



Curso de Bacharelado em Biblioteconomia na Modalidade a Distância

José de França Bueno

Métodos Quantitativos,
Qualitativos e Mistos de Pesquisa

Semestre

7

Curso de Bacharelado em Biblioteconomia na Modalidade a Distância

José de França Bueno

Métodos Quantitativos, Qualitativos e Mistos de Pesquisa

Semestre

7

Brasília, DF



Rio de Janeiro

Faculdade de Administração
e Ciências Contábeis
Departamento
de Biblioteconomia



Permite que outros remixem, adaptem e criem a partir do seu trabalho para fins não comerciais, desde que atribuam o devido crédito ao autor e que licenciem as novas criações sob termos idênticos.

Presidência da República

Ministério da Educação

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)

Diretoria de Educação a Distância (DED)

Sistema Universidade Aberta do Brasil (UAB)

Universidade Federal do Rio de Janeiro (UFRJ)

Núcleo de Educação a Distância (NEAD)

Faculdade de Administração e Ciências Contábeis (FACC)

Departamento de Biblioteconomia

Leitor

Maria Imaculada Cardoso Sampaio

Comissão Técnica

Célia Regina Simonetti Barbalho

Helen Beatriz Frota Rozados

Henriette Ferreira Gomes

Marta Lúcia Pomim Valentim

Comissão de Gerenciamento

Mariza Russo (*in memoriam*)

Ana Maria Ferreira de Carvalho

Maria José Veloso da Costa Santos

Nadir Ferreira Alves

Nysia Oliveira de Sá

Equipe de apoio

Eliana Taborda Garcia Santos

José Antonio Gameiro Salles

Maria Cristina Paiva

Miriam Ferreira Freire Dias

Rômulo Magnus de Melo

Solange de Souza Alves da Silva

Coordenação de

Desenvolvimento Instrucional

Cristine Costa Barreto

Desenvolvimento instrucional

Bruno Peixoto

Diagramação

Patrícia Seabra

Revisão de língua portuguesa

Mariana Caser

Projeto gráfico e capa

André Guimarães de Souza

Patrícia Seabra

Normalização

Dox Gestão da Informação

B928m Bueno, José de França.

Métodos quantitativos, qualitativos e mistos de pesquisa / José de França Bueno ; [leitora] Maria Imaculada Cardoso Sampaio. - Brasília, DF : CAPES : UAB ; Rio de Janeiro, RJ : Departamento de Biblioteconomia, FACC/ UFRJ, 2018.

192p. : il.

Inclui bibliografia.

ISBN 978-85-85229-30-6 (brochura)

ISBN 978-85-85229-31-3 (e-book)

1. Análise quantitativa. 2. Análise qualitativa. I. Sampaio, Maria Imaculada Cardoso. II. Título.

CDD 001.4

CDU 001.891

Caro leitor,

A licença CC-BY-NC-AS, adotada pela UAB para os materiais didáticos do Projeto BibEaD, permite que outros remixem, adaptem e criem a partir desses materiais para fins não comerciais, desde que lhes atribuam o devido crédito e que licenciem as novas criações sob termos idênticos. No interesse da excelência dos materiais didáticos que compõem o Curso Nacional de Biblioteconomia na modalidade a distância, foram empreendidos esforços de dezenas de autores de todas as regiões do Brasil, além de outros profissionais especialistas, a fim de minimizar inconsistências e possíveis incorreções. Nesse sentido, asseguramos que serão bem recebidas sugestões de ajustes, de correções e de atualizações, caso seja identificada a necessidade destes pelos usuários do material ora apresentado.

LISTA DE FIGURAS

Figura 1 –	Pesquisadores.....	18
Figura 2 –	Previsibilidade nas viagens	21
Figura 3 –	Balança: objetividade e previsibilidade	22
Figura 4 –	Etapas de uma pesquisa quantitativa	23
Figura 5 –	Abordagens de pesquisa: diferenças	33
Figura 6 –	Esquema para classificação das variáveis	41
Figura 7 –	Tipos de desenhos de pesquisas quantitativas	44
Figura 8 –	Exemplo 1 da simbologia dos desenhos experimentais.....	45
Figura 9 –	Exemplo 2 da simbologia dos desenhos experimentais.....	46
Figura 10 –	Estudos experimentais: pré-, quase- e experimentos “puros”	46
Figura 11 –	Simbologia usada em desenhos experimentais.....	48
Figura 12 –	Fotografia do fluxo de uma pesquisa transversal	51
Figura 13 –	Subdivisão dos estudos transversais	51
Figura 14 –	Amostragem de alimentos.....	54
Figura 15 –	Ideia de amostragem	55
Figura 16 –	Amostragem estratificada.....	60
Figura 17 –	Amostragens aleatórias: simples, sistemática, por conglomerado e estratificada.....	61
Figura 18 –	Intervalos de confiança para uma distribuição normal	65
Figura 19 –	Diagrama de dispersão: tempo de acesso (min.) x taxa de recuperação da informação (%).....	77
Figura 20 –	Coeficiente de correlação linear de <i>Pearson</i>	78
Figura 21 –	Diagrama de dispersão e coeficiente de correlação linear de <i>Pearson</i>	78
Figura 22 –	Tela 1	79
Figura 23 –	Tela 2	80
Figura 24 –	Tela 3	80
Figura 25 –	Tela 4	80
Figura 26 –	Diagrama de dispersão para consultas x matrículas.....	82
Figura 27 –	Valores críticos e regiões de rejeição em um teste de hipóteses.....	87
Figura 29 –	Relações entre as disciplinas métricas.....	96
Figura 30 –	Intersecção entre as “metrias”	96
Figura 31 –	Formulação gráfica original da Lei de <i>Bradford</i>	105
Figura 32 –	Lei de <i>Bradford</i>	106
Figura 33 –	Tela com <i>softwares</i> bibliométricos da USP São Carlos.....	117

Figura 34 – Estágios da SSM	156
Figura 35 – <i>Continuum</i> no desenho de pesquisa	179
Figura 36 – Esquema do desenho exploratório sequencial	182
Figura 37 – Esquema do desenho explicativo sequencial	182
Figura 38 – Esquema do desenho transformador sequencial	183
Figura 39 – Esquema do desenho de triangulação concomitante.....	183
Figura 40 – Esquema de desenhos incrustados concomitantes de modelo dominante	184
Figura 41 – Esquema do desenho incrustado concomitante de vários níveis	184
Figura 42 – Esquema de desenho transformador concomitante.....	185

LISTA DE QUADROS

Quadro 1 – Símbolos utilizados na representação de estudos experimentais	45
Quadro 2 – Parâmetros e estimadores.....	64
Quadro 3 – Exemplo 1 da escala de diferencial semântico	73
Quadro 4 – Exemplo 2 da escala de diferencial semântico	73
Quadro 5 – Quadro de exemplo: dados de universidades, centros universitários e faculdades, 2015.....	75
Quadro 6 – Erros dos tipos I e II nos testes de hipóteses	86
Quadro 7 – Modo de consultar a tabela qui-quadrado	90
Quadro 8 – Bibliometria, infometria, cientometria e webometria.....	118
Quadro 9 – Limitações das estratégias: grupos focais, observação e entrevistas	148
Quadro 10 – Comparação entre a Análise do conteúdo e a Análise do discurso	153
Quadro 11 – Literatura estrangeira com uso do “pensar alto”	154

LISTA DE TABELAS

Tabela 1 –	Valores de confiança alfa e respectivos valores críticos para $Z_{\alpha/2}$	66
Tabela 2 –	Disponibilidade de bibliotecas públicas e acervo do município de São Paulo, Subprefeitura de Pinheiros e distritos municipais, 2006	76
Tabela 3 –	Tempo de acesso ao sistema de bibliotecas x taxa de recuperação da informação	77
Tabela 4 –	Número de alunos matriculados x número de consultas à biblioteca.....	81
Tabela 5 –	Dados fictícios de apoio aos projetos A, B e C, com relação ao gênero (em valores absolutos)	88
Tabela 6 –	Dados fictícios de apoio aos projetos A, B e C, com relação ao gênero (em %).....	89
Tabela 7 –	Matriz de frequências esperadas para o exemplo do teste qui-quadrado	89
Tabela 8 –	Dados utilizados por <i>Bradford</i> em seu levantamento.....	106
Tabela 9 –	Exemplo para a Lei de <i>Zipf</i>	109

SUMÁRIO

1	UNIDADE 1: QUAIS MÉTODOS E PARA QUÊ?	15
1.1	OBJETIVO GERAL	15
1.2	OBJETIVOS ESPECÍFICOS	15
1.3	MAIS UMA DISCIPLINA DE MÉTODOS DE PESQUISA?	17
1.4	PERGUNTAS DIFERENTES EXIGEM ABORDAGENS DISTINTAS	19
1.4.1	Pesquisas quantitativas: aspectos principais	19
1.4.1.1	<i>Características da abordagem quantitativa</i>	20
1.4.2	Pesquisas qualitativas: aspectos principais	24
1.4.3	Atividade	26
1.4.4	Pesquisas mistas: aspectos principais	27
1.4.5	Atividade	28
1.4.6	Qual abordagem é a mais indicada?	28
1.4.6.1	<i>Situações nas quais se aplicam pesquisas quantitativas</i>	29
1.4.6.2	<i>Situações nas quais se aplicam estudos qualitativos</i>	29
1.4.6.3	<i>Situações nas quais se aplicam estudos mistos</i>	30
1.4.7	Críticas às pesquisas quantitativas	30
1.4.7.1	Reduccionismo dos métodos quantitativos	30
1.4.8	Atividade	31
1.4.9	Atividade	31
1.4.10	Atividade	32
1.5	CONCLUSÃO	33
	RESUMO	34
	SUGESTÃO DE LEITURA	35
	REFERÊNCIAS	35
2	UNIDADE 2: MÉTODOS QUANTITATIVOS: SEMPRE FUI RUIM EM MATEMÁTICA, TEREI QUE FAZER CÁLCULOS?	37
2.1	OBJETIVO GERAL	37
2.2	OBJETIVOS ESPECÍFICOS	37
2.3	O PROBLEMA DE UMA PESQUISA QUANTITATIVA	39
2.3.1	A redação do problema de uma pesquisa quantitativa	40
2.3.2	Entendendo bem o que são variáveis	40
2.3.3	Atividade	42
2.4	O DELINEAMENTO DA PESQUISA: SUA “ESTRUTURA”	43
2.4.1	Estudos experimentais	44
2.4.1.1	<i>Notação para estudos experimentais</i>	44
2.4.2	Estudos não experimentais	49
2.4.2.1	<i>Estudos não experimentais transversais</i>	50
2.4.2.2	<i>Estudos não experimentais longitudinais</i>	52
2.4.2.3	<i>Estudos não experimentais correlacionais</i>	52

2.4.3	Atividade.....	52
2.5	AMOSTRAGENS: VOCÊ JÁ FEZ EXAME DE SANGUE? AMOSTRAGEM E VÍCIO	53
2.5.1	Definição de população	55
2.5.2	Definição de amostra	55
2.5.3	Amostras não probabilísticas.....	56
2.5.4	Amostras probabilísticas	57
2.5.5	Atividade.....	58
2.5.6	Atividade.....	62
2.6	DETERMINANDO O TAMANHO DA AMOSTRA	63
2.6.1	Tamanho da amostra para amostra aleatória simples	63
2.6.2	Tamanho da amostra para médias	67
2.6.3	Tamanho da amostra no caso de proporções	68
2.7	COLETANDO OS DADOS	69
2.7.1	Escalas: <i>Likert</i> , diferencial semântico e escala de <i>Guttman</i>	72
2.7.1.1	Escala de Likert	72
2.7.1.2	Escala de diferencial semântico (ou de Osgood)	73
2.8	TABELAS DE FREQUÊNCIA: ORGANIZANDO E VISUALIZANDO OS DADOS.....	74
2.9	ANÁLISE DOS DADOS	76
2.9.1	Correlação e regressão	77
2.9.2	Atividade.....	79
2.9.3	Reta de regressão	81
2.9.4	Atividade.....	83
2.10	TESTES DE HIPÓTESES	84
2.10.1	O formato geral de um teste de hipóteses.....	85
2.10.2	Teste qui-quadrado (associação entre dados categóricos)	88
2.11	CONCLUSÃO	90
	RESUMO	91
	REFERÊNCIAS	92
3	UNIDADE 3: MÉTODOS QUANTITATIVOS ESPECÍFICOS DA BIBLIOTECONOMIA: COMO QUANTIFICAR COMPORTAMENTOS CULTURAIS?	93
3.1	OBJETIVO GERAL	93
3.2	OBJETIVOS ESPECÍFICOS	93
3.3	BIBLIOMETRIA, INFOMETRIA, CIBERMETRIA, WEBOMETRIA, CIENTOMETRIA, PATENTOMETRIA	95
3.3.1	As várias “metrias”	95
3.3.2	Bibliometria.....	97
3.3.3	Cientometria	98
3.3.4	Infometria	99
3.3.5	Webometria	101
3.3.6	Atividade.....	101
3.3.7	Cibernetria	102
3.3.7	Atividade.....	102
3.3.8	Atividade.....	103

3.4	AS TRÊS LEIS DA BIBLIOMETRIA.....	104
3.4.1	Lei de <i>Bradford</i> (Lei da dispersão).....	104
3.4.2	Atividade.....	107
3.4.3	Leis de <i>Zipf</i> (Lei do mínimo esforço)	107
3.4.4	Lei de <i>Lotka</i>	111
3.4.4.1	Lei do elitismo, de <i>Price</i>	112
3.5	ÍNDICES E INDICADORES.....	112
3.5.1	Fator de impacto.....	112
3.5.2	H-índice (<i>h-index</i>)	113
3.5.3	Fator de impacto na <i>web</i>	113
3.5.4	Outros indicadores	114
3.5.5	Limitações dos indicadores bibliométricos	114
3.6	AS BASES DE DADOS UTILIZADAS EM PESQUISAS BIBLIOMÉTRICAS.....	115
3.6.1	<i>Web of Science</i>	115
3.6.2	<i>SciELO</i>	116
3.7	SOFTWARES BIBLIOMÉTRICOS	117
3.7.1	Atividade.....	117
3.8	CONCLUSÃO.....	119
	RESUMO	120
	SUGESTÃO DE LEITURA	120
	REFERÊNCIAS	120
4	UNIDADE 4: MÉTODOS QUALITATIVOS	123
4.1	OBJETIVO GERAL	123
4.2	OBJETIVOS ESPECÍFICOS	123
4.3	FORMULAÇÃO DE PROBLEMAS EM PESQUISAS QUALITATIVAS.....	125
4.4	DESENHOS DO PROCESSO EM PESQUISAS QUALITATIVAS.....	127
4.4.1	Teoria fundamentada (<i>Grounded Theory</i>).....	127
4.4.1.1	Coleta dos dados.....	128
4.4.1.2	Amostragem.....	129
4.4.1.3	Saturação.....	129
4.4.1.4	Codificação.....	129
4.4.1.5	Registro (<i>memos</i>).....	130
4.4.1.6	Redação da teoria.....	130
4.4.2	Etnografia	131
4.4.2.1	Coleta de dados	132
4.4.2.2	Notas de campo	132
4.4.2.3	Triangulação na pesquisa etnográfica.....	132
4.4.2.4	Análise de dados etnográficos.....	132
4.4.2.5	Dificuldades e eventuais limitações da pesquisa etnográfica.....	133
4.4.3	Estudo de caso	133
4.4.4	Histórias de vida ou desenhos narrativos	134
4.4.4.1	Coleta de dados	135
4.4.4.2	Análise das entrevistas narrativas.....	135

4.4.5	Pesquisa-ação	135
4.4.5.1	<i>Modalidades de pesquisa-ação</i>	136
4.4.5.2	<i>As fases na pesquisa-ação</i>	136
4.4.5.3	<i>Limitações da pesquisa-ação</i>	136
4.5	A AMOSTRAGEM EM PESQUISAS QUALITATIVAS	137
4.5.1	A coleta de dados na pesquisa qualitativa	138
4.5.1.1	<i>Observação</i>	139
4.5.1.2	<i>Entrevistas</i>	140
4.5.1.3	<i>Grupos focais (focus groups)</i>	141
4.5.1.3.1	O roteiro de questões em um grupo focal	143
4.5.2	Atividade	144
4.5.3	Análise de documentos e outros materiais	145
4.5.4	Pesquisa qualitativa na internet	145
4.5.5	Netnografia	146
4.5.6	Outras abordagens	147
4.5.6.1	<i>Triangulação de métodos de coleta de dados</i>	147
4.5.7	Atividade	148
4.6	A ANÁLISE DE DADOS NA PESQUISA QUALITATIVA	149
4.6.1	Análise do discurso	151
4.6.2	Protocolo verbal	153
4.6.3	Soft Systems Methodology	155
4.6.3.1	<i>Limitações ou restrições à abordagem SSM</i>	158
4.6.4	Discurso do sujeito coletivo	158
4.6.5	Análise do conteúdo	159
4.6.5.1	<i>Categorias na AC</i>	160
4.6.5.2	<i>As fases na AC</i>	160
4.6.6	Método Delphi	162
4.6.6.1	<i>Etapas para aplicação da técnica Delphi</i>	163
4.6.6.2	<i>A técnica Delphi eletrônica</i>	164
4.6.6.3	<i>Limites da técnica Delphi</i>	164
4.6.7	Atividade	164
4.6.8	Atividade	165
4.6.9	Atividade	165
4.6.10	Atividade	166
4.6.11	Atividade	166
4.6.12	Atividade	167
4.6.13	Atividade	167
4.7	CONCLUSÃO	168
	RESUMO	169
	SUGESTÃO DE LEITURA	169
	REFERÊNCIAS	170
5	UNIDADE 5: MÉTODOS MISTOS	177
5.1	OBJETIVO GERAL	177

5.2	OBJETIVOS ESPECÍFICOS	177
5.3	MÉTODOS MISTOS DE PESQUISA.....	179
5.4	A SEQUÊNCIA DOS MÉTODOS NAS PESQUISAS MISTAS.....	180
5.5	DESENHOS DE PESQUISA MISTA	180
5.5.1	Desenho exploratório sequencial.....	182
5.5.2	Desenho explicativo sequencial	182
5.5.3	Desenho transformador sequencial	182
5.5.4	Desenho de triangulação concomitante	183
5.5.5	Desenho incrustado concomitante de modelo dominante	183
5.5.6	Desenho incrustado concomitante de vários níveis.....	184
5.5.7	Desenho transformador concomitante	184
5.5.8	Atividade.....	185
5.6	CONCLUSÃO.....	187
	RESUMO.....	187
	REFERÊNCIAS.....	188
	CONCLUSÃO DA DISCIPLINA	188
	BIBLIOGRAFIA DA DISCIPLINA.....	188

UNIDADE 1

QUAIS MÉTODOS E PARA QUÊ?

1.1 OBJETIVO GERAL

Apresentar a disciplina Métodos Quantitativos, Qualitativos e Mistos de Pesquisa, sua importância para o aprimoramento profissional, e mostrá-la como ferramenta de trabalho. Afinal, o bibliotecário vem sendo requerido para atuar como consultor na geração da produção científica.

1.2 OBJETIVOS ESPECÍFICOS

Esperamos que, ao final desta Unidade, você seja capaz de:

- a) listar semelhanças e diferenças entre cada uma das abordagens de pesquisa;
 - b) descrever características dos métodos quantitativos, qualitativos e mistos de pesquisa;
 - c) apresentar critérios de escolha para cada uma dessas abordagens;
 - d) estabelecer relações entre os métodos de pesquisa e os estudos em bibliotecas.
-

1.3 MAIS UMA DISCIPLINA DE MÉTODOS DE PESQUISA?

Iniciamos deixando como sugestão para você, aluno(a), um vídeo sobre o método científico, disponível em: <https://www.youtube.com/watch?v=uZ_vdGFMbBA> (acesso em: 26 jun. 2021). Nele, você terá a oportunidade de aprofundar sua compreensão sobre o método e os procedimentos científicos. Vale a pena!

Nesta disciplina, estudaremos métodos quantitativos, qualitativos e mistos de pesquisa. Você deve estar se perguntando: “Qual é a necessidade de estudar esses métodos, se já cursei Metodologia da Pesquisa I e Metodologia da Pesquisa II? Mais uma disciplina de Metodologia? Os conteúdos irão se repetir?”

Calma, cara aluna, caro aluno. Não, os conteúdos não se repetirão. Guarde bem as informações aprendidas aqui, você vai buscar por elas se for atuar em bibliotecas universitárias e em centros de pesquisa. Se for atuar na consultoria de publicações científicas, também.

Vamos nos situar?

Certamente, você já teve um problema para resolver em sua vida “real” (aquela sua vida antes de virar estudante universitário, lembra dela?). Esse problema pode ter sido de natureza financeira, ou um problema de relacionamento interpessoal com seus colegas de trabalho. Você pode, ainda, ter vivido dilemas como: pedir ou não demissão do emprego; fazer este ou aquele curso superior; de que forma encontrar um parceiro ou parceira para a vida conjugal? Pode ter se perguntado: Por que as pessoas preferem cervejas de uma marca em detrimento de outras? Essa preferência tem relação com a idade? Com o local onde vivem? Com sua renda?

O que esperamos que você tenha compreendido no parágrafo anterior? Que a pesquisa, a investigação e, mais especificamente, a pesquisa científica, não são devaneios de um cientista de avental branco, instalado em sua inatingível torre de marfim, totalmente distante das necessidades dos profissionais da atualidade. Sejam médicos, contadores, engenheiros químicos ou bibliotecários, enfim, todos nós, profissionais, temos o dever de resolver problemas novos e desafiadores em nossa atividade profissional e de torná-los públicos em eventos e revistas científicas e de divulgação científica. Veja: os problemas de pesquisa são frutos da curiosidade do sujeito pesquisador. Inclusive, o seu progresso profissional e nas áreas de Biblioteconomia e Ciência da Informação no Brasil também dependem de sua capacidade de pesquisar, inovar e resolver problemas!

Semestre

7



Multimídia

Viver é resolver problemas! Entenda um pouco mais sobre a *ciência como atividade de solucionar problemas* assistindo ao vídeo homônimo, que está disponível no *link*: <<https://prezi.com/wvsel9n7tvwu/copy-of-a-ciencia-como-atividade-de-solucionar-problemas/>>. Acesso em: 26 jun. 2021.



Multimídia

Você acha que só gênios conseguem fazer ciência? Então, leia o texto contido no *link*: <<https://cienciahoje.org.br/coluna/profissao-cientista/>>. Último acesso em: 6 out. 2021.

O que você acha que é importante para ser pesquisador? Todos podemos pesquisar? O que o artigo destaca como sendo os aspectos importantes da pesquisa científica (e que foge do padrão divulgado pela “grande mídia”)?

Figura 1 – Pesquisadores



Fonte: *News Hour* (20--?).¹

¹ Disponível em: <http://newshour.s3.amazonaws.com/photos/2012/04/10/StudentdClassroom_business_desk.jpg>.

1.4 PERGUNTAS DIFERENTES EXIGEM ABORDAGENS DISTINTAS

No decorrer de nossas vidas pessoais e/ ou profissionais, todos nos deparamos com problemas de várias naturezas para resolver. Com tantas dificuldades circulando no mundo “real”, parece bastante razoável que existam ferramentas (a que nós chamaremos, aqui, de abordagens ou métodos de pesquisa) que sejam mais adequadas para determinadas classes de problemas.

Em comparação com as já estudadas Metodologia da Pesquisa I e II, nesta disciplina, você aprenderá, com mais detalhes, algumas ferramentas para que desenvolva sua pesquisa na prática. Em outras palavras, Métodos Quantitativos, Qualitativos e Mistos de Pesquisa é uma disciplina mais operacional, prática e (um pouco) menos teórica. Certamente, teremos teoria, mas nosso foco será usá-las nas ferramentas de pesquisa.

Estudaremos as abordagens quantitativa, qualitativa e mista. Todas elas lançam mão de aspectos comuns no seu percurso metodológico:

- a) observação dos fenômenos;
- b) desenvolvimento de hipóteses/ suposições decorrentes dessa observação;
- c) utilização de técnicas específicas para fundamentar tais hipóteses/ suposições;
- d) revisão das hipóteses/ suposições, a partir do que foi observado, e proposição de novas investigações, para compreender o objeto de estudo.

Apresentadas as semelhanças, a partir da seção seguinte, você começará a ver as diferenças entre as três abordagens de pesquisa e, especificamente, alguns detalhes da abordagem quantitativa.

1.4.1 Pesquisas quantitativas: aspectos principais

Antes de abordarmos a pesquisa quantitativa, vamos revisitar parte do quadro geral do conhecimento científico. Nos últimos séculos, o acúmulo de conhecimento científico e tecnológico aumentou exponencialmente. Convivemos com máquinas elétricas, motores a combustão, aviões, radares, computadores, produtos farmacêuticos, químicos e biológicos produzidos pela inteligência humana. No final do século XIX, momento de grandes descobertas na Física, na Matemática, na Química e na Astronomia, bem como de expansão crescente da industrialização, a abordagem filosófica denominada *Positivismo* ganhou preponderância: as Ciências Exatas e a matematização eram vistas como os modelos a serem seguidos no sentido de demonstrar como a humanidade poderia vir a conhecer a realidade. Ganhava destaque o chamado conhecimento positivo, baseado nos fenômenos naturais, de modo que positivo era



tudo que pudesse ser observado/ comprovado/ medido por instrumentos. Essa concepção filosófica postulava, ainda, que positivo era o que poderia passar por uma prova científica. Os positivistas acreditavam que era possível reduzir (reducionismo) os fenômenos sociais e humanos às leis da ciência (de preferência, utilizando, para isso, mensurações e expressões matemáticas). É característica do Positivismo a proposição de que há uma única verdade objetiva a ser descoberta pela ciência. É justamente o sucesso das ciências como a Física, a Química e a Matemática (além da Estatística e suas inúmeras aplicações) que influenciou pesquisadores das Ciências Humanas e Sociais a buscarem “transferir” ferramentas da Estatística e da Matemática para quantificar fenômenos sociais.



Atenção

E como a corrente filosófica positivista se expressou na Biblioteconomia?

No campo específico da Biblioteconomia, a concepção positivista é patente nos estudos de usuários realizados na década de 1940 (ARAÚJO, 2010). Por exemplo, registros de estatísticas de empréstimos e a segmentação de tarefas e necessidades em tipos. Tais estudos não consideravam como objeto de estudo o ser humano (o usuário, afinal de contas), com suas características pessoais, sua subjetividade etc.

1.4.1.1 Características da abordagem quantitativa

Como você deve ter observado, foram algumas disciplinas e a filosofia do Positivismo que aproximaram a abordagem quantitativa das Ciências Humanas e Sociais.

Uma palavra-chave, em pesquisas de natureza quantitativa, é a *objetividade*. Esperamos que as pesquisas quantitativas valorizem a objetividade e excluam as avaliações subjetivas. Já nas pesquisas qualitativas, a subjetividade é aceita e, muitas vezes, valorizada. As crenças, valores, julgamentos e pré-conceitos do pesquisador não devem interferir nos resultados da pesquisa. Você se lembra do parágrafo no qual destacamos a influência das Ciências Físicas, Químicas, Biológicas e Matemáticas na história das Ciências Humanas e Sociais? Uma analogia que poderíamos fazer, relativa à objetividade em pesquisas quantitativas, seria a de um médico utilizando um termômetro para avaliar se uma criança está com febre. Não gostaríamos que a opinião do médico interferisse na leitura do termômetro, não é verdade? A conclusão sobre se a criança está com febre ou não deve ser objetiva, não subjetiva, certo?

Outro aspecto de destaque nas pesquisas quantitativas é seu caráter fortemente estruturado. Há uma sequência-padrão de etapas no processo desse tipo de pesquisa que deve ser seguida sempre. Nesse sentido, dizemos que as pesquisas quantitativas apresentam um processo *previsível* em suas etapas. Claro, isso não significa que o resultado da pesquisa seja previsível!

Nas pesquisas quantitativas, destaca-se também a busca pela *generalização*. Para isso, são utilizadas técnicas de amostragem estruturadas, definidas antes do processo de coleta de dados, de forma a permitir a generalização dos resultados para determinada população.

Pesquisas quantitativas buscam conclusões ou resultados que possam gerar *leis gerais* ou *relações de causa e efeito* que, de alguma forma, sejam reproduzidas em outros contextos.

Você se lembra do que estudou de Física no Ensino Médio? Lembra-se da Lei da Gravitação Universal de *Newton*? Era uma lei geral ou universal. É esse padrão de busca de leis gerais que as pesquisas quantitativas, empregadas nas Ciências Sociais Aplicadas, buscam realizar, adaptando cada contexto.

Veja que a busca por leis gerais e relações de causalidade (relações de causa e efeito) objetivam construir teorias que produzam explicações e a *previsibilidade* dos fenômenos estudados. Continuando com a lembrança da Física do Ensino Médio: se disparamos uma bala de canhão, com massa conhecida, a determinado ângulo e a determinada velocidade, queremos conhecer a “lei geral” que nos permite prever a qual distância a bala irá cair e quanto tempo isso levará para acontecer.

Figura 2 – Previsibilidade nas viagens



Fonte: Pixabay (2014).²

Assim como utilizamos os conhecimentos exatos da Física para sabermos quanto tempo demoraremos para chegar a um local, com a pesquisa quantitativa, esperamos chegar a dados contáveis e exatos.



Atenção

E na Biblioteconomia? Encontramos “leis gerais”? Talvez você já tenha ouvido falar na Lei de *Bradford* (referente à dispersão de artigos de periódicos), na Lei de *Zipf* e na Lei de *Lotka*. Nos próximos capítulos desta disciplina, você entrará em contato com essas e outras ferramentas da Bibliometria, da Cientometria e da Infometria.

² Disponível em: <<http://pixabay.com/pt/paisagem-natureza-montanhas-estrada-540122/>>.

Figura 3 – Balança: objetividade e previsibilidade



Fonte: Pixabay (2014).³

A objetividade e o poder de generalização dos métodos quantitativos podem ser representados com a imagem de uma balança. Não queremos que a mensuração da massa de um objeto seja subjetiva, tampouco queremos que, a cada nova pesagem, o objeto apresente um valor diferente de massa: essa medida deve ser previsível.

Pesquisas quantitativas utilizam o *raciocínio dedutivo* (do geral para o particular). Partem de uma teoria, levantam hipóteses e, na sequência, testam se essas hipóteses são válidas.

As pesquisas quantitativas pressupõem que existe uma “*realidade externa*” ao pesquisador e que essa realidade pode ser conhecida. Na situação em que as hipóteses (aquilo em que se acredita) do pesquisador que utiliza a abordagem quantitativa coincidem com essa *realidade objetiva*, sua teoria é comprovada. Nos casos em que as hipóteses são rejeitadas, o pesquisador ou outros pesquisadores farão mudanças nelas (ou seja, mudarão suas crenças) ou na teoria.

A abordagem quantitativa é constituída de uma sequência de passos bem definidos. Observe-os e aceite que não é possível “saltar” etapas:

- a) o problema de pesquisa deve circunscrever (delimitar) de forma explícita o objeto de estudo. Caso contrário, posteriormente, você não conseguirá mensurar ou quantificar as suas variáveis de estudo. Escreva seu problema de estudo com palavras claras e objetivas;
- b) suas perguntas de pesquisa também devem ser bastante específicas e estar escritas em linguagem clara;
- c) a partir das perguntas de pesquisa, são formuladas uma ou mais hipóteses a serem testadas. Essas hipóteses são definidas antes da coleta dos dados;
- d) as variáveis também precisam estar bem definidas no início do estudo. Variável é todo elemento a ser estudado;
- e) antes de sair a campo para a coleta de dados, defina o tamanho mínimo da amostra e os critérios de seleção dos elementos que participarão da amostra. Os procedimentos de determinação do tamanho mínimo da amostra são muito bem definidos por equações

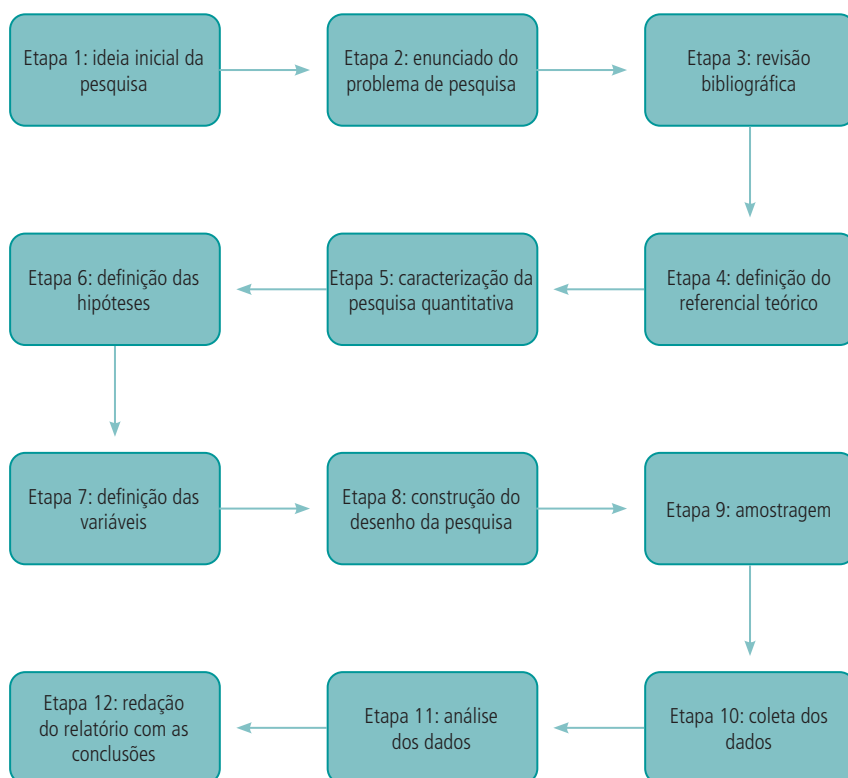
³ Disponível em: <<http://pixabay.com/pt/horizontal-balan%C3%A7a-de-cozinha-477492/>>.

matemáticas e devem ser apropriados para o seu estudo. Além disso, a ponderação com respeito aos participantes de pesquisas deve obedecer a critérios de renda, sexo, escolaridade e outros previamente definidos;

- f) a medição ou mensuração das variáveis deve seguir padrões também estabelecidos, com escalas de mensuração explícitas no seu trabalho. Os instrumentos para coleta de dados devem, preferencialmente, ser validados por pesquisas anteriores;
- g) a coleta de dados ocorre por meio de questionários ou levantamentos com perguntas objetivas, de modo a permitir posterior tabulação estatística. A coleta de dados também pode ocorrer por consulta a bancos de dados;
- h) antes de construir um questionário (instrumento de pesquisa), procure, na literatura, se já não existe um similar validado em pesquisas anteriores. Esse cuidado dará maior validade ao seu estudo;
- i) a análise dos dados ocorre por meio de métodos estatísticos ou indicadores e índices aceitos pela comunidade científica;
- j) existem vários procedimentos de testes de hipóteses e você deve adotar o mais apropriado para o seu estudo.

Entendemos por pesquisas quantitativas aquelas que fazem relação entre variáveis, testando hipóteses objetivas de forma que valores numéricos possam ser avaliados por meio de técnicas estatísticas. Com tais métodos, busca-se, a partir de uma amostra representativa da população, desenvolver inferências válidas para dada população (generalização).

Figura 4 – Etapas de uma pesquisa quantitativa



Fonte: Adaptado de Sampieri; Collado; Lucio (2013).

1.4.2 Pesquisas qualitativas: aspectos principais



Multimídia

Antes de apresentar as características da abordagem qualitativa, gostaríamos de sugerir que você consulte o texto “A problemática dos métodos quantitativos e qualitativos em biblioteconomia e documentação: uma revisão de literatura”. Leia-o antes de continuar estudando esta seção acessando o *link*: <<http://revista.ibict.br/ciinf/article/view/312>>. Último acesso em: 6 out. 2021.

Nas pesquisas qualitativas, busca-se a compreensão do significado que as pessoas atribuem a algum problema específico. Destaca-se o *caráter subjetivo* da análise e interpretação dos dados. Trabalha-se com informações não traduzíveis em números, mas em sentimentos ou percepções. O papel do pesquisador como intérprete é central nas pesquisas qualitativas. Outro aspecto é que tais pesquisas possuem caráter *descritivo*, que, aqui, faz referência ao fato de que os pesquisadores qualitativos têm o propósito de descrever e compreender fenômenos por meio das interpretações e dos significados gerados a partir do ponto de vista dos sujeitos da pesquisa. Em comparação, a pesquisa quantitativa busca descrever (e prever, sempre que possível) relações de causa e efeito dos fenômenos.

O estudo do tipo qualitativo se preocupa com a constante reformulação dos pressupostos. O conhecimento nunca está pronto e acabado, sempre pode ser alterado por outro estudo. O problema é focalizado em seu aspecto total e não pode ser fragmentado. O contexto e os fatores externos podem interferir na apreensão e interpretação dos resultados.

Nas pesquisas qualitativas, temos a possibilidade de refinar as perguntas durante o processo da pesquisa. Uma palavra-chave importante para a abordagem qualitativa é *flexibilidade*. Pesquisas qualitativas possuem uma natureza exploratória, que permite o aprofundamento de aspectos inicialmente não explícitos ou conscientes.

Outra palavra-chave em pesquisas qualitativas é *diversidade*: há muitas técnicas de coleta de dados (entrevistas, grupos focais, pesquisa documental, entre outras).

