não se pode comparar as variâncias diretamente, portanto, a comparação é feita pelo Coeficiente de Variação (CV). Observa-se que o CV da AC2|3 × 14 é aproximadamente 2,6 vezes menor do que o da AC2|6 × 7 e 1,6 vezes menor do que o da AAS. Apesar de possuírem o CV menor que a AAS, suas margens de erro ficaram maiores e até ultrapassaram ligeiramente o limite de 15%. Isso ocorre devido aos graus de liberdade usados na Equação (3.7.4). No caso da AAS, isso faz com que a distribuição *t-Student* se aproxime de uma Normal ($t \approx 1,96$ para IC de 95%) enquanto que para a AC, os graus de liberdade são menores, resultando em quantis de t maiores.

Também é interessante observar os $Deff$ de cada plano (Tabela 5.8), que mostram que a AC2|6 × 7 possui maior variabilidade em relação à AAS, diferentemente da AC2|3 × 14, que possui menor variabilidade em relação à AAS.

Tabela 5.8: $Deff$ s da AC2|6 × 7 e AC2|3 × 14

Amostra	$Deff$
AC2 6 × 7	1,3265
AC2 3 × 14	0,3417

Já era esperado que menos conglomerados e mais livros analisados em cada conglomerado resultasse em estimativas melhores, pois a natureza da distribuição dos livros nos módulos os torna homogêneos entre si, resultado demonstrado pelas variâncias descritas na Tabela 5.9.

Tabela 5.9: Variâncias específicas da AC2|6 × 7 e AC2|3 × 14

Amostra	S_1^2	S_2^2
AC2 6 × 7	0,0299	0,1508
AC2 3 × 14	0,0051	0,2161

O tempo médio de coleta foi de 30 minutos para a amostra AC2|3 × 14, levando em consideração que cada módulo leva em torno de 11 minutos para ser amostrado e a verificação das avarias de um livro demanda em média 49 segundos (mais 42 segundos para anotar o código do livro, totalizando 91 segundos ou 1,5167 minutos). Assim, considerando os mesmos parâmetros da AAS para a classe 2, e considerando $\bar{M} = 170,86$ além dos custos calculados de $c_1 = 11,7167$ e $c_2 = 1,5167$ em minutos, é possível calcular o tamanho amostral por meio das Equações (3.8.3) e (3.8.4).

A vantagem de possuir uma coleta mais rápida é a diminuição dos riscos de movimentação do acervo durante o processo. Possivelmente, para classes muito grandes, o tempo gasto na preparação da pesquisa seja superior ao necessário para o levantamento via AAS. No entanto, a AAS necessita que o banco de dados seja constantemente atualizado.

5.5 Custos Estimados

Para ter uma ideia básica do custo monetário para a pesquisa, sem considerar os custos de planejamento, consultou-se alunos de extensão para determinar o valor da mão de obra. O custo de anotar as classificações iniciais e finais, por módulo, foi definido em R\$3,00 reais. Já a verificação dos livros para identificar avarias foi estabelecida em R\$4,00 por livro. Considerando que o Acervo Geral da BCE atualmente possui 444.072 exem-

plares nas condições estabelecidas, levaria cerca de 252 dias ininterruptos para verificar todos os livros e suas avarias, conforme o cálculo abaixo:

$$\text{Duração do censo em dias} = \frac{444.072 \times 49}{60 \times 60 \times 24} = 251,85 \approx 252 \quad (5.5.1)$$

considerando o tempo de 49 segundos estimado para verificar a avaria, por livro, encontrado na AC2. Caso o tempo de análise seja de 8 horas diárias, então o censo duraria 756 dias ou pouco mais de 2 anos. O custo para este levantamento poderia ser estimado em 1,77 milhão de reais, calculado como:

$$\text{Custo do censo} = 444.072 \times 4 = 1.776.288 \quad (5.5.2)$$

Em relação à pesquisa por AC2, considerando a classe 3, por exemplo, que é a maior classe e possui 767 módulos, e considerando o tempo de 47 segundos para realizar o levantamento da classificação inicial e final de um módulo, o custo e o tempo necessários para essa etapa da AC2, seriam calculados como:

$$\text{Custo módulos classe 3} = 767 \times 3 = 2.301 \quad (5.5.3)$$

$$\text{Tempo módulos classe 3} = \frac{767 \times 47}{60 \times 60} = 10,014 \quad (5.5.4)$$