São características das pesquisas qualitativas:

- a) o objetivo precisa ser compreendido por seus *significados* e a partir da *vivência* do pesquisador;
- b) *não têm como objetivo testar e refutar hipóteses*, com base em uma amostra estatisticamente definida que permita generalizações;
- c) a amostragem, coleta e análise de dados podem ser desenvolvidas *ao mesmo tempo*. Não precisam seguir etapas sucessivas como nas pesquisas quantitativas;

- d) os dados podem estar contidos em textos, fotos, imagens, documentos, entre outras formas de registro. Além disso, existem dados observados e analisados a partir da abordagem qualitativa que sequer estão registrados (por exemplo, informações que um pesquisador pode obter por meio da técnica da memória oral, a partir da interação com o sujeito participante). Existem situações, como na pesquisa participante e na pesquisa-ação, nas quais o pesquisador atua no contexto, dialoga com seus demais participantes e recolhe dados a partir dessas interações;
- e) como o pesquisador que adota a abordagem qualitativa pode coletar dados em variadas formas, a análise destes também é bastante diversificada;
- f) valoriza-se a interpretação e a subjetividade do pesquisador;
- g) é utilizada a *lógica indutiva*: parte-se do particular para o geral;
- h) as *amostras* em pesquisas qualitativas são *relativamente pequenas*, diferentemente das amostras em pesquisas quantitativas, que costumam ser maiores;
- i) não costumam ser utilizadas técnicas estatísticas sofisticadas, mas contagens simples;
- j) estão interessadas em estudar fenômenos e vivências em profundidade, em obter uma compreensão deles, em examinar seus detalhes.

Vamos analisar, no exemplo a seguir, a flexibilidade das pesquisas qualitativas.

Suponha que, em um primeiro momento, um pesquisador desenhe seu projeto de pesquisa com a coleta de dados baseada em entrevistas de grupos focais (estudaremos a técnica de grupos focais na Unidade 4). Durante as entrevistas, os próprios entrevistados comentam sobre a existência de documentos importantes para o assunto em foco. A partir dessa informação, o pesquisador descobre documentos que enriquecerão sua pesquisa e verifica a necessidade de modificar a etapa de coleta de dados. Em consequência, também a análise de dados será modificada. Esse é um exemplo da flexibilidade possível nesse tipo de pesquisa.

Quer ver um exemplo no qual a mera produção de tabelas e gráficos não parece ser a melhor abordagem de pesquisa? Vamos imaginar que você é um pesquisador da área de Enfermagem e pretende estudar as reações emocionais que os pacientes apresentam ao realizar uma cirurgia bariátrica (redução do estômago). É possível que indivíduos nessa situação apresentem reações emocionais bastante negativas. Observar uma tabela e gráficos com contagens de casos pode não ser útil para compreender de forma profunda a situação vivenciada pelos pacientes e seus significados.

Um exemplo de pesquisa qualitativa (voltado à área de Biblioteconomia) pode ser encontrado em LEITÃO, 2005. Nesse trabalho, a autora apresenta o uso de grupos focais na avaliação de uma biblioteca universitária.



1.4.3 Atividade

Esta atividade atende ao objetivo a: “listar semelhanças e diferenças entre cada uma das abordagens de pesquisa”.

Suponha que você tenha sido incumbido de realizar uma pesquisa para identificar fatores que influenciam na aquisição de um automóvel. Você pode realizar uma pesquisa quantitativa, uma pesquisa qualitativa ou adotar procedimentos mistos de pesquisa. Quais as principais diferenças entre tais abordagens que você elencaria para o gestor do projeto?

Resposta comentada

Consideremos, inicialmente, a pesquisa quantitativa: pesquisas quantitativas buscam obter relações de *causa e efeito* ou *leis gerais* (no nosso exemplo, poderíamos obter relações estatísticas entre a faixa de renda, a classe social e o tipo de veículo adquirido, por exemplo). Com isso, teríamos as características de *previsibilidade* e *generalização*, tão marcantes nas pesquisas quantitativas. Destaque-se aqui também a característica da *objetividade*: para chegar a tais resultados, os pesquisadores quantitativos são obrigados a utilizar procedimentos estatísticos padronizados e que permitem pouca flexibilidade. Para efetuar os testes de hipóteses estatísticas, pela própria estrutura desses testes, as hipóteses têm que ser definidas previamente. Isso obriga a que as variáveis sob estudo estejam muito bem definidas.

Por outro lado, se adotássemos uma abordagem qualitativa, poderíamos utilizar grupos focais (apenas como um dos exemplos possíveis de coleta de dados nas abordagens qualitativas) para nos aprofundarmos, sob o *ponto de vista* da *subjetividade* do consumidor, em quais são suas *emoções* e *percepções* ao dirigir o veículo A ou B de determinada montadora. Obter tais informações pode ser extremamente útil para as equipes de marketing da indústria automobilística (e, diga-se de passagem, é assim mesmo que funciona na vida real: as montadoras desenvolvem pesquisas qualitativas com grupos focais, com o objetivo de descobrir as percepções mais profundas dos clientes).

Ainda na questão da *coleta* de *dados* nas abordagens qualitativas: veremos que há grande *diversidade* nas estratégias de recolha de informações, bem como nas técnicas de análise de dados associadas. Você estudará esse aspecto com mais detalhes na unidade relativa aos métodos qualitativos.

Diferentemente da abordagem quantitativa, as *hipóteses não precisam* estar previamente definidas no início do estudo, pois há uma *flexibilidade* bastante característica nos estudos qualitativos. O objetivo da abordagem qualitativa não é refutar uma hipótese, mas obter uma *compreensão profunda* dos *significados* e *percepções* dos sujeitos de pesquisa a partir das vivências do pesquisador.

As pesquisas qualitativas adotam uma *lógica indutiva*, enquanto que as pesquisas quantitativas adotam uma *lógica dedutiva*.

E se for adotada uma abordagem de pesquisa mista? Como as pesquisas com métodos mistos, que utilizam as duas abordagens (em alguns casos, com preponderância de métodos quantitativos;

em outros, com preponderância de métodos qualitativos e, ainda em outros, com equilíbrio entre ambas as abordagens), são, forçosamente, métodos de pesquisa mais complexos, muitas vezes, elas exigem mais tempo e planejamento para serem realizadas. Você verá, na última Unidade deste curso, algumas das muitas combinações possíveis, que são utilizadas nos métodos de pesquisa mistos.

1.4.4 Pesquisas mistas: aspectos principais

Nas pesquisas de natureza mista, características dos métodos qualitativos e quantitativos são integradas. Projetos de pesquisa mistos podem ser de natureza sequencial (em primeiro lugar, efetua-se uma pesquisa quantitativa e, posteriormente, uma pesquisa qualitativa, ou o inverso) ou concomitante (as coletas de dados ocorrem ao mesmo tempo).

Segundo *Creswell* (2010, p. 38), “o conceito de misturar diferentes métodos originou-se em 1959, quando *Campbell* e *Fiske* utilizaram múltiplos métodos para estudar a validade de traços psicológicos”. Com o tempo, outros pesquisadores foram se sentindo encorajados a misturar técnicas qualitativas com técnicas quantitativas.

Essas pesquisas utilizam estratégias de coleta, tratamento e análise de dados afeitos tanto aos procedimentos qualitativos quanto aos quantitativos. Fazem uso de questões abertas e fechadas e os resultados podem ser apresentados em forma de filmes, imagens e textos, não apenas com dados numéricos. No caso de pesquisas mistas, informações obtidas por determinado instrumento podem ser aprofundadas utilizando-se observações.

Nas últimas duas décadas, a integração entre métodos quantitativos e qualitativos atingiu grande convergência. Cada vez mais são construídos bancos de dados comuns, nos quais os resultados oriundos de métodos quantitativos são utilizados em conjunto com resultados obtidos por métodos qualitativos, de forma que um dá mais força para o outro, e vice-versa. Para o estudo de usuários, combinar as pesquisas quantitativas com as qualitativas é recomendado, conforme veremos na Unidade 5.

Vejamos, a seguir, dois exemplos de estudos com métodos mistos de pesquisa.

Tanto na área de Marketing quanto em pesquisas eleitorais são utilizados, de forma rotineira, métodos mistos de pesquisa: são realizadas pesquisas quantitativas, com questionários respondidos em locais de grande passagem de pessoas, e efetuados estudos qualitativos com grupos focais. Tais grupos são constituídos de amostras selecionadas de poucas pessoas (entre seis e doze). Esses indivíduos são convidados a apresentar suas percepções dos produtos à venda (ou dos candidatos à eleição) para um mediador treinado.

Como outro exemplo, temos uma empresa automobilística que pode estar interessada em saber a opinião de um grupo selecionado de potenciais compradores de seus veículos. Ela quer conhecer quais são as emoções relatadas por essas pessoas ao dirigirem um automóvel.

Como exemplo voltado à área de Biblioteconomia, podemos ter um gestor de uma rede de bibliotecas públicas interessado em investigar a imagem de sua instituição perante o público (usuário e não usuário). Ele poderá utilizar

uma combinação de estudos quantitativos e qualitativos para investigar os fatores que afetam a imagem institucional da rede de bibliotecas.



1.4.5 Atividade

Esta atividade atende ao objetivo *b*: “descrever características dos métodos quantitativos, qualitativos e mistos de pesquisa”.

Utilize a internet para buscar relatos sobre três pesquisas (qualitativas e/ou quantitativas) e quais problemas foram resolvidos por elas.

Liste o resultado de sua pesquisa. Identifique as diferentes abordagens de pesquisa encontradas, suas semelhanças e diferenças. Quais as justificativas apresentadas pelos autores dos trabalhos para adotar a abordagem utilizada?

Resposta comentada

Esperamos, com esta atividade, que você encontre uma razoável diversidade de exemplos. É esperado que uma parte dos alunos apresente relatos de artigos científicos/ dissertações de Mestrado/ teses de Doutorado relacionados com a área de Biblioteconomia.

Esperamos, ainda, que você, aluno(a), leia com atenção a discussão metodológica e as justificativas da abordagem utilizada nos trabalhos científicos da área de Biblioteconomia. Queremos que você identifique, em tais pesquisas, as características dos métodos quantitativo, qualitativo e misto, bem como leia com atenção as justificativas utilizadas pelos autores. É desejável que você compare as características apresentadas nas sub-seções 1.4.1 (“Pesquisas quantitativas”) e 1.4.2 (“Pesquisas qualitativas”) com o que você encontrou nos três artigos.

1.4.6 Qual abordagem é a mais indicada?

Em primeiro lugar, devemos destacar que uma abordagem não exclui a outra, e que elas tampouco são opostas. Um entendimento mais equilibrado é o de que, na verdade, ambas se complementam. Com a abordagem quantitativa, investigamos relações de causalidade entre as variáveis estudadas. Por outro lado, com a abordagem qualitativa, estamos interessados em interpretar e conhecer o fenômeno em seus significados mais profundos. Um estudo conduzido sem rigor, independentemente de sua abordagem, terá baixa qualidade. O que define a abordagem a ser utilizada é o objetivo da pesquisa e o que já se conhece anteriormente daquele objeto.

Existe uma noção equivocada de que, por lidarem com números e apresentarem resultados em tabelas e gráficos, as pesquisas quantitativas têm maior rigor e “seriedade” quando comparadas com a abordagem qualitativa. Na entanto, a suposta objetividade da pesquisa quantitativa não impede que a visão de mundo do pesquisador se projete sobre a

forma como as perguntas do questionário serão redigidas ou sobre como ele obterá suas conclusões. A subjetividade é inerente ao ser humano.

Antes de prosseguir, vamos apresentar duas definições que serão utilizadas ao longo deste texto:

- a) *variáveis independentes*: são as variáveis que alteram os resultados a serem investigados no estudo. Essas variáveis podem modificar de alguma forma as variáveis dependentes;
- b) *variáveis dependentes*: são aquelas variáveis que mudam de valor como consequência ou resultado das variáveis independentes.

Como exemplo, considere que estudantes foram submetidos a um teste; o desempenho deles é a variável dependente. Essa variável pode depender de por quanto tempo os estudantes se prepararam para o teste (número de horas de estudo), da renda familiar ou do tipo de escola na qual fizeram o Ensino Médio. Nesse exemplo, o número de horas de estudo, a renda familiar e o tipo de escola são variáveis independentes.

1.4.6.1 Situações nas quais se aplicam pesquisas quantitativas

As pesquisas quantitativas, nas palavras de Cervi (2009, p. 2), “[...] têm como principal finalidade a medição de quantidades e quantificação de qualidades”. Esse autor destaca ainda que a adoção ou não da abordagem quantitativa deve ser pautada na delimitação do objeto de estudo (CERVI, 2009):

- para obtermos perfis populacionais, sociais ou econômicos que possam resultar em generalizações;
- para mensurar (medir) preferências, comportamentos, atitudes etc;
- para testar hipóteses;
- para identificarmos quais variáveis afetam um desempenho ou variável-resposta e como essas variáveis se relacionam;
- para avaliar a eficácia de uma intervenção ou “tratamento”. Por exemplo, imagine uma situação na qual tenha sido realizada uma campanha de educação de uso da biblioteca local. O poder público pode estar interessado em avaliar a eficácia dessa campanha e a forma como ela foi recebida pelo público. O “tratamento” (intervenção), aqui, é a campanha de educação ou campanha publicitária.

1.4.6.2 Situações nas quais se aplicam estudos qualitativos

Quando existe pouca pesquisa sobre um fenômeno, necessitamos de pesquisas exploratórias. Precisamos aprofundar nossos conhecimentos do fenômeno para poder conhecer as variáveis intervenientes no problema. Não é possível, em tais circunstâncias, saber de antemão quais as variáveis dependentes e independentes relevantes para o problema estudado. Nessas situações de pouca informação, teremos dificuldades em definir uma amostra representativa para atender aos requisitos dos testes de hipóteses da estatística. Pesquisas qualitativas são aplicadas quando o pesquisador não sabe se existem os atributos/ características dos objetos em foco. São úteis para gerar hipóteses e detectar variáveis.

1.4.6.3 Situações nas quais se aplicam estudos mistos

Existem situações nas quais os métodos quantitativos e qualitativos se complementam. Os métodos quantitativos não conseguem, por meio de sua aplicação, fornecer uma compreensão mais profunda do fenômeno estudado, mas explicam as relações entre as variáveis dependentes e independentes do problema. Por outro lado, técnicas qualitativas (grupos focais, por exemplo) permitem entender, em profundidade, aquilo que os dados quantitativos não explicitam com a tradicional “frieza dos números”. Não é recomendável adotar uma abordagem qualitativa em um estudo para o qual pretendemos que o viés do observador seja eliminado.

1.4.7 Críticas às pesquisas quantitativas

Embora os modelos matemáticos/ estatísticos adotados nas pesquisas quantitativas partam do pressuposto de não interferência da subjetividade do pesquisador na construção desses modelos, é justamente ele quem define quais serão esses modelos. A subjetividade já entrou aí.

1.4.7.1 Reduccionismo dos métodos quantitativos

Uma crítica frequente aos métodos quantitativos é a de que, na busca por generalizações, eles conduzem a uma simplificação das relações complexas entre os indivíduos, o ambiente e as relações estabelecidas entre ambos. Métodos quantitativos podem ocultar questões que, por vezes, em uma mesma situação, podem apresentar-se multifacetadas, com “várias verdades” ou diversas narrativas válidas. Assim, a diversidade do real nem sempre pode ser bem representada por números, quantidades ou fórmulas matemáticas. Na verdade, números, tabelas ou expressões matemáticas podem, inclusive, colaborar para ocultar aspectos mais profundos da realidade em estudo.



Multimídia

Um exemplo bastante didático do debate envolvendo métodos quantitativos x métodos qualitativos pode ser encontrado no artigo do Prof. Dr. *Carlos Alberto Araújo*, da *Universidade Federal de Minas Gerais (UFMG)*, que documenta o trabalho com diferentes abordagens de pesquisa, feito com seus alunos da disciplina de “Usuários da Informação”. É interessante que você leia as conclusões dos alunos e as confronte com a nossa discussão sobre métodos quantitativos e qualitativos. O *link* para o artigo é: <http://www.uel.br/revistas/uel/index.php/informacao/article/viewFile/6485/6995>. Acesso em: 26 jun. 2021.

Entre as décadas de 1960 e 1980, predominaram os estudos de usuários de técnicas quantitativas (BATISTA; CUNHA, 2007 apud ARAÚJO, 2012). Os autores apontam que houve maior presença de estudos de usuários de natureza qualitativa a partir de meados dos anos 1980 pois

as pesquisas com métodos quantitativos não contribuíram para a identificação das necessidades individuais e para a implementação de sistemas de informação adequados a essas necessidades (BATISTA; CUNHA, 2007, p. 170 apud ARAÚJO, 2012).

Vamos verificar o que foi visto até agora? A seguir, você encontrará três atividades que avaliam algumas das questões abordadas nesta Unidade.



1.4.8 Atividade

Esta atividade atende ao objetivo c: “apresentar critérios de escolha para cada uma das abordagens”.

Suponha que você tenha sido convidado a realizar uma pesquisa objetiva, e não subjetiva, sobre os usuários de uma biblioteca. A pergunta de pesquisa é: a renda familiar e a distância da moradia à biblioteca afetam o número de vezes que o usuário acessa o equipamento?

Sob essas condições, sua recomendação será a de que sejam adotados métodos quantitativos, qualitativos ou mistos de pesquisa? Por quê?

Resposta comentada

Com a solicitação de que a pesquisa seja objetiva e por termos uma delimitação bastante clara e explícita do objeto de estudo, somos direcionados a optar por abordagens quantitativas.



1.4.9 Atividade

Esta atividade também atende ao objetivo c: “apresentar critérios de escolha para cada uma das abordagens”.

Neste novo projeto de pesquisa, você foi apresentado a uma nova solicitação: seu supervisor deseja saber quais são as percepções dos usuários do ambiente da biblioteca. Ele espera obter uma compreensão do fenômeno social *uso da biblioteca*. Seu supervisor espera que você apresente descrições do fenômeno uso da biblioteca a partir do ponto de vista dos usuários. Busca-se, com



esta pesquisa, o significado que os usuários atribuem ao uso da biblioteca e não se tem como objetivo testar hipóteses.

Neste caso, você recomendaria a adoção de qual método de pesquisa: quantitativo, qualitativo ou misto?

Resposta comentada

Aqui, temos a valorização de aspectos subjetivos e não temos hipóteses previamente definidas a serem testadas. O fenômeno sob estudo também não é facilmente traduzível em números. Não há uma obrigatoriedade ou rigidez quanto à forma de coleta de dados ou variáveis previamente definidas. Todas essas características nos direcionam a adotar métodos qualitativos de pesquisa.



1.4.10 Atividade

Esta atividade atende ao objetivo d: “estabelecer relações entre os métodos de pesquisa e os estudos em bibliotecas”.

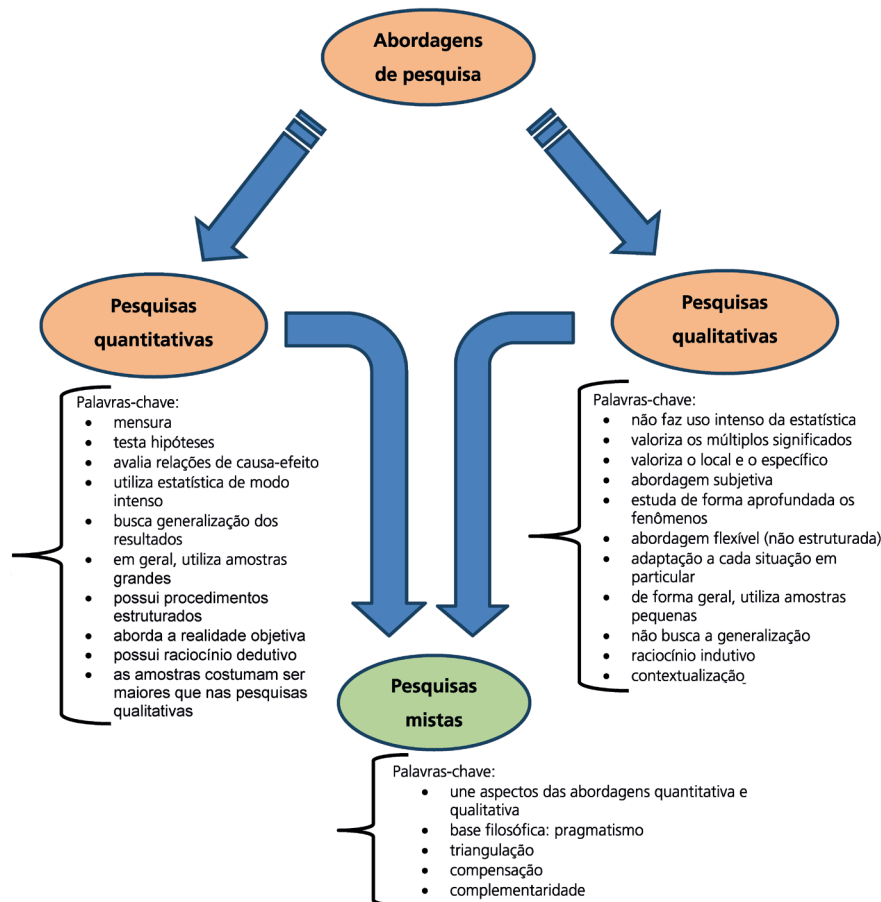
Suponha que você deseja estudar os usuários de bibliotecas de uma universidade. Você possui informações como: total de usuários atendidos, total de empréstimos, total de consultas, quantidade de livros, bem como teses e periódicos nas várias bibliotecas do sistema. Você também sabe a quantidade de bibliotecários e auxiliares na rede de bibliotecas. Além disso, você possui a informação de investimentos realizados, para cada ano, em aquisição de livros e periódicos nos formatos impresso e digital. Apresente sua avaliação sobre os alcances e limites das informações que você possui para a tomada de decisões.

Resposta comentada

Neste caso, não temos uma única resposta certa. Espera-se que você avalie que, embora a coleta de dados quantitativos seja importante para a tomada de decisões, é de extrema relevância o número de usuários para cada uma das bibliotecas desta rede quanto à categorização dos usuários, bem como à avaliação dos investimentos realizados ao longo do tempo. Tais informações não auxiliam no conhecimento sobre como se dá a interação entre o usuário e o sistema, tampouco auxiliam a que se conheça o ponto de vista do significado que esse usuário atribui ao acesso à informação obtida. Assim, uma abordagem meramente quantitativa não facilita a construção do diálogo entre as bibliotecas e os usuários. Para mais detalhes com respeito a esse aspecto, por favor, consulte Leitão (2005, p. 42-44).

Vamos resumir? Na Figura 5, a seguir, encontram-se palavras-chave de cada uma das abordagens de pesquisa:

Figura 5 – Abordagens de pesquisa: diferenças



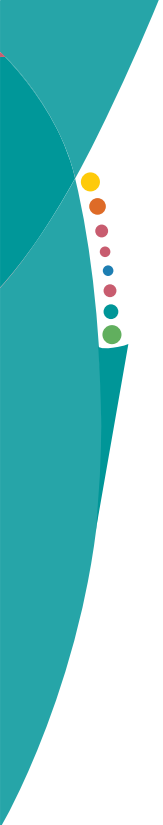
Fonte: Produção do próprio autor (2017).

Na próxima Unidade, vamos começar a estudar técnicas quantitativas gerais e suas aplicações nas pesquisas em Biblioteconomia e Ciência da Informação.

1.5 CONCLUSÃO

Compreender os alcances e limites dos diferentes métodos de pesquisa é fundamental para a tomada de decisão sobre qual deles deve ser utilizado em cada situação de pesquisa. Esperamos que você, agora, consiga descrever as características dos métodos quantitativos (objetividade, busca por leis gerais e relações de causa e efeito, previsibilidade, presunção de uma realidade externa ao observador, entre outras), qualitativos (subjetividade, flexibilidade, compreensão profunda do significado dos fenômenos sociais do ponto de vista, entre outras) e mistos (combinação de métodos quantitativos com métodos qualitativos, de natureza sequencial ou concomitante).

Para responder conscientemente a questões sobre em quais contextos de pesquisa deve-se adotar este ou aquele método, você necessita conhecer as semelhanças e diferenças entre cada abordagem e os critérios de escolha para cada um dos métodos.



No nosso entendimento, não há uma “superioridade” entre métodos quantitativos, qualitativos ou mistos, mas há, muito mais, uma complementaridade entre eles. Desse ponto de vista, não cabe a pergunta sobre qual método é o melhor, mas qual é o mais adequado para esta ou aquela situação. Há algumas décadas, existia uma visão de que os métodos quantitativos e qualitativos eram opostos e acreditava-se que não seria possível uma pesquisa que utilizasse os dois. Atualmente, essa é uma visão ultrapassada.

RESUMO

Nesta Unidade, estudamos as características dos métodos quantitativos, qualitativos e mistos.

Vejamos, a seguir, as principais características dos métodos quantitativos:

- a) objetividade;
- b) lógica dedutiva;
- c) busca por leis gerais, relações de causa e efeito;
- d) busca por leis ou expressões matemáticas/ estatísticas que permitam previsibilidade;
- e) generalização;
- f) pressuposição de uma realidade externa ao pesquisador (e de que essa realidade pode ser conhecida);
- g) as perguntas de pesquisa devem ser específicas e claras;
- h) as hipóteses são definidas antes da coleta de dados;
- i) as escalas de mensuração das variáveis e devem seguir padrões bem estabelecidos.

Na sequência, temos as principais características dos métodos qualitativos:

- a) subjetividade;
- b) lógica indutiva;
- c) flexibilidade do pesquisador;
- d) grande diversidade de técnicas de coleta e análise de dados;
- e) pesquisas qualitativas não têm como objetivo testar e refutar hipóteses;
- f) diferentemente das pesquisas quantitativas, nas quais amostragem, coleta e análise de dados devem ser realizadas em etapas sucessivas, nas pesquisas qualitativas, elas podem ocorrer ao mesmo tempo;
- g) os dados podem ser constituídos por fotos, imagens, figuras, documentos ou outros registros;
- h) as amostras não precisam ser numericamente grandes para responder a determinadas margens de erro, como nas pesquisas

quantitativas (muitas vezes, a pesquisa só poderá ser realizada com amostras pequenas);

- i) busca uma compreensão profunda de fenômenos e vivências, além do significado atribuído a eles.

Quanto aos métodos mistos de pesquisa, suas principais características são:

- a) combinação, de forma sequencial ou concomitante, de métodos quantitativos e qualitativos;
- b) grande diversidade de estratégias de coleta, tratamento e análise de dados;
- c) nos últimos anos, a integração entre métodos quantitativos e qualitativos é cada vez maior.



Sugestão de Leitura

CAMPBELL, D. T.; FISKE, D. Convergent and discriminant validation by the multitrait-multimethod matrix. **Psychological Bulletin**, [S.l.], n. 56, p. 81-105, 1959.

REFERÊNCIAS

ARAÚJO, C. A. Á. Abordagem interacionista de estudos de usuários da informação. **Ponto de Acesso**, [S.l.], v. 4, n. 2, 2010. Disponível em: <<http://www.portalseer.ufba.br/index.php/revistaici/article/view/3856>>. Acesso em: 15 out. 2014.

CERVI, E. U.; HEDLER, A. P. Métodos quantitativos na produção de conhecimento sobre jornalismo: uma abordagem alternativa ao fetichismo dos números e ao debate com qualitativistas. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 32., 2009, Curitiba. **Anais...** Curitiba: Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, 2009. Disponível em: <http://www.unisc.br/portal/upload/com_arquivo/metodos_quantitativos_na_producao_de_conhecimento_sobre_jornalismo_abordagem_alternativa.pdf>. Acesso em: 21 fev. 2015.

CRESWELL, J. W. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. 3. ed. Porto Alegre: Artmed, 2010.

LEITÃO, B. J. M. **Avaliação qualitativa e quantitativa numa biblioteca universitária**. Rio de Janeiro: Interciência; Niterói: Intertexto, 2005.



UNIDADE 2

MÉTODOS QUANTITATIVOS: SEMPRE FUI RUIM EM MATEMÁTICA, TEREI QUE FAZER CÁLCULOS?



2.1 OBJETIVO GERAL

Apresentar técnicas quantitativas gerais utilizadas pelos profissionais de Ciências Sociais Aplicadas, em particular, de Biblioteconomia.

2.2 OBJETIVOS ESPECÍFICOS

Esperamos que, ao final desta Unidade, você seja capaz de:

- a) redigir seu problema de pesquisa de forma precisa;
 - b) delinear pesquisas experimentais e não experimentais;
 - c) descrever as técnicas básicas de amostragem utilizadas em estudos quantitativos;
 - d) descrever as ferramentas básicas de coleta de dados em pesquisas quantitativas;
 - e) utilizar as principais ferramentas de análise de dados em pesquisas quantitativas;
 - f) utilizar as funções estatísticas básicas do *software Microsoft Excel*®.
- 

2.3 O PROBLEMA DE UMA PESQUISA QUANTITATIVA

Nesta Unidade e na próxima, aprenderemos a usar ferramentas quantitativas em pesquisas na área de Biblioteconomia. Você pode estar pensando que não se lembra de mais nada das aulas de Matemática do Ensino Médio e que (talvez) sempre tenha tido dificuldades com essa disciplina.

Veja, certamente, teremos alguns cálculos, mas eles não serão monstruosos. Felizmente, os computadores foram inventados e eles são calculadoras poderosas! Atualmente, os computadores efetuam a parte mais extensa e trabalhosa dos cálculos. O que é atribuição do pesquisador ao usar ferramentas estatísticas e modelos matemáticos é a *interpretação* dos resultados numéricos fornecidos pelos computadores. O pesquisador deve, a todo momento, se perguntar: “estes valores numéricos fazem sentido no contexto em que estou trabalhando? Será que o uso desta ferramenta estatística, neste contexto, é adequado?”.

Outro aspecto importante para justificar seu aprendizado de noções de Estatística e métodos quantitativos de pesquisa consiste no fato de que, para acompanhar a evolução técnica que ocorre nas várias áreas do conhecimento, os profissionais têm que participar de congressos e consultar revistas especializadas. Parte significativa desses trabalhos utiliza ferramentas da Estatística e/ ou métodos quantitativos de pesquisa.

Outra questão importante: não deixe para consultar um estatístico somente nos estágios finais do seu trabalho! A sua interação com o profissional estatístico deve ocorrer logo nos primeiros momentos de estudo, antes mesmo da definição das variáveis. Não colete centenas de dados para depois os apresentar para o estatístico, solicitando-lhe que os analise. É possível que, já no processo de amostragem, você introduza um viés (também denominado *tendenciosidade* – discutiremos isso mais tarde, nesta mesma Unidade) que não permita que efetue generalizações a partir dos dados coletados. Além disso, os vários testes estatísticos exigem que certas premissas estejam satisfeitas para que seus resultados sejam considerados válidos. Cada teste possui um conjunto específico de premissas a ser atendido. Se você não verificar se terá como atender a essas condições antes de iniciar o trabalho, todo o seu esforço poderá ser desperdiçado.

Suas preocupações com o uso da Estatística devem ser muito mais direcionadas para a questão da correta verificação do atendimento dos pressupostos da ferramenta (ou seja, uma questão de análise e interpretação) do que para a questão dos cálculos.

Ok? Combinado? Então vamos lá!

Por favor, retome a Figura 4 da Unidade 1, que demonstra as etapas sucessivas de uma pesquisa quantitativa. Na presente Unidade, seguiremos as etapas descritas naquela figura, sem entrar, entretanto, nos detalhes de cada “caixinha”. Assim, veremos os seguintes conteúdos:



- a) classificação de variáveis;
- b) tabelas de frequência;
- c) visualização dos dados;
- d) amostragem;
- e) correlação e regressão;
- f) testes de hipóteses.

2.3.1 A redação do problema de uma pesquisa quantitativa

Certamente, o primeiro passo na pesquisa quantitativa é conceber a ideia inicial, ou seja, decifrar o problema de estudo. Quando você for construir o texto para o seu problema de pesquisa quantitativa, lembre-se de deixar explícitas as variáveis relevantes para seu trabalho, bem como as relações entre essas variáveis. Mas, afinal, o que são variáveis? Por favor, estude a subseção 2.3.2, logo a seguir, para entender o que são elas, ok?

O *problema de pesquisa* é sempre uma pergunta, por exemplo: “a biblioteca necessita diminuir o acesso aos periódicos para reduzir custos?”. Dessa pergunta sai o *objetivo geral* do estudo: “identificar os periódicos mais consultados, de acordo com a categoria de usuários”. A *justificativa* poderia ser: “necessidade de limitar o uso dos periódicos para determinada categoria de usuários, por uma questão de redução de custo de acesso”. Em tempos de crise, os bibliotecários certamente se depararão com problemas dessa natureza.



Explicativo

Em uma investigação sobre gestão de pessoas, um gerente de uma rede de bibliotecas universitárias pode estar interessado nos fatores que possam ter influência no desempenho dos seus funcionários. Lembre-se: ao construir o projeto de pesquisa, a redação do problema deve deixar totalmente explícitas as variáveis (dependentes e independentes) que serão investigadas, bem como as relações a serem testadas entre elas. A redação do problema de pesquisa deve conter o objetivo do estudo. Como sugestão de texto para um modelo de declaração de objetivo quantitativo, consulte Creswell (2010, p. 148-151).

2.3.2 Entendendo bem o que são variáveis

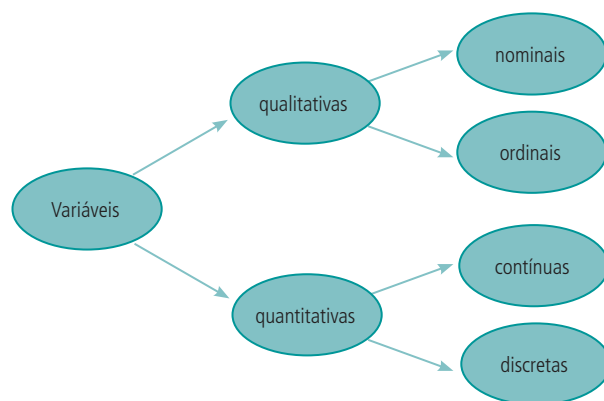
Você se recorda de quando, na Unidade 1, comentamos que uma das características da abordagem qualitativa era o fato de ela testar hipóteses sobre relações entre variáveis? Entender sobre variáveis, hipóteses e teste de hipótese é muito importante. Então, vamos nos aprofundar nesses assuntos.

Quando utilizamos métodos quantitativos, efetuamos contagens e/ou mensuramos algo. Mas contamos/ mensuramos o quê? Podemos medir a temperatura de um paciente e registrar esse valor numérico. Por exemplo, 36,4° C. Se nosso centro médico está, no momento do registro das temperaturas dos pacientes, com 22 internados, teremos 22 valores de temperatura para registrar. Mas também podemos estar interessados, apenas, em classificar os pacientes em estado febril e não febril. Assim, nossos registros terão as palavras “sim”, para indicar os pacientes com febre, e “não”, para indicar os sem febre. Nesta nova situação, não estamos registrando números, mas a “qualidade” do paciente. Veja que podemos contar quantos “sim” e quantos “não” temos por dia.

Vamos imaginar outra situação: poderíamos estar trabalhando na área ortopédica. Estaríamos interessados em classificar e ordenar a escolaridade de nossos pacientes com acidente vascular cerebral, quais possuem o Ensino Fundamental, o Ensino Médio e o Ensino Superior. Veja que, neste caso, é possível ordenar essa classificação: o Ensino Fundamental vem antes do Ensino Médio e este é anterior ao Superior.

As situações apresentadas anteriormente representam as possibilidades de classificação das variáveis que podemos encontrar nas abordagens quantitativas. Veja a Figura 6, a seguir:

Figura 6 – Esquema para classificação das variáveis



Fonte: Produção do próprio autor (2017).

- a) *Variáveis qualitativas nominais*: são variáveis que descrevem qualidades, nomeando categorias. Você está usando uma variável qualitativa nominal quando separa seus elementos em “sacolas” ou conjuntos. Por exemplo: se você fizer um levantamento, entre os leitores de uma biblioteca, quanto ao time de futebol preferido deles, estará trabalhando com uma variável qualitativa nominal. As respostas podem ser: Flamengo, Vasco, Palmeiras, Santos, Corinthians etc. Apesar daquilo que alguns torcedores mais entusiasmados possam julgar, não podemos ordenar os times de futebol (não podemos classificar os times em 1.º, 2.º, 3.º, excetuando a tabela do campeonato). Isso também vale para a variável sexo: só podemos classificar os respondentes (por enquanto) de uma entrevista em masculino ou feminino. Outros exemplos de variáveis qualitativas nominais são: cor dos olhos (azul, verde, castanho); cor do cabelo (loiro, preto, castanho, ruivo); religião (católica, espírita, evangélica...). A divisão das áreas do conhecimento em Ciências Exatas, Ciências Humanas e Ciências Biológicas também é um exemplo de variável qualitativa nominal.

- b) *Variáveis qualitativas ordinais*: são variáveis cujos valores possíveis admitem algum tipo de ordenação. São exemplos de variáveis qualitativas ordinais: uma classificação do tamanho de bibliotecas em pequenas, médias e grandes; classes sociais (A, B, C, D, E) e categorias de classificação de peso (magro, normal, sobrepeso, obeso, obesidade mórbida).
- c) *Variáveis quantitativas contínuas*: usualmente, quando efetuamos *medidas*, os resultados são variáveis quantitativas contínuas. Por exemplo: o peso (em kg) de uma pessoa, a distância entre duas cidades (em km) e o tempo dispendido para a execução de determinada tarefa.
- d) *Variáveis quantitativas discretas*: são variáveis resultantes de *contagens*. Tipicamente, mas não obrigatoriamente, são números inteiros. São exemplos de variáveis quantitativas discretas: o número de livros em uma biblioteca, sua quantidade de usuários, o número de funcionários de uma empresa e o número de gols tomados pelo Corinthians no último campeonato brasileiro.

Você deve estar se perguntando: e por que precisamos perder tempo classificando variáveis? Qual a importância disso? Calma, caro aluno e cara aluna! A classificação das variáveis é importante para estruturarmos nossos questionários ou investigações em bancos de dados. Além disso, você verá a importância da classificação das variáveis quando estudarmos os testes de hipóteses. Existem testes de hipóteses que só podem ser aplicados para variáveis quantitativas discretas, enquanto outros só podem ser aplicados a variáveis qualitativas.



2.3.3 Atividade

Esta atividade atende aos objetivos a e f: “redigir seu problema de pesquisa de forma precisa” e “utilizar as principais ferramentas de análise de dados em pesquisas quantitativas”.

Atividade de classificação: classifique as variáveis seguintes em quantitativas discretas, quantitativas contínuas, qualitativas nominais ou qualitativas ordinais.

Variável	Classificação
Número de usuários em uma biblioteca por semana	
Escolaridade dos usuários (Superior incompleto, Superior completo, Mestrado, Doutorado, Pós-Doutorado)	
Sexo (masculino/ feminino)	
Renda em R\$ (reais)	

Resposta comentada

Número de usuários em uma biblioteca por semana: resulta de contagens, sendo, portanto, variável quantitativa discreta.

Escolaridade dos usuários: a escolaridade é uma qualidade ou atributo dos usuários. Assim, é variável qualitativa. Como podemos ordenar a escolaridade (Ensino Fundamental < Ensino Médio < Ensino Superior), é variável qualitativa ordinal.

Sexo: o sexo constitui uma qualidade ou atributo. Como é um nome (como se fosse um “rótulo”), é uma variável qualitativa nominal.

Renda: é uma variável na qual o salário é mensurado. Assim, trata-se de variável quantitativa contínua.

Observe que categorizar corretamente suas variáveis será fundamental para que, posteriormente, você faça uso dos *softwares* estatísticos. Por exemplo, o *Statistical Package for Social Sciences (SPSS)*, um *software* estatístico largamente utilizado no mundo inteiro, exige que você defina os tipos de variáveis antes de iniciar seu trabalho.

A definição de qual ferramenta de análise estatística você utilizará depende, entre outras questões, do tipo de variável que está sendo investigado. Por exemplo, como veremos em seções mais à frente, ainda nesta Unidade, o teste de hipóteses qui-quadrado é voltado para variáveis qualitativas.

2.4 O DELINEAMENTO DA PESQUISA: SUA “ESTRUTURA”

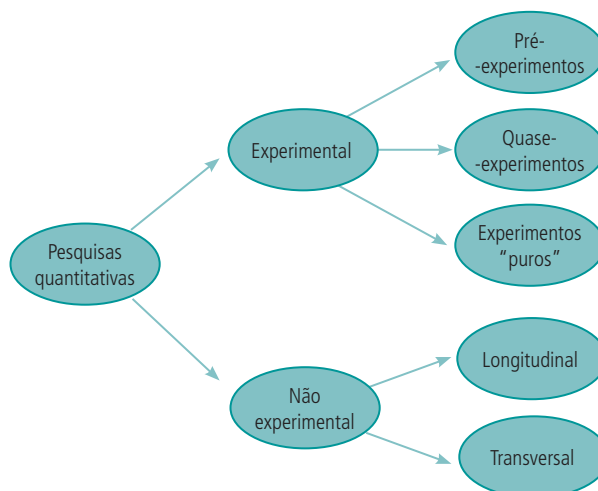
Você já formulou o seu problema de pesquisa. Mas como ela será desenvolvida? Como é a “estrutura” da sua pesquisa quantitativa? É para responder a essas perguntas que discutiremos o desenho ou delineamento da pesquisa.

As pesquisas quantitativas podem ser divididas em dois tipos, a que frequentemente denominamos de desenhos de pesquisa ou delineamentos de pesquisa: as pesquisas experimentais e as não experimentais. Observe que não cabe afirmar que os desenhos experimentais são “superiores”, “melhores” ou mais “científicos” que os não experimentais. Você vai utilizar cada desenho (Figura 7) em função do seu problema de pesquisa, dos objetivos que tem com ela e das suas hipóteses.

Semestre

7

Figura 7 – Tipos de desenhos de pesquisas quantitativas



Fonte: Sampieri (2003).

2.4.1 Estudos experimentais

Estudos *experimentais* realizam experimentos (se eu não te contasse isso, você não teria percebido, não é verdade?), ou seja, você vai alterando os valores de uma ou mais variáveis independentes, *de forma controlada*, para, posteriormente, analisar o resultado (mensurando-o, de alguma forma) em uma ou mais variáveis dependentes. Isso é um *experimento*. Estudos experimentais são adequados quando temos uma forte suspeita de relação causa-efeito entre as variáveis independentes (causa) e as variáveis dependentes (efeito). *Campbell e Stanley* (1966 apud SAMPIERI, 2010) apresentaram uma clássica categorização dos desenhos de pesquisas experimentais. Estudos experimentais são indicados para avaliar a eficiência de diferentes tratamentos ou ações preventivas. Nesta seção, seguiremos a proposta desses autores.



Multimídia

Este *link* apresenta um vídeo que faz a distinção entre estudos experimentais e observacionais: <https://www.facebook.com/PortalAction/?ref=page_internal>.

Vamos assisti-lo?

2.4.1.1 Notação para estudos experimentais

Para sintetizar os estudos experimentais, os pesquisadores acabaram por convencionar um conjunto de símbolos específicos. Veja, no Quadro 1, os símbolos utilizados em desenhos experimentais:

Quadro 1 – Símbolos utilizados na representação de estudos experimentais

Grupo 1 = G1; Grupo 2 = G2 ...	São os grupos nos quais a amostra é dividida
X	Representa o tratamento (pode ser a presença de algum estímulo ou valor da variável independente)
O	Representa a observação ou medição efetuada nos sujeitos de um grupo
–	Representa o nível “base” ou zero da variável independente (é a ausência de estímulo)
R	Indica seleção aleatória (aleatorizada, ou seja, por sorteio)

Fonte: Produção do próprio autor (2017).

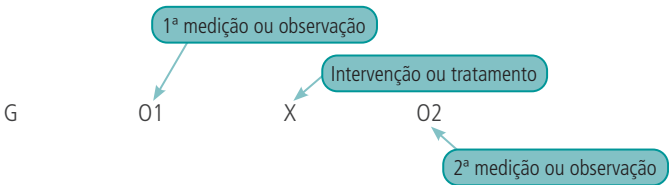
G1 pode representar o grupo dos estudantes de Pós-Graduação que utilizam determinada biblioteca e G2 pode representar o grupo dos professores da universidade.

X pode representar a presença (ou ausência) de ruído em uma biblioteca, em um estudo de qualidade de serviço. Em outras palavras, a notação X indica a presença do tratamento ou causa sob investigação (será que a “aplicação” do ruído na biblioteca afeta a qualidade do serviço?).

O pode representar a medição de índice de acertos em um teste de atenção, sob condições de presença/ ausência de ruído. Se existe mais de uma medição ou observação, cada uma dessas medições/ observações estará associada a um número. Por exemplo, indica-se o pré-teste como O1 e o pós-teste como O2. A seguir, você verá alguns exemplos de uso dessa notação.

- a) Exemplo 1 da simbologia dos desenhos experimentais (Figura 8): neste exemplo, vamos considerar a situação de um pré-teste (com apenas um grupo) seguido de pós-teste:

Figura 8 – Exemplo 1 da simbologia dos desenhos experimentais

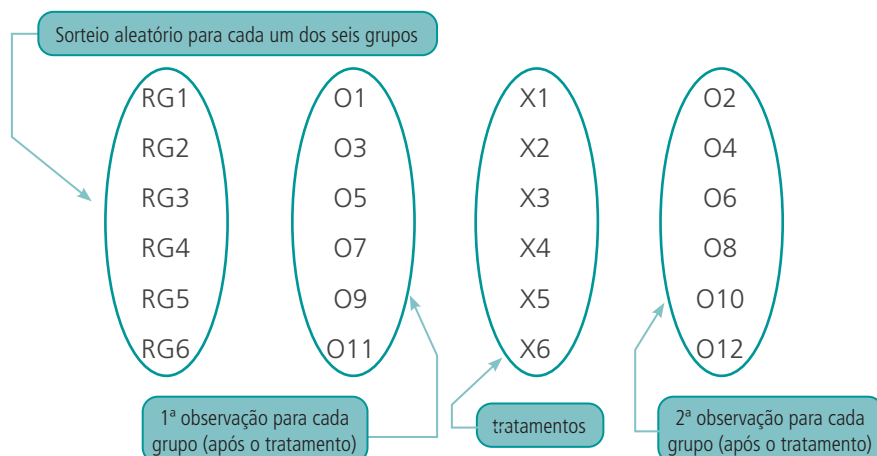


Fonte: Produção do próprio autor (2017)

Neste caso, o grupo G é submetido a um tratamento ou intervenção (que se representa pela letra X) e, posteriormente, é realizada uma observação (medição) da variável (ou variáveis) de interesse para o estudo.

- b) Exemplo 2 da simbologia dos desenhos experimentais (Figura 9): neste outro exemplo, supomos que temos seis grupos e seis intervenções; com observações pré-teste e pós-teste, o diagrama fica assim:

Figura 9 – Exemplo 2 da simbologia dos desenhos experimentais



Fonte: Produção do próprio autor (2017).



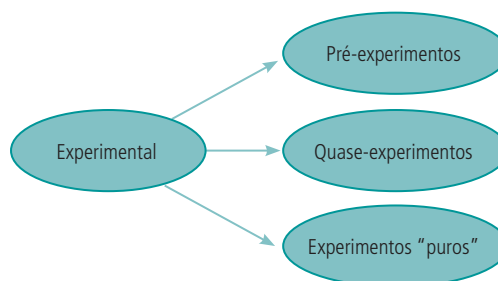
Explicativo

Não temos espaço, neste texto, para apresentar todos os possíveis desenhos de pré-experimentos, quase-experimentos ou experimentos. Caso você tenha necessidade de se aprofundar nesse item, por favor, consulte *Cano* (2004) e *Mallhotra* (2006), ou ainda *Sampieri, Collado e Lucio* (2014).

Seguindo a classificação de *Campbell e Stanley* (1966 apud *SAMPIERI, 2013*), os desenhos experimentais são subdivididos em três categorias: pré-experimentos, experimentos "puros" e quase-experimentos.

A seguir, temos uma figura apresentando a classificação dos tipos de estudos experimentais:

Figura 10 – Estudos experimentais: pré-, quase- e experimentos "puros"



Fonte: Produção do próprio autor (2017).

Na sequência, apresentamos uma rápida caracterização de cada um desses tipos de desenhos experimentais.

- a) *Pré-experimento*: em um estudo pré-experimental, o grau de controle das variáveis é mínimo. Por exemplo, uma situação na qual temos apenas um único grupo com pós-teste é considerada um pré-experimento. Não temos um grupo de comparação.

O desenho de grupo único com pós-teste é assim simbolizado:

X O

Aplica-se o tratamento e, na sequência, acontece a observação. Note que associamos a realização de cada evento da pesquisa ao tempo, da esquerda para a direita: em primeiro lugar, efetuamos o tratamento (X) e, depois, a observação (O). Recomenda-se observar os resultados obtidos em estudos pré-experimentais com precaução. Tais estudos são considerados vulneráveis no que diz respeito ao estabelecimento de relações causais. Segundo *Sampieri*, “alguns autores consideram que (estudos pré-experimentais) devem ser utilizados apenas como treino de outros experimentos com controle maior” (2010, p. 156. Inserção nossa).

- b) *Quase-experimento*: em um quase-experimento, o investigador não possui o controle total sobre o tempo ou o lugar dos eventos. Em tais estudos, não temos aleatorização e, de forma geral, não se utilizam as expressões para determinação do tamanho da amostra. O investigador encaminha o estudo como uma experiência, dividindo os sujeitos de pesquisa em grupos. É a dificuldade de se efetuar um sorteio aleatório dos sujeitos nos grupos que costuma ser associada com um quase-experimento. Os grupos são formados *a priori*, independentemente do estudo. Um exemplo de um quase-experimento seria um estudo no qual os grupos de usuários de uma biblioteca estão separados por grandes áreas, por exemplo, Ciências Exatas, Ciências Humanas e Ciências Biológicas.
- c) *Experimentos “puros”*: os experimentos denominados puros contam com duas características:
- apresentam grupos de comparação;
 - têm equivalência entre os grupos.

Vamos entender um pouco melhor esses dois itens.

- Por que os grupos de comparação são necessários?

Em um experimento, precisamos de grupos de comparação. São necessários, pelo menos, dois deles. Com apenas um grupo, não conseguiríamos saber se alguma variável independente (ou, eventualmente, desconhecida) estaria ou não gerando “ruído” nas nossas conclusões. A comparação é importante para provar se um grupo influencia o outro.

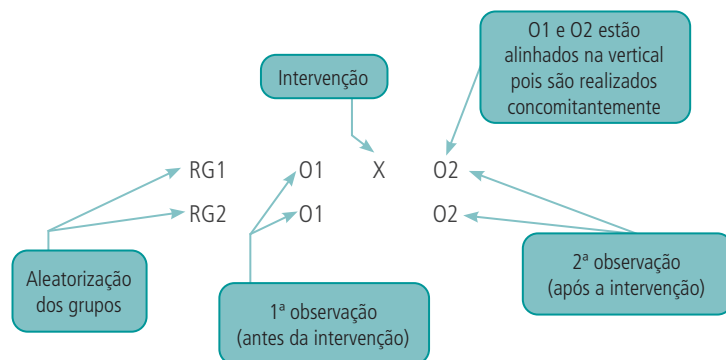
- Como garantir a equivalência entre os grupos?

A equivalência entre eles é obtida com grupos que sejam similares em tudo, exceto na manipulação da(s) variável(is) independente(s). O pesquisador “controla” tudo, menos a variável de intervenção. Assim, se você for testar a eficiência de determinada ferramenta de recuperação de informação, o mais adequado é que separe os usuários entre os de Graduação e aqueles de Pós-Graduação (os dois grupos, nitidamente, não são similares – os estudantes deste nível já têm maior experiência que os daquele).

A equivalência entre os grupos, no início do experimento, pode ser realizada por aleatoriedade (sorteio) ou por emparelhamento. No primeiro caso, a equivalência é realizada por meio de um sorteio (ou seleção aleatória) dos elementos participantes do experimento. É essa aleatoriedade que vai garantir que os dois (ou mais) grupos sejam similares. Para efetuar esse sorteio, utilizaremos as técnicas de amostragem, que serão vistas mais adiante. No segundo caso, a partir de alguma variável específica, serão construídos pares em cada um dos grupos.

A seguir (Figura 11), você tem um exemplo de desenho experimental: é um desenho com pré-teste, pós-teste e grupo controle. Trata-se de um desenho bastante frequente, em que temos aleatoriedade e observações realizadas antes e após a intervenção. A letra *R* indica aleatorização.

Figura 11 – Simbologia usada em desenhos experimentais



Fonte: Produção do próprio autor (2017).

O pesquisador vai avaliar se há mudança significativa entre o grupo experimental (aquele que recebe a intervenção *X*) e o controle (que não recebe a intervenção – veja que não existe a letra *X* na 2ª linha). Em princípio, a única diferença entre os dois grupos é a presença-ausência da intervenção *X*. Se houver, a diferença será atribuída à intervenção. Estamos assumindo que a randomização dos sujeitos de pesquisa garante que a única diferença entre os grupos seja devida à intervenção.

Um profissional de Tecnologia da Informação pode estar interessado em estudar se determinado modelo de construção de *sites* proporciona melhor recuperação da informação. Por exemplo, um modelo pode ser o *site* com determinado conjunto de *links* de acesso no canto superior esquerdo da tela (seria uma das intervenções); outro modelo de *site* seria este conjunto de *links* no canto inferior direito da tela. Podem, ainda, existir outros modelos de disposição da informação. O pesquisador está interessado em saber a proporção de acesso que recupera a informação desejada mais rapidamente, mensurando os tempos de resposta dos grupos participantes da pesquisa. Ele pode formular um desenho de pesquisa na forma do último que vimos, só que com grupos em números proporcionais para executar a tarefa de recuperação da informação.

Como sugestão de leitura complementar sobre pesquisa experimental, você pode consultar *Malhotra* (2006).

Uma crítica feita aos métodos experimentais é a de que o ambiente de laboratório nem sempre reproduz fielmente como os fenômenos ocorrem na vida real. Outro problema com os métodos experimentais ocorre

quando lidamos com seres vivos (principalmente quando os sujeitos da pesquisa são seres humanos!), já que a questão ética está envolvida nos experimentos. Por exemplo, é ético estudar a influência da baixa ingestão de alimentos em seres humanos, deixando, deliberadamente, pessoas sem alimentos? É ético avaliar o impacto de determinados fármacos sobre doenças, infectando deliberadamente os sujeitos da pesquisa? Ou ainda, no caso de bibliotecas: não capacitar o usuário para o acesso à informação, em algumas fontes de informação, para comparar os resultados dos trabalhos, é um gesto ético?



Curiosidade

Quando o pesquisador está interessado em examinar as mudanças nas atitudes dos entrevistados individuais, é interessante avaliar a utilização do *estudo de quatro grupos de Solomon*. Nesse tipo de desenho de pesquisa, as limitações dos desenhos de grupos de controle pré-teste e pós-teste e dos grupos-controle apenas pós-teste são ultrapassadas. Você poderá encontrar mais informações sobre o delineamento de quatro grupos de *Solomon* em *Sampieri, Collado e Lucio* (2013, p. 161).



Explicativo

Quer exemplos de pesquisa quantitativa experimental? A área de Marketing possui pesquisas bastante interessantes, facilmente encontradas na internet. Por favor, consulte o *link* a seguir: <http://www.anpad.org.br/abrir_pdf.php?e=MTY3MTY=>>. Acesso em: 17 out. 2021.

O artigo apresenta o arcabouço de uma pesquisa experimental. Nele, você terá contato com alguns conceitos que veremos mais à frente: teste ANOVA, p-valor de um teste estatístico, testes de hipóteses e rejeição de uma hipótese.

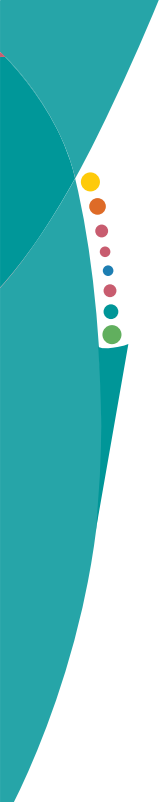
Outro exemplo de uma pesquisa experimental, ainda no campo do Marketing, é o estudo de *Lima* (2012).

2.4.2 Estudos não experimentais

Claro que estudos *não experimentais* não realizam variações *controladas* das variáveis independentes. Os estudos não experimentais costumam ser classificados em longitudinais (isto é, a coleta de dados é realizada ao longo do tempo) ou transversais (a coleta é realizada em um único momento no tempo). Uma analogia para estudos transversais é a

Semestre

7



ideia de uma fotografia do movimento das moléculas de água em um rio. No estudo transversal, efetuamos um “corte” do “rio” em determinado momento. Os estudos transversais ainda costumam ser subdivididos em: exploratórios, descritivos ou correlacionais.

Nos estudos não experimentais, observamos e mensuramos as variáveis da forma como elas ocorrem em campo, sem a modificação de seus valores de modo intencional por parte do pesquisador. Certamente, esta é uma restrição dos estudos não experimentais – é possível observamos a associação entre duas variáveis, mas pode existir uma outra variável independente que seja aquela verdadeiramente responsável pelo efeito na variável dependente. Veja: você observa (é uma pesquisa observacional!) que, em determinada cidade, aumentou o número de cegonhas. Você também observa que aumentou o número de nascimentos nessa cidade. Podemos concluir que cegonhas causam nascimentos? Ou seria mais razoável imaginar que, nos meses anteriores à chegada das cegonhas, naquela cidade, por ter a economia fortemente correlacionada com a safra de milho, soja ou algum outro produto, os maridos voltam de seus trabalhos temporários em outras regiões do país? Fica como sugestão que você pesquise mais sobre o que são correlações espúrias.

Vejamos essa questão de maneira mais aprofundada, com exemplos.

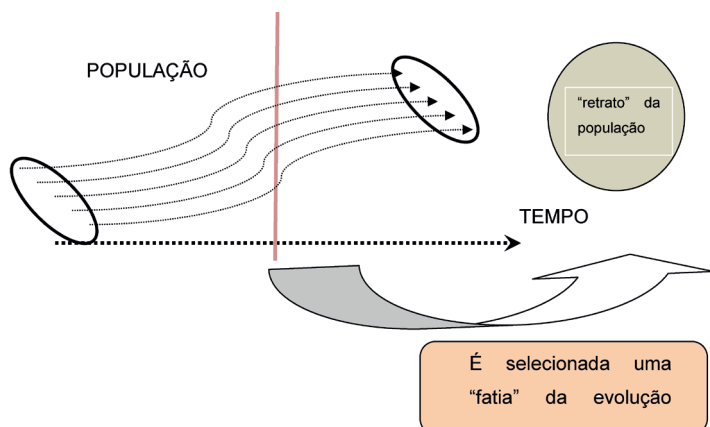
- a) *Primeiro exemplo*: um exemplo frequente para esta situação seria a mensuração da temperatura em uma sala por dois termômetros: um graduado na escala *Celsius* e outro, na escala *Fahrenheit*. As duas medições estão fortemente correlacionadas. No entanto, não é a temperatura medida na escala *Celsius* que causa aumento na temperatura medida na outra escala, ou vice-versa. Na verdade, há uma outra variável, a temperatura da sala, que causa a variação nos dois termômetros.
- b) *Segundo exemplo*: aumentou o número das vendas de determinado doce. Acredita-se, inicialmente, que a variável causadora dessa mudança tenha sido a propaganda. Mas antes de se afirmar isso, é necessário que outras explicações sejam excluídas. As vendas de doce podem ter aumentado devido a fatores como: menos propaganda da concorrência, mudança de preferência dos consumidores ou uma melhor rede de distribuição do produto.
- c) *Terceiro exemplo*: digamos que aumentou a consulta a determinado título em uma das bibliotecas de uma universidade (suponha que tenha sido um título de literatura brasileira da primeira metade do século XX). Ao mesmo tempo, aumentou o número de alunos matriculados no curso de Letras. Pode-se imaginar que a variável responsável pelo aumento no número de consultas seja o aumento do número de aluno matriculados no curso de Letras. Contudo, outras variáveis podem explicar o fenômeno: os títulos de literatura, antes dispersos pela rede de bibliotecas da universidade, foram todos centralizados em uma única biblioteca, ou, alternativamente, os professores que lecionam determinada disciplina alteraram as referências bibliográficas solicitadas aos alunos.

2.4.2.1 Estudos não experimentais transversais

Para ilustrar a ideia de um estudo transversal, veja a Figura 12, a seguir. Os comportamentos da população vão variando ao longo do tempo (são

as flechas se modificando). Então, em determinado momento, nós “tiramos uma fotografia” da população. Será nosso estudo transversal. Em outras palavras: em um estudo transversal, a informação é coletada em um único instante no tempo.

Figura 12 – Fotografia do fluxo de uma pesquisa transversal

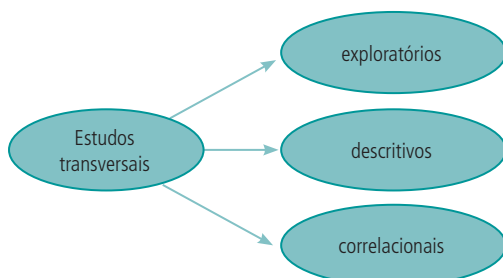


Fonte: Produção do próprio autor (2017).

Quais são as vantagens de um estudo transversal? Entre elas, podem ser citadas: custo relativamente baixo, realização mais simples dos levantamentos em relação a outros delineamentos e coleta de dados mais rápida. Por outro lado, estudos transversais não mostram a influência da variável tempo no que está sendo estudado (já que estamos medindo em um mesmo instante). Também não são úteis para a investigação causal. Mesmo assim, estudos transversais são proveitosos para a formulação de hipóteses. Contudo, como os estudos transversais são uma “fotografia” de determinado instante no tempo, eles não permitem descobrir se a intervenção veio antes ou depois da causa. Estudos transversais são utilizados, na área de Saúde, para fins epidemiológicos. Assim, é bastante razoável que, por analogia, bibliotecários utilizem estudos transversais para seus estudos “epidemiológicos”. Por exemplo, a “propagação” de determinado comportamento cultural pode ser entendida como uma “epidemia”. Um bibliotecário poderia obter um retrato dos seus usuários em relação ao conhecimento de determinadas fontes de informação para definir um programa de capacitação dirigido. Nesse caso, um delineamento transversal seria adequado.

Segundo *Sampieri* (2010), estudos transversais podem ser subdivididos em: exploratórios, descritivos ou correlacionais (Figura 13):

Figura 13 – Subdivisão dos estudos transversais



Fonte: Produção do próprio autor (2017).

2.4.2.2 Estudos não experimentais longitudinais

Já nos estudos longitudinais, é feito um acompanhamento ao longo do tempo. O objetivo desses estudos é aprender como as variáveis mudam no decorrer do tempo. É como se “embarcássemos” em uma daquelas flechas e a seguíssemos ao longo de um intervalo de tempo (que é a duração do estudo). Uma amostra da população é medida várias vezes no decorrer do tempo, coletando-se dados ao menos em dois instantes. Dessa forma, estudos longitudinais permitem avaliar (e quantificar) o comportamento de uma ou mais variáveis no período. Esses estudos podem ser classificados, ainda, em prospectivos (quando os sujeitos do estudo são acompanhados ao longo do tempo) e retrospectivos (os dados são provenientes de períodos passados). Analisar a frequência dos alunos de um curso de Graduação à biblioteca seria um exemplo de estudo longitudinal retrospectivo nesse ambiente. A medida poderia ser feita analisando um determinado dia por ano, ao longo dos quatro anos do curso.

2.4.2.3 Estudos não experimentais correlacionais

Neste tipo de estudo descritivo, buscamos avaliar a força de relação entre variáveis quantitativas ou qualitativas. Por exemplo, uma informação correlacional é a de que a renda média das pessoas aumenta (está positivamente correlacionada) com os seus anos de estudo. Comumente, utiliza-se o diagrama de dispersão para apresentar graficamente a informação da correlação. Mais à frente, trataremos de diagramas de dispersão com detalhes.



2.4.3 Atividade

Esta atividade atende aos objetivos a e b: “redigir seu problema de pesquisa de forma precisa” e “delinear pesquisas experimentais e não experimentais”.

Você é um dos bibliotecários em uma rede de bibliotecas universitárias, onde está sendo discutida a possibilidade de implantação de novos processos de gestão. Após reuniões, implantou-se um novo *software* para a gestão da rede de bibliotecas. Construa uma sequência de diálogos justificando a adoção de um desenho de pesquisa não experimental longitudinal para avaliar o desempenho das equipes da rede de bibliotecas, a partir da adoção do novo *software*.

Assuma dois pontos de vista na construção do seu diálogo: o do autor da proposta de pesquisa e o do gestor dos recursos financeiros na universidade. Como autor da proposta, você deve buscar convencer o gestor de recursos financeiros da universidade de que uma pesquisa quantitativa não experimental longitudinal apresentará subsídios adequados para a tomada de decisão concernente ao problema da universidade, justificando o investimento. Sua argumentação deve ser técnica e fundamentada nos elementos expostos nesta Unidade, ou em elementos que você encontrar na literatura técnica. Já no papel de gestor dos recursos financeiros, você vai questionar o autor da proposta acerca das vantagens e desvantagens de outros desenhos de pesquisa para o fim proposto. Procure, no seu diálogo,

responder às questões: qual a dificuldade para se adotar um desenho experimental? Por que um desenho longitudinal seria adequado? Qual seria a vantagem, neste caso, de um desenho longitudinal, em comparação com um desenho transversal?

Resposta comentada

Espera-se que seu diálogo apresente comparações entre desenhos experimentais e longitudinais. Um desenho experimental exigirá que você selecione as pessoas a serem observadas, efetue sorteios aleatórios dos grupos (assegurando-se da não tendenciosidade das amostras), efetue o experimento (tratamento) e observe seus resultados. Tudo isso exige tempo, dinheiro e uma equipe condutora do experimento. Em um desenho longitudinal prospectivo, você poderia instalar o *software* em um ou dois terminais mantidos em cada uma das bibliotecas-piloto (assim você conseguiria obter respostas de usuários de todas as bibliotecas) e acompanhar as reações dos usuários ao longo de um semestre. A equipe exigida, nesse caso, seria bem menor, com consequentes reduções de tempo e custos. Com os relatos das reações dos usuários, melhorias e adaptações no *software* poderão ser desenvolvidas com a participação dos próprios usuários.

Esses aspectos devem ser levados em conta quando você estiver redigindo seu problema de pesquisa. Pergunte-se: será que este problema de pesquisa poderá ser respondido no tempo, com a equipe e os recursos financeiros e de infraestrutura de que disponho? Quais são os elementos críticos (gargalos) para o sucesso no desenvolvimento da sua pesquisa?

2.5 AMOSTRAGENS: VOCÊ JÁ FEZ EXAME DE SANGUE? AMOSTRAGEM E VÍCIO

Para começar esta discussão, propomos, ainda, outra provocação: Você já foi entrevistado em alguma pesquisa de opinião? Por exemplo, em uma pesquisa eleitoral? Uma crítica rotineira feita pelas pessoas que não conhecem a Estatística e suas técnicas é desqualificar as pesquisas com alegações como: “eu nunca fui entrevistado por nenhum instituto de pesquisas!”. Outra alegação que talvez você já tenha ouvido é “todos os meus amigos do *Facebook* não vão votar no candidato X para presidente. Logo, é impossível ele estar à frente nas pesquisas!”. Qual é a crítica à última alegação? Se você for uma mulher, na faixa dos 20-30 anos, cursando nível Superior, na zona urbana da Região Sul do Brasil, grande parte de seus contatos no *Facebook* serão outras mulheres com perfil parecido. Em

Semestre

7

outras palavras: em termos estatísticos, todas vocês são “a mesma pessoa”. Não é necessário entrevistar todas essas mulheres com perfil parecido. Não precisamos entrevistar todas as pessoas com características semelhantes. Em média, você e suas amigas devem apresentar respostas muito parecidas nos questionários.

Outras situações nas quais fazemos amostragens são os exames de sangue, aos quais nos referimos na abertura desta parte do curso, ou os de urina. Uma pessoa que não aceita as técnicas de amostragem para pesquisas eleitorais deveria solicitar ao seu laboratório de análises clínicas que não seja extraída uma “amostra” de seu sangue, mas que seja retirado todo ele (brincadeira, é claro!).

Figura 14 – Amostragem de alimentos



Fonte: Pixabay (2014).⁴

Quando trabalhamos com amostragens de preparos culinários, extraímos uma amostra (parte) do conteúdo da panela para inferir a adequação do tempero ao prato como um todo.

Ao efetuarmos pesquisas, é frequente a utilização de amostras, com o intuito de generalizarmos os resultados para um conjunto mais amplo de indivíduos, sem termos que efetuar medições de todos os elementos da população. Uma das razões que justificam a utilização de amostras em vez de se entrevistar a população toda é o custo: é muito mais barato entrevistar apenas dois ou quatro por cento da população do que entrevistá-la em sua totalidade. Mas, mesmo que você tivesse recursos econômicos para entrevistar toda a população, outro aspecto deveria ser considerado: o tempo de processamento dos questionários e a tabulação dos dados é muito maior quando lidamos com a população. Claro que devemos levar em conta o tamanho da nossa população-alvo. Se ela é constituída apenas de 100 indivíduos (ou seja, é um grupo relativamente pequeno), é mais simples trabalharmos com o grupo inteiro do que fazer amostragem.

Quando examinamos todos os indivíduos de uma população ou universo, estamos fazendo um censo, e não uma amostragem. Cabe ressaltar: em que situações se recomenda um censo (em detrimento da amostragem)? Quando temos populações relativamente pequenas ou quando se exige uma precisão extremamente alta nos dados (o que acabaria por nos conduzir a amostras relativamente grandes em relação à população). Um exemplo famoso de censo são os censos demográficos que o *Instituto Brasileiro de Geografia e Estatística (IBGE)* realiza a cada dez anos.

⁴ Disponível em: <<http://pixabay.com/pt/kebab-de-espeto-de-carne-417994/>>.

A partir do senso comum, boa parte das pessoas acredita que censos são sempre mais precisos que amostras, mas isso não é verdade! Uma amostragem bem conduzida pode produzir resultados mais confiáveis que um censo mal feito. Amostragens costumam necessitar de menos tempo para que os dados sejam processados e o relatório seja publicado, permitindo a análise pública das informações.

2.5.1 Definição de população

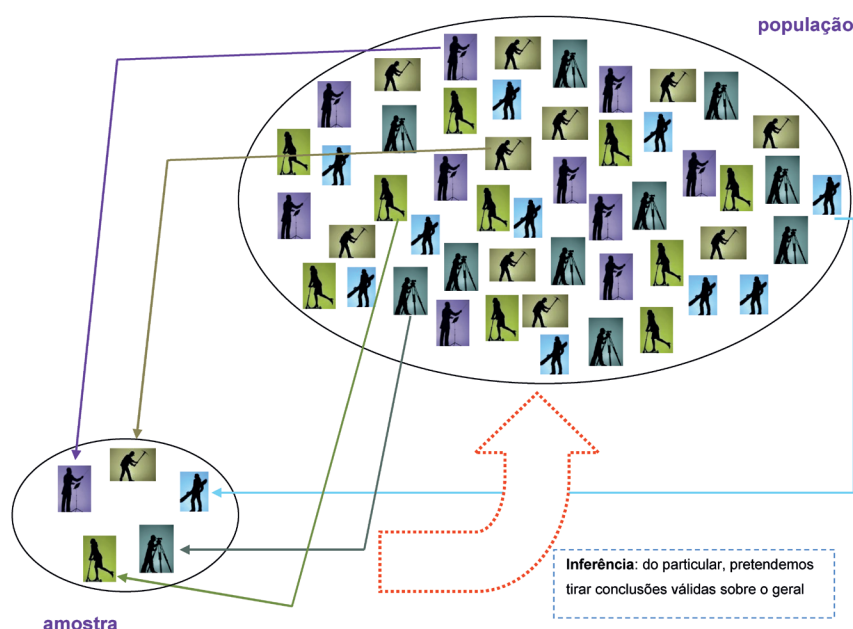
Uma população é um conjunto de indivíduos com uma ou mais características em comum, que se pretende estudar. Indivíduos, aqui, podem ser pessoas, mas também podem ser livros de uma biblioteca ou o conjunto de artigos publicados em periódicos Qualis A.

Podemos falar na população de alunas universitárias brasileiras ou na população de operações de cadastramento de usuários em uma biblioteca. As características que definem a população devem estar descritas no seu trabalho de pesquisa. Assim, se você estiver interessado em realizar uma pesquisa sobre como crianças matriculadas em escolas públicas de regiões com vulnerabilidade social respondem a determinado projeto de inclusão com bibliotecas de bairro, a delimitação da população do seu estudo deve estar descrita no trabalho. Observe que o trabalho de pesquisa não será tão melhor quanto maior for a população, mas sua qualidade estará na clara delimitação da população, tendo em vista o problema de pesquisa em estudo.

2.5.2 Definição de amostra

Uma amostra é qualquer parte da população. O objetivo da amostragem em pesquisas quantitativas é obter conjuntos de elementos representativos tais que permitam a generalização para a população toda. A dificuldade é conseguir que as amostras sejam representativas da população. As amostras são classificadas em dois tipos: amostras probabilísticas e amostras não probabilísticas.

Figura 15 – Ideia de amostragem



Fonte: Produção do próprio autor (2017).

Amostra representativa

As proporções da população devem ser as mesmas da amostra. Assim, se a população contiver 13% de mulheres com Ensino Superior incompleto, a amostra também deverá apresentar essa proporção. Se a população apresentar 18% de homens negros com renda individual acima de quatro salários mínimos, a amostra deverá ser representativa desse aspecto da população. Por favor, consulte o verbete amostra tendenciosa, que está algumas páginas à frente. Leia, também, o estudo de caso referente ao problema de amostra tendenciosa das eleições para presidente dos Estados Unidos da América (EUA), em 1936.

Inferência estatística

O objetivo da inferência estatística é tirar conclusões válidas para a população como um todo, a partir de uma amostra representativa da população.

Semestre

7

2.5.3 Amostras não probabilísticas

Nas amostras não probabilísticas, a seleção dos elementos constituintes da amostra não depende de probabilidade, mas de algum aspecto ou característica da pesquisa ou de quem a realiza. As técnicas de amostragem não probabilística mais utilizadas são: amostragem por julgamento, casual (ou por conveniência) e amostragem por quotas.

- *Amostragem não probabilística por julgamento*: nesse tipo de amostragem, dependemos da capacidade de avaliação do próprio pesquisador ou de algum especialista contratado para selecionar a amostra por julgamento. Esse especialista, com base em sua prévia experiência profissional, vai selecionar aqueles elementos da população que julga serem mais adequados para o propósito da pesquisa. Exceto se você estiver trabalhando com uma população inteiramente homogênea, esta técnica de amostragem não permitirá a generalização dos resultados.
- *Amostragem não probabilística casual ou por conveniência*: neste tipo de amostragem, os elementos são selecionados em função de estarem acessíveis ou disponíveis no instante e no local em que a pesquisa tenha sido realizada. Exemplo: imaginemos que você vai entrevistar as pessoas que estão passando em frente a uma biblioteca, na hora do almoço, em um dia de semana. Claramente, você vai obter uma amostra enviesada. Em outras palavras: nesta técnica de amostragem, você não vai obter respostas daquelas pessoas que pouco frequentam a biblioteca da sua cidade – pessoas que trabalham nos bairros mais afastados do centro, aposentados, crianças ou jovens em idade escolar (tipicamente, existem poucas escolas nos centros das cidades). Uma variante da amostragem por conveniência é a amostragem por autosseleção, caso típico das atuais pesquisas pela internet. São os respondentes que definem se vão participar ou não da pesquisa; eles se “autosselecionam”. Dificilmente, pesquisas com autosseleção serão representativas da população.

Amostra enviesada (tendenciosa)

Um processo de amostragem apresenta viés (ou tendência) se, sistematicamente, tende a extrair elementos da população com determinado perfil, e a não extrair elementos da população com outros perfis.



Um estudo de caso famoso sobre viés na amostragem é a pesquisa realizada pela revista norte-americana *Literary Digest*, para as pesquisas presidenciais dos Estados Unidos da América de 1936. Tal estudo conduziu uma pesquisa telefônica no país e concluiu (equivocadamente) que *Alfred Landon* (do Partido Republicano) derrotaria *Franklin D. Roosevelt* (do Partido Democrata). A amostra utilizada foi uma grande listagem de participantes, que incluía pessoas que possuíam linha telefônica, pessoas que assinavam alguma revista e, ainda, pessoas inscritas em associações profissionais ou algum clube.