obtendo o valor de aproximadamente R\$2.301,00 reais para a anotação de todas as classificações iniciais e finais, e levando um total de 10,014 horas sem interrupção para concluir a tarefa. Para o levantamento de uma amostra na AC2, considerando as variâncias estimadas para o caso AC2|3 × 14 (Tabela 5.9), $\gamma = 0,95$, e uma margem de erro de 8%, $\epsilon = 0,08$, o tamanho amostral seria de:

$$m = \sqrt{\frac{0,2161}{0,0051 - \frac{0,2161}{170}} \frac{11,7167}{1,5167}} \approx 21 \quad (5.5.5)$$

$$a = \frac{0,0051 - \frac{0,2161}{170} + \frac{0,2161}{21}}{\left(\frac{0,08}{1,96}\right)^2 + \frac{0,0051}{767}} \approx 8, \quad (5.5.6)$$

ou seja, 8 módulos e 21 livros por módulo, e gastaria-se em torno de R\$672,00 reais para a verificação das avarias em 4,247 horas seguidas de trabalho, conforme cálculos abaixo:

$$\text{Custo de verificação classe 3 (AC2)} = 4 \times 8 \times 21 = 672 \quad (5.5.7)$$

$$\text{Tempo de verificação classe 3 (AC2)} = \frac{8 \times 21 \times 1,5167}{60} = 4,247 \quad (5.5.8)$$

O custo final para a classe 3 por AC2 ficaria por volta de $R\$2.301,00 + R\$672,00 = R\$2.973,00$ reais e levaria $10,014 + 4,261 = 14,26$ horas para ser concluído. Assim, considerando 8 horas diárias, levaria 1,75 dias de trabalho para realizar o levantamento.

Para a AAS, utilizando o mesmo tamanho de amostra $n = am = 8 \times 21 = 168$ livros, o custo seria apenas da verificação das avarias nos livros, estabelecido em $168 \times R\$4,00 = R\$672,00$ reais e o tempo seria calculado como função do deslocamento, localização e anotação das avarias. Como a classe 3 possui mais estantes, logo, mais corredores para percorrer (conforme a Figura 2.12), foi estabelecida uma relação linear entre o tempo total, quantidade de módulos e quantidade de livros. Se na classe 2, com 36 módulos, foram necessários 71 minutos para se analisar 42 livros, então temos que:

$$\text{Tempo de verificação classe 3 (AAS) por livro} = \frac{767 \times 71}{36 \times 42} = 36,01 \text{min} \quad (5.5.9)$$

$$\text{Tempo de verificação classe 3 (AAS)} = \frac{8 \times 21 \times 36,01}{60} = 100,828 \text{h}, \quad (5.5.10)$$

logo, o levantamento AAS levaria em torno de 101 horas. Considerando o tempo de análise de 8 horas diárias, esse levantamento demandaria 12,6 dias de trabalho. Para realizar o plano AAS no mesmo tempo da AC2, seria necessária a contratação de 24 colaboradores, conforme o cálculo abaixo:

$$\text{Número de colaboradores} = \frac{100,828}{4,247} \approx 24 \quad (5.5.11)$$

Assim, apesar do custo monetário da AAS ser menor, demanda em torno de 4,4 vezes mais tempo do que a AC2. O custo da AC2 pode ser ainda menor ao suprimir a etapa de levantamento do número de livros por módulo, que é responsável por aproximadamente 77,4% do custo monetário e 70,2% do custo de tempo total. Isso é possível porque, mesmo que se possa selecionar as UPA's por PPT, uma vez que no segundo estágio o número de livros é fixo (m) para todos os módulos a fim de facilitar o processo de amostragem, a utilização de um valor médio para M_i (\bar{M}) ainda mantém o plano autoponderado, conforme a Equação (3.5.6). A utilização de \bar{M} é possível porque, ao observar o Acervo Geral, os módulos possuem baixa variabilidade no quantitativo de livros.

Capítulo 6

Conclusão

Este trabalho teve como objetivo desenvolver um plano de amostragem para estimar a proporção de livros com necessidade de restauração no Acervo Geral da Biblioteca Central da Universidade de Brasília. A partir do mapeamento da área de estudo e do levantamento dos dados referentes aos exemplares da BCE, foi possível estudar e comparar duas técnicas de amostragem: amostragem aleatória simples e amostragem por conglomerado em dois estágios.