Foi gerada uma base com aproximadamente 10 milhões de nomes (uma amostra bastante grande). Para cada pessoa da lista, a revista enviou uma cédula fictícia de votação, solicitou-lhe que anotasse sua preferência na cédula e que a retornasse para a revista. Com essa pesquisa, a *Literary Digest* concluiu que a vitória seria de *Landon*, com 57% dos votos, contra 43% de *Roosevelt*. Na verdade, *Roosevelt* ganhou de *Landon* de 62% a 38%. Em função desse enorme erro estatístico, a credibilidade da revista despencou, o que acabou por afetar suas vendas de tal forma que ela encerrou suas atividades.

De forma notável, *George Gallup* (o mesmo do *Instituto de Pesquisas Gallup*) previu a eleição de *Roosevelt* com uma “pequena” amostra de 50.000 pessoas, cerca de 200 vezes menor que a amostra utilizada pela *Literary Digest*. Lição importante: nem sempre uma amostra grande significa uma amostra boa!

Quais foram os erros de amostragem da revista *Literary Digest*?

Essa amostra selecionou pessoas que possuíam linha telefônica (certamente, poucas pessoas possuíam renda suficiente para manter uma linha telefônica – apenas aquelas com uma renda mais elevada, fato que já introduz um viés na amostra). Lembre-se de que, em 1936, os EUA (e grande parte do mundo, com exceção dos países socialistas), ainda sofriam os efeitos do desemprego e da recessão econômica de 1929, quando quebrou a *Bolsa de Nova Iorque*. Assim, assinar revistas e participar de clubes eram atividades apenas dos mais ricos, de modo que a amostra da *Literary Digest* não era representativa da população! Essa amostra apresentava forte *viés de seleção*. Além desse problema, a revista também apresentou *viés de não resposta*: o pesquisador deve saber que, entre todas as pessoas para quem ele envia um questionário, nem todas responderão à sua pesquisa. Muitas vezes, quem não responde são aquelas pessoas que não possuem uma posição bem definida a respeito do assunto. No caso da pesquisa da *Literary Digest*, o viés de não resposta foi elevado: cerca de 24% das pessoas que receberam a cédula não a devolveram preenchida.

- *Amostragem não probabilística por quotas (ou proporcional)*: neste tipo de amostragem, o especialista responsável divide a amostra em quotas proporcionais, segundo determinado critério. De alguma forma, o pesquisador possui informações prévias sobre a população em questão (por exemplo, ele pode utilizar dados do censo demográfico do IBGE), acerca das proporções de homens e mulheres, relacionando a população urbana x a rural e as proporções na população, seja por faixas de renda, seja de escolaridade. Suponha que o pesquisador tenha a informação de que a porcentagem de mulheres em determinado município é de 52,3%. Então, ele utilizará essa informação na construção de suas amostras, que devem conter 52,3% de mulheres. Na realidade, em muitas situações, o pesquisador deverá dividir sua amostra em mulheres (e homens) por faixa de renda e escolaridade. Anteriormente, você conheceu um estudo de caso sobre a eleição presidencial dos EUA em 1936, na qual foi utilizada, sem sucesso, a amostragem por quotas. Você verá que é bastante problemático conduzir amostras por esse método. Devemos levar em conta o sexo, a renda, a escolaridade, o bairro,... Até qual variável poderemos ir? Para pesquisas epidemiológicas, o referido tipo de estudo, também chamado de estudo sociodemográfico, é essencial.

2.5.4 Amostras probabilísticas

Nas amostras probabilísticas, a seleção dos elementos é baseada na probabilidade de seleção. Nelas, os elementos da população possuem, *a priori*, probabilidade conhecida e não nula de fazerem parte da amostra. Os quatro tipos de amostragem probabilística mais utilizados são: amostragem aleatória simples, amostragem sistemática, amostragem estratificada e amostra probabilística por conglomerados.

- a) *Amostragem aleatória simples*: na amostragem aleatória simples, utilizamos algum procedimento aleatório para escolher uma amostra com certo número de elementos pré-determinados “n”. As amostras aleatórias simples são abreviadas por AAS e todas possuem igual probabilidade de serem sorteadas. As AAS podem ser feitas com reposição do elemento extraído ou sem reposição desse elemento.

Como escolhemos uma amostra aleatória simples?

Uma forma bastante fácil para se entender o processo de amostragem aleatória simples é imaginar que você colocou os nomes (ou algum código que identifique cada um dos elementos da população) em uma urna e sorteou, aleatoriamente, a quantidade “n” para constituir sua amostra. É a mesma ideia do sorteio de um bingo ou loteria.

Certamente, se o valor da população for muito grande, esse procedimento não será prático. Nesses casos, utilizamos um gerador de números aleatórios ou uma tabela deles. Os números aleatórios podem ser gerados computacionalmente, com o uso de *softwares* estatísticos, como o *IBM SPSS Statistical Package* ou o *Excel*, da *Microsoft*.

Você pode encontrar facilmente tabelas de números aleatórios na internet ou nos livros de Estatística. Você deve associar cada elemento de sua população a um número. Se sua população possui 1.000 elementos, os números começariam em 000 e iriam até 999. A partir de qualquer ponto da tabela, não importa se na horizontal ou na vertical, você vai agrupando os números de três em três (pois sua população possui 1.000 elementos, associados a números com três dígitos). Você deve prosseguir até selecionar o tamanho desejado de elementos para sua amostra.

A planilha eletrônica *Excel* disponibiliza uma função para gerar números aleatórios. São as funções *ALEATÓRIO* e *ALEATÓRIO ENTRE* (inferior; superior). Para acessar essas funções, você deve clicar na aba *Fórmulas* → *inserir Função* e, na categoria *Estatística*, deve selecionar a função *ALEATÓRIO*, ou a função *ALEATÓRIO ENTRE* (inferior; superior). A função *ALEATÓRIO* apresenta um número aleatório entre 0 e 1, distribuído uniformemente. Já a função *ALEATÓRIO ENTRE* (inferior; superior) retorna um número aleatório no intervalo específico. Experimente você mesmo. Assim, você vai adquirindo maior familiaridade com as funções estatísticas do *Excel*.



2.5.5 Atividade

Esta Atividade atende ao objetivo *f*: “utilizar as funções estatísticas básicas do *software Microsoft Excel*”.

Pesquise, no *software Microsoft Excel*, para que servem as seguintes funções estatísticas: *CONT.SE*, *CONT.SES*, *CURT*, *DESPAD.A*, *MED*, *MÉDIA*.

Resposta comentada

Para descobrir o que cada função faz, você deve clicar na aba "Fórmulas". Depois, clique em "Inserir função". Vá em "Selecione uma categoria", abra a listagem e escolha "Estatística". Você vai clicando nas funções dadas. Vão aparecer os textos abaixo:

- a) CONT.SE: "Calcula o número de células não vazias em um intervalo que corresponde a uma determinada condição" (texto do Excel).

Quando você deve usar esta função? Por exemplo, quando quiser contar, nas células de A1 até A100, quantas vezes aparece a sigla da *Universidade Federal do Rio de Janeiro (UFRJ)*: CONT.SE (A1:A100; "=UFRJ").

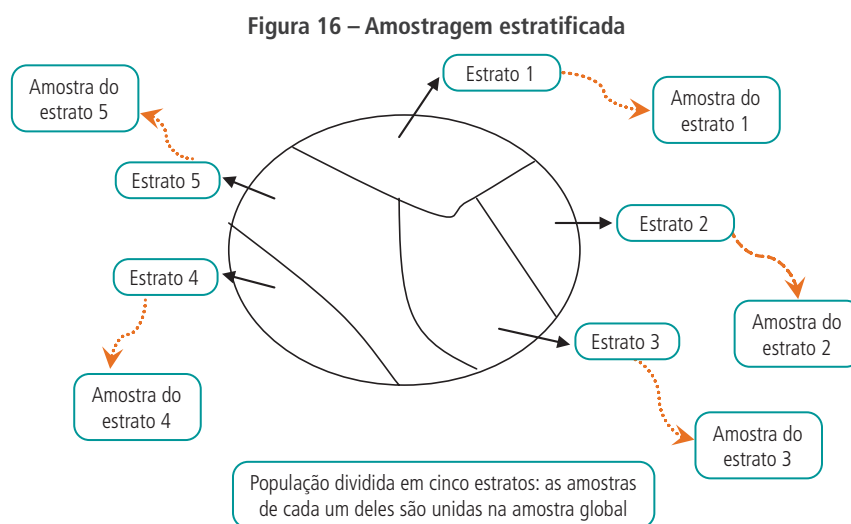
Observe que, se a coluna estiver preenchida com os valores *UFRJ* e *ufrj*, a função vai retornar o valor 2;

- b) CONT.SES: você vai usar esta função para contar se vários critérios, em várias colunas (ou linhas), estão satisfeitos simultaneamente. Suponha que as colunas A, B e C contenham siglas de universidades. Considere a função =CONT.SES (A1:A100; "UFRJ"; B1:B100; "UFMG"; C1:C100; "UFPR"). Esta função vai retornar um valor adicional para cada linha das colunas A, B e C que contenha a sequência *UFRJ*, *UFMG* e *UFPR* (nessa sequência). Se, nas colunas A, B e C, aparecerem seis vezes a sequência *UFRJ*, *UFMG* e *UFPR*, então a função, tal como escrita anteriormente, vai retornar o valor 6;
- c) CURT: retorna a curtose de um conjunto de dados. A curtose é uma medida estatística de achatamento de uma distribuição contínua de dados, comparada com a distribuição normal. Curtoses positivas indicam que a distribuição possui uma elevação maior que a normal. Curtoses negativas indicam que a distribuição é mais "achatada" que a distribuição normal;
- d) DESVPA.A: esta função calcula o desvio-padrão amostral, ignorando os valores lógicos e o texto da amostra;
- e) MED: retorna a mediana, que é o valor que deixa 50% dos seus dados abaixo dele e 50% dos dados acima dele;
- f) MÉDIA: calcula a média de um conjunto de valores numéricos.

- b) *Amostragem probabilística sistemática*: neste tipo de amostragem, temos o cadastro dos entrevistados em uma lista. É uma amostragem que "pula" um certo número de elementos de forma sistemática. Temos o ponto de início da amostragem e o seu intervalo. Por exemplo, se o intervalo for 10, vamos escolher para nossa amostra um elemento a cada 10 da população. Como se obtém o intervalo de amostragem? Dividimos o tamanho da população pelo tamanho pretendido para a nossa amostra. No contexto dado, para definir o ponto de início da amostragem, escolhemos um número qualquer entre 1 e 10. Por exemplo, suponhamos que nossa população seja a das casas em uma avenida, um total de 400 casas. Se foi definido que nossa amostra deve ter tamanho igual a 20, então o intervalo

será $400/20 = 20$. Agora, sorteamos um número qualquer entre 1 e 20. Digamos que o número sorteado tenha sido o 7. Então, a casa de número 7 será a primeira da nossa amostra. A próxima casa é a casa de número $7 + 20 = 27$. Prosseguindo, nossa amostra será constituída das residências de números {7, 27, 47, 67, 87, 107, 127, 147, 167, 187, 207, 227, 247, 267, 287, 307, 327, 347, 367, 387}.

- c) *Amostragem probabilística estratificada*: na amostragem probabilística estratificada (Figura 16), divide-se a população em estratos (grupos com características semelhantes). Os estratos podem ser definidos por critérios demográficos ou geográficos (zona rural e zona urbana; cidades pequenas, médias e grandes; renda; faixa etária; sexo; escolaridade etc). Os estratos devem apresentar grande homogeneidade, também podendo ser subdivididos em estratos menores. Após dividirmos a população nos vários estratos, é possível efetuar uma amostra aleatória simples em cada um deles. Em seguida, unimos todas essas AAS para constituir a amostra geral.



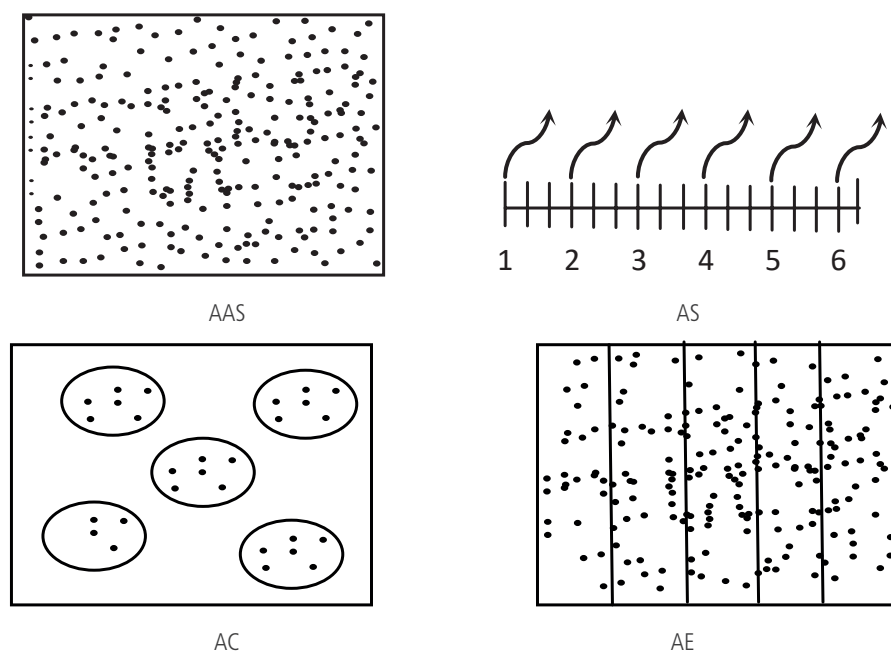
Fonte: Produção do próprio autor (2017).

- d) *Amostragem probabilística por conglomerados*: a amostragem por conglomerados é adequada para situações nas quais a população pode ser separada em “pacotes”, grupos ou conglomerados semelhantes. Retiramos amostras aleatórias simples dos conglomerados em vez das unidades da população. Cada um dos conglomerados é “semelhante” ao outro. Por exemplo, considere a seleção de domicílios em um município; os conglomerados podem ser os bairros da cidade. Com um mapa, selecionamos uma AAS de 2% dos bairros da cidade. Após isso, escolhemos os domicílios no bairro. Observe que amostras estratificadas trabalham com cada um dos estratos (queremos efetuar comparações entre eles). Já na amostragem por conglomerados, o que fazemos é trabalhar com amostras dos conglomerados, não com todos eles. Em estudos sobre a satisfação dos usuários de bibliotecas públicas municipais, poderíamos selecionar uma

amostragem aleatória simples e, daí, coletar as informações dos usuários nas unidades sorteadas. A amostragem por conglomerados é mais econômica que a amostragem aleatória simples quando a população está espalhada em uma área geográfica bastante extensa. Não é necessário que tenhamos uma lista de toda a população (como na AAS). Precisamos, sim, ter uma lista dos conglomerados e outra dos elementos, dentro de cada conglomerado sorteado.

A Figura 17 busca apresentar, de forma gráfica, as ideias principais associadas às amostragens aleatória simples (AAS), sistemática (AS), por conglomerado (AC) e estratificada (AE):

Figura 17 – Amostragens aleatórias: simples, sistemática, por conglomerado e estratificada



Fonte: Adaptado de Agresti; Finlay (2012).



Explicativo

Para você conhecer alguns problemas que surgem na obtenção da amostra para uma pesquisa, sugerimos a consulta a esta tese:

WALTER, M. T. **Bibliotecários no Brasil**: representações da profissão. 2008. 345 f. Tese (Doutorado) – Departamento de Ciência da Informação e Documentação, Universidade de Brasília, 2008. p. 131-135.



2.5.6 Atividade

Esta atividade atende ao objetivo c: “descrever as técnicas básicas de amostragem utilizadas em estudos quantitativos”.

Suponha que, em determinada cidade, a porcentagem de homens seja de 48%. Sabe-se, a partir de um censo realizado recentemente, que a população desse local se divide em cinco classes econômicas, da seguinte forma: classe A, com 5% da população; classe B, com 12% da população; classe C, com 23% da população; classe D, com 40% da população, e classe E, com o restante da população da cidade. Suponha que cada classe econômica se divide, no tocante ao sexo, mantendo a proporção de 48% para 52%. Suponha ainda que a amostra contere 0,0345% da população da cidade. Determine quantos homens e mulheres, de cada classe econômica, deverão constar na amostra.

Resposta comentada

Lembre-se de que, em primeiro lugar, precisamos determinar o tamanho total da amostra, para só então determinarmos, a partir das proporções 48% para homens e 52% para mulheres, as respectivas quantidades de homens e mulheres nela presentes. Em seguida, determinamos as quantidades de indivíduos selecionados por amostra (para homens e para mulheres). É importante que você efetue somas parciais para verificar seus cálculos. Assim, o resultado da soma *homens da classe A + homens da classe B + homens da classe C + homens da classe D + homens da classe E* tem que ser igual ao total de homens. Não é incomum descobrirmos falhas nos cálculos por meio desses procedimentos simples de verificação.

A amostra terá $0,0345\% \times 2.200.000 = 759$ entrevistados. Destes, serão $48\% \times 759 = 364,32$ homens e 394,68 mulheres. Multiplicando a quantidade de homens pelas respectivas porcentagens para cada classe socioeconômica, teremos: homens da classe A = $37,95 = 5\% \times 48\% \times 759 = 37,95$; homens da classe B = 91,08; homens da classe C = 174,57; homens da classe D = 303,6; homens da classe E = 151,8. Para as mulheres, os números são: classe A = 19,734; classe B = 47,36; classe C = 90,77; classe D = 157,87; classe E = 78,93.

2.6 DETERMINANDO O TAMANHO DA AMOSTRA

Após determinar o tipo de amostragem a ser utilizado, devemos definir o tamanho da amostra. Veremos, a seguir, como fazer isso com uma amostra aleatória simples e com uma amostra aleatória estratificada. A determinação do tamanho da amostra é um problema complexo, que não se restringe às questões internas da pesquisa. Muitas vezes, não temos como efetuar a amostra desejável, por questões financeiras ou de prazo. O pesquisador deve levar em consideração essas outras dimensões, obedecendo às restrições do ponto de vista estatístico, é claro. No entanto, é importante saber que um dos fatores que atribui validade a um estudo é a forma como a amostra foi definida e o tamanho que ela tem, bem como se esse tamanho atende aos critérios exigidos para a pesquisa.

2.6.1 Tamanho da amostra para amostra aleatória simples

Primeiramente, precisamos definir alguns conceitos:

- a) *parâmetro*: é alguma característica que descreve um aspecto da *população*. Podem ser tomados como parâmetros: a média da renda, a mediana das alturas, o desvio-padrão (desvio-padrão é uma medida do quanto os dados estão dispersos com relação à média) do número de livros retirados de uma biblioteca, entre outros. Os parâmetros representam os valores verdadeiros da população, mas, em geral, desconhecemos esses valores. Nosso objetivo com a inferência estatística é, justamente, a partir de uma amostra, estimar um valor para determinado parâmetro da população;
- b) *estatística*: refere-se a alguma característica na amostra. Podemos falar na média da amostra, mediana da amostra, desvio-padrão amostral etc;
- c) *estimador*: é a “fórmula” que operacionaliza o cálculo. É a sequência de cálculos para estimar o valor do parâmetro, feita com os dados da amostra;
- d) *estimativa*: é o valor numérico do estimador de um parâmetro. É o resultado dos cálculos, usando os dados da amostra, na expressão do estimador. Tradicionalmente, usamos letras gregas para descrever parâmetros em uma população e letras latinas na descrição de estimadores em uma amostra. Veja no Quadro 2, a seguir:



Quadro 2 – Parâmetros e estimadores

Estatística	Parâmetro populacional	Estimador (amostra)
Média	μ	$\bar{X}$
Desvio-padrão	σ	s
Proporção	π	p
Diferença entre a média de duas populações	$\mu_1 - \mu_2$	$\bar{X}_1 - \bar{X}_2$

Fonte: Produção do próprio autor (2017).

- e) *proporção*: proporção é apenas outra palavra para porcentagem. Por exemplo, se falamos que a proporção de alunos de Pós-Graduação entrevistados em uma pesquisa é de 1 para 4, estamos falando na porcentagem de 25% de alunos de Pós-Graduação em relação ao total. A proporção amostral é a fração dos indivíduos com determinada propriedade. Em símbolos:

$$p = \frac{\text{quantidade de indivíduos com a característica estudada}}{\text{tamanho da amostra}}$$

- f) *erro amostral*: é a diferença entre o valor verdadeiro do parâmetro (valor desconhecido na vida real) e o valor calculado pelo estimador. O pesquisador utiliza um limite aceitável (ou tolerável) no erro amostral para determinar o tamanho mínimo da amostra. Em outras palavras: determina-se o tamanho da amostra de tal forma que o erro amostral fique dentro do que se considera aceitável.

As estimativas podem ser de duas naturezas: estimativas pontuais ou estimativas intervalares.

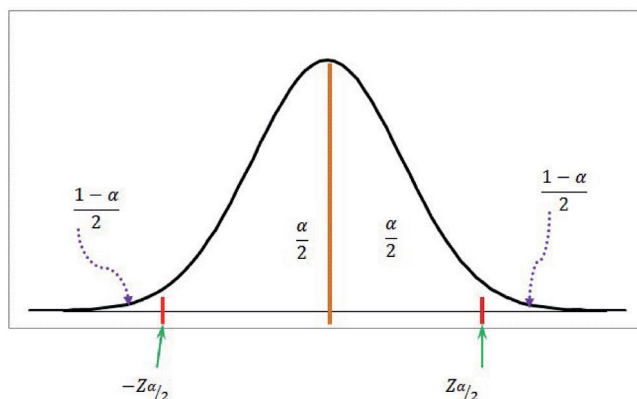
- a) *Estimativas pontuais*: são aquelas estimativas efetuadas apenas para um único valor (ou ponto). Tais estimativas têm um único valor numérico como resultado.
- b) *Estimativas intervalares*: construímos um intervalo de confiança que, com uma probabilidade definida pelo pesquisador, irá conter o valor verdadeiro do parâmetro. Diferentemente das estimativas pontuais (nas quais temos apenas um valor numérico único), nas estimativas intervalares temos um conjunto infinito de pontos, que contém, com probabilidade específica, o valor verdadeiro do parâmetro. É bastante frequente (apesar de não ser obrigatório) que os valores de confiança adotados por pesquisadores sejam de 1%, 5% ou, no máximo, 10%. O pesquisador é livre para escolher seu nível de confiança aceitável. Mas quanto mais confiança ele desejar no seu intervalo, maior terá que ser o tamanho da amostra. Consequentemente, mais cara e demorada será a pesquisa. Há um “balanço” a ser definido pelo pesquisador.
- c) *Nível de precisão*: em alguns textos de Estatística, você encontrará, no capítulo sobre intervalos de confiança, simplesmente *erro*: é o tamanho do intervalo de confiança. Em outras palavras: é a distância máxima admitida pelo pesquisador entre o valor estimado e o valor do parâmetro.

- d) *Coeficiente de confiança*: chamamos de coeficiente de confiança a probabilidade de um intervalo de confiança aleatoriamente selecionado conter o parâmetro da população, se o processo de amostragem for repetido diversas vezes. Vamos utilizar a letra grega α (alfa) para representar a confiança.
- e) *Nível de confiança*: vejamos, nos três casos a seguir, como construímos intervalos de confiança. Na sequência, veremos como estimar o tamanho de uma amostra.

No primeiro caso, partiremos da hipótese de que temos uma amostra com mais de 30 elementos e que conhecemos o desvio-padrão populacional.

- *Caso 1: intervalo de confiança para a média (com desvio-padrão populacional σ conhecido)*: considere uma população com distribuição de probabilidade normal (você estudou a distribuição de probabilidade normal na disciplina “Estatística Básica”), com média populacional igual a μ e desvio-padrão populacional conhecido σ . Define-se o erro-padrão da média como $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$. Vamos supor, para facilitar nossos exemplos numéricos, que a confiança adotada pelo pesquisador seja $\alpha = 95\%$. Isso significa que vamos construir um intervalo em que, caso repetamos a amostragem 100 vezes, o esperado é que em 95 delas o intervalo contenha o valor verdadeiro do parâmetro.

Figura 18 – Intervalos de confiança para uma distribuição normal



Fonte: Produção do próprio autor (2017).

Associados ao valor de confiança $\alpha = 95\%$, temos os valores $Z_{\alpha/2}$ e $-Z_{\alpha/2}$. Esses dois valores são localizados a partir da tabela de distribuição normal padronizada, buscando-se, no corpo da tabela, o valor $\alpha/2 = 0,4750$. O valor de $Z_{\alpha/2}$ será 1,96. Assim, a expressão para o intervalo de confiança para a média será:

$$IC(\text{média}; \alpha) = \left[\bar{x} - Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}; \bar{x} + Z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right].$$

O intervalo de confiança para a média pode ser escrito também na forma: $\bar{x} \pm Z_{\alpha/2} \sigma_{\bar{x}}$. Vamos apresentar um exemplo numérico: Você entrevistou 130 pessoas, com o objetivo de obter uma estimativa para a renda média mensal. Por pesquisas anteriores,

você foi informado de que o desvio-padrão nessa população é igual a R\$70,00. A média obtida na sua amostra (média amostral) foi igual a R\$832,00. Apresente o intervalo de confiança para a média, com 95% de confiança.

Resolução: determinamos o erro-padrão da média:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{70}{\sqrt{130}} = 6,1394.$$

Como foi solicitada confiança de 95%, teremos $Z_{\alpha/2} = 1,96$. Então, nosso intervalo de confiança será:

$$IC(média; 95\%) = [832 - 1,96 \cdot 6,1394; 832 + 1,96 \cdot 6,1394] = [819,96; 844,03].$$

Observe que a amplitude desse intervalo é de $844,03 - 819,96 = 24,07$. Ou seja, a máxima distância entre o valor estimado e o valor real (e desconhecido) da média populacional é metade da amplitude do intervalo, que totaliza 12,035.

Como os valores de confiança $\alpha = 99\%$, 95% e 90% são bastante frequentes na prática, verifique a Tabela 1:

Tabela 1 – Valores de confiança alfa e respectivos valores críticos para $Z_{\alpha/2}$

α	$Z_{\alpha/2}$
90%	1,65
95%	1,96
99%	2,56

- *Caso 2: intervalo de confiança para a média (amostra maior que 30 elementos), com desvio-padrão populacional, σ desconhecido:* como faremos se não soubermos o valor do desvio-padrão populacional? Nessa situação, calculamos o desvio-padrão amostral a partir dos dados da amostra. O intervalo de confiança fica assim:

$$IC(média; \alpha) = \left[\bar{x} - Z_{\alpha/2} \frac{S}{\sqrt{n}}; \bar{x} + Z_{\alpha/2} \frac{S}{\sqrt{n}} \right],$$

onde $S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$ é o desvio-padrão amostral.

Para a determinação do intervalo de confiança para a média, em uma situação em que você estiver trabalhando com uma amostra com menos de 30 elementos e o desvio-padrão não for conhecido, você não poderá mais trabalhar com a distribuição normal. Terá que utilizar a distribuição *t-Student*. Consulte outros textos de Estatística para saber como proceder nessa situação. Sugerimos dois títulos: *Levine; Berenson; Stephan* (2000) e *Dancey* (2006).

- *Caso 3: intervalo de confiança para proporções (com proporção populacional conhecida):* a expressão matemática para o intervalo de confiança para proporções é:

$$IC(\text{proporção}; \alpha) = \left[p - Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}; p + Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}} \right],$$

onde p é a proporção da característica estudada na amostra.

Veja um exemplo numérico: considere que uma empresa prestadora de serviços está de posse de uma amostra aleatória com 320 arquivos. A empresa identificou que 21 deles foram produzidos com algum tipo de erro. Construa um intervalo de confiança (com 99% de confiança) para estimar a proporção de arquivos com algum tipo de erro na população.

Nesse exemplo, temos: $n = 320$, $p = 21/320 = 0,06563$, $\alpha = 99\%$ (e, portanto, $Z_{\alpha/2} = 2,56$). O intervalo de confiança fica assim:

$$IC(\text{proporção}; \alpha) = \left[p - Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}; p + Z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}} \right] =$$

$$\left[0,06563 - 2,56 \cdot \sqrt{\frac{0,06563(1-0,06563)}{320}}; 0,06563 \right.$$

$$\left. + 2,56 \cdot \sqrt{\frac{0,06563(1-0,06563)}{320}} \right] =$$

$$[0,06563 - 2,56 \cdot 0,01384; 0,06563 + 2,56 \cdot 0,01384] = [0,03019; 0,10106].$$

A amplitude desse intervalo é de 0,07087. A distância máxima entre o valor estimado e o valor populacional corresponde à metade da amplitude do intervalo estimado.

2.6.2 Tamanho da amostra para médias

Retome o caso 1 (intervalo de confiança com uma amostra de tamanho 130, desvio-padrão populacional igual a R\$70,00 e média amostral igual a R\$832,00). A amplitude do intervalo de confiança foi igual a 24,07. Suponha que você tenha decidido que aceita um erro (nível de precisão) de, no máximo, R\$6,00. Qual terá que ser o tamanho mínimo da amostra para que o nível de precisão seja de R\$6,00? É o que veremos agora.

Esse erro ou nível de precisão é a distância máxima entre o valor da média amostral e o valor do parâmetro na população. Neste exemplo, esse valor é a média populacional μ (desconhecida).

Queremos impor a restrição: $|\bar{x} - \mu| \leq 6$. A expressão para a distribuição normal padronizada, com a fórmula para o erro-padrão da média, é:

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}. \text{ Teremos, então, que (isolando o "n", tamanho da amostra):}$$

$$Z_{\alpha/2} = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \Rightarrow \frac{\sigma}{\sqrt{n}} = \frac{\bar{x} - \mu}{Z_{\alpha/2}} \Rightarrow \frac{\sqrt{n}}{\sigma} = \frac{Z_{\alpha/2}}{\bar{x} - \mu} \Rightarrow \sqrt{n} = \sigma \cdot \frac{Z_{\alpha/2}}{\bar{x} - \mu}.$$



Finalmente, elevando os dois lados da igualdade ao quadrado, temos:

$$n = \sigma^2 \cdot \frac{\left(Z_{\alpha/2}\right)^2}{(\bar{x} - \mu)^2}.$$

Se escrevermos o nível de precisão (erro máximo admitido) como $E = \bar{x} - \mu$, a fórmula para o tamanho mínimo da amostra, no caso de um intervalo de confiança para a média (com valor do desvio-padrão populacional conhecido), será

$$n = \sigma^2 \cdot \frac{\left(Z_{\alpha/2}\right)^2}{E^2}.$$

Vamos resolver nosso exercício?

Temos que o desvio-padrão populacional é igual a R\$70,00, o erro máximo admitido é $E = 12$ e o nível de confiança exigido continua sendo de 95%. Então, $Z_{\alpha/2} = 1,65$. Substituindo os valores na fórmula, temos:

$$n = \sigma^2 \cdot \frac{\left(Z_{\alpha/2}\right)^2}{E^2} = 70^2 \cdot \frac{(1,65)^2}{6^2} = 370,56.$$

Assim, a amostra agora deverá ter, no mínimo, 371 elementos.

Perceba que o tamanho da amostra terá que ser maior à medida que o desvio-padrão da população aumentar (ou seja, quanto maior a dispersão da população, mais difícil é a obtenção de estimativas precisas). O tamanho da amostra terá que ser maior à medida que aumentarmos o nível de confiança (para maiores níveis de confiança, precisamos de amostras maiores). Como o erro = E está no denominador da expressão para o tamanho da amostra, quanto menor o erro admissível, maior será o tamanho da amostra.

Se não soubermos o desvio-padrão populacional, basta substituí-lo pela estimativa do desvio-padrão, que é obtida a partir da amostra

$$n = S^2 \cdot \frac{\left(Z_{\alpha/2}\right)^2}{E^2}.$$

2.6.3 Tamanho da amostra no caso de proporções

O procedimento de cálculo do tamanho da amostra para proporções é parecido com o anterior:

$$n = \frac{p(1-p)\left(Z_{\alpha/2}\right)^2}{E^2}.$$

Considere o exemplo do caso 3: $p = 21/320 = 0,06563$, $\alpha = 99\%$, portanto, $Z_{\alpha/2} = 2,56$. Agora, entretanto, suponha que queiramos determinar o valor de n tal que o erro máximo admitido seja de 0,04. Substituindo os valores numéricos na expressão anterior, temos:

$$n = \frac{0,06563(1-0,06563)(2,56)^2}{0,04^2} = 207,585.$$

Assim, a amostra mínima terá que conter 208 elementos.

Sempre que o tamanho da amostra mínima (tanto para o caso da média quanto para o caso de proporções) for maior que 10% da população, teremos que utilizar o fator de correção de população finita:

$$n_{\text{corrigido}} = \frac{n \cdot N}{N + n - 1},$$

onde N = tamanho da população, n = tamanho da amostra obtida, sem a correção de população finita, $n_{\text{corrigido}}$ é o tamanho da amostra com a correção de população finita.

Se o tamanho da amostra for menor que 10% da população, não será necessário efetuar ajustes nas fórmulas de estimativa para o erro-padrão da média. Mas, se nossa amostra for maior que 10% da população, será necessário multiplicar o erro-padrão da média pelo fator de correção de população finita:

$$\text{FCP finita} = \sqrt{\frac{N - n}{N - 1}},$$

onde N é o tamanho da população e n é o tamanho da amostra. O erro-padrão com esse ajuste fica assim:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{N - n}{N - 1}}.$$

2.7 COLETANDO OS DADOS

Já vimos como classificar as variáveis e definir as amostras. Mas como o pesquisador quantitativo pode obter os valores das variáveis? Em outras palavras: como são os processos de coleta de dados em pesquisas quantitativas? É o que aprenderemos agora.

Em uma biblioteca, podemos aplicar um questionário entre os usuários para conhecer seu perfil socioeconômico. Por exemplo, se estivermos em uma biblioteca universitária, uma categorização importante em uma pesquisa é a que indica se nosso entrevistado é aluno de Graduação, Pós-Graduação, professor ou usuário externo. Dependendo do local de inserção da biblioteca (pode ser a biblioteca da universidade de uma cidade pequena), podemos estar interessados em investigar a idade e os hábitos de uso da internet de nosso público.

Como podemos coletar as respostas para essas questões? Basicamente, usando questionários. Antes de começar a montar seu questionário, vamos listar algumas preocupações, a seguir.

- a) O questionário é um instrumento de medição, certo? Como assegurar que ele mede o que promete medir?



- b) Régua, termômetros, balanças e relógios também são instrumentos de medição. Todos usam escalas. Como são as escalas em um questionário?
- c) Antes de usar um instrumento de medição “na vida real”, devemos efetuar um teste-piloto ou pré-teste, para antecipar eventuais problemas com a nossa “régua”.

Provavelmente, o instrumento de coleta de dados mais utilizado é o questionário. Um questionário é constituído por uma sequência de perguntas com referências às variáveis do estudo em questão. Na construção do seu questionário, você deve, a todo momento, retomar o seu problema de pesquisa, bem como suas hipóteses (com as relações esperadas entre as variáveis), de forma que ele apresente consistência em relação ao seu projeto de pesquisa. Caso contrário, você corre o risco de inserir perguntas inúteis ou redundantes para seu problema de pesquisa. Os questionários podem apresentar perguntas abertas e perguntas fechadas.

Em uma pergunta aberta, o entrevistado, em geral, pode redigir uma resposta livremente. É fácil imaginar que perguntas abertas apresentam maior dificuldade no momento da tabulação. Por outro lado, acabam fornecendo mais informações ao pesquisador.

Mas e quanto às perguntas fechadas? Veja:

- perguntas fechadas com duas possibilidades (dicotômicas). Exemplo:

Você usa o sistema de consulta à biblioteca pela internet?

() Sim () Não

- perguntas fechadas com várias possibilidades de resposta. Podem ser construídas perguntas fechadas para as quais é possível marcar uma única opção, ou perguntas fechadas nas quais é possível marcar mais de uma opção. Exemplo:

Assinale o serviço da biblioteca que você usou no mês passado (você pode marcar mais de uma opção):

() Não utilizei nenhum serviço.

() Utilizei o espaço físico da biblioteca para estudar sozinho.

() Utilizei o espaço físico da biblioteca para estudar em grupo.

() Efetuei retirada de livros.

() Consultei livros, mas sem retirá-los.

() Consultei periódicos.

() Utilizei os computadores da biblioteca para acessar a internet.

() Outros (especifique) _____.

No caso de perguntas fechadas, as respostas devem apresentar todas as opções possíveis;

- perguntas fechadas com pontuação. Você pode estar interessado em hierarquizar as categorias. Para isso, solicita ao respondente que pontue as alternativas de resposta. Exemplo:

Dê uma nota, de 0 (péssimo) a 10 (excelente), para os seguintes aspectos da biblioteca que você utiliza:

- () Qualidade do atendimento
 - () Qualidade dos móveis (mesas, cadeiras)
 - () Silêncio (0 = ambiente com excesso de ruídos e 10 = ambiente silencioso)
 - () Temperatura adequada do ambiente
 - () Quantidade de computadores no ambiente da biblioteca
 - () Horário de atendimento
- é possível construir perguntas fechadas nas quais o respondente é solicitado a anotar um valor numérico. Exemplo:
- Escreva quantas horas semanais você utiliza a internet:
-

Como pudemos perceber, perguntas fechadas são mais fáceis de se responder, tabular e analisar.

Algumas recomendações gerais sobre questionários:

- a) é bastante recomendável que você faça um pré-teste antes de levar sua pesquisa a campo. Avalie quanto tempo os respondentes estão gastando para responder o questionário. Questionários muito longos acabam levando os respondentes a “chutarem” as respostas, por falta de tempo ou interesse;
- b) observe se surgem dúvidas quanto à linguagem/ vocabulário/ redação das questões. A linguagem deve ser adequada ao público da pesquisa. Pode parecer óbvio e caricato, mas um questionário dirigido para crianças e jovens usuários de uma gibiteca é bastante diferente daquele dirigido a professores-pesquisadores em um instituto de pesquisas;
- c) evite perguntar mais de uma coisa em uma mesma questão: você gosta de livros de autoajuda e de livros de romances? Fica impossível analisar se a resposta dada foi afirmativa para a primeira parte da pergunta, para a segunda parte ou para ambas. Divida essas perguntas;
- d) evite perguntas muito longas (redija algo em torno de 20 a 30 palavras, no máximo). Divida as perguntas em questões menores;
- e) perguntas de natureza sociodemográfica devem fazer parte de seu questionário para permitir a caracterização do entrevistado sob esse aspecto. São questões que buscam identificar sexo, idade, escolaridade, classe socioeconômica, profissão etc;
- f) verifique com cuidado se suas perguntas não induzem uma resposta;
- g) evite utilizar negativas nas perguntas. Isso acaba por confundir o entrevistado e o induz a responder negativamente.



Atenção

Como exemplo de pesquisa quantitativa e qualitativa na Biblioteconomia, sugerimos ler o livro de *Leitão* (2005).

2.7.1 Escalas: *Likert*, diferencial semântico e escala de *Guttman*

Em pesquisas quantitativas, estamos interessados na mensuração de atitudes. Como operacionalizar essa mensuração? Um instrumento adequado para isso são as escalas. Veremos as ferramentas mais utilizadas para a mensuração de atitudes.

2.7.1.1 Escala de *Likert*

Na escala de *Likert* (desenvolvida por *Rensis Likert*, em 1932), temos três, cinco ou sete itens para o entrevistado indicar o quanto concorda/discorda de um conjunto de afirmações. A seguir, temos um exemplo de escala de *Likert* com cinco itens:

- a) opções de escolha (da pesquisa):
- () Concordo totalmente.
 - () Concordo.
 - () Discordo.
 - () Discordo totalmente.
 - () Não concordo nem discordo.

A escala de *Likert* é largamente utilizada, com sucesso, em inúmeras pesquisas, há décadas. É simples para os respondentes indicarem concordância/discordância de uma afirmação. Por outro lado, ela possui críticas. Quando usamos escalas com muitos pontos, pode ser difícil para o respondente identificar a diferença entre os conceitos. Veja:

- a) opções de escolha (da pesquisa):
- () Discordo totalmente.
 - () Discordo fortemente.
 - () Discordo parcialmente.
 - () Indeciso.
 - () Concordo parcialmente.
 - () Concordo fortemente.
 - () Concordo totalmente.

2.7.1.2 Escala de diferencial semântico (ou de Osgood)

O criador da escala de diferencial semântico foi *Charles Osgood*. Nela, são atribuídos números às intensidades de uma variável. Os extremos da variável são a base para a escala. São utilizados adjetivos para qualificar o objeto. Por exemplo (Quadro 3):

Quadro 3 – Exemplo 1 da escala de diferencial semântico

O ambiente da biblioteca é								
Barulhento	1	2	3	4	5	6	7	Silencioso

Fonte: Produção do próprio autor (2017).

Alguns exemplos de adjetivos “extremos”: forte x fraco, organizado x desorganizado, puro x impuro, barato x caro, limpo x sujo. A escala de diferencial semântico também pode variar entre -3 e +3 (Quadro 4).

Quadro 4 – Exemplo 2 da escala de diferencial semântico

O atendimento na biblioteca é								
Passivo	-3	-2	-1	0	+1	+2	+3	Ativo

Fonte: Produção do próprio autor (2017).

A escala de diferencial semântico é utilizada em pesquisas de Marketing para comparação de marcas e imagem de empresas, bem como para o desenvolvimento de estratégias de propaganda. Consultar *Malhotra* (2006).



Atenção

Conheça uma situação real, do ponto de vista de uma pesquisa na área de Biblioteconomia, e saiba como foi o processo de criação do questionário, pré-teste e questões abertas x questões fechadas. Para isso, consulte a tese de *Maria Tereza Machado Teles Walter*, referenciada anteriormente, no que diz respeito ao instrumento questionário; leia o conteúdo das páginas 158 a 164.



Atenção

Outras escalas: pode ser interessante, para seu questionário, investigar mais sobre escalas de mensuração de atitudes. Além das já mencionadas, existem as escalas de *Guttman*, *Stapel* (seu criador foi *Jan Stapel*), escalas contínuas, escalas multi-itens (ou de múltiplos itens), escalas balanceadas e escalas não balanceadas, escalas comparativas... Enfim, há muito mais do que podemos abordar no espaço deste texto. Além disso, o livro *Pesquisa de Marketing* é um texto que aprofunda a discussão. Ele é de autoria de Costa (2012).

2.8 TABELAS DE FREQUÊNCIA: ORGANIZANDO E VISUALIZANDO OS DADOS

Neste ponto do trabalho, você já definiu seu problema de pesquisa, concebeu seu delineamento, definiu as variáveis dependentes e independentes, montou seu questionário, efetuou o pré-teste, repensou a quantidade de questões, efetuou a revisão de língua, eliminou a dubiedade na redação das questões, foi a campo realizar as entrevistas e está com os questionários preenchidos na sua pasta de trabalho.

O que fazer agora?

Vamos começar a organizar os dados para prepará-los para a etapa de análise. Não se esqueça de apresentar a fonte dos dados, que deve estar logo abaixo da sua tabela. Esta, por sua vez, deve apresentar um título que contenha: o momento (mês, ano) ao qual a tabela se refere, o local no qual o evento descrito ocorreu e o fenômeno que é descrito.

De acordo com o IBGE (1993, p. 9), tabela “é uma forma não discursiva de apresentar informações, das quais o dado numérico se destaca como informação central, na sua forma identificam-se espaços e elementos”. O título da tabela deve ser digitado acima dela, centralizado. Devem estar em maiúsculas apenas a inicial da frase e os nomes próprios. O título da tabela deve ser digitado logo após a palavra *Tabela*, e eles se separam por hífen.

Deve ser destacada a diferença entre quadros e tabelas. Os quadros possuem os quatro lados fechados, com traços horizontais e verticais para separar as informações. Nas tabelas, os lados direito e esquerdo são abertos e as partes superior e inferior são fechadas. Além disso, não se utilizam traços horizontais nem verticais para a separação dos números. As tabelas devem estar numeradas de acordo com a sua ocorrência no texto.

Tanto de acordo com o IBGE (1993, p. 9) quanto com a NBR 14724:2011, deve ser utilizada a forma *tabela* quando a informação a ser apresentada for numérica. Para as demais informações, recomenda-se a forma *quadro*. Destacamos que a diferença entre quadros e tabelas não consiste simplesmente na formatação (“com as bordas fechadas” x “com as bordas abertas”). Quadros também devem apresentar o título acima deles e a fonte, abaixo.

A seguir (Quadro 5) apresentamos um exemplo de quadro:

Quadro 5 – Quadro de exemplo: dados de universidades, centros universitários e faculdades, 2015

Universidade/ Faculdade	Categoria institucional	Possui Mestrado/ Doutorado?	Nota
Universidade ABC	Pública	Sim	A
Faculdade XYZ	Privada	Não	A
Centro Universitário XPTO	Privada	Não	D
Centro Universitário do Litoral Norte	Pública	Não	C
Faculdades do Sudoeste do Livramento	Privada	Sim	D
Universidade EPSILON	Pública	Não	B
Universidade Pinga Fogo	Privada	Sim	B
Faculdade DEF	Pública	Não	C
Centro Universitário ALPHABETA	Privada	Não	A
Faculdade Pena Branca e Xavantinho	Privada	Sim	B
Universidade GAMMADELTA	Pública	Não	D

Fonte: Produção do próprio autor (2017), com dados fictícios, apenas para fins didáticos.

Não se esqueça de apresentar a fonte dos dados das suas tabelas. Vamos ver o exemplo de uma tabela (Tabela 2)?

Tabela 2 – Disponibilidade de bibliotecas públicas e acervo do município de São Paulo, Subprefeitura de Pinheiros e distritos municipais, 2006

Unidades territoriais	Equipamentos						
	Área urbana (km²)	Requerido ^(f)	Disponíveis	Carência	Acervo	Pop. 15 ou +	Ac./pop
MSP	849,6	120,1	54	68	4.552.034	8.258.471	0,55
SP – Pinheiros	33,7	4,8	1,0	3,8	28.157	236.584	0,12
Alto de Pinheiros	7,7	1,1	-	1,1	-	37.824	-
Itaim Bibi	10,7	1,5	-	1,5	-	70.514	-
Jardim Paulista	7,0	1,0	-	1,0	-	73.524	-
Pinheiros	8,3	1,2	1,0	0,2	28.157	54.721	0,51

Fonte: Secretaria Municipal de Planejamento (SEMPLAI DIPRO) (2008).

Em caso de dúvidas, consulte a *NBR 14724:2011: Informação e documentação – Trabalhos acadêmicos – Apresentação*, ou versão mais recente, se houver.

A organização dos dados é realizada com as tabelas de frequência. Como você estudou esse tópico na disciplina “Estatística Básica”, não o abordaremos aqui. É interessante, neste ponto, você utilizar algum *software* (pode ser a Planilha Excel[®]) para ajudá-lo na construção das tabelas de frequência.

2.9 ANÁLISE DOS DADOS

Após organizarmos os dados em tabelas de frequência (simples, acumulada, relativa e relativa acumulada) e produzirmos os gráficos, se for o caso, podemos aplicar algumas medidas-resumo aos nossos dados. As medidas-resumo costumam ser divididas em duas grandes classes: medidas de tendência central (média, moda e mediana) e medidas de dispersão (variância, desvio-padrão, separatrizes e coeficiente de variação). Com essas medidas-resumo, podemos construir o gráfico *box-plot*.

Você já estudou as medidas de tendência central (média, moda, mediana) na disciplina “Estatística Básica”, no início do curso. Também já estudou as medidas de dispersão: variância, desvio-padrão, coeficiente de variação e separatrizes (quartis, decis e percentis). Portanto, não repetiremos esse conteúdo nesta disciplina.

2.9.1 Correlação e regressão

É comum, nas pesquisas, estarmos interessados em investigar a “força” da associação entre duas variáveis quantitativas. Por exemplo, podemos estar interessados em avaliar o tempo, em minutos, de acesso ao sistema de bibliotecas e a taxa de recuperação da informação desejada. Suponha que a seguinte tabela foi obtida a partir de uma amostragem realizada em uma biblioteca. Os dados são fictícios.

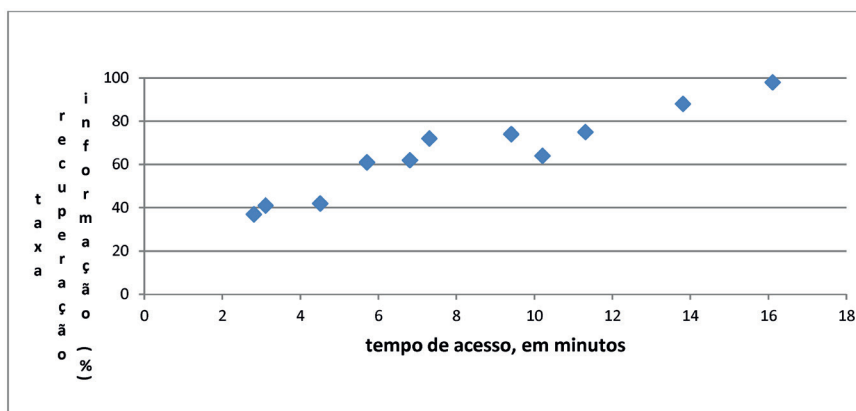
Tabela 3 – Tempo de acesso ao sistema de bibliotecas x taxa de recuperação da informação

Tempo de acesso ao sistema de bibliotecas (em min.)	Taxa de recuperação da informação (em %)
3,1	26
5,7	69
4,5	42
10,2	87
16,1	98
6,8	62
7,3	85
9,4	82
2,8	32
11,3	62
13,8	74

Fonte: Produção do próprio autor (2017), com dados fictícios apenas para fins didáticos.

Em primeiro lugar, vamos construir um diagrama de dispersão. Nele, temos uma das variáveis disposta no eixo das abscissas e a outra no eixo das ordenadas. Veja a Figura 19:

Figura 19 – Diagrama de dispersão: tempo de acesso (min.) x taxa de recuperação da informação (%)

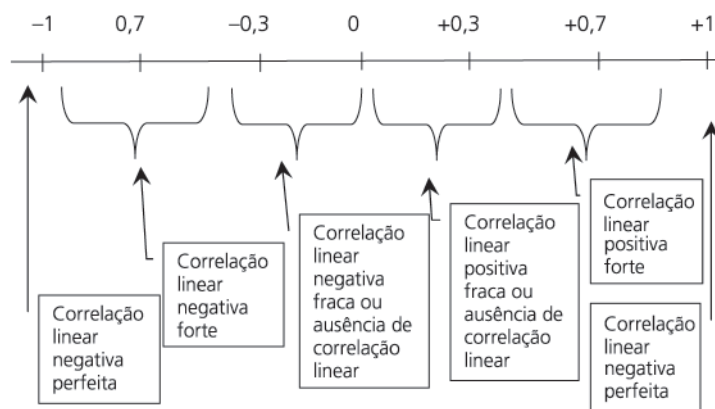


Fonte: Produção do próprio autor (2017).

Observe que as medidas devem ser realizadas sobre os mesmos indivíduos. Cada par de medidas corresponde a um ponto (x, y) no plano cartesiano. Nesse exemplo, temos 11 medidas. É usual que, no eixo das abscissas, sejam marcados os valores da variável independente e, no eixo das ordenadas, os valores da variável dependente. Após construir o diagrama de dispersão, devemos observar como se dispõe a “nuvem de pontos” e se existem pontos discrepantes (possíveis *outliers*).

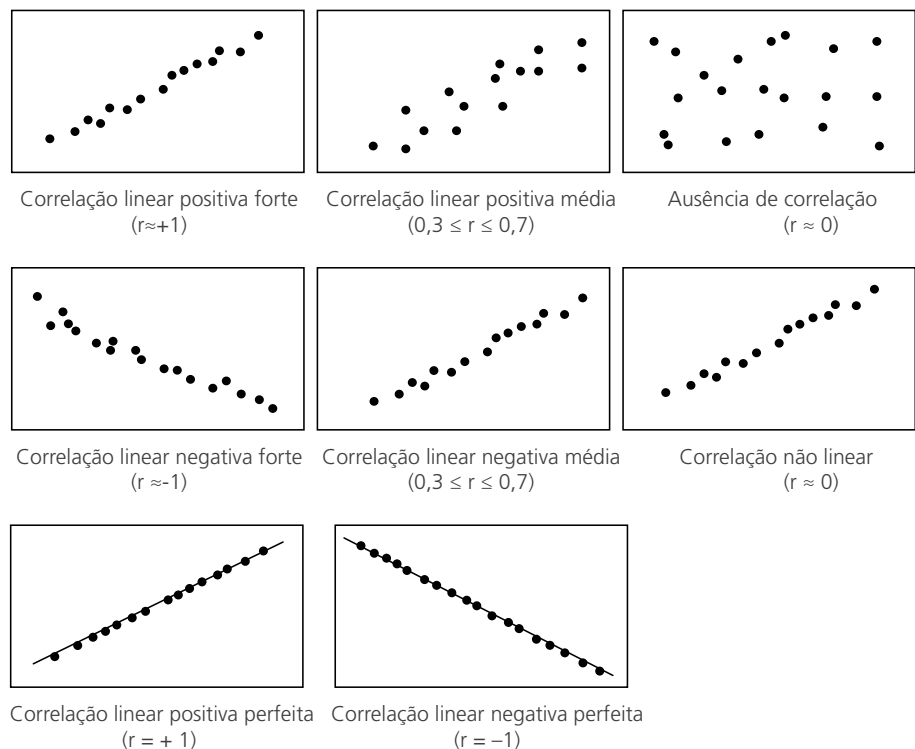
Mas não podemos tirar conclusões com base apenas na análise gráfica. Para fundamentar nossa avaliação sobre a associação entre as variáveis x e y , existe o coeficiente de correlação linear de *Pearson* (utiliza-se a letra “ r ” para representar esse coeficiente). Ele “mede” a força ou intensidade da correlação entre as variáveis x e y . Veja as Figuras 20 e 21, a seguir:

Figura 20 – Coeficiente de correlação linear de *Pearson*



Fonte: Produção do próprio autor (2017).

Figura 21 – Diagrama de dispersão e coeficiente de correlação linear de *Pearson*



Fonte: Produção do próprio autor (2017).



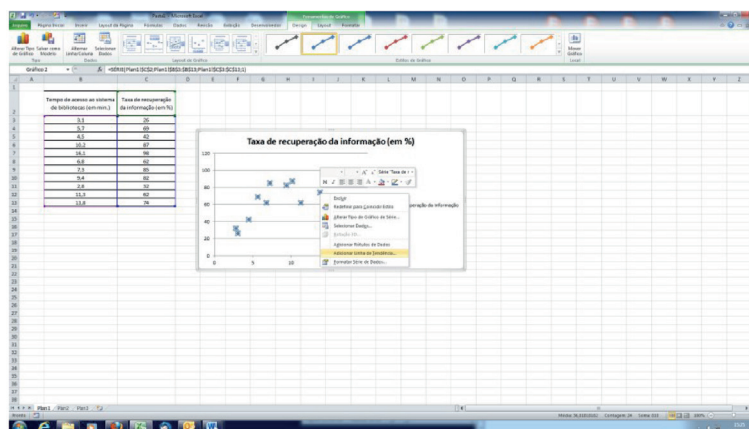
Você também pode calcular o coeficiente linear de correlação utilizando a função CORREL do Excel. Vá em Fórmulas → Inserir Função → Estatística e procure pela função CORREL.

Após inserir os dados na planilha do *Excel*, você deverá seguir o caminho: Inserir → Gráfico → Dispersão (Figura 22). Depois de aparecer o gráfico de dispersão, se você clicar com o botão direito do *mouse* em qualquer dos pontos (Figura 23), aparecerá a janela da Figura 24. Clique em ok. Vai aparecer o conteúdo da Figura 25.

[illegible]

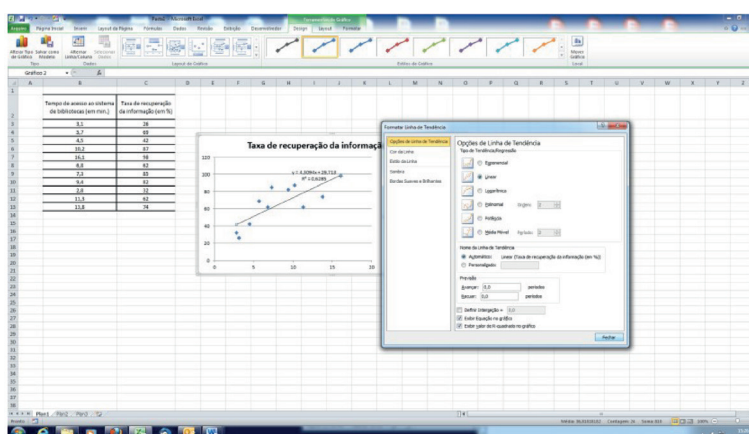
Semestre
7

Figura 23 – Tela 2



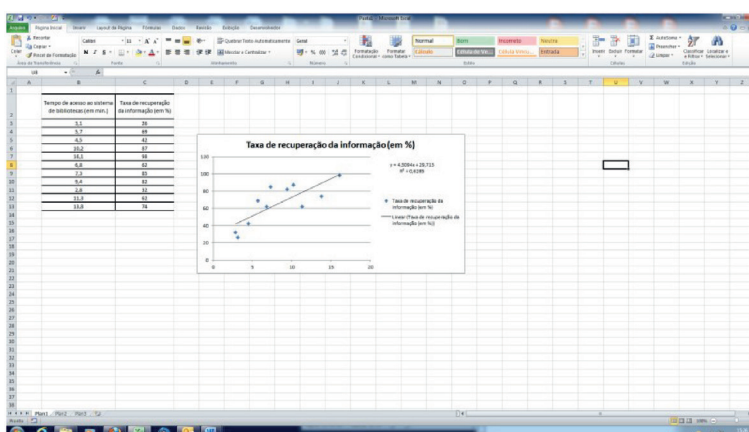
Fonte: Produção do próprio autor (2017).

Figura 24 – Tela 3



Fonte: Produção do próprio autor (2017).

Figura 25 – Tela 4



Fonte: Produção do próprio autor (2017).

Observe que o próprio Excel fornece $r^2 = 0,6285$. Extraindo a raiz quadrada, obtemos o valor do coeficiente de correlação linear: $r = 0,79278$.

Note bem: esse gráfico ainda não possui identificação das variáveis representadas nos eixos!

Para utilizar a função CORREL, digite em qualquer célula do Excel: =CORREL(B2:B12;C2:C12) e aperte ENTER.

Você terá que obter como resposta: 0,7928. A diferença entre os dois valores deve-se a arredondamento.

2.9.3 Reta de regressão

Após construir o diagrama de dispersão, o passo natural é verificar se é possível construir uma reta de regressão bem ajustada aos seus dados. O que é a reta de regressão? É uma reta que aproxima melhor os seus pontos no diagrama de dispersão. O “melhor” significa que essa aproximação ocorre devido a um critério: o de minimizar a soma dos quadrados das distâncias dos pontos até a reta.

O propósito da análise de regressão é avaliar o quanto uma variável resposta (também denominada de *variável dependente y*) se modifica conforme alteramos os valores da variável independente *x*. Em situações nas quais é possível construir a equação de uma “boa” reta de regressão, a equação pode ser usada para simular a variação da variável resposta *y*, de acordo com a variação de valores para a variável independente *x*. Com a equação de regressão, podemos avaliar impactos na variável resposta *y*, em associação com os valores que escolhermos para a variável independente *x*. É a mesma lógica usada em treinamento de simuladores de voo pelas companhias de aviação. Buscamos evitar um acidente ou imprevisto, antecipando-nos aos eventos. Vamos ver um exemplo?

Digamos que o número de acessos a determinada biblioteca universitária seja dependente do número de matrículas na universidade. Temos os dados fictícios a seguir, na Tabela 4, coletados para dez universidades:

Tabela 4 – Número de alunos matriculados x número de consultas à biblioteca

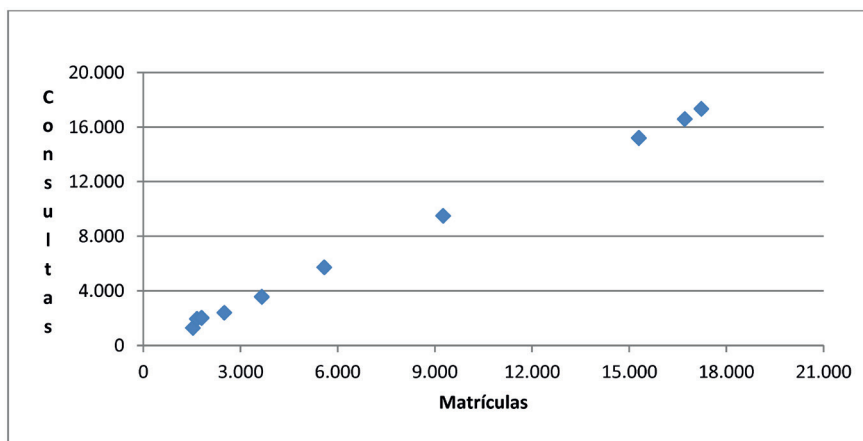
Número de alunos matriculados	Número de consultas à biblioteca
1.523	1.701
1.647	1.804
1.792	1.920
2.492	2.416
3.651	3.583
5.578	5.402
9.249	9.743
15.291	14.965
16.711	16.556
17.225	17.321

Fonte: Produção do próprio autor (2017), com dados fictícios, apenas para fins didáticos.



Assim, se a relação matemática for válida, tendo o número de matrículas na universidade, o bibliotecário terá condições de estimar, facilmente, o número de consultas à biblioteca. O diagrama de dispersão para os dados anteriores é (Figura 26):

Figura 26 – Diagrama de dispersão para consultas x matrículas



Fonte: Produção do próprio autor (2017).

Agora, vejamos como se constrói essa reta de regressão.

A expressão para uma função de 1.º grau (cujo gráfico é uma reta – lá do Ensino Médio, lembra?) é dada por: $y = ax + b$.

A equação da reta de regressão é dada pela expressão:

$$y_i = \alpha + \beta x_i + e_i,$$

onde:

$$\alpha = \bar{y} - \beta \bar{x}$$

$$\beta = \frac{\sum_{i=1}^n x_i y_i - n \left(\frac{\sum_{i=1}^n x_i}{n} \right) \left(\frac{\sum_{i=1}^n y_i}{n} \right)}{\sum_{i=1}^n x_i^2 - n \left(\frac{\sum_{i=1}^n x_i}{n} \right)^2}.$$

$\bar{y} = \frac{\sum_{i=1}^n y_i}{n}$ e $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ representam a média dos y_i e dos x_i , respectivamente.

Os e_i representam os erros aleatórios supostos, independentes e normalmente distribuídos, com $i = 1, \dots, n$.

A variável y_i é denominada variável dependente ou variável resposta. Ela varia conforme são alterados os valores da variável independente x_i .

A variável x_i é denominada de variável independente ou variável explicativa.

No caso do exemplo anterior, como temos dez universidades na nossa amostra, $n = 10$. Os valores para x_i são dados pela primeira coluna, número de alunos matriculados. Os valores da variável dependente y_i são dados pela segunda coluna, número de consultas à biblioteca.

Substituindo os valores nas equações dadas anteriormente para β e α , teremos os seguintes valores: $y = 0,99x + 100,62$.

Vamos experimentar essa equação? Se o número de matrículas em uma universidade for $x = 5.000$, então, com essa equação, é possível estimar que o número de consultas à biblioteca será, aproximadamente $y = 0,99 * 5.000 + 100,62 = 5050,62$. É claro que o assunto regressão não se encerra aqui. Ainda temos que verificar se essa equação de regressão se ajusta adequadamente aos dados. Para isso, existem testes de hipótese relativos à reta de regressão. Por sair do nível desejado para este texto, esses tópicos não serão abordados aqui, mas você pode encontrar esses assuntos em *Levine, Berenson e Stephan* (2000), ou em *Dancey e Reidy* (2006).



2.9.4 Atividade

Esta Atividade atende ao objetivo e: “utilizar as principais ferramentas de análise de dados em pesquisas quantitativas”.

A região da Baixada Fluminense, no estado do Rio de Janeiro, é composta por 13 municípios de sua região metropolitana: Belford Roxo, Duque de Caxias, Guapimirim, Itaguaí, Japeri, Magé, Mesquita, Nilópolis, Nova Iguaçu, Paracambi, Queimados, São João de Meriti e Seropédica.

É possível estabelecer alguma correlação entre o *Índice de Desenvolvimento da Educação Básica (IDEB)* e a renda familiar em um município?

Faremos duas questões a partir deste artigo:

OLIVEIRA, F. A. O. Desempenho educacional e renda domiciliar: análise do IDEB dos municípios da Baixada Fluminense. **Revista Eletrônica de Extensão da URI: Vivências**, [S.l.], v. 8, n. 14, p. 92-99, 2012. Disponível em: <http://www2.reitoria.uri.br/~vivencias/Numero_014/artigos/artigos_vivencias_14/n14_09.pdf>. Acesso em: 7 jul. 2021.

Apesar de estarmos fornecendo a informação necessária para responder às questões adiante, sugerimos que você consulte o artigo para observar o uso de ferramentas estatísticas na análise de questões sociais.

Nesse texto, você encontrará a informação de que, para os 13 municípios da Baixada Fluminense (dados relativos ao *IDEB* do *Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP* 2010), o coeficiente de correlação linear entre as variáveis *IDEB* e renda é igual a 0,702. Assim, qual é o significado desse valor para o coeficiente de correlação? O que podemos concluir a partir dele?

Além disso, o referido artigo apresenta que a reta de regressão entre *IDEB* e renda é $M_{IDEB} = 3,5 + 0,002Renda$. Sendo assim, essa reta cresce ou decresce à medida que a renda aumenta? Qual o significado de essa reta ser crescente ou decrescente?



Resposta comentada

Com $r = 0,702$, podemos concluir que existe correlação linear positiva entre as variáveis *IDEB* e renda. Assim, quanto maior a renda, maior o resultado esperado no *IDEB* (em média).

A partir da equação da reta de regressão entre renda e *IDEB*, como o coeficiente angular da reta (o valor 0,002) é positivo, podemos afirmar que essa reta é crescente. Ou seja: à medida que a renda aumenta, também observamos um aumento no desempenho no *IDEB*. Mais especificamente: para cada R\$100,00 de aumento na renda, esperamos um aumento no *IDEB* municipal de 0,2 ponto.

Não se assuste com os cálculos! Os *softwares* fazem todas as contas. Existem vídeos no *Youtube* que ensinam a construir a equação da reta de regressão, a calcular o coeficiente de correlação e muito mais. Procure e você verá que é muito mais simples do que parece!

2.10 TESTES DE HIPÓTESES

Nas pesquisas quantitativas, fazemos uma ou mais hipóteses, levantamos dados e testamos (estatisticamente) as hipóteses a partir dos dados levantados. Existe um grande número de testes de hipóteses. Algumas situações:

- a) queremos testar se determinada população possui como valor médio um valor específico. Neste caso, é melhor utilizar o teste Z ou o teste *t-Student*, se os pressupostos estiverem satisfeitos;
- b) queremos testar se duas populações possuem a mesma média. Podemos utilizar o teste Z ou o teste *t-Student* (dependendo da situação particular do caso);
- c) queremos testar se um número maior que duas populações possui mesma média. Neste caso, podemos pensar em usar o teste ANOVA (análise de variância), se os pressupostos para utilizá-lo estiverem satisfeitos;
- d) queremos testar a hipótese de igualdade de variâncias entre duas populações;
- e) queremos testar se existe independência entre duas variáveis categóricas (qualitativas). Por exemplo, um gerente quer testar se determinado funcionário apresenta desempenho que varia conforme o dia da semana. Ou ainda, queremos testar se os dados observados estão bem representados por determinado modelo. Neste caso, pensaríamos em um teste qui-quadrado de aderência.

Observe que são inúmeras as situações. Aqui, apresentamos apenas algumas delas. Em outras palavras: você deve avaliar com cuidado qual hipótese pretende testar antes de ir a campo coletar dados, inclusive para

efetuar uma estimativa adequada do tamanho da amostra, para que seu teste de hipóteses seja válido, ok?

Vamos apresentar a lógica geral de um teste de hipóteses. Consultando a literatura, seu orientador ou uma assessoria de Estatística, você vai, com seu projeto de pesquisa, avaliar qual, entre os vários testes de hipóteses que existem, é o mais adequado para o seu trabalho.

2.10.1 O formato geral de um teste de hipóteses

O teste de hipóteses é uma prova que permite que sejam realizadas inferências sobre um parâmetro da população, tomando por base valores observados em uma amostra. Nossa decisão, rejeitando ou não uma hipótese inicial, ocorre em função das diferenças entre os resultados observados na amostra e o valor assumido como verdadeiro pela hipótese inicial (também chamada de hipótese nula).

Todo teste de hipóteses necessita que você escreva duas delas: a nula (H_0) e a alternativa (que é a negação lógica da hipótese nula). A hipótese alternativa é representada por H_1 .

Para desenvolver um teste de hipóteses, você deverá seguir uma “receita” com os seguintes passos:

- a) formule as hipóteses H_0 (hipótese nula) e H_1 (hipótese alternativa);
- b) defina o valor do nível de significância α (nível aceitável para o erro tipo I);
- c) apresente os critérios de tomada de decisão (definindo as regiões de rejeição e não rejeição de H_0);
- d) calcule o valor da estatística de teste (que é dada pela “fórmula” daquele teste específico com o qual você está trabalhando). Cada teste de hipóteses possui a própria estatística de teste;
- e) efetue a tomada de decisão;
- f) analise os resultados e escreva suas conclusões.

Vamos exemplificar. Suponha que um gestor esteja interessado em testar se o tempo que seus funcionários levam para responder às demandas do público está sendo cumprido. Imagine que o tempo médio de espera aceitável para tal atendimento seja o de 23 minutos. O gestor decide selecionar uma amostra de 49 atendimentos, observar os tempos utilizados e verificar se a hipótese de que o atendimento médio de 23 minutos está sendo cumprido deve ser rejeitada ou não.

A quais respostas o gestor pode chegar? O atendimento médio está ocorrendo em 23 minutos e o gestor não precisa atuar sobre o sistema de atendimento. Nesse caso, a hipótese nula não terá sido rejeitada. Se a hipótese nula de que o atendimento médio está sendo realizado em 23 minutos tiver sido rejeitada, então o gestor deve atuar de alguma maneira, investigando o que pode estar ocorrendo. É possível que o atendimento médio esteja levando menos de 23 minutos ou mais que esse tempo. Ele deve avaliar qual é a consequência em cada um dos casos.

Vamos escrever a hipótese nula:

H_0 : o tempo médio de atendimento é de 23 minutos.

Poderíamos ter escrito assim: $H_0: \mu = 23$.



Lembre-se: a letra grega μ representa o tempo médio da população. Nos testes de hipóteses, queremos testar alguma afirmação sobre o valor de determinado parâmetro da população (pode ser um valor médio, um valor do desvio-padrão, da mediana etc). A hipótese nula é sempre a hipótese que é testada.

Como hipótese alternativa, temos:

H1: o tempo médio de atendimento é diferente de 23 minutos.

Poderíamos ter escrito assim:

H1: $\mu \neq 23$

A hipótese alternativa é sempre o oposto lógico de H0.

Vamos adotar como nível de significância do teste o valor $\alpha = 5\%$. Esse nível de significância de 5% está associado ao erro tipo I. É o erro que cometemos quando a hipótese nula é rejeitada sendo verdadeira e, portanto, não deveria ser rejeitada. O nível de significância é a probabilidade de cometermos um erro tipo I.

Importante! O nível de significância é especificado ANTES de o analista realizar o teste de hipóteses.

Não há um teorema na Matemática ou na Estatística que afirme que somos obrigados a usar determinado valor para o nível de significância. Valores tradicionalmente utilizados para o nível de significância são 1%, 5% e 10%. Valores mais elevados para o erro tipo I não são aceitáveis para a maior parte das aplicações. Por outro lado, para reduzir o erro tipo I abaixo de 1%, teríamos um custo (em tempo, dinheiro, aumento do tamanho da amostra) que pode não ser aceitável para o patrocinador da pesquisa. Observe que é impossível tornar o valor de α igual a zero. O valor $1 - \alpha$ é denominado *coeficiente de confiança*: é a probabilidade de que H0 não seja rejeitada quando, de fato, for verdadeira, não devendo ser rejeitada.

O erro tipo II é identificado pela letra grega β . É o erro que cometemos ao não rejeitarmos H0 quando esta é falsa e deveria ser rejeitada. A *eficácia de um teste* de hipóteses é o valor $1 - \beta$: é a probabilidade de se rejeitar H0, quando H0 for falsa, devendo ser rejeitada.

Quadro 6 – Erros dos tipos I e II nos testes de hipóteses

Decisão estatística	Situação efetiva	
	H0 verdadeiro	H0 falso
Não rejeitar H0	Confiança ($1 - \alpha$)	Erro tipo II (β)
Rejeitar H0	Erro tipo I (α)	Eficácia ($1 - \beta$)

Fonte: Produção do próprio autor (2017).

A probabilidade de acontecer um erro tipo II varia inversamente ao nível de significância α do teste. Quanto maior o tamanho da amostra, menor a probabilidade de erro tipo II, dado um nível α de erro tipo I.

Voltemos ao nosso exemplo. Queremos testar a hipótese de o tempo médio de atendimento ser igual a 23 minutos. As hipóteses ficam assim:

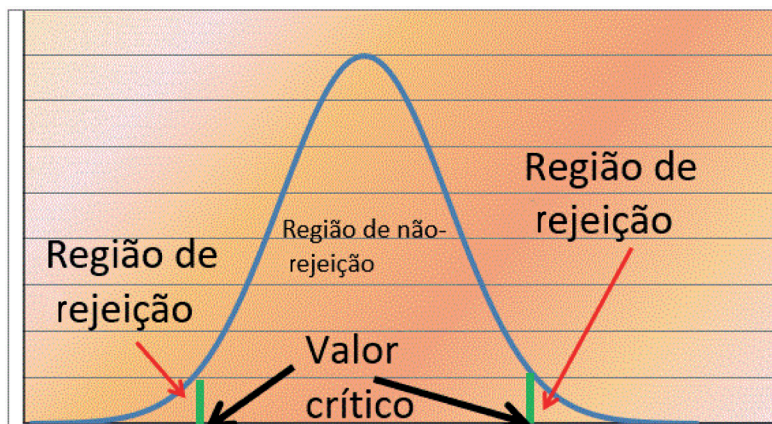
$H_0: \mu = 23$

$H_1: \mu \neq 23$

Definimos o nível de significância como $\alpha = 5\%$.

Vamos construir a região de rejeição (Figura 27):

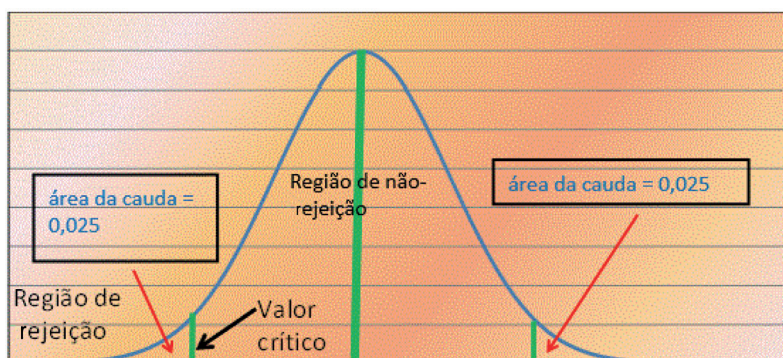
Figura 27 – Valores críticos e regiões de rejeição em um teste de hipóteses



Fonte: Produção do próprio autor (2017).

Como adotamos $\alpha = 5\%$, teremos 2,5% para cada "cauda" da distribuição normal (Figura 28):

Figura 28 – Valores críticos e área das caudas na distribuição normal



Fonte: Produção do próprio autor (2017).

Uma cauda da distribuição normal com área $0,025 = 2,5\%$ corresponde a uma área de $0,475 = 47,5\%$ entre a média da hipótese e o valor crítico. Da tabela da distribuição normal padronizada, determinamos que o valor crítico é igual a $+1,96$ e $-1,96$.