A amostragem aleatória simples, apesar de parecer uma escolha elementar, apresentou algumas dificuldades no cenário observado: a seleção de uma amostra por meio do banco de dados pode trazer exemplares que não estão fisicamente disponíveis, como foi o caso de um livro da amostra do estudo desse trabalho. Além disso, o tempo de coleta por livro é relativamente alto: deve-se localizá-lo na estante por seu código e nem sempre os livros estarão perfeitamente ordenados, além do tempo de deslocamento ser considerável, ficando mais evidente em classes com muitas estantes, como a classe 3. E por último, existe a necessidade do banco de dados estar constantemente atualizado. Um ponto positivo é que suas estimativas são facilmente calculadas e seu custo de planejamento é menor.

A amostragem por conglomerados em dois estágios é mais complexa em suas estimativas e demanda conhecimentos mais aprofundados dos métodos utilizados. A coleta dos dados, via amostragem sistemática, é mais simples que a da AAS e mostrou-se mais eficiente, principalmente por não precisar localizar livros, apenas contar suas posições. Além disso, o tempo de coleta é muito menor do que o método AAS, diminuindo problemas com a movimentação do acervo e evitando o problema de exemplares ausentes.

A proporção de livros com necessidade de restauração na classe estudada variou conforme a mudança do plano amostral: 30% para a AAS, 19% para a AC2|6 × 7 e 28,5%

para a AC2|3 × 14. O coeficiente de variação mostrou também valores distintos, mesmo para as estimativas próximas: 0,23 para AAS, 0,37 para AC2|6 × 7 e 0,14 para AC2|3 × 14.

Por fim, ao observar os resultados de ambos os planos amostrais, é possível concluir que a AC2 se mostrou superior em vários aspectos: eliminação da possibilidade de exemplares ausentes, menor tempo da coleta e menor CV. Porém, com a margem de erro maior do que a AAS (14,58% para AAS, 19,05% para AC2|6 × 7 e 18,16% para AC2|3 × 14). Apesar disso, é possível concluir que, dentre os planos amostrais sugeridos, levando em consideração os custos e a precisão, a amostragem por conglomerados em dois estágios seria o planejamento mais indicado.

Referências Bibliográficas

- Antunes, M. A. (2010). *Pequenos Reparos em material bibliográfico*. Secretaria de Estado da Cultura de São Paulo.
- BCE (2022). Biblioteca central da universidade de Brasília. relatório de atividades 2022. Technical report, Universidade de Brasília. Disponível em: URL <https://repositorioadm.bce.unb.br/relatorios/relatorio-de-atividades-2022/>. Acesso em 13 out. 2023.
- BCE (2023). Biblioteca central da universidade de Brasília. história da BCE. Disponível em: URL <https://bce.unb.br/sobre-a-bce/historia-da-bce/>. Acesso em 04 out. 2023.
- Bolfarine, H. & Bussab, W. O. (2005). *Elementos de Amostragem*. Editora Blucher.
- Brasil (1962). Decreto nº 1.872, de 12 de dezembro de 1962. *Diário Oficial da República Federativa do Brasil*.
- Bussab, W. & Morettin, P. (2017). *Estatística Básica*, (9 ed.). Saraiva.
- CEPLAN (2017). Projetos. Disponível em: URL http://ceplan.unb.br/index.php?option=com_phocadownload&view=category&id=3&Itemid=684&limitstart=20. Acesso em 27 nov. 2023.
- Cochran, W. G. (1977). *Sampling Techniques*, (3rd ed.). John Wiley & Sons.
- FID (1987). *Federação Internacional de Documentação. Classificação Decimal Universal: edição média em língua portuguesa*, (2 ed.). Instituto Brasileiro de Informação em Ciência e Tecnologia.
- Fonseca, E. N. (1973). Biblioteca Central da Universidade de Brasília: história com um pouco de doutrina e outro tanto de memórias. *Revista de Biblioteconomia de Brasília*, 1(1). Disponível em: URL <https://periodicos.unb.br/index.php/rbbsb/article/view/19870/36177>. Acesso em 04 out. 2023.

- Froner, Y.-A. & Souza, L. A. C. (2008). *Preservação de bens patrimoniais: conceitos e critérios*. Escola de Belas Artes UFMG.
- Kish, L. (1965). *Survey Sampling*. John Wiley & Sons.
- Lohr, S. L. (2010). *Sampling: design and analysis*, (2nd ed.). Cengage Learning.
- Paglione, C. Z. (2017). *Glossário Visual de Conservação: Um Guia de Danos Comuns em Papéis e Livros*. Biblioteca Brasileira Guita e José Mindlin.
- RES (2024). Setor de Conservação e Restauração, Biblioteca da Universidade de Brasília.
- SAS (2024). TRANTAB procedure. Disponível em: URL https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.5/n1sref/p140vpknze29j0n1rbfs5fcxemu3.htm. Acesso em 15 mai. 2024.