Assim, se a estatística de teste der um valor acima de $+1,96$ ou abaixo de $-1,96$, estaremos na região de rejeição e a hipótese nula será rejeitada com nível de significância igual a 5%.

Vamos calcular a estatística de teste. Suponhamos que, na amostra extraída, a média obtida foi igual a 24 minutos. Esse é o valor de $\bar{x} = 24$.

Vamos supor (neste exemplo) que o desvio-padrão da população seja conhecido e que $\sigma = 5$. No caso de que estamos tratando (teste de hipóteses para a média, com desvio-padrão populacional conhecido), a estatística de teste é:

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}},$$

onde:

$\bar{x}$: é a média amostral;

μ : é a média da população;

σ : é o desvio-padrão da população;

n : é o tamanho da amostra;

z : é o valor padronizado, para consulta na tabela da distribuição normal.

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{24 - 23}{\frac{5}{\sqrt{49}}} = \frac{1}{\frac{5}{7}} = 1,4.$$

Como o valor da estatística de teste não caiu na região de rejeição, não descartamos a hipótese nula e consideramos que o gestor não terá que realizar ajustes em sua equipe.



Multimídia

A página <<https://www.facebook.com/PortalAction>> é a de uma consultoria em Estatística que oferece bastante informação gratuita. Apresenta conteúdo estatístico (desde o básico até assuntos bastante avançados) e macros que “rodam” usando o Excel®.

Outro *site* interessante é: <<http://www.galileu.esalq.usp.br/>>. Acesso em: 7 jan. 2022.

2.10.2 Teste qui-quadrado (associação entre dados categóricos)

Dados categóricos são os de variáveis qualitativas. No exemplo a seguir, temos a variável gênero (mulheres e homens) e a variável apoio ao projeto. Foram apresentados três projetos a um grupo de 2.846 homens e mulheres. O apoio a cada um deles, por gênero, está indicado nas Tabelas 5 e 6:

Tabela 5 – Dados fictícios de apoio aos projetos A, B e C, com relação ao gênero (em valores absolutos)

Gênero	Apoio a cada projeto			Total
	Projeto A	Projeto B	Projeto C	
Mulheres	816	415	312	1.543
Homens	186	718	399	1.303
Total	1.002	1.133	711	2.846

Fonte: Produção do próprio autor (2017), com dados fictícios, apenas para fins didáticos.

Tabela 6 – Dados fictícios de apoio aos projetos A, B e C, com relação ao gênero (em %)

Gênero	Apoio a cada projeto			Total
	Projeto A	Projeto B	Projeto C	
Mulheres	0,528	0,268	0,202	1
Homens	0,142	0,551	0,306	1
Total	0,352	0,398	0,249	1

Fonte: Produção do próprio autor (2017), com dados fictícios, apenas para fins didáticos.

O teste qui-quadrado é um teste de hipóteses para variáveis categóricas (qualitativas) de associação e a pergunta que ele tenta responder é: será que existe alguma associação entre o apoio aos projetos e o sexo dos respondentes? Em outros termos, de forma ainda mais clara: estamos interessados em saber se o apoio aos projetos depende do gênero do respondente.

As hipóteses são:

- a) H0: não existe relação entre o apoio a cada projeto e gênero;
- b) H1: existe relação entre o apoio a cada projeto e gênero.

A estatística de teste para o teste qui-quadrado de independência é:

$$\chi^2 = \sum \frac{(f_{obs} - f_{esp})^2}{f_{obs}}$$

Nela, f_{obs} é a frequência observada (resultado da coleta de dados) e f_{esp} é a frequência esperada (obtida na hipótese de independência entre as variáveis).

Para calcular a frequência esperada em cada célula, efetuamos o cálculo a seguir (neste exemplo, $n = 2.846$):

$$f_{esp} = \frac{\text{soma das linhas} \times \text{soma das colunas}}{n}$$

Para essa equação, a matriz de frequências esperadas será (Tabela 7):

Tabela 7 – Matriz de frequências esperadas para o exemplo do teste qui-quadrado

Gênero	Apoio a cada projeto (frequências esperadas)			
	Projeto A	Projeto B	Projeto C	Total
Mulheres	543	614	385	1.543
Homens	459	519	326	1.303
Total	1.002	1.133	711	2.846

Fonte: Produção do próprio autor (2017), com dados fictícios, apenas para fins didáticos.

Agora, fazemos a substituição na expressão da estatística de teste:

$$\chi^2 = \sum \frac{(f_{obs} - f_{esp})^2}{f_{obs}}. \text{ O valor da estatística de teste será: } \chi^2 = 91,1 + 95,6 + 17,3 + 399,9 + 55,3 + 13,53 = 672,9.$$

No caso do teste qui-quadrado, existe o conceito de graus de liberdade.
Graus de liberdade = (número de linhas – 1).(número de colunas – 1).
Neste exemplo: graus de liberdade = (2 – 1).(3 – 1) = 1.2 = 2.

Adotamos nível de confiança $\alpha = 5\%$.

O critério de decisão será: rejeitar H_0 se

$X^2 > X^2_{\text{tabelado}}$, com 2 graus de liberdade.

Apresentamos apenas um trecho do quadro da distribuição qui-quadrado no Quadro 7:

Quadro 7 – Modo de consultar a tabela qui-quadrado

Graus de liberdade	Áreas da cauda superior alfa											
	0,995	0,99	0,975	0,95	0,9	0,75	0,25	0,1	0,05	0,025	0,01	0,005
1			0,001	0,004	0,016	0,102	1,323	2,706	3,841	5,024	6,635	7,879
2	0,01	0,02	0,051	0,103	0,211	0,575	2,773	4,605	5,991	7,378	9,21	10,597
3	0,072	0,115	0,216	0,352	0,584	1,213	4,108	6,251	7,815	9,348	11,345	12,838

Fonte: Levine; Berenson; Stephan (2000, p. 740).

Como temos dois graus de liberdade, consultamos o encontro da linha 2 com a coluna 5%. O valor crítico que define a região de rejeição de H_0 é 5,991. Como a estatística de teste foi igual a 672,9 e esse valor é maior que o valor crítico (na verdade, é extremamente superior a ele!), rejeitamos a hipótese nula a um nível de significância de 5%.

Conclusão: o apoio aos projetos, muito provavelmente, não é independente do gênero.

2.11 CONCLUSÃO

Nesta Unidade, você estudou o desenho de projetos de pesquisa quantitativa e noções de amostragem não probabilística e probabilística.

Você também aprendeu a determinar o tamanho da amostra (para algumas situações).

Observe que, na prática, você deverá identificar se sua situação é semelhante à que apresentamos aqui. Se não for, busque outras técnicas para determinar a amostra.

Você também estudou intervalos de confiança e testes de hipóteses, que são ferramentas clássicas da Estatística, utilizadas em inúmeras áreas da ciência (tanto nas Ciências Exatas quanto nas Biológicas ou nas Humanas).

Como você pode perceber ao longo desta Unidade, com as pesquisas quantitativas, podemos efetuar análises e correlações entre variáveis, bem como testar hipóteses a respeito delas.

Em contrapartida, devemos verificar cuidadosamente se o nosso contexto de pesquisa cumpre com as premissas exigidas pelas ferramentas estatísticas e/ ou matemáticas que pretendemos utilizar.

Nesta Unidade, ainda, vimos exemplos de utilização de ferramentas quantitativas aplicadas a questões sociais (por exemplo, na atividade sobre *IDEB* e renda na Baixada Fluminense). Assim, você pode observar exemplos do alcance no uso de ferramentas quantitativas no estudo de fenômenos sociais.

Contudo, é importante evitar a “ideologia da certeza” associada aos números e modelos matemáticos e estatísticos. Aqui, chegamos à questão dos limites das ferramentas quantitativas: devemos reconhecer que, com elas, não conseguimos compreender a subjetividade associada aos fenômenos sociais.

Outra questão relevante é que, com a facilidade e a maior disponibilidade de computadores e *softwares* estatísticos, devemos nos lembrar de que o programa faz os cálculos que o usuário ordena que sejam feitos. A interpretação, análise e avaliação das premissas para o uso das ferramentas estatísticas são responsabilidade do pesquisador/analista.

RESUMO

Nesta Unidade, aprendemos a classificar variáveis em pesquisas quantitativas (quantitativas contínuas e discretas, qualitativas nominais e ordinais), estudamos como efetuar o delineamento de pesquisas quantitativas (experimentais e não experimentais) e, na sequência, discutimos a importante questão da amostragem, apresentando um exemplo histórico de erro em amostragem com viés de seleção (o famoso caso da pesquisa para a eleição de *Franklin Roosevelt*).

Você também foi apresentado a algumas funções estatísticas do *software Excel* (CONT.SE, CONT.SES, DESVPAD.A, MED e MÉDIA). Logo após, estudamos como determinar o tamanho mínimo da amostra no caso de amostra aleatória simples.

Vimos também como efetuar a coleta de dados em pesquisas quantitativas, além de recomendações importantes sobre a elaboração de questionários e escalas (*Likert* e diferencial semântico).

Estudamos tabelas de frequência, uma ferramenta simples, mas relevante, para a organização de dados quantitativos.

Com respeito à análise de dados quantitativos, estudamos o coeficiente de correlação linear e a reta de regressão.

Finalmente, concluímos a Unidade com uma introdução aos testes de hipóteses.



REFERÊNCIAS

AGRESTI, A.; FINLAY, B. **Métodos estatísticos para as Ciências Sociais**. 4. ed. Porto Alegre: Penso, 2012.

CANO, I. **Introdução à avaliação de programas sociais**. 2. ed. Rio de Janeiro: FGV, 2004.

COSTA, Maira Murrieta. As bibliotecas brasileiras em 2018: resultados da técnica de delfos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 17, n. 1, p. 74-93, jan./ mar. 2012.

CRESWELL, J. W. **Projeto de pesquisa**: métodos qualitativo, quantitativo e misto. 3. ed. Porto Alegre: Artmed, 2010.

DANCEY, C. P.; REIDY, J. **Estatística sem Matemática para Psicologia**: usando SPSS para Windows. Porto Alegre: Artmed, 2006.

LEITÃO, B. J. M. **Avaliação qualitativa e quantitativa numa biblioteca universitária**. Rio de Janeiro: Interciência; Niterói: Intertexto, 2005.

LEVINE, D. M.; BERENSON, L. M.; STEPHAN, D. **Estatística**: teoria e aplicações: usando Microsoft Excel português. Rio de Janeiro: LTC, 2000.

LIMA, P. A. **Impacto dos formatos tradicional e interativo de propaganda na atitude e intenção de compra dos consumidores**. Dissertação (Mestrado em Administração) – Faculdade de Economia, Administração e Contabilidade (FACE), Universidade de Brasília, Brasília, 2012. Disponível em: <http://repositorio.unb.br/bitstream/10482/12297/1/2012_PatriciaAlbuquerqueLima.pdf>. Acesso em: 21 fev. 2015.

MALHOTRA, N. **Pesquisa de marketing**: uma orientação aplicada. Porto Alegre: ArtMed, 2006.

OLIVEIRA, Joelma Gualberto. **Processo de avaliação do INEP/ MEC de bibliotecas universitárias pertencentes às instituições de educação superior privadas de Belo Horizonte/ MG**. 2010. Dissertação (Mestrado em Ciência da Informação) – UFMG, Belo Horizonte, 2010.

SAMPIERI R. H.; COLLADO, C. F.; LUCIO, M. P. B. **Metodologia da pesquisa**. 5. ed. PortoAlegre: McGraw-Hill/ Penso, 2013.

VANTI, N. A. P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência da Informação**, Brasília, v. 31, n. 2, p. 152-162, 2002. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/171/150>>. Acesso em: 25 fev. 2015.

UNIDADE 3

MÉTODOS QUANTITATIVOS ESPECÍFICOS DA BIBLIOTECONOMIA: COMO QUANTIFICAR COMPORTAMENTOS CULTURAIS?



3.1 OBJETIVO GERAL

Apresentar técnicas quantitativas específicas utilizadas pelos profissionais de Biblioteconomia, bem como as ferramentas apropriadas a elas.

3.2 OBJETIVOS ESPECÍFICOS

Esperamos que, ao final desta Unidade, você seja capaz de:

- a) descrever o objeto de estudo, as variáveis, os métodos e os objetivos da bibliometria, infometria, cientometria e webometria;
 - b) definir indicadores bibliométricos, reconhecendo algumas de suas limitações;
 - c) descrever as Leis de *Bradford*, *Lotka* e *Zipf*;
 - d) efetuar a análise bibliométrica das bases de dados *Web of Science*, *Scopus* e *SciELO*;
 - e) operar *softwares* bibliométricos.
- 

3.3 BIBLIOMETRIA, INFOMETRIA, CIBERMETRIA, WEBOMETRIA, CIENTOMETRIA, PATENTOMETRIA

Na Unidade anterior, estudamos ferramentas quantitativas gerais, que são aplicadas ao levantamento e análise de dados de pesquisas e à própria Biblioteconomia.

Nesta Unidade, você entrará em contato com algumas ferramentas quantitativas que foram desenvolvidas especificamente para o bibliotecário e/ ou cientista da informação.

Inicialmente, precisamos delimitar as várias subáreas quantitativas relacionadas com a Biblioteconomia.

Vamos lá?

A partir dos investimentos massivos realizados em Ciência e Tecnologia, pelos diversos países envolvidos na Segunda Guerra Mundial, a gestão e a produção do conhecimento mudaram de escala, atingindo níveis sem precedentes na história. A Ciência e a Tecnologia passam a ser produzidas por grandes equipes de pesquisadores, exigindo enormes montantes de recursos e laboratórios, que, por sua vez, também exigem grandes recursos para serem administrados. A produção de artigos científicos, patentes e livros, o surgimento de novas subáreas do conhecimento e, mais recentemente, da internet, bem como a criação de periódicos científicos eletrônicos, eventos e congressos virtuais, elevaram a complexidade da avaliação da produção científica e tecnológica. Tudo isso aconteceu dentro de uma mesma área do conhecimento e em um mesmo conjunto de países.

3.3.1 As várias “metrias”

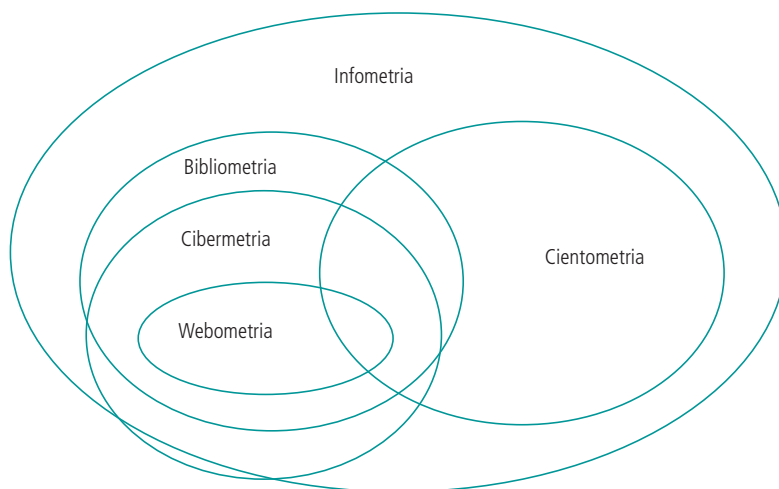
Não há um consenso entre os especialistas das fronteiras entre as várias “metrias”. Seguindo a definição proposta por *Bjorneborne* e *Ingwersen* (2004), citados por *Lucas, Zorita* e *Casado* (2013, p. 257. Grifos dos autores), a

[...] *Cientometria* diz respeito aos estudos da atividade científica; a *Bibliometria*, aos estudos da produção da informação; os *Estudos de usuários* (área de estudos também chamada de *Bibliotecometria*) reportam aos estudos de consumo de informação e das necessidades de informação; a *Patentometria* diz respeito aos estudos da inovação; a *Cibernetria* aos estudos de internet; e a *Webometria*, os estudos de sites.



Na citação anterior, não é apresentado o termo infometria, que estudaremos nas próximas páginas. As Figuras 29 e 30, a seguir, extraídas dos respectivos artigos citados, apresentam duas propostas para o entendimento das relações e fronteiras entre as várias “metrias”:

Figura 29 – Relações entre as disciplinas métricas

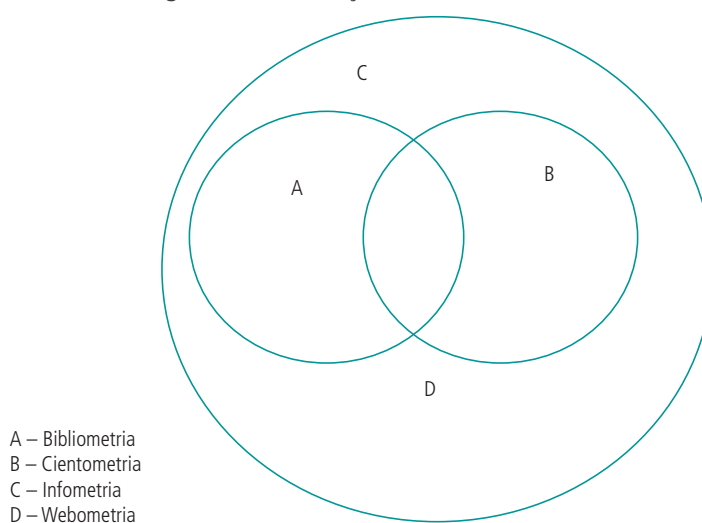


Fonte: Lucas; Zorita; Casado (2013, p. 258).

Observe, na Figura 30, que há uma intersecção entre o foco de estudo da bibliometria e da cientometria, mas essas duas disciplinas não se confundem. Por essa figura, a infometria abrange todas as outras “metrias”. Não há consenso entre os autores sobre as fronteiras entre as várias “metrias”. Para alguns deles (por exemplo, *Bjorneborn* e *Ingwersen, op. cit.*), a webometria é tratada como uma área dentro da cibermetria. Já para *Bar-Ilan* (2008), *cibermetria* e *webometria* são considerados sinônimos.

De *Vanti* (2002) extraímos o diagrama ilustrativo (Figura 30):

Figura 30 – Intersecção entre as “metrias”



Fonte: Vanti (2002, p. 10).

Na Figura 30, não está incluída a cibermetria. Com esse diagrama, você pode observar melhor as relações de intersecção entre as “metrias” apresentadas.

De acordo com *Le Coadic* (2005) (apud SANTIN, 2011, p. 109), as variáveis mais utilizadas nas análises métricas são: título, palavras-chave e nome dos autores. Com as palavras-chave, o pesquisador pode efetuar inferências a respeito dos temas em destaque em uma área do conhecimento. Ainda segundo o autor,

a ocorrência de dois termos num mesmo documento ou base de dados representa a associação entre os assuntos, permitindo destacar as tendências de investigação e os principais centros de interesse dos cientistas (LE COADIC, 2005 apud SANTIN, 2011, p. 109).

Na sequência, você terá um detalhamento de cada uma das “metrias”: bibliometria, cientometria, infometria e webometria.

3.3.2 Bibliometria

Do inglês, *bibliometrics*. De acordo com *Chaviano* (2004), o primeiro estudo bibliométrico foi realizado em 1917. *Cole* e *Eales*, em 1917, efetuaram uma análise estatística de publicações sobre Anatomia Comparada, entre 1550 e 1860. Em 1923, *Edward Hulme* apresentou uma análise estatística da História da Ciência, tendo introduzido o termo *statistical bibliography* (bibliografia estatística) (ARAÚJO, 2006; VANTI, 2002). Segundo *Guedes* e *Borschiver* (2005), o termo *statistical bibliography* ficou esquecido por 22 anos, voltando a ser utilizado em 1944, por *Gosnell* (1944). Ressurge, em 1962, empregado por *L. M. Raisig*. Em 1927, *Gross* e *Gross*, com o objetivo de fornecer subsídios para determinar quais periódicos deveriam ser adquiridos, fizeram uso da técnica de contagem das referências citadas no periódico *Journal of the American Chemical Society* (ARAÚJO; ALVARENGA, 2011). Em 1934, *Paul Otlet* foi o primeiro autor a apresentar o termo “bibliometria”, em sua obra *Traité de Documentation* (ARAÚJO, 2006).

A disseminação do termo bibliometria ocorre, contudo, a partir de 1969, com a publicação de um artigo de *Allan Pritchard* sobre a questão: “bibliografia estatística ou bibliometria?” (ARAÚJO, 2006). Nesse artigo, a bibliometria é caracterizada como área de estudo que usa métodos matemáticos e estatísticos para investigar e quantificar os processos de comunicação escrita. Assim, a bibliometria não se restringe à avaliação da produção científica, nem foca seus esforços unicamente na *web* (distinguindo-se, assim, da webometria).

Segundo *Figueiredo*, que é citado por *Araújo* (2006, p. 13),

a bibliometria desde sua origem é marcada por uma dupla preocupação: a análise da produção científica e a busca de benefícios práticos imediatos para bibliotecas (desenvolvimento de coleções, gestão de serviços bibliotecários) (FIGUEIREDO, 1977 apud ARAÚJO, 2006, p. 13).



Silva, Hayashi e Hayashi (2011, p. 114) esclarecem que, “a partir da década de 1990, a Bibliometria se tornou uma ferramenta padrão da política científica”. Tendo se tornado uma valiosa forma de medir a comunicação científica, o que podemos observar é que, desde então, reforçou-se o papel do profissional de Ciência da Informação e Biblioteconomia no planejamento e na gestão de políticas científicas.

A análise de citação e a de cocitação são duas técnicas da análise bibliométrica. Na subseção “Softwares bibliométricos”, você conhecerá uma lista de programas utilizados nas diversas “metrias”.

A bibliometria faz análises sobre aspectos quantitativos de autores, grupos de pesquisa, área de conhecimento e outros indicadores da produção científica.

As três principais leis da bibliometria são:

- c) Lei de *Lotka* (que foca na produtividade de autores);
- d) Lei de *Zipf* (Lei do mínimo esforço, que analisa a frequência de palavras);
- e) Lei de *Bradford* (ou Lei de dispersão, que estuda a produtividade de periódicos).

Em seção posterior, estudaremos em detalhes cada uma dessas leis bibliométricas.



Multimídia

No endereço <<http://www.feg.unesp.br/~fmarins/seminarios/Bibliometria/>>, você encontrará mais informações sobre bibliometria, com *links* para outros artigos e textos sobre o tema. Vale a pena a visita!

3.3.3 Cientometria

A Cientometria (em inglês, *scientometrics*) quantifica a produção científica (artigos, revistas e livros). Ela avalia, quantitativamente, a produção do conhecimento científico de pesquisadores, núcleos de pesquisa, universidades, instituições e até mesmo países. O planejamento e a avaliação das políticas científicas, as decisões sobre investimentos financeiros em determinadas áreas do conhecimento nas universidades e as verbas para pesquisadores individuais são, muitas vezes, baseados em dados cientométricos. Observe que, enquanto a bibliometria estuda questões mais gerais da informação, a cientometria foca na produção científica, mas claramente há uma intersecção entre elas.

De acordo com *Vanti* (2002), o termo cientometria – ou cientometria – foi criado na antiga União Soviética (URSS) e Europa Oriental. O autor explica que o termo alcançou notoriedade, no final da década de 1970, com a publicação da revista *Scientometrics*, editada originalmente na Hungria e, atualmente, na Holanda.



Multimídia

Acesse o *link* para a revista *Scientometrics*: <<http://www.springer.com/computer/database+management+%26+information+retrieval/journal/11192>>. Acesso em: 15 jul. 2021.

Entre as variáveis estudadas na cientometria, *Santin* (2011) destaca a comunicação entre os cientistas e os fatores de diferenciação das subdisciplinas. *Macias-Chapula* (1998) e *Spinak* (1998) apresentam, como objetos de estudo da cientometria,

- a) as várias áreas da ciência que evoluem em termos quantitativos;
- b) o estudo de como as várias subáreas científicas evoluem quantitativamente;
- c) as relações entre ciência e tecnologia;
- d) as relações entre o desenvolvimento científico e o crescimento econômico.

Na subseção sobre “Indicadores bibliométricos e cientométricos”, você conhecerá os principais indicadores utilizados nessas “metrias”.

Com a fundação, por *Eugene Garfield*, do *Institute for Scientific Information (ISI)*, em 1958, foi criado um dos principais centros para a quantificação e geração de insumos na área da produção científica.

3.3.4 Infometria

A infometria (em inglês, *informetrics*) ou informetria, para alguns autores, como área de estudo, surge após a constituição da bibliometria e da cientometria. É considerada a “metria” mais abrangente. Segundo *Vanti* (2002), *Otto Nacke* foi o primeiro a apresentar o termo infometria, em 1979. Ela não se restringe a estudar a produção científica e, assim, extrapola os limites da cientometria, possuindo, com ela, uma intersecção. Por outro lado, a infometria é o estudo dos aspectos quantitativos da informação em qualquer formato, e não apenas em registros catalográficos ou bibliografias, tampouco é restrita aos cientistas (VANTI, 2002).

Convém destacar que a infometria

[...] utiliza-se de unidades definidas, tais como palavras, documentos, textos, fontes ou bases de dados, como focos de análise, podendo priorizar variáveis como recuperação, a relevância, a revocação ou outras características da informação que possam ser consideradas relevantes. (BUFREM; PRATES, 2005, p. 11).

Wormell (1998), em um importante artigo, destaca as novas perspectivas que a análise infométrica apresenta, seja para investigação das bases de dados, seja como instrumento de análise. Nele, a autora relata três experiências do *Centro de Estudos Infométricos (CIS)*, criado em

Semestre

7

1997, e que faz parte da *Real Escola de Biblioteconomia e Informação*, em inglês, *Royal School of Library and Information Studies (RSLIS)*, situada em Copenhague e em Aalborg, na Dinamarca. O *CIS* aplica os estudos bibliométricos, além da avaliação da produção em ciência e tecnologia, nos estudos do impacto sobre as relações sociais e econômicas. Dessa forma, a infometria estende sua aplicação para além de ferramenta unicamente quantitativa.

Tague-Sutckiffe, citado por *Vanti* (2002), destaca, como aspecto adicional de distinção entre infometria, bibliometria e cientometria, a relação delas com os objetos e sujeitos de estudo. A infometria não tem como escopo apenas a informação registrada. Estuda, também, fenômenos de comunicação informal (comunicação oral, inclusive) e de grupos sociais diversos (não apenas aqueles pertencentes às elites intelectuais).



Atenção

Segundo *Araújo* (2006), os principais grupos de pesquisa em bibliometria são:

- a) *BIRG* (Sidney, Austrália);
- b) *CEST* (Berna, Suíça);
- c) *CINDOC* (Madrid, Espanha);
- d) *CIS* (Copenhague, Dinamarca);
- e) *CRRM* (Marselha, França);
- f) *CWTS* (Leiden, Holanda);
- g) *FhG-ISI* (Karlsruhe, Alemanha);
- h) *Inforsk* (Umea, Suécia);
- i) *OST* (Montreal, Canadá);
- j) *OST* (Paris, França);
- k) *SPRU* (Sussex, Inglaterra);
- l) *ISI Research Service Group* (Filadélfia, EUA).



Atenção

Para uma visão mais completa da evolução histórica das “metrias”, sugerimos que você consulte a linha do tempo dos principais eventos na história dos estudos métricos da informação, em *Lucas, Zorita e Casado* (2013).

3.3.5 Webometria

Esta “metria” distingue-se por utilizar as técnicas da bibliometria na mensuração de processos de comunicação em *sites* na *web*. A partir dessas técnicas, é possível avaliar a evolução de uma ou mais áreas do conhecimento ao longo do tempo, efetuar a análise da produção de indivíduos, universidades ou instituições, regiões geográficas ou mesmo países. A webometria estuda a *web* para avaliar as quantidades de padrões de acessos, *hiperlinks* e *sites*.

A webometria (*webmetrics*, em inglês), conforme definição apresentada por *Almind* e *Ingwersen* (1997, p. 404), é a aplicação de métodos infométricos à WWW. *Vanti* (2002, p. 157) ensina:

Entre as medições que podem ser realizadas no campo da webometria, encontra-se, por exemplo, aquela que diz respeito à frequência de distribuição das páginas no ciberespaço. Esta medição aponta para o estudo ou análise comparativa da presença dos diversos países na rede, das proporções de páginas pessoais, comerciais e institucionais.

A webometria produz “mapas de fluxos de informação” baseados na internet e não em outros suportes eletrônicos ou em papel, como aqueles utilizados pela cientometria. As análises webométricas também podem ser feitas por domínio (por país ou por instituições).



3.3.6 Atividade

Esta atividade atende aos objetivos *a* e *b*: “descrever o objeto de estudo, as variáveis, os métodos e os objetivos da bibliometria, infometria, cientometria e webometria” e “definir indicadores bibliométricos, reconhecendo algumas de suas limitações”.

Para desenvolvê-la, consulte *Vanti* (2002), e verá que a webometria é tratada em seção específica. A partir dessa consulta, faça o que se pede:

- a) apresente três exemplos de mensurações que podem ser utilizadas na webometria;
- b) demonstre quais inconsistências lógicas são apresentadas pelos motores de busca, de tal forma que tornam mais complexa a tarefa de mensuração dos dados obtidos por essa via.

Resposta comentada

- a) Apresentamos, a seguir, algumas das mensurações que podem ser realizadas no campo da webometria, elencadas em *Vanti* (2002):
 - frequência de distribuição de páginas no ciberespaço;
 - classificações sobre o tipo de páginas (setor público x setor privado, por exemplo);
 - efetuar a medição dos mesmos elementos em momentos diferentes no tempo;



- pode-se mensurar o tamanho médio de uma página (expressa em *bytes*) ou em relação ao número médio de *links* por página;
- um indicador relevante na área da webometria é o fator de impacto na *web* (*web impact factor* – *WIF*).

b) Um problema bastante relevante é a reprodutibilidade das mensurações. Na página 159 do referido artigo, de acordo com *Judit Bar-Ilan* e *Thelwall*, destacamos:

Os motores costumam perder informação: URLs relevantes recuperados em um determinado momento por um determinado motor de busca não são encontrados por esse mesmo motor em um momento posterior (apesar de continuarem existindo e sendo relevantes) (BAR-ILAN; THELWALL apud VANTI, 2002, p. 159).

Outro ponto destacado no artigo diz respeito ao fato de que URL recuperados em um levantamento posterior nem sempre trazem informação idêntica à da primeira pesquisa, inclusive tendo baixa sobreposição de resultados entre os vários buscadores. Além da questão das inconsistências no que tange à contabilização, existe o importante problema do processamento das sintaxes de busca. Os motores de busca podem trazer resultados espúrios, além do fato de que a hierarquização dos documentos é função de análises subjetivas.

3.3.7 Cibermetria

Do inglês, *cibermetrics*. Como já mencionamos anteriormente, enquanto alguns autores utilizam cibermetria e webometria como sinônimos, outros, como *Bjorneborn* e *Ingwesen*, citados por *Lucas, Zorita* e *Casado* (2013, p. 260), definem cibermetria como o “estudo dos aspectos quantitativos da construção e dos usos dos recursos de informação, estruturas e tecnologias em internet, de acordo com as aproximações bibliométricas e informétricas”.

Em *Bufrem* e *Prates* (2005), você encontrará um quadro que apresenta a diversidade de conceituações para os termos bibliometria, biblioteconometria, cientometria, infometria e webometria, com as respectivas referências.



3.3.7 Atividade

Esta atividade atende ao objetivo *a* da Unidade, a saber: “descrever o objeto de estudo, as variáveis, os métodos e os objetivos da bibliometria, infometria, cientometria e webometria”.

Vanti (2002) apresenta possibilidades de aplicação das técnicas bibliométricas, cientométricas e infométricas. Apresente cinco aplicações, dentre as citadas no artigo.

Resposta comentada

Certamente, é relevante conhecer as definições teóricas das várias “metrias”, seus alcances e limites. Atingimos melhor compreensão desses métodos conhecendo suas possibilidades de aplicação. A seguir, citamos as aplicações apresentadas em *Vanti* (2002, p. 155):

- identificar as tendências e o crescimento do conhecimento em uma área;
- identificar as revistas do núcleo de uma disciplina;
- prever tendências de publicação;
- medir o grau e os padrões de colaboração entre autores;
- analisar os processos de citação e cocitação.

O referido artigo ainda apresenta outras oito possibilidades de aplicação.



3.3.8 Atividade

Esta atividade atende ao objetivo a: “descrever o objeto de estudo, as variáveis, os métodos e os objetivos da bibliometria, infometria, cientometria e webometria”.

Estudos de usuários e de seu comportamento frente aos serviços ofertados devem ser realizados, seja na fase de planejamento de um novo serviço, seja após sua implantação. A bibliometria proporciona ferramentas para análises quantitativas de grupos de pesquisa, autores e outros indicadores de produção científica.

Com isso em mente, faça o que se pede:

- apresente a caracterização da bibliometria segundo *Allan Pritchard*;
- Figueiredo* (1977) destaca que a bibliometria, desde sua origem, apresenta dupla preocupação. Apresente quais as duas preocupações explicitadas pela autora;
- A respeito da segunda preocupação explicitada por *Figueiredo* (1977) no item *b*, dê exemplos específicos de aplicações “práticas” da bibliometria para bibliotecas.

Resposta comentada

- De acordo com *Pritchard*, a bibliometria é caracterizada como uma área de estudo que usa métodos matemáticos e estatísticos para investigar e quantificar os processos de comunicação escrita.
- Segundo *Figueiredo*, citada por *Araújo* (2006, p. 13),

a bibliometria desde sua origem é marcada por uma dupla preocupação: a análise da produção científica e a busca de benefícios práticos imediatos para

Semestre

7

bibliotecas (desenvolvimento de coleções, gestão de serviços bibliotecários (FIGUEIREDO, 1977 apud ARAÚJO, 2006, p. 13).

- c) Como aplicações práticas voltadas a bibliotecas, citamos: efetuar estimativas de crescimento dos acervos e efetuar avaliações quantitativas de suas características.

3.4 AS TRÊS LEIS DA BIBLIOMETRIA

Sem dúvida, as três principais leis bibliométricas são:

- a) Lei de *Bradford* (relativa à produtividade de periódicos);
- b) Lei de *Lotka* (voltada ao estudo da produtividade de autores);
- c) Lei de *Zipf* (associada aos estudos sobre frequência de palavras).

Cada uma dessas leis está associada a determinados “padrões” ou regularidades encontrados nos seus respectivos focos de estudo. Alguns desses padrões apresentam pequenas variações nos coeficientes ou expoentes numéricos, em função da área do conhecimento dos dados.

3.4.1 Lei de *Bradford* (Lei da dispersão)

A importância prática da Lei de *Bradford* consiste no seu uso para identificação de quais são os periódicos (ou outros tipos de publicação) a serem adquiridos (ou descartados) pelas bibliotecas. Os periódicos responsáveis pela produção de maior quantidade de artigos de determinada área do conhecimento são denominados “núcleo de periódicos”.

Os estudos de *Bradford* resultaram na publicação, em 1934, da sua Lei de dispersão. O conjunto de periódicos da área de Geofísica utilizados por ele foi constituído de um núcleo menor, com periódicos mais próximos ao assunto, e um outro conjunto maior, menos próximo do assunto. *Bradford* observou que, à medida que a quantidade de periódicos aumentava, a produtividade dos pesquisadores caía. Em sua análise, foram utilizados 326 periódicos. Desses, nove continham 429 artigos, 59 continham 499 artigos e 258 continham 404 artigos. A Lei de *Bradford* foi proposta nos seguintes termos, tal como citado por *Coutinho* (1988, p. 219):

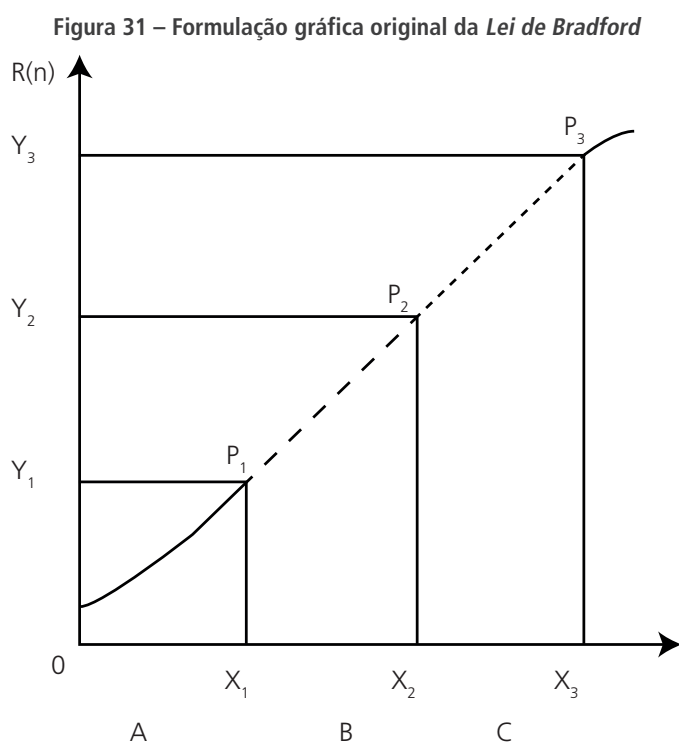
se os periódicos forem ordenados em ordem de produtividade decrescente de artigos sobre determinado assunto, poderão ser distribuídos num núcleo de periódicos mais particularmente dedicados a esse assunto e em diversos grupos ou zonas contendo o mesmo número de artigos que o núcleo, sempre que o número de periódicos e das zonas sucessivas for igual a $1:n:n^2$.

No seu trabalho original, *Bradford* encontrou três zonas de produtividade. Ele utilizou um gráfico semilogarítmico para apresentar graficamente sua lei. Esse gráfico apresenta uma leve curvatura quando nos aproximamos da origem, tornando-se, praticamente, uma reta, a partir do ponto denotado por P1.

Pinheiro (1983) alerta para uma das “armadilhas” da Lei de *Bradford*: o núcleo de periódicos determinado por ela não corresponde ao dos mais particularmente devotados ao assunto, mas aos periódicos mais produtivos em determinado período. *Coutinho* (1988) também alerta para fatores como a periodicidade e o número de artigos publicados por fascículo do periódico, que podem resultar em mudanças na posição de sua produtividade.