Apêndice A

Tabela da CDU para dois dígitos

- | | |
|---|--|
| 0 GENERALIDADES. CIÊNCIA E CONHECIMENTO. ORGANIZAÇÃO. INFORMAÇÃO; DOCUMENTAÇÃO. BIBLIOTECONOMIA. INSTITUIÇÕES | 08 E CONTROLE; GERÊNCIA COMERCIAL E EDITORIAL. JORNALISMO. POLIGRAFIAS. OBRAS EM COLABORAÇÃO. OBRAS DE AUTORIA COLETIVA |
| 00 PROLEGÔMENOS. FUNDAMENTOS DO CONHECIMENTO E DA CULTURA | 09 MANUSCRITOS. OBRAS RARAS E NOTÁVEIS. OBRAS COM CARACTERÍSTICAS EXTRÍNSECAS ESPECIAIS, TAIS COMO A ESCRITA OU TIPO, ENCADERNAÇÃO, ILUSTRAÇÃO E O MATERIAL EM QUE SÃO ESCRITAS OU IMPRESSAS |
| 01 BIBLIOGRAFIA E BIBLIOGRAFIAS. CATÁLOGOS | |
| 02 BIBLIOTECONOMIA | |
| 03 OBRAS GERAIS DE REFERÊNCIA (COMO ASSUNTO) | |
| 05 PUBLICAÇÕES PERIÓDICAS. PERIÓDICOS (COMO ASSUNTO) | |
| 06 ORGANIZAÇÕES E OUTRAS FORMAS DE COOPERAÇÃO. ASSOCIAÇÕES. CONGRESSOS. EXPOSIÇÕES. MUSEUS | |
| 07 JORNAIS, IMPRENSA: SIGNIFICADO, PROPRIEDADE | |
| | 1 FILOSOFIA. PSICOLOGIA |
| | 11 METAFÍSICA |
| | 12 METAFÍSICA ESPECIAL |
| | 13 FILOSOFIA DA MENTE E DO ESPÍRITO. METAFÍSICA DA VIDA ESPIRITUAL |
| | 14 SISTEMAS E PONTOS DE VISTA FILOSÓFICOS. |

- 15 PSICOLOGIA
- 16 LÓGICA. EPISTEMOLOGIA. TEORIA DO CONHECIMENTO. METODOLOGIA DA LÓGICA
- 17 FILOSOFIA MORAL. ÉTICA. FILOSOFIA PRÁTICA
- 2 RELIGIÃO. TEOLOGIA**
- 21 TEOLOGIA NATURAL. TEO-DICÉIA. DE DEO. TEOLOGIA RACIONAL. FILOSOFIA RELIGIOSA
- 22 A BÍBLIA. SAGRADAS ESCRITURAS
- 23 TEOLOGIA DOGMÁTICA
- 24 TEOLOGIA PRÁTICA
- 25 TEOLOGIA PASTORAL
- 26 IGREJA CRISTÃ EM GERAL (NATUREZA E CARACTERÍSTICAS)
- 27 HISTÓRIA GERAL DA IGREJA CRISTÃ
- 28 IGREJAS CRISTÃS. SEITAS
- 29 RELIGIÕES NÃO CRISTÃS
- 3 CIÊNCIAS SOCIAIS: ESTATÍSTICA. POLÍTICA. ECONOMIA. COMÉRCIO. DIREITO. ADMINISTRAÇÃO PÚBLICA. FORÇAS ARMADAS. ASSISTÊNCIA SOCIAL. SEGURO. EDUCAÇÃO. FOLCLORE**
- 30 TEORIAS, METODOLOGIA E MÉTODOS NAS CIÊNCIAS SOCIAIS EM GERAL. SOCIOGRAFIA
- 31 DEMOGRAFIA. SOCIOLOGIA. ESTATÍSTICA
- 32 POLÍTICA
- 33 ECONOMIA. CIÊNCIA ECONÔMICA
- 34 DIREITO. JURISPRUDÊNCIA
- 35 ADMINISTRAÇÃO PÚBLICA. GOVERNO
- 36 ASSISTÊNCIA SOCIAL. PREVIDÊNCIA SOCIAL. SEGURIDADE SOCIAL
- 37 EDUCAÇÃO. ENSINO. INSTRUÇÃO. LAZER
- 39 ANTROPOLOGIA CULTURAL. ETNOLOGRAFIA. USOS E COSTUMES. TRADIÇÕES. MODO DE VIDA. FOLCLORE
- 4 VAGA**
- 5 MATEMÁTICA E CIÊNCIAS NATURAIS**
- 50 GENERALIDADES SOBRE AS CIÊNCIAS PURAS
- 51 MATEMÁTICA
- 52 ASTRONOMIA. ASTROFÍSICA. PESQUISA ESPACIAL. GEODÉSIA
- 53 FÍSICA
- 54 QUÍMICA. CIÊNCIAS MINERALÓGICAS

- 55 CIÊNCIAS DA TERRA. GEOCIÊNCIA. GEOLOGIA. METEOROLOGIA ETC.
- 56 PALEONTOLOGIA
- 57 CIÊNCIAS BIOLÓGICAS EM GERAL
- 58 BOTÂNICA
- 59 ZOOLOGIA
- 6 CIÊNCIAS APLICADAS. MEDICINA. TECNOLOGIA**
- 60 BIOTECNOLOGIA
- 61 CIÊNCIAS MÉDICAS
- 62 ENGENHARIA. TECNOLOGIA EM GERAL
- 63 AGRICULTURA E CIÊNCIAS E TÉCNICAS AFINS. SILVICULTURA. AGRONOMIA. ZOOTECNIA. EXPLORAÇÃO DA VIDA SELVAGEM
- 64 ECONOMIA DOMÉSTICA. CIÊNCIA DOMÉSTICA
- 65 ORGANIZAÇÃO E ADMINISTRAÇÃO DA INDÚSTRIA, DO COMÉRCIO E DOS TRANSPORTES
- 66 TECNOLOGIA QUÍMICA. INDÚSTRIAS QUÍMICAS E AFINS
- 67 INDÚSTRIAS, ARTES INDUSTRIAIS E OFÍCIOS DIVERSOS
- 68 INDÚSTRIAS, ARTES E OFÍCIOS DE ARTIGAS ACABADOS OU MONTADOS
- 69 INDÚSTRIA DA CONSTRUÇÃO. MATERIAIS DE CONSTRUÇÃO. PRÁTICA E PROCESSOS DE CONSTRUÇÃO
- 7 ARTE. BELAS ARTES. RECREAÇÃO. DIVERSÕES. ESPORTES**
- 71 PLANEJAMENTO TERRITORIAL, FÍSICO. PLANEJAMENTO REGIONAL, URBANO E RURAL. PAISAGENS, PARQUES, JARDINS. PAISAGISMO
- 72 ARQUITETURA
- 73 ARTES PLÁSTICAS. ESCULTURA. NUMISMÁTICA. TRABALHOS ARTÍSTICOS EM METAL
- 74 DESENHO. DESENHO ARTÍSTICO. ARTES INDUSTRIAIS. ARTES E OFÍCIOS APLICADOS
- 75 PINTURA
- 76 ARTES GRÁFICAS
- 77 FOTOGRAFIA E PROCESSOS SEMELHANTES
- 78 MÚSICA
- 79 RECREAÇÃO. DIVERSÕES. JOGOS. ESPORTES
- 8 LINGUAGEM. LINGUÍSTICA. LITERATURA**
- 80 QUESTÕES GERAIS REFERENTES À LINGUÍSTICA E À LITERATURA. FILOLOGIA

-
- 81 LINGUÍSTICA E LÍNGUAS
82 LITERATURA EM GERAL
- 9 GEOGRAFIA. BIOGRAFIA.
HISTÓRIA**
- 90 ARQUEOLOGIA. PRÉ-
HISTÓRIA. VESTÍGIOS CUL-
TURAIS. ESTUDOS REGIO-
NAIS
- 91 GEOGRAFIA. EXPLORAÇÃO
DA TERRA E DOS DIVERSOS
PAÍSES. VIAGENS. GEOGRA-
FIA REGIONAL
- 92 ESTUDOS BIOGRÁFICOS.
GENEALOGIA. HERÁLDICA.
BANDEIRAS
- 93 CIÊNCIA HISTÓRICA. HIS-
TORIOGRAFIA
- 94 HISTÓRIA MEDIEVAL E MO-
DERNA EM GERAL
- 95 HISTÓRIA DA ÁSIA EM GE-
RAL
- 96 HISTÓRIA DA ÁFRICA
- 97 HISTÓRIA DAS AMÉRICAS
DO NORTE E CENTRAL
- 98 HISTÓRIA DA AMÉRICA DO
SUL
- 99 HISTÓRIA DA OCEA-
NIA, DOS TERRITÓRIOS
ÁRTICOS, DA ANTÁRTICA