Brooks (1969), citado por *Lousada et al.* (2012), apresenta condições para a aplicação adequada da Lei de *Bradford*:

- o tema da bibliografia deve ser definido com clareza;
- todos os periódicos relevantes devem estar incluídos;
- o período do levantamento deve ser longo o bastante para permitir maior representatividade.



$\log(n)$, n é o número de periódicos.

Fonte: *Pinheiro* (1983).

$R(n)$ é o número acumulado de artigos sobre um assunto e n é o número acumulado de periódicos contendo aqueles artigos.

Na Tabela 8, a seguir, extraída de *Aymard* (1980), temos os dados utilizados por *Bradford* em seu levantamento original e, na sequência, o gráfico com os valores $\log_{10}(n)$, no eixo das abscissas, e $\sum_{i=1}^m (R_i \cdot A_i)_{\text{observado}}$, no eixo das ordenadas.

Tabela 8 – Dados utilizados por *Bradford* em seu levantamento

R_i	A_i	$n = \sum_{i=1}^{i=m} R_i$	$R_i \cdot A_i$	$\text{Log}_{10}(n)$	$\sum_{i=1}^{i=m} (R_i \cdot A_i)$ observado
1	22	1	22	0	22
1	18	2	18	0,30103	40
1	15	3	15	0,477121	55
2	13	5	26	0,69897	81
2	10	7	20	0,845098	101
1	9	8	9	0,90309	110
3	8	11	24	1,041393	134
3	7	14	21	1,146128	155
1	6	15	6	1,176091	161
7	5	22	35	1,342423	196
2	4	24	8	1,380211	204
13	3	37	39	1,568202	243
25	2	62	50	1,792392	293
102	1	164	102	2,214844	395

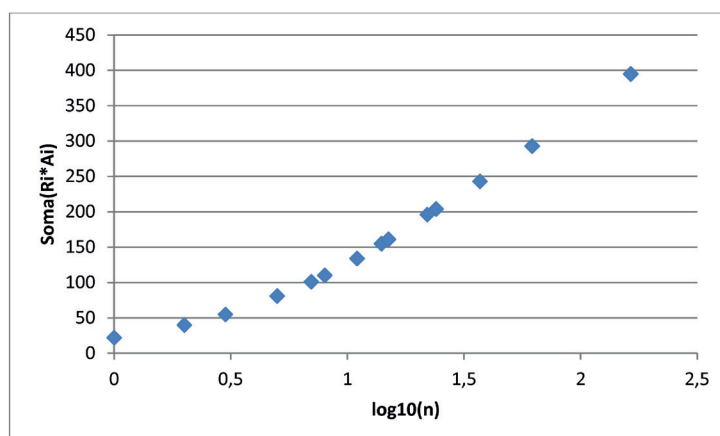
Fonte: Aymard (1980).

R_i é a quantidade de periódicos que produziram uma quantidade correspondente de artigos.

A_i é a quantidade correspondente de artigos durante o período da pesquisa.

Na Figura 31, temos, no eixo x, o logaritmo na base 10 do número de periódicos ($\log_{10}(n)$). No eixo y, temos a soma do produto associada. Já na Figura 32, adiante, pode ser observada a tendência linear à medida que o número de periódicos aumenta.

Figura 32 – Lei de *Bradford*



Fonte: Aymard (1980).

Para mais detalhes matemáticos sobre a aplicação da Lei de *Bradford*, consultar *Aymard* (1980) e *Nunez* (2014, p. 61).



3.4.2 Atividade

Esta atividade atende ao objetivo c: “descrever as Leis de *Bradford* e *Lotka*”.

Atividade: a Lei de *Bradford* é uma das leis bibliométricas. A respeito dela, faça o que se pede:

- a) descreva a Lei de *Bradford*;
- b) *Pinheiro* (1983) destaca possíveis “armadilhas” da Lei de *Bradford*. Por sua vez, *Brooks* (1969) destaca condições para a adequada aplicação da Lei de *Bradford*. Quais são essas condições?

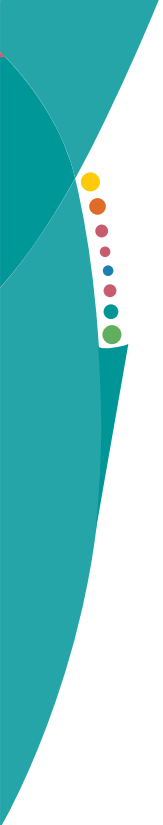
Resposta comentada

- a) A Lei de *Bradford* foi formulada em 1934. Em seu artigo original, *Bradford* utilizou artigos da área de Geofísica. Ele apresentou três zonas de produtividade, cada uma com uma porcentagem do número total de artigos relacionados com o assunto em investigação. *Bradford* identificou que 50% ou mais dos artigos constam de um número bastante pequeno de periódicos.
- b) De acordo com *Brooks* (1969), as condições a seguir devem estar satisfeitas para que a Lei de *Bradford* seja adequadamente aplicada:
 - a definição do tema da bibliografia deve ser efetuada com clareza;
 - não devem ser excluídos da aplicação os periódicos relevantes da área em estudo;
 - o levantamento deve ser desenvolvido por período suficientemente longo para assegurar representatividade.

3.4.3 Leis de *Zipf* (Lei do mínimo esforço)

A Lei de *Zipf* é associada aos estudos sobre frequência de palavras-chave, sendo utilizada para efetuar estimativas de frequências de palavras em um texto técnico. De acordo com a Lei de *Zipf*, um grupo relativamente pequeno de palavras apresenta frequência muito maior que um grande número de palavras.





Foi *George Zipf*, linguista da *Universidade de Harvard*, que, em 1935, pela primeira vez destacou que as palavras apresentam determinada frequência nos textos. *Zipf* observou que há uma relação entre a frequência com que uma palavra aparece em um texto suficientemente longo e sua posição (ordem de série, ou *rank*, em inglês) na lista de palavras ordenadas de acordo com a frequência com que as palavras aparecem no texto.

A primeira Lei de *Zipf* afirma que, se ordenarmos a frequência com que as palavras aparecem nos textos e multiplicarmos o número de ordem pela frequência, obteremos um valor constante:

$$r.f = c,$$

onde *r* é a ordem de série da palavra, *f* é a frequência e *c*, uma constante.

Como um exemplo simples, extraímos de *Maia* (1973, p. 99-100) o texto a seguir:

A palavra *informação* é usada, na linguagem cotidiana, de diferentes maneiras. Falamos de informação "útil", "valiosa", "real", "fidedigna", "precisa", "verdadeira", "relevante"; mas nenhuma dessas expressões ocorre na teoria matemática da comunicação, que descreve a informação tão somente como uma probabilidade estatística de sinais provindos de uma determinada fonte. Apesar de todos sabermos o que é informação, não há maneira de defini-la, exceto através de suas características e propriedades, estudadas pela Ciência da Informação. Dentre as propriedades, a representação da informação é um dos campos mais investigados, em consequência de uma premissa antiga, na qual se identifica o bom funcionamento de um sistema de informação com a representação adequada desta informação – e isto acontece frequentemente. Assim, os diferentes tipos de representação se refletem nas bibliotecas, através do sistema de classificação adotado, dos cabeçalhos de assunto e da catalogação; nos sistemas de recuperação da informação, através das linguagens de indexação; nos sistemas que utilizam o computador, através das linguagens de programação e sua codificação; nos sistemas de transmissão de sinais, através de sua codificação.

Contando as palavras diferentes e ordenando-as, construímos a Tabela 9, a seguir:

Tabela 9 – Exemplo para a Lei de Zipf

	F	R	R * F = C
de	19	1	19
maneiras	11	2	22
informação	8	3	24
através	5	4	20
da	5	5	25
e	4	6	24
a	3	7	21
é	3	7	21
na	3	7	21
nos	3	7	21
o	3	7	21
que	3	7	21
sistemas	3	7	21
uma	3	7	21
codificação	2	8	16
das	2	8	16
diferentes	2	8	16
dos	2	8	16
linguagens	2	8	16

Fonte: Produção do próprio autor (2017).

A frequência de ocorrência de uma palavra, como função da sua posição n em uma lista ordenada, é dada pela expressão $f = \frac{b}{n^a}$, onde b e a são constantes.

Zipf apresentou uma segunda lei (aplicada às palavras de baixa frequência). Em 1967, *Booth* apresentou uma revisão e modificação para esta segunda Lei de *Zipf* (GUEDES; BORSCHIVER, 2005). Por isso, encontramos, na literatura especializada, referências à segunda Lei de *Zipf* como Lei de *Zipf-Booth*.

A segunda Lei de *Zipf* propõe que, para as palavras de baixa frequência em um texto, várias delas apresentarão a mesma frequência. A apresentação desta segunda lei em símbolos matemáticos por *Booth* resultou na expressão:

$$\frac{I_1}{I_n} = \frac{n(n+1)}{2}, \text{ em que:}$$

I_1 representa a quantidade de palavras com frequência igual a 1;

I_n representa a quantidade de palavras com frequência igual a n ;

2 é a constante válida para a língua inglesa.

Ora, se a constante da segunda Lei de *Zipf* para a língua inglesa é 2, qual seria essa constante na língua portuguesa? *Maia* (1973) apresenta pesquisa realizada com três textos em língua portuguesa para buscar essa resposta. Ela determinou o valor 1,5 para a segunda Lei de *Zipf* em língua portuguesa.

Guedes e Borschiver (2005) explicam que *Goffman* observou que a primeira Lei de *Zipf* era válida apenas para a região de palavras de alta frequência de ocorrência. Por outro lado, a segunda Lei de *Zipf* descreve o comportamento de palavras de baixa frequência de ocorrência. Parece razoável imaginar que existe uma região crítica na qual ocorre a mudança no comportamento das palavras de alta frequência para a região de baixa frequência. O número de palavras com frequência “n” que demarca a região de transição é denominado ponto de transição T de *Goffman*. Para o autor (GOFFMAN, 1966), nessa região de transição estão as palavras com elevado conteúdo semântico (são os termos de indexação e as palavras-chave). A quantidade de palavras com alta frequência apresentaria $I_n = 1$. Temos a equação:

$$\frac{I_1}{1} = \frac{n(n+1)}{2}.$$

Resolvendo a equação de 2.º grau: $n^2 + n - 2I_1 = 0$, teremos a raiz positiva

$$n = \frac{-1 + \sqrt{1 + 8I_1}}{2}.$$

Denomina-se esse valor “n” de ponto de transição T de *Goffman*. Esse valor identifica o ponto de transição das palavras de baixa frequência para as de alta frequência (GUEDES; BORSCHIVER, 2005). A importância das Leis de *Zipf* está na construção de processos automáticos de indexação. Deve-se destacar, além da questão da aplicação, a importância específica das Leis de *Zipf*: elas se relacionam a estudos sobre a representação da informação.



Atenção

Existem outros conceitos e linhas de investigação na Bibliometria, além dos que foram mencionados aqui, mas está fora do escopo deste texto apresentá-los. Uma sugestão para você se aprofundar nesse tema é a consulta ao artigo de *Guedes e Borschiver* (2005). Nele, são abordados: as frentes de pesquisa, colégios invisíveis, acoplamento bibliográfico e cocitação, teoria epidêmica de *Goffman*, obsolescência e vida média da literatura.

Outra sugestão é o artigo de *Araújo* (2006) que discute a obsolescência da literatura e sua vida média. O propósito no desenvolvimento das referidas leis foi descrever a queda da validade ou utilidade de informações com o decorrer do tempo. Assim, com os estudos realizados a partir delas, disciplinas como Matemática, Geologia e Botânica apresentam

[...] forte componente de literatura clássica, enquanto outras são quase exclusivamente compostas por literatura efêmera (Física, Engenharia), sendo algumas de caráter intermediário (Fisiologia, Química) (ARAÚJO, 2006, p. 21).

3.4.4 Lei de Lotka

Um estudo sobre a produtividade de autores foi o ponto de partida para a Lei de Lotka, formulada em 1926. Lotka estava interessado em estudar a produtividade científica dos pesquisadores e estudou a produção científica de químicos e físicos (com publicações no *Chemical Abstracts* entre 1909 e 1916 e no *Auerbach's Geschichtstafeln der Physik*, a partir do início da publicação deste periódico, em 1900). Segundo a relação apresentada por Lotka, pequena quantidade de autores é responsável por uma porcentagem bastante grande da produção científica. A relevância do resultado de Lotka está no fato de ele observar que existe uma distribuição desigual na produção científica.

A Lei de Lotka é uma lei de potência: se x representa o número de trabalhos publicados, c é uma constante que varia de acordo com as áreas do conhecimento científico e y é o número de autores que publicaram a quantidade x de artigos então,

$$y = \frac{C}{x^2}.$$

Em termos matemáticos, dizemos que a Lei de Lotka segue a lei do inverso do quadrado.

Talvez você não se lembre do que vem a ser uma lei que segue o padrão “inverso do quadrado”. É quando uma variável (“ y ”, no caso anterior) apresenta associação com uma constante dividida por uma variável (“ x ”, no caso anterior) elevada ao quadrado. Só isso. Ok?

A descoberta de Lotka significa que um número pequeno de autores produz grande parte dos artigos científicos e uma grande quantidade de autores publica relativamente poucos artigos. A importância da Lei de Lotka, segundo Guedes e Borschiver (2005), está na avaliação da produtividade de pesquisadores, na identificação dos centros de pesquisa mais desenvolvidos, em dada área de assunto, e no reconhecimento da “solidez” de uma área científica.

Segundo a Lei de Lotka, para determinado período com um número “ n ” de artigos, a quantidade de pesquisadores que são autores de dois deles é igual a $\frac{1}{4}$ da quantidade de pesquisadores que publicaram um artigo. A quantidade de pesquisadores autores de três artigos é igual a $\frac{1}{9}$ dos pesquisadores que escreveram um artigo.

Desde a publicação do trabalho de Lotka, muitos pesquisadores publicaram refinamentos na lei que leva seu nome, enquanto outros apontaram questionamentos a ela. Um deles é a Lei do elitismo, de Price.



3.4.4.1 Lei do elitismo, de Price

Segundo Araújo (2006, p. 14), “o número de membros da elite corresponde à raiz quadrada do número total de autores e a metade do total da produção é considerada o critério para se saber se a elite é produtiva ou não”. Em símbolos matemáticos,

$$N_{elite} = \sqrt{N_{autores}}.$$

Observe que uma contagem que leva em consideração os coautores dos trabalhos pode acabar resultando em um valor para o expoente distinto da contagem que exclui os coautores. Uma crítica frequente à Lei de Lotka é a ausência, em muitos estudos, de testes estatísticos para provar a hipótese assumida e poder validar o ajuste, ou não, a uma distribuição do tipo Lotka. A esse respeito, consultar o trabalho de Sobrino, Caldes e Guerrero (2008).

3.5 ÍNDICES E INDICADORES

Silva (2001) separa os indicadores bibliométricos em duas categorias: (1) indicadores quantitativos da atividade científica, mensurados pelo número de publicações e (2) indicadores de impacto (consideram o número de citações realizadas sobre cada artigo ou produção científica). Os indicadores de produção científica indicam quais instituições, departamentos, países e pesquisadores possuem maior produção. Contudo, de acordo com o autor, esses indicadores não conseguem apresentar uma estimativa da importância (ou impacto) do trabalho científico produzido. Dois indicadores foram desenvolvidos para aferir esse aspecto: o número de citações de um artigo e o fator de impacto do periódico no qual ele foi publicado. Deve-se, adicionalmente, lembrar que a qualidade de um trabalho científico não é a mesma coisa que o seu impacto.

3.5.1 Fator de impacto

O fator de impacto (FI) calcula a relação entre o número de citações, nos dois anos anteriores à publicação, e o número de artigos publicados nesses dois anos. Em símbolos (LOPES et al., 2012):

$$FatorImpacto_{ano} = \frac{citações_{ano-2} + citações_{ano-1}}{artigos_{ano-2} + artigos_{ano-1}}.$$

Lopes et al. (2012) destacam outro ponto de análise com o FI: diferentes áreas científicas possuem taxas distintas para a chamada vida média de citação (mediana da antiguidade das referências), bem como diferentes taxas de citação.

O FI pode ser utilizado para duas funções:

- a) analisar o impacto de um periódico, para fazer avaliação de coleção e aquisição de um novo título;
- b) avaliar o impacto dos pesquisadores a partir do número de citações aos seus trabalhos.

Atualmente, já não é mais importante apenas publicar; os autores devem ser citados. Nessa questão, o FI é uma ferramenta de muito valor.

Como alerta *Garfield* (2004, p. 4),

o termo fator de impacto evoluiu gradativamente, especialmente na Europa, para descrever tanto o impacto do periódico quanto do autor. Essa ambiguidade frequentemente causa problemas. Uma coisa é usar fatores de impacto para comparar periódicos, e outra bem diferente é usá-los para comparar autores.

Observe que *Garfield* (2004) alerta para o fato de que é bem diferente utilizar fatores de impacto para comparar periódicos e autores. Os FIs para periódicos são calculados a partir de uma “população” razoavelmente grande de artigos e citações. Por outro lado, quando observamos um autor individual, ele produzirá, em geral, números menores.

Garfield também alerta para uma crença equivocada: o tamanho da comunidade que um periódico atende alteraria, de forma significativa, o seu FI. De fato, o elemento determinante no impacto não é nem o número de autores, nem o de artigos em uma área do conhecimento, mas a densidade de citações e a idade da literatura citada. Ou seja, o pesquisador bibliométrico, cientométrico ou webométrico deve estar consciente dos alcances, das limitações e especificidades (que variam de acordo com as áreas de conhecimento) de indicadores como o FI.

3.5.2 H-índice (*h-index*)

Em 2005, *J. Hirsch* propôs a quantificação do impacto e da produtividade dos pesquisadores considerando os artigos mais citados. Suponha que um pesquisador possua h-índice igual a 10. Então, de todos os artigos publicados por esse pesquisador, 10 deles possuem, no mínimo, 10 citações.

O h-índice também pode ser aplicado a países, grupos de pesquisa e periódicos. *Lopes et al.* (2012) alertam para o fato de que, utilizando o h-índice para avaliar investigadores em início de carreira, com baixa produção e/ ou com uma estratégia de publicação mais seletiva, podemos incorrer em alguns erros. Os autores explicam que, entre as críticas ao *h-index*, estão a longevidade da carreira do investigador, sua estratégia de publicação e as diferenças existentes entre as áreas científicas (*LOPES et al.*, 2012).

3.5.3 Fator de impacto na web

Segundo *Ingwersen* apud *Vanti* (2002), o FI na web (ou *Web Impact Factor – WIF*) é dado pela expressão adiante:

$$\text{Fator Impacto Web} = \frac{n.^{\circ} \text{ da página que linka determinado site}}{n.^{\circ} \text{ de páginas do site linkado}}$$



O propósito desse indicador é avaliar a atratividade e a importância que os vários *sites* apresentam na *web*. Vanti (2002) esclarece que os indicadores da *web* podem diminuir (e até mesmo deixar de existir), comportamento diferente do das citações científicas tradicionais.

Destacamos este trecho de Rodrigues (2002, p. 51), citando Velho (1989), além de Meis e Leta (1996):

Há uma série de questionamentos quanto aos critérios adotados pelo *ISI*, o principal deles é que sua abrangência não é total, ou seja, um significativo contingente de publicações é desconsiderado, sobretudo aquelas de países periféricos como os latino-americanos [...] (VELHO, 1989; MEIS; LETA, 1996 apud RODRIGUES 2002, p. 51).

3.5.4 Outros indicadores

Existem outros indicadores bibliométricos (que, por questões de espaço, não abordaremos neste texto), como apontam Lopes et al. (2012): o *Eigenfactor™ Metrics (EF)*, o *Article Influence™ Score (AI)* e o *SCImagoJournalRank (SJS)*.

O *Google Acadêmico*, ou *Google Scholar*, criou uma interessante ferramenta para medir o *h-index*, ou índice-h, dos autores. A vantagem é que todos os trabalhos do autor que receberam citações são contabilizados, diferentemente das bases de dados seletivas, que limitam as citações apenas às revistas indexadas. Considerando a seletividade de bases de dados como a *Web of Science* ou *Scopus*, que você vai estudar na disciplina de “Fontes de Informação”, o *Google Acadêmico* promete ser olhado com atenção pelos avaliadores da produção científica. A *Fundação de Amparo à Pesquisa (FAPESP)* solicita aos pesquisadores que submetam projetos ao *h-index*, de acordo com as medidas do *Google Scholar*.

Spinak (1998) apresenta dezenas de indicadores de publicação, indicadores de citação e técnicas de mapeamento. Já o trabalho de Santos e Kobashi (2009) apresenta técnicas tradicionais e inovadoras de visualização de informação: diagramas de dispersão, gráficos de setores, pictogramas, *tree-map*, dendogramas, grafos ou redes de relacionamentos e rede de relacionamentos dos descritores de registros de uma base de dados. Sugerimos fortemente que você, estudante, consulte esses artigos, para ter uma visão mais detalhada da “floresta” de indicadores bibliométricos existentes.

3.5.5 Limitações dos indicadores bibliométricos

Nem tudo são flores. Há inúmeras críticas aos indicadores bibliométricos:

- a) o número de publicações não é uma medida de qualidade da produção científica. Sugere-se a utilização de indicadores bibliométricos em conjunto com a avaliação por pares;
- b) a produtividade das áreas do conhecimento não é idêntica. Não é razoável supor que Humanidades, Engenharias, Artes Plásticas e Matemática possuam, todas, os mesmos valores numéricos para os diversos indicadores, ou as mesmas tradições de publicação.

Lopes et al. (2012) destacam que, geralmente, as áreas clínicas se caracterizam por menor produtividade, em relação às áreas básicas ou experimentais.

Indicadores de impacto possuem limitações tais como a existência de artigos (raros, mas existem) com muitas citações, que, contudo, são negativas. Por vezes, os autores apresentam referências em seus trabalhos que não foram diretamente utilizadas neles (LOPES et al., 2012). Também existe a barreira da língua: artigos publicados em inglês ou francês tendem a ter mais citações que aqueles publicados em coreano ou russo;

- c) é reconhecido que artigos de revisão (review) e artigos metodológicos apresentam número elevado de citações. Isso também é aceito para periódicos de revisão;
- d) as autocitações (quando um autor cita a si mesmo) podem distorcer os resultados;
- e) Lopes et al. (2012) apontam que o FI é calculado apenas com os dados registrados na *Web of Science*. Como a base de dados é extremamente seletiva, o número de periódicos indexados é reduzido, fazendo com que um grupo pequeno de artigos possa ser analisado pela ferramenta do FI. O *Scielo Citation Index* foi uma alternativa do *Scielo* para minimizar os efeitos de distorção do FI da *Web of Science*. Vamos falar sobre a *Web of Science* e *SciELO* logo adiante.

3.6 AS BASES DE DADOS UTILIZADAS EM PESQUISAS BIBLIOMÉTRICAS

Todos os indicadores bibliométricos apresentados seriam inúteis se não fosse possível conseguir os dados para estudo. Existem bases de dados que são utilizadas nos estudos bibliométricos. As mais usadas atualmente são a *Web of Science*, *Scopus* e, mais recentemente, o *Google Scholar Metrics*, de acordo com Lopes et al. (2012).

3.6.1 *Web of Science*

Em 1955, *Eugene Garfield* apresentou um índice de citação da ciência baseando-se no *Shepard's Citations*, que é um índice de citações de recursos legais utilizado nos Estados Unidos desde 1873. As causas jurídicas se relacionam umas com as outras por meio de uma codificação específica. O sistema então proposto por *Garfield* permite a recuperação de um artigo por meio de outros que o citam. A partir disso, *Garfield* criou o *Institute for Scientific Information (ISI)* e as bases de dados *Arts &*

Humanities Citation Index, Social Science Citation Index, Science Citation Index e Conference Proceedings Citation Index. Todas essas bases estão, na atualidade, na base de dados *Web of Science* (Editora Thomson Scientific). A *Web of Science* indexa milhares de periódicos, nas diversas áreas da ciência, com dados que datam desde o início do século XX. O *Journal Citation Reports (JCR)* é considerado, por Lopes et al. (2012, p. 5), a “ferramenta líder de mercado dos *rankings* de periódicos”.



Multimídia

Sugerimos fortemente que você consulte o *SCImago*, cujo *link* é: <<http://www.scimagojr.com/index.php>>. Por favor, explore a grande diversidade de produção de *rankings* e mapas que o *site* oferece.

No gráfico *Shape of Science* do *SCImago*, você pode acessar indicadores bibliométricos dos periódicos da base de dados do *SJR SCImago Journal & Country Rank*. Nossa sugestão é a de que você consulte este *site* e o explore para observar a potencialidade de mapas e *rankings* que lá podem ser desenvolvidos.

Por exemplo, você poderá produzir um “mapa de bolhas” (*bubblemap*). As variáveis nos eixos x e y podem ser escolhidas entre:

- a) *documents*;
- b) *citable documents*;
- c) *cites*;
- d) *self cites*;
- e) *cites per doc*;
- f) *h-index*;
- g) *% cited documents*.

3.6.2 SciELO

O *Scientific Electronic Library Online (SciELO)* (<<http://www.scielo.org/php/index.php?lang=pt>>) oferece acesso a indicadores bibliométricos (de publicação, de coleção e de citação). No *link* <http://statbiblio.scielo.org/stat_biblio/index.php?lang=en>, você pode consultar dados de citação de periódicos (fator de impacto para dois anos, fator de impacto para três anos, meia-vida, citações recebidas e citações concedidas). Vale lembrar que o *SciELO* foi a primeira iniciativa de publicação de revistas em acesso aberto no mundo e o primeiro a lançar dados de citações dessas publicações.

3.7 SOFTWARES BIBLIOMÉTRICOS

Existe grande variedade de *softwares* para facilitar os trabalhos de pesquisa na bibliometria e cientometria. No Portal da Escrita Científica da *Universidade de São Paulo (USP)* da cidade de São Carlos, você encontrará uma relação de *softwares* bibliométricos e de colaboração científica. Veja a tela copiada do *site* na Figura 33:

Figura 33 – Tela com *softwares* bibliométricos da USP São Carlos



Fonte: *Escrita Científica* (c2012).⁵

Alguns desses *softwares* são de uso gratuito e outros, não. Outros *softwares* utilizados em pesquisas bibliométricas são: *Ucinet* e *Sitkis*. O *link* de acesso ao *Ucinet*, que é um *software* voltado à análise de redes sociais, é: <<https://sites.google.com/site/ucinetsoftware/home>>. Você pode efetuar o *download* (gratuito) do *Ucinet* e usá-lo por um período de 90 dias. Como outros *softwares*, existem tarifas especiais para estudantes. O *Sitkis* é uma ferramenta gratuita para análises bibliométricas e pode ser acessado no *link*: <<https://sites.google.com/site/sitkisbibliometricanalysis/>>. Você não terá dificuldades em encontrar tutoriais e ajuda na *web* para cada um desses *softwares*.



3.7.1 Atividade

Esta atividade atende ao objetivo a: “descrever o objeto de estudo, as variáveis, os métodos e objetivos da bibliometria, infometria, cientometria e webometria”.

Para realizar esta atividade, considere a Figura 30, Intersecção entre as “metrias”, e o Quadro 8, a seguir, ambos extraídos de *Vanti* (2002).

⁵ Disponível em: <<http://www.escritacientifica.sc.usp.br/gerenciadores-bibliometricos/>>.

Quadro 8 – Bibliometria, infometria, cientometria e webometria

Tipologia/ subcampo	Bibliometria	Cientometria	Infometria	Webometria
Objeto de estudo	Livros, documentos, revistas, artigos, autores, usuários	Disciplinas, assuntos, áreas e campos científicos e tecnológicos, patentes, dissertações e teses	Palavras, documentos, bases de dados, comunicações informais (inclusive em âmbitos não científicos), <i>homepages</i> na WWW	Sítios na WWW (URL, título, tipo, domínio, tamanho e <i>links</i>), motores de busca
Variáveis	Número de empréstimos (circulação) e de citações, frequência de extensão de frases	Fatores que diferenciam as subdisciplinas, como os cientistas se comunicam	Difere da cientometria no propósito das variáveis, por exemplo, medir a recuperação, a relevância, a revocação	Número de páginas por sítio, de <i>links</i> por sítio, de <i>links</i> que remetem a um mesmo sítio e de sítios recuperados
Métodos	<i>Ranking</i> , frequência, distribuição	Análise de conjuntos e de correspondência, co-ocorrência de termos, expressões, palavras-chave	Modelo vetor-espaco, modelos booleanos de recuperação, modelos probabilísticos, linguagem de processamento, abordagens baseadas no conhecimento, tesouros	Fator de impacto da <i>web</i> (FIW), densidade dos <i>links</i> , "situações", estratégias de busca
Objetivos	Alocar recursos: pessoas, tempo, dinheiro etc.	Identificar domínios de interesse. Onde os assuntos estão concentrados. Compreender como e quanto os cientistas se comunicam	Melhorar a eficiência da recuperação da informação, identificar estruturas e relações dentro de diversos sistemas de informação	Avaliar o sucesso de determinados sítios, detectar a presença de países, instituições e pesquisadores na rede e melhorar a eficiência dos motores de busca na recuperação das informações

Fonte: Vanti (2002).

- Apresente a distinção dos objetos de estudo da infometria da cientometria;
- identifique, de acordo com a Figura 30, qual das "metrias" pode ser considerada como a mais ampla e justifique sua resposta;
- de acordo com o Quadro 8, quais os principais métodos da bibliometria?;
- de acordo com a Figura 30, existem intersecções entre os campos da bibliometria e da webometria. Apresente exemplos nos quais temos essa intersecção;
- de acordo com a Figura 30, existem intersecções entre os campos da cientometria e da webometria. Apresente exemplos nos quais temos essa intersecção.

Resposta comentada

- a) Enquanto a cientometria tem como objeto de estudo as disciplinas, assuntos, áreas e campos científicos e tecnológicos, a infometria trata de palavras, documentos e bases de dados (mesmo que não científicos).
- b) A Figura 30 está de acordo com análises efetuadas por *Russel* apud *Vanti* (2002, p. 10) e *Tague-Sutcliffe* apud *Vanti* (2002, p. 10), segundo as quais a infometria compreende a bibliometria, a cientometria e a webometria. A infometria é o termo considerado “guarda-chuva”, abrangendo os conceitos de bibliometria, cientometria e webometria.
- c) De acordo com o Quadro 8, os principais métodos da bibliometria são *ranking*, frequência e distribuição.
- d) No caso da intersecção entre bibliometria e webometria, em particular no que diz respeito aos objetivos, um pesquisador pode estar interessado na alocação de recursos, pessoas, tempo e dinheiro, que se relaciona com o sucesso de determinados sítios.
- e) No caso da intersecção entre cientometria e webometria, em particular no que diz respeito às variáveis, um pesquisador pode estar interessado em estudar fatores que diferenciam as subdisciplinas e em como os cientistas se comunicam avaliando o número de *links* que remetem a um mesmo sítio.

3.8 CONCLUSÃO

Concluimos esta Unidade destacando que os métodos quantitativos são de grande importância, seja ela teórica (ao propor leis e padrões de dispersão), seja ela prática (ao possibilitar estimativas e previsões sobre acervos e coleções).

Contudo, nem todos os valores numéricos dos parâmetros das leis bibliométricas são idênticos entre as várias áreas ou línguas (por exemplo, lembremo-nos da diferença entre a constante da segunda Lei de *Zipf* para a língua inglesa e para outras línguas).

De forma análoga à discussão sobre os alcances e limitações que fizemos na Unidade anterior, com relação aos métodos quantitativos gerais, temos também a discussão sobre a limitação dos indicadores bibliométricos. Em particular, não podemos comparar a produtividade das diferentes áreas de conhecimento. Não podemos esperar que Artes Plásticas, Física Nuclear e Direito possuam, todas, os mesmos valores numéricos e de produtividade para os diferentes indicadores quantitativos. É possível encontrar artigos que apresentam citações que não foram, efetivamente, utilizadas nos trabalhos. Outro aspecto: é natural esperar que periódicos voltados a artigos de revisão (*review*) e artigos metodológicos apresentem elevado número de citações. Você deve se lembrar, ainda, de ressalvas quanto a eventuais inconsistências lógicas apresentadas pelos motores de buscas.



RESUMO

Nesta Unidade, você estudou alguns dos principais instrumentos quantitativos utilizados pelos bibliotecários e cientistas da informação para análise da produção científica e avaliação de coleções, bem como para efetuar mensurações no registro e difusão da informação e do conhecimento. Você também estudou as Leis de *Bradford*, *Zipf* e *Lotka*. Conheceu o ponto T de *Goffman* e a Lei do elitismo, de *Price*.

Você estudou as diversas “metrias” da Biblioteconomia: a bibliometria, a cientometria, a infometria e a mais recente delas, a webometria, abordando suas semelhanças e diferenças com relação a objetos de estudo, objetivos, variáveis e métodos em cada uma delas.

Entrou em contato com os principais índices e indicadores utilizados na Biblioteconomia: fator de impacto, h-índice e fator de impacto na *web*.

Também viu o quão fundamentais à Biblioteconomia e à Ciência da Informação são a *Web of Science*, o *Scopus*, *SCImago*, *Scielo* e o *Google Acadêmico*.



Sugestão de Leitura

TAGUE-SUCKFFE, J. An introduction to informetrics. **Information Processing & Management**, [S.l.], v. 28, n. 1, p. 1-3, 1992.

REFERÊNCIAS

ALMIND, T. C.; INGWEERSEN, P. Informetric analyses on the World Wide Web: methodological approaches to “webometrics”. **Journal of Documentation**, [S.l.], v. 53, n. 4, p. 404-426, 1997.

ARAÚJO, C. R. Bibliometria: evolução histórica e questões atuais. **Em Questão**, Florianópolis, v. 12, n. 1, p. 11-32, 2006.

ARAÚJO, R. F.; ALVARENGA, L. A. Bibliometria na pesquisa científica da pós-graduação brasileira de 1987 a 2007. **Encontros Bibli**, Florianópolis, v. 16, n. 31, p. 51-70, 2011.

AYMARD, M. A Lei de dispersão bibliográfica de Bradford. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 13, n. 3-4, p. 147-156, 1980.

BAR-LLAN, J. Informetrics at the beginning of the 21st century: a review. **Journal of Informetrics**, [S.l.], v. 2, p. 1-52, 2008.

BUFREM, L.; PRATES, Y. O saber científico registrado e as práticas de mensuração da informação. **Ciência da Informação**, Brasília, v. 34, n. 2, p. 9-25, 2005. Disponível em: <<http://www.scielo.br/pdf/ci/v34n2/28551.pdf>>. Acesso em: 21 fev. 2015.

CHAVIANO, O. G. Algunas consideraciones teórico-conceptuales sobre las disciplinas métricas. **Acimed**, Ciudad de La Habana, v. 12, n. 5, 2004. Disponível em: <http://bvs.sld.cu/revistas/aci/vol12_5_04/aci07504.htm>. Acesso em: 21 fev. 2015.

COUTINHO, E. As armadilhas da Lei de Bradford. **Revista de Biblioteconomia**, Brasília, v. 16, n. 2, p. 217-225, 1988. Disponível em: <<http://www.brapci.ufpr.br/documento.php?dd0=0000002566&dd1=2410e>>. Acesso em: 21 fev. 2015.

FIGUEIREDO, Nice. **Tópicos modernos em Bibliometria**. Brasília: Associação dos Bibliotecários do Distrito Federal, 1977.

GOFFMAN, W. Mathematical approach to the spread of scientific ideas: the history of mast cell research. **Nature**, [S.l.], v. 212, p. 449-452, 1966.

GOSNELL, C. Obsolescence of books in college library. **College & Research Libraries**, [S.l.], v. 5, p. 115-125, 1944.

GUEDES, V. L.; BORSCHIVER, S. Bibliometria: uma ferramenta estatística para a gestão da informação e do conhecimento, em sistemas de informação, de comunicação e de avaliação científica e tecnológica. In: ENCONTRO NACIONAL DE CIÊNCIA DA INFORMAÇÃO, 6., 2005, Salvador. **Anais...** Salvador: [s.n.], 2005.

LOPES, S. et al. A Bibliometria e a avaliação da produção indicadores e ferramentas. In: CONGRESSO NACIONAL DE BIBLIOTECÁRIOS, ARQUIVISTAS E DOCUMENTALISTAS, 11., 2012, Lisboa. **Actas...** Lisboa: [s.n.], 2012.

LOUSADA, M. et al. Produção científica sobre gestão do conhecimento e gestão da informação no âmbito da Ciência da Informação: uma aplicação da Lei de Bradford. **Anales de Documentación**, Murcia, v. 15, n. 2, 2012. Disponível em: <<http://eprints.rclis.org/18034/>>. Acesso em: 21 fev. 2015.

LUCAS, E. O.; ZORITA, J. C. G.; CASADO, E. S. Evolução histórica de investigação em informetria: ponto de vista espanhol. **Liinc em Revista**, Rio de Janeiro, v. 9, n. 1, p. 255-270, 2013. Disponível em: <<http://www.ibict.br/liinc>>. Acesso em: 21 dez. 2014.

MAIA, E. L. S. Comportamento bibliométrico da língua portuguesa, como veículo de representação da informação. **Ci. Inf.**, Rio de Janeiro, v. 2, n. 2, p. 99-138, 1973.

MEIS, L. J. **O perfil da ciência no Brasil**. Rio de Janeiro: Ed. UFRJ, 1996. p. 13-37.



NUNEZ, Z. A. G. **A produção científica brasileira em Medicina Tropical indexada nas bases de dados Web of Science e Scopus entre os anos de 2005 e 2012**. 2014. Dissertação (Mestrado em Comunicação e Informação) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2014. Disponível em: <<https://www.lume.ufrgs.br/bitstream/handle/10183/95387/000917848.pdf?sequence=1>>. Acesso em: 21 fev. 2015.

PINHEIRO, L. V. R. Lei de Bradford: uma reformulação conceitual. **Ciência da Informação**, Brasília, v. 12, n. 2, p. 59-80, 1983. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1498/1116>>. Acesso em: 21 fev. 2015.

RODRIGUES, N. A. **Pesquisa agropecuária oficial do Rio Grande do Sul: a temática da produção técnico-científica no período 1990/1998**. 2002. Dissertação (Mestrado em Comunicação e Informação) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2002.

SANTIN, D. M. Avanços e perspectivas da infometria e dos indicadores multidimensionais na análise de fluxos da informação e estruturas do conhecimento. **Encontros Bibli**, Florianópolis, v. 16, n. 32, p. 107-122, 2011.

SANTOS, R. N. M.; KOBASHI, N. Y. Bibliometria, Cientometria, Infometria: conceitos e aplicações. **Pesquisa Brasileira em Ciência da Informação**, João Pessoa, v. 2, n. 1, p. 155-172, 2009. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/view/21/43>>. Acesso em: 21 fev. 2015.

SILVA, J. A.; BIANCHI, M. L. P. Cientométrica: a métrica da Ciência. **Paideia**, Ribeirão Preto, v. 11, n. 20, p. 5-10, 2001. Disponível em: <<http://www.scielo.br/pdf/paideia/v11n21/02.pdf>>. Acesso em: 21 fev. 2015.

SILVA, M. R.; HAYASHI, C. R.; HAYASHI, M. C. P. I. Análise bibliométrica e cientométrica: desafios para especialistas que atuam em campo. **InCID: Revista de Ciência, Informação e Documentação**, Ribeirão Preto, v. 2, n. 1, p. 110-129, 2011.

SOBRINO, M. I. M.; CALDES, A. I. P.; GUERREIRO, A. P. Lei de Lotka aplicada à produção científica da área de Ciência da Informação. **Brazilian Journal of Information Science**, Marília, v. 2, n. 1, p. 16-32, 2008. Disponível em: <<http://www.bjis.unesp.br>>. Acesso em: 21 fev. 2015.

SPINAK, E. Indicadores cientométricos. **Ciência da Informação**, Brasília, v. 27, n. 2, p. 141-148, 1998. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/349/1690>>. Acesso em: 21 fev. 2015.

VANTI, N. A. P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência da Informação**, Brasília, v. 31, n. 2, p. 152-162, 2002. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/171/150>>. Acesso em: 25 fev. 2015.

VELHO, L. Avaliação acadêmica: a hora e a vez do “baixo clero”. **Ciência e Cultura**, Campinas, v. 41, n. 10, p. 57-68, out. 1989.

UNIDADE 4

MÉTODOS QUALITATIVOS

4.1 OBJETIVO GERAL

Apresentar as características da pesquisa qualitativa, bem como as principais ferramentas de coleta e análise de dados.

4.2 OBJETIVOS ESPECÍFICOS

Esperamos que, ao final desta Unidade, você seja capaz de:

- a) formular problemas para pesquisas qualitativas;
 - b) definir a amostra inicial em pesquisas qualitativas;
 - c) coletar os dados entrando no ambiente da pesquisa (vamos ao campo?);
 - d) analisar os dados em pesquisas qualitativas;
 - e) aplicar *softwares* nas pesquisas qualitativas.
-

4.3 FORMULAÇÃO DE PROBLEMAS EM PESQUISAS QUALITATIVAS

Uma característica preponderante nas pesquisas qualitativas é que elas não são baseadas em um arcabouço teórico-metodológico unificado: existe uma extensa variedade de métodos e abordagens e grande valorização da subjetividade do pesquisador. Há também uma preponderância do tema de pesquisa sobre os métodos, além da flexibilidade e abertura do pesquisador, de quem se exige uma atividade permanente de reflexão.

Após o pesquisador decidir a ideia geral do seu estudo de pesquisa qualitativa, é necessário que ele formule o problema dela. *Creswell* (2010) denomina esse item do projeto de pesquisa de “declaração de objetivo de pesquisa”. A formulação de um problema de pesquisa qualitativa usualmente apresenta, conforme *Sampieri, Collado e Lucio* (2013):

- a) os objetivos da pesquisa;
- b) as perguntas da pesquisa;
- c) justificativa e viabilidade da pesquisa;
- d) avaliação de deficiências sobre o que se conhece daquele problema;
- e) caracterização do contexto/ local e do ambiente de pesquisa.

Nesta etapa, você deve estar preocupado em responder por que pretende fazer determinada pesquisa e em dizer quais são seus objetivos com ela.

Algumas recomendações, para quando você for escrever seu problema de pesquisa qualitativa, são: os objetivos devem estar redigidos com clareza e precisão, deixando nítido ao leitor as delimitações entre cada um. Cada objetivo deve estar registrado em frases distintas, para certificar que cada um deles foi cumprido ao longo da pesquisa. É comum, conforme você vai desenvolvendo seu projeto, que, em paralelo, você revise suas questões de pesquisa, de forma que elas fiquem mais bem definidas e sucintas.

As perguntas de pesquisa são as questões que, ao final do projeto, devem estar respondidas. Você deve se certificar de que elas estão relacionadas com os objetivos. *Creswell* (2010) sugere que as questões de pesquisa comecem com as palavras “o que” ou “como”, de forma a transmitir “um projeto aberto e emergente”. O autor entende que o uso frequente do *por quê* implica em que o pesquisador esteja tentando explicar o motivo pelo qual algo ocorre, e isso sugere um pensamento do tipo causa-efeito que pode ser associado à pesquisa quantitativa, em vez de à postura mais aberta e emergente da pesquisa qualitativa (*CRESWELL*, 2010).

A justificativa deve explicitar os motivos da pesquisa. Esse item será particularmente relevante quando for tomado como elemento de decisão por eventuais órgãos financiadores da pesquisa.



Quanto à questão da viabilidade, não deixe de avaliar os recursos necessários à pesquisa, como:

- a) tempo para tabulação dos dados;
- b) entrevistas;
- c) apoio técnico;
- d) equipamentos;
- e) competências necessárias para a condução do projeto.

Pesquisadores novatos não costumam considerar que, para uma entrevista, não necessitamos apenas do tempo que a atividade leva. Precisamos levar em conta o tempo necessário para localizar o entrevistado, conseguir autorização para entrevistá-lo e a chegada ao local da entrevista. Após a atividade, é necessário tempo (que não é pouco) para que seja efetuada a transcrição da entrevista. *Morse apud Flick (2009, p. 133)* sugere que

[...] para aqueles que transcrevem com maior agilidade, que o tempo de duração da fita que contém o registro da entrevista seja multiplicado por quatro. No entanto, ao incluir-se aí o tempo necessário à verificação da transcrição em comparação com a fita, o tempo de duração da fita deve ser multiplicado por seis.

No que diz respeito à avaliação das deficiências da literatura especializada sobre o problema, você deverá explicitar em que medida sua pesquisa vai contribuir para o estado atual do conhecimento sobre o tema. Espera-se que, nessa revisão da literatura, você:

- a) identifique conceitos relevantes para o seu projeto de pesquisa, mas que você desconhecia;
- b) avalie estratégias de coleta e análise de dados para pesquisas qualitativas, investigando qual o mais adequado para o seu caso;
- c) identifique problemas e equívocos que tenham ocorrido em situações análogas;
- d) aprofunde sua compreensão acerca do problema de pesquisa.

Na caracterização do contexto, devem ser expostos os grupos de pessoas que serão estudados e os locais. Por exemplo:

- a) grupos:
 - professores do Ensino Médio?;
 - estudantes de Pós-Graduação?;
 - mulheres com nível superior?
- b) locais:
 - instituição;
 - empresa;
 - departamento;
 - órgão.

O local deverá ser descrito com detalhes, para que o leitor compreenda as especificidades da pesquisa.



Atenção

Creswell (2010, p. 145) apresenta um roteiro com “as principais palavras e ideias de uma declaração de objetivos de pesquisa qualitativa”.

Diferentemente dos estudos quantitativos, as hipóteses em estudos qualitativos não serão testadas estatisticamente. Elas não costumam ser estabelecidas antes de o pesquisador visitar o local da pesquisa. Segundo Sampieri, Collado e Lucio (2013, p. 382) “as hipóteses de trabalho qualitativas são gerais ou amplas, emergentes, flexíveis e contextuais, adaptam-se aos dados e às mudanças no decorrer da pesquisa”.

Para jovens pesquisadores mais habituados às pesquisas quantitativas, isso pode criar dificuldades. Segundo Creswell (2010, p. 162), “na pesquisa qualitativa, a intenção é explorar o conjunto complexo de fatores que envolvem o fenômeno central e apresentar as perspectivas ou os significados variados dos participantes”.

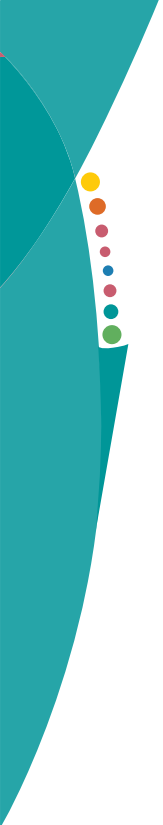
4.4 DESENHOS DO PROCESSO EM PESQUISAS QUALITATIVAS

Na seção 2.4 da Unidade 2, tratamos do delineamento das pesquisas com abordagens quantitativas. Nesta seção, faremos o equivalente com pesquisas qualitativas. Discutiremos a Teoria fundamentada, a Etnografia, o Estudo de caso e as Histórias de vida (ou desenhos narrativos).

4.4.1 Teoria fundamentada (*Grounded Theory*)

A Teoria fundamentada ou *Grounded Theory* (também conhecida como Teoria fundamentada nos dados) privilegia os dados e a pesquisa de campo em relação às suposições teóricas. A metodologia da Teoria fundamentada não prevê a aplicação de teorias ao sujeito em estudo. Em vez disso, esta abordagem assume que as teorias sejam formuladas e desenvolvidas a partir dos dados e informações obtidos do campo, extraindo-lhes o significado a partir do que é vivenciado pela pessoa e produzindo conhecimentos mais profundos sobre determinado contexto.





A Teoria fundamentada nos dados fornece um procedimento sistemático de investigação qualitativa que abre possibilidades para a constituição de uma compreensão mais profunda para fenômenos ainda pouco estudados. Esse novo corpo de conhecimentos possui relevância para a prática da Biblioteconomia, já que possibilita a emergência de teorias a partir do estudo de interações humanas e o conjunto de significados que pode ser extraído delas.

Nos anos 1960, os métodos quantitativos de pesquisa apresentavam cada vez mais sofisticação e relevância nos EUA. Os sociólogos norte-americanos *Anselm L. Strauss* e *Barney G. Glaser* publicaram, em 1967, *The Discovery of grounded theory: strategies for qualitative research*. As estratégias metodológicas sistemáticas desenvolvidas pelos referidos sociólogos permitem a elaboração de explicações teóricas que não fazem uso de conceitos e ideias pré-existentes. A Teoria fundamentada apresenta um conjunto de métodos que permitem a construção de uma teoria a partir dos dados investigados em determinado contexto, fazendo uso de categorias conceituais organizadas. Assim, possibilita a explicação do fenômeno estudado. Na área da Ciência da Informação, a Teoria fundamentada foi adotada a partir da década de 1980 (BASTOS, 2013).

A Teoria fundamentada é considerada uma abordagem variante do Interacionismo simbólico. Ambos enfatizam a compreensão do fenômeno e a forma como ele pode ser lido a partir dos dados, e não na sua fundamentação a partir de teorias adotados pelo pesquisador.

Na Teoria fundamentada, a revisão da literatura não consiste no primeiro passo da pesquisa. Os próprios dados e sua análise, por parte do pesquisador, orientarão quais direções deverão ser adotadas na revisão da literatura. Antes e ao longo da análise, o uso da literatura é limitado, de forma que novos aspectos do fenômeno estudado possam ser “impedidos” de serem visualizados pelo pesquisador.

A criação das hipóteses não ocorre antes de o pesquisador ir a campo (esta é uma diferença fundamental entre as abordagens qualitativas e as quantitativas). A criação das hipóteses é realizada após o início da coleta e análise de dados. Uma importante distinção entre a Teoria fundamentada e outras metodologias de pesquisa é a oposição entre a testagem de hipóteses (utilizada em outras metodologias) e a noção de emergência (surgimento), desenvolvida pela Teoria fundamentada. Aqui, as hipóteses “emergem” dos dados de campo e não são “trazidas prontas” pelo pesquisador. Como *Glaser* sugere, existem dois principais critérios de avaliação da adequação de uma teoria emergente: se ela se adequa à situação e se ela funciona (em outras palavras: se é útil para as pessoas, para compreenderem dada situação e administrá-la melhor).

4.4.1.1 Coleta dos dados

A Teoria fundamentada admite diversas técnicas de coleta de dados: observação participante, entrevistas, discursos, cartas, biografias, autobiografias, pesquisas na biblioteca (DICK, 2005 apud PINTO, 2012). Além disso, na Teoria fundamentada, existe um processo concomitante de coleta e análise de dados e de construção de categorias analíticas tendo como ponto de partida os dados obtidos em campo (e não a partir de hipóteses definidas antes da ida ao campo). Também são consideradas comparações a todo o tempo e em cada etapa do processo de pesquisa.

Há uma circularidade no método (diz-se também que o método é circular). À medida que o pesquisador coleta mais dados, ele pode adotar novos pontos de atenção (novos focos).

4.4.1.2 Amostragem

Ao selecionar o objeto de estudo, a Teoria fundamentada preconiza que aqueles de maior relevância para o estudo sejam escolhidos, em oposição à abordagem quantitativa, que exige amostras estatisticamente representativas da população.

Na Teoria fundamentada adota-se o conceito de *amostragem teórica*: nesse processo de amostragem, o pesquisador, concomitantemente à coleta de dados, também os codifica e analisa. Em seguida, decide quais dados serão coletados na sequência, de forma a contribuir com a construção da teoria emergente. O controle do processo é dado pela teoria que está sendo elaborada.

4.4.1.3 Saturação

A noção de saturação significa que os dados são coletados até o momento em que o pesquisador consegue identificar que dados com relevância ou informação nova não possam ser localizados (começa a ocorrer a repetição de padrões). Nesse ponto, os especialistas da Teoria fundamentada utilizam a expressão “sensibilidade teórica”: é a capacidade do olhar do pesquisador para com os dados. Ele consegue distinguir os dados relevantes para a pesquisa dos não relevantes.

4.4.1.4 Codificação

Na codificação, são utilizados procedimentos para categorizar os dados coletados. Nela, fazemos uso de comparação constante: você codifica cada entrevista comparando-a com as entrevistas anteriores. *Charmaz* (2009) aponta que a codificação pode ser feita palavra a palavra, linha por linha, incidente por incidente ou segmento por segmento. A cada instante você se pergunta: o que pode estar acontecendo aqui? Como esta pessoa administra tal situação? Quais categorias são sugeridas por esta afirmação, feita por este indivíduo?

Ao mesmo tempo que o pesquisador codifica, ele deve estar atento a eventuais teorias que venham à sua mente, devendo anotá-las quando surgirem.

Existem os procedimentos de codificação aberta, codificação axial e codificação seletiva. O ponto, aqui, é que cada uma dessas codificações não apresenta fronteiras claramente definidas. Também, nada obriga a que elas estejam distantes (no tempo) entre si:

- a) *codificação aberta*: na definição de *Strauss* e *Corbin* (1990, apud *GASQUE*, 2007, p. 97), a “codificação aberta é o processo analítico pelo qual os conceitos são identificados e desenvolvidos em relação às suas propriedades e dimensões”. Com a codificação aberta, o pesquisador divide, examina, compara e classifica os dados em categorias. Situações e ocorridos que apresentem semelhança devem ser agrupados nas mesmas categorias, para a rotulagem. Tal como definido por *Sampieri*, *Collado* e *Lucio* (2013, p. 458), “categorias são conceitos, experiências, ideias, fatos relevantes e com significado”;

- b) *codificação axial*: na codificação axial, o pesquisador refina as categorias obtidas na codificação aberta. Aquelas categorias de maior importância são escolhidas para que possam ser avaliadas relações entre as categorias e subcategorias. É na codificação axial que o pesquisador terá condições de identificar semelhanças e diferenças entre as categorias, passando de uma forma de pensar indutiva para uma dedutiva;
- c) *codificação seletiva*: o objetivo desta última etapa do processo de codificação é a integração e o refinamento das categorias com maior grau de abstração.

Se sua categoria central (*core category*, em inglês) e as categorias relacionadas saturam, então é sinal de que as coletas não são mais necessárias. A categoria central deve ser ampla o suficiente para abarcar (e relacionar) as categorias secundárias (GASQUE, 2007).

Este é um sinal de que é o momento de se deslocar para ordenação, categorização, sequenciamento ou hierarquização. Você agrupa os seus memos em categorias com aqueles que se assemelham e cria uma sequência ou algum tipo de ordenação para tornar sua teoria mais clara.

4.4.1.5 Registro (*memos*)

A Teoria fundamentada faz uso do registro por meio de memorandos (ou *memos*): com tais registros, são construídas categorias, efetuadas relações entre elas e eventuais discrepâncias são identificadas.

Os *memos* (memorandos) são registros, e se dividem em notas teóricas, notas metodológicas e notas de observação. Essas notas são efetuadas ao longo da coleta e análise de dados. Com os *memos*, o pesquisador tem uma ferramenta para produzir reflexão a partir dos dados e conjecturas. O memorando representa ponto fundamental na pesquisa, veja:

A redação dos memorandos permitiu elaborar anotações analíticas para explicar posteriormente as codificações elaboradas, refletir sobre os dados e descobrir ideias sobre eles, identificar pontos a serem checados, identificar relações, elaborar comparações, descobrir lacunas na coleta de dados e produzir material para os capítulos de análise e interpretação de dados (PAVÃO, 2014, p. 132).

De acordo com Gasque (2007, p. 106), os memorandos podem ser subdivididos em notas teóricas, notas de códigos (registram os três tipos de codificação) e notas operacionais.

A forma como você organizou os seus *memos* vai criar uma estrutura (um “esqueleto”) para sua pesquisa. A partir disso, você começa o processo de redação do seu trabalho.

4.4.1.6 Redação da teoria

Finalmente, o pesquisador tem um conjunto de materiais e instrumentos para produzir a teoria: relações entre as categorias e memorandos redigidos, entre outros. De acordo com Strauss e Corbin (1990, p. 229, apud GASQUE, 2007), o pesquisador deve ter construído “a especificação clara das relações entre as categorias, também com o esclarecimento dos níveis de conceitualização”.

Além desse aspecto, o pesquisador deverá ter clareza de dois procedimentos:

- a) uma história analítica que não deixe espaço para ambiguidades, a partir dos memorandos e diagramas;
- b) um arcabouço não definitivo que, durante o próprio processo de redação da escrita, evoluirá até englobar demais aspectos do estudo.

4.4.2 Etnografia

O significado literal da palavra etnografia é “descrição de um povo” (ANGROSINO, 2009, p. 16). A importância da etnografia e da observação participante é muito grande no desenvolvimento das pesquisas qualitativas em geral. *Malinowski*, com a obra *Argonauts of the Western Pacific* (*Argonautas do Pacífico Ocidental*), é considerado o responsável por profunda transformação metodológica na Antropologia, ao destacar o contato do pesquisador com o campo e a observação direta (o “viver como vivem os nativos”) (AMARAL, 2013).

Nos seus primórdios, a etnografia, originada da Antropologia, investigou culturas distantes. Mais recentemente, voltou-se a grupos mais próximos do nosso cotidiano (FLICK, 2009). Com o trabalho desenvolvido por sociólogos da *Universidade de Chicago*, a partir dos anos 1920, a pesquisa etnográfica passou a ser instrumento utilizado na investigação em grupos sociais “modernos” (ANGROSINO, 2009). Por exemplo, a etnografia é utilizada para investigar o comportamento social em praticantes de atividades esportivas (triatlo) em Berlim (FLICK, 2009, p. 111), ou o de idosos em cursos para navegação na internet (TEZZA; BONIA, 2010). Dessa forma, não apenas a Antropologia, mas também a Ciência da Informação, a Sociologia, a Educação, a Enfermagem, a Administração e outras áreas do conhecimento se beneficiam das técnicas e procedimentos desenvolvidos ao longo das últimas décadas pela etnografia.

Na etnografia, o pesquisador não faz uso de conceitos para descrever o contexto, mas descreve eventos (TAYLOR; BOGDAN, 1997).



Atenção

Há uma ampla gama de orientações teóricas na abordagem etnográfica: funcionalista, interacionismo simbólico, feminismo, marxismo, etnometodologia, teoria crítica, estudos culturais e pós-modernismo. Ao leitor interessado em mais detalhes sobre cada uma dessas orientações, é sugerido consultar *Angrosino* (2009, p. 19-30).

A vivência prolongada do pesquisador etnográfico entre os “nativos” e sua interação no dia a dia são aspectos fundamentais na metodologia etnográfica.



Na etnografia, o pesquisador não formula hipóteses após a definição do problema de pesquisa: elas são produzidas na conclusão do estudo (SPRADLEY, 1980, apud AMARAL, 2013).

4.4.2.1 Coleta de dados

Na pesquisa etnográfica, o pesquisador possui flexibilidade sobre como e quando deve ser a coleta de dados. A observação participante é o principal instrumento para isso, sendo também utilizadas a entrevista, a análise de documentos e a história de vida.

A entrevista frente a frente, no ambiente natural do entrevistado, é característica típica da pesquisa etnográfica. O pesquisador vai até a “ilha” do pesquisado; o pesquisado não é trazido ao “laboratório” do pesquisador.

Também podem ser utilizadas, na pesquisa etnográfica, fotografias (é a chamada fotoetnografia) e filmagens, como fontes de informação. Nesse caso, a fotografia é recurso utilizado para a construção de narrativas, deixando de ser mera ilustração.

4.4.2.2 Notas de campo

A seguir, destacamos algumas das recomendações de Angrosino (2009) acerca das notas de campo do pesquisador etnográfico. Elas devem possuir identificação de data, local e hora da observação. O pesquisador deve produzir anotações literais, utilizando as próprias palavras e termos dos entrevistados. Os registros devem obedecer à sua sequência temporal. Os entrevistados devem ter seu anonimato garantido pelo pesquisador.

Na pesquisa etnográfica, a amostragem é função do grupo estudado. É possível, em determinados contextos de pesquisa, que o pesquisador entreviste, em profundidade, um único sujeito.

4.4.2.3 Triangulação na pesquisa etnográfica

A triangulação de métodos consiste na utilização de duas ou mais metodologias na busca de uma abordagem mais completa do problema de pesquisa.

Na etnografia, o uso da triangulação tem sido considerado um aspecto relevante para a promoção da qualidade na pesquisa (FLICK, 2009). Assim, para a obtenção de dados, são considerados diversos métodos: análises de documentos, observação, entrevistas e obtenção de materiais de arquivo.

Como exemplo de triangulação na etnografia, recomendamos a leitura de Flick (2009, p. 107-118).

4.4.2.4 Análise de dados etnográficos

Após a coleta de dados, o pesquisador iniciará o processo de analisá-los. Uma grande massa de dados foi produzida no longo tempo de interação do pesquisador com o grupo estudado. Uma das suas tarefas é efetuar a transcrição das gravações e/ ou filmagens.

Um dos objetivos do pesquisador é a identificação de padrões nas entrevistas e observações. Como lemos em Angrosino (2009, p. 46), “um

verdadeiro padrão é aquele que é partilhado pelos membros de um grupo (seu comportamento real) e/ ou que se acredita desejável, legítimo ou correto pelo grupo (seu comportamento ideal)".

Não existem "receitas de bolo" ou procedimentos específicos sobre análise de dados para o pesquisador etnográfico. A análise de dados é desenvolvida "caso a caso". Mesmo assim, existem etapas frequentemente adotadas na análise de dados etnográfica:

uma cuidadosa administração da enorme massa de dados produzida pelo trabalho de campo; a realização das denominadas leituras panorâmicas para rememoração de detalhes do campo; a explicitação das categorias utilizadas. (ANGROSINO, 2009, p. 47).

4.4.2.5 Dificuldades e eventuais limitações da pesquisa etnográfica

Como outras abordagens, a etnografia também apresenta dificuldades e limitações. A pesquisa etnográfica exige do pesquisador "um esforço intenso para minimizar os riscos de omissão ou da revelação de dados distorcidos por parte do grupo investigado" (VERGARA, 2012, p. 66).

Como é uma característica da pesquisa etnográfica que o pesquisador "vá a campo" por um período razoavelmente prolongado, esse tipo de investigação exige esse tempo para sua execução.

O pesquisador não deve se esquecer de que pode encontrar severas dificuldades para se aproximar do grupo estudado.

Que tal fazer um pequeno exercício (adaptado de FLICK, 2009, p. 217)?

Você pode praticar um pouco de observação participante indo a alguns locais de seu ambiente de estudo (ou de trabalho). Por exemplo, a lanchonete, cyber cafés ou bibliotecas. Como são os grupos nesses ambientes? Como é a forma de relacionamento entre as pessoas nesses locais? Você observou situações de discriminação (de gênero, de raça ou outras)? Como as pessoas se aproximam e interagem? Desenvolva notas de campo a partir de suas observações e redija um relatório ao final do processo. Releia sobre como conduzir uma observação participante, tomar notas de campo e produzir um relato etnográfico. Como seria desenvolvida a triangulação neste seu caso? Você conduziu triangulação entre observação e entrevistas?

4.4.3 Estudo de caso

No estudo de caso, é realizada uma investigação aprofundada de um fenômeno em seu próprio contexto real. É uma abordagem bastante utilizada nas Ciências Sociais. O propósito, ao se adotar a abordagem de estudo de caso, é generalizar teorias (generalização analítica) e não efetuar uma generalização estatística (tal como concebido nas pesquisas quantitativas).

Nos estudos de caso, podemos estar interessados em estudar uma única pessoa, um projeto, grupos de pessoas, um departamento ou gerência de uma organização ou mesmo a organização como um todo. Não há impedimento para a utilização de dados quantitativos nos estudos de caso.

O estudo de caso pode ser classificado em estudo de caso único (com apenas uma unidade de análise) ou estudos de casos múltiplos (são investigadas várias unidades de análise). O caso a ser estudado deve ser relevante e de interesse público.

Yin (2005) ensina que os estudos de caso podem ser classificados em:

- a) explanatórios, quando são estabelecidas relações causais entre duas variáveis para a explicação de um fenômeno;
- b) descritivos, quando ocorre a descrição de um fenômeno;
- c) exploratórios, utilizados quando o fenômeno em avaliação não possui um conjunto de resultados claro.



Curiosidade

Mueller (2007), em capítulo sobre estudos de caso, apresenta dissertações de mestrado e teses de doutorado na área de Biblioteconomia e Ciência da Informação que utilizaram a referida estratégia de pesquisa. Vale a pena consultar.

4.4.4 Histórias de vida ou desenhos narrativos

Neste tipo de desenho, o pesquisador utiliza relatos narrativos dos indivíduos para buscar as experiências de vida de pessoas, interpretá-las e analisá-las. O pesquisador tem como foco o próprio indivíduo (ou grupos de indivíduos), o ambiente circundante e as pessoas que interagem com aquele indivíduo (ou grupo). Creswell (2005 apud SAMPIERI; COLLADO; LUCIO, 2013, p. 509) afirma que

o desenho narrativo muitas vezes é um esquema de pesquisa, mas também é uma espécie de intervenção, pois ao contar uma história ajuda a pensar em questões que não eram claras ou conscientes. Geralmente é usado quando o objetivo é avaliar uma sucessão de acontecimentos.

Como fontes de dados, o pesquisador pode utilizar: documentos, entrevistas, biografias ou autobiografias, diários ou cartas.

O pesquisador efetua uma escolha de indivíduos para participarem de pesquisas com esse desenho. O que ele busca é ouvir as narrativas, analisá-las individualmente e através dos participantes, até que uma metanarrativa ou história geral é sintetizada e recontada pelo pesquisador (DRIESSNACK; SOUSA; MENDES, 2007). De acordo com Flick (2009, p. 307), “as narrativas são estimuladas e coletadas na entrevista narrativa, com a finalidade de reconstruir processos biográficos”.

O pesquisador pode optar por recolher a história de vida toda ou efetuar uma história de vida temática. A história de vida pode abordar um fenômeno, um período específico ou um evento.

4.4.4.1 Coleta de dados

O pesquisador pode utilizar entrevistas (que devem estar narradas em primeira pessoa), análise de documentos ou depoimentos para a coleta dos dados.

A narrativa inclui a cronologia dos eventos e fatos. O pesquisador utilizará várias fontes (registros na imprensa, gravações em vídeo ou áudio e documentos da instituição à qual o entrevistado estava vinculado) para validar os eventos, sua sequência e cronologia.

O pesquisador deverá incluir aspectos contextualizadores de local e tempo na narrativa, apresentar dados e características dos outros envolvidos nela, ações desenvolvidas pelo narrador (com justificativas e pensamentos), o problema que tentavam resolver e a solução adotada.

Por outro lado, o pesquisador não pode deixar de lado eventuais elementos que possam invalidar aspectos da narrativa, tais como “dados falsos, acontecimentos distorcidos, exageros e esquecimentos provocados por traumas ou pela idade” (SAMPLER; COLLADO; LUCIO, 2013, p. 512).

4.4.4.2 Análise das entrevistas narrativas

Tendo em vista a reconstrução de eventos a partir da análise das entrevistas narrativas, a literatura da área apresenta várias possibilidades de análise. Você poderá encontrar a descrição de algumas dessas possibilidades, com mais detalhes, em *Flick* (2009, p. 307).

4.4.5 Pesquisa-ação

A pesquisa-ação é uma pesquisa participativa, que exige elevada interação entre pesquisador e sujeitos da pesquisa, além de uma autocrítica e reflexão criteriosas, que envolvem a sequência reflexão-ação-reflexão. Além disso, esse desenho de pesquisa parte de problemas que existem na prática e que estão situados em determinado ambiente.

Logo no início do seu texto, *Tripp* (2005, p. 444) alerta que

[...] como resultado do grande aumento de sua popularidade e da amplitude de sua aplicação, a pesquisa-ação tornou-se atualmente um termo aplicado de maneira vaga a qualquer tipo de tentativa de melhora ou de investigação da prática.

Tripp (2005) vê a pesquisa-ação como uma das vertentes inseridas na investigação-ação. Ao mesmo tempo, a pesquisa-ação não deve ser confundida com uma consultoria: na pesquisa-ação, há intencionalidade para a construção e busca do conhecimento teórico.

É indicada para a investigação em comunidades/ organizações/ grupos relativamente pequenos, sendo contraindicada para investigação em escalas muito grandes. Também, como adverte *Vergara* (2012, p. 192), esse tipo de pesquisa “não é recomendado para aplicação em organizações que apresentem alta rotatividade de dirigentes e demais funcionários, o que afeta os propósitos e a continuidade da pesquisa”.

A pesquisa-ação possibilita a análise da teoria durante o processo de investigação. É um desenho de pesquisa necessariamente coletivo,



que envolve debates e discussões entre o pesquisador e os membros do grupo, comunidade ou movimento social envolvidos.

4.4.5.1 Modalidades de pesquisa-ação

Tripp (2005) apresenta cinco modalidades de pesquisa-ação:

- a) pesquisa-ação técnica;
- b) pesquisa-ação prática;
- c) pesquisa-ação política;
- d) pesquisa-ação socialmente crítica;
- e) pesquisa-ação emancipatória.

A pesquisa-ação técnica apresenta essa denominação no sentido de que o pesquisador transfere uma prática que já ocorre em um local e para a própria prática. O termo “técnica”, aqui, se refere ao sentido “mecânico” dessa transferência.

A pesquisa-ação prática é caracterizada por Grundy (apud TRIPP, 2005, p. 457) como distinguindo-se da pesquisa-ação técnica “pelo fato de que o pesquisador escolhe ou projeta as mudanças feitas”.

A pesquisa-ação política está associada “à mudança da cultura institucional e/ ou de suas limitações” (TRIPP, 2005, p. 457).

Já a pesquisa-ação socialmente crítica é um tipo específico de pesquisa-ação política. Nesse caso, a investigação não é apenas para melhorar algo que já é realizado, mas visa o “como tornar o seu pedaço do mundo um lugar melhor em termos de justiça social” (TRIPP, 2005, p. 458).

Por fim, na pesquisa-ação emancipatória, espera-se que os participantes não apenas consigam resolver problemas, mas que ocorra uma mudança social de fundo, por meio da pesquisa.

4.4.5.2 As fases na pesquisa-ação

A pesquisa-ação se desenvolve em três fases:

- a) observar, ou seja, desenvolver um esquema do problema e efetuar a coleta de dados;
- b) pensar, o que envolve análise e interpretação dos dados;
- c) agir, que consiste na resolução dos problemas e implementação das propostas construídas.

Essas fases serão repetidas de forma cíclica, até que seja alcançada uma solução para o problema.

4.4.5.3 Limitações da pesquisa-ação

Gray (2012) apresenta algumas limitações da pesquisa-ação: alguns projetos com pesquisa-ação acabam exigindo mais tempo que outras abordagens, acarretando rotatividade das pessoas envolvidas no projeto e, por vezes, atrapalhando o seu desenvolvimento adequado.

Outra limitação é que, como os resultados de muitas pesquisas-ação são práticos, nem sempre eles podem ser tornados públicos.

O autor ainda relata a questão da generalizabilidade: como vários projetos de pesquisa-ação são bastante específicos, nem sempre conseguimos generalizar os resultados (GRAY, 2002).

4.5 A AMOSTRAGEM EM PESQUISAS QUALITATIVAS

Tal como na formulação do problema, a definição da amostra nas pesquisas qualitativas apresenta grandes diferenças. Na pesquisa qualitativa, não definimos o tamanho da amostra *a priori*, nem a amostra é probabilística. A intenção do processo de amostragem em pesquisas qualitativas não é efetuar generalizações (como ocorre na pesquisa quantitativa). Outro aspecto a se destacar, na amostragem em pesquisas qualitativas, é sua flexibilidade: a amostra pode ser alterada ou modificada durante a pesquisa.

Acredita-se que, a esta altura de sua pesquisa, você teve mais de um contato com seu ambiente de pesquisa (você já foi “a campo” mais de uma vez). Além disso, você possui alguma ideia do cotidiano nesse ambiente e já conversou com algumas pessoas.

Você pode começar a definir uma amostra “provisória”. As suas perguntas na definição da amostra em uma pesquisa quantitativa são: quais os elementos que são mais importantes para responder ao problema de pesquisa? Em qual ambiente ou contexto posso encontrar esses elementos/ indivíduos?

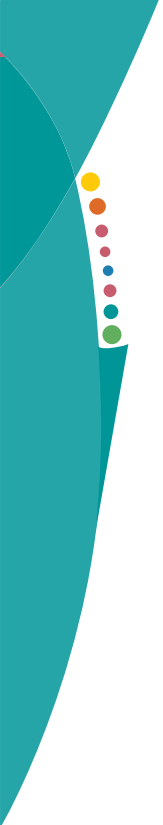
Sampieri, Collado e Lucio (2013) destacam três fatores para sugerir o tamanho da amostra qualitativa:

- a) você possui as condições operacionais de coleta e análise? De que adiantaria coletar centenas de casos se seu prazo para entregar a monografia se encerra ao final do próximo semestre? Você não terá tempo ou equipamento para analisar todos estes casos;
- b) o tamanho da amostra qualitativa está associado ao entendimento do fenômeno. Mais à frente, voltaremos a este tópico com a expressão “saturação de categorias”;
- c) o fenômeno que você está estudando apresenta casos facilmente detectáveis? Quanto tempo é necessário para efetuarmos a coleta dos dados?

Um ponto que *Creswell* (2010, p. 212) aponta é a *intencionalidade* na seleção dos indivíduos para o estudo: é “a seleção intencional dos participantes ou dos locais (ou dos documentos, ou ainda do material visual) que mais ajudará o pesquisador a entender o problema e a questão de pesquisa”.

Sampieri, Collado e Lucio (2013) destacam quatro tipos de participantes em pesquisas qualitativas: participantes voluntários, especialistas





(por exemplo, entrevistas com bibliotecários atuantes em bibliotecas universitárias), casos típicos e amostra por cotas. No caso da utilização da abordagem dos grupos focais (estudaremos este tópico nas próximas seções), é recomendado um máximo de 10 pessoas por grupo. Para os estudos etnográficos, a Teoria fundamentada, entrevistas e observações, os autores sugerem amostras entre 30 e 50 casos. Quando o estudo é do tipo história de vida, a amostra deve contar com o sujeito de estudo bem como com a maior quantidade possível de pessoas que tenham tido contato com ele (inclusive seus críticos).

É claro que o padrão de amostra que você irá adotar será em função da sua intenção e do ambiente e/ ou contexto de campo de sua pesquisa. Assim, se sua intenção é investigar várias perspectivas, você deve buscar uma amostra com a maior variedade possível. Por outro lado, se sua intenção é descrever um grupo com detalhes, você deve buscar uma amostra com a menor heterogeneidade possível. Se o que você pretende é descrever determinado padrão que se repete (um padrão típico), sua amostra deve se concentrar nesses casos.

Um princípio geral sobre amostragem nas pesquisas qualitativas, tal como apontado por *Flick* (2009, p. 122), é a seleção gradual da amostra:

Ela (a amostra) será desenvolvida gradualmente, durante a coleta e análise dos dados, sendo então complementada por novas dimensões, ou limitada a dimensões e a campos determinados (*FLICK*, 2009, p. 122)

4.5.1 A coleta de dados na pesquisa qualitativa

É na coleta de dados das pesquisas qualitativas que encontramos uma enorme diversidade de técnicas e abordagens. Você pode coletar seus dados utilizando: observação, entrevistas, grupos focais, análise de documentos, registros audiovisuais e outros materiais, além de histórias de vida/ biografias. Diferentemente de pesquisas quantitativas, você não irá “medir” variáveis na coleta de dados em uma pesquisa qualitativa. Sua coleta de dados, nesse tipo de investigação, vai ocorrer no ambiente no qual estão os seus objetos de estudo. Se você estiver efetuando um estudo de usuários em bibliotecas, ele se desenvolverá em bibliotecas ou nas plataformas virtuais de acesso a elas. Não se esqueça do alerta apresentado por *Sampieri, Collado e Lucio* (2013): o principal instrumento de coleta de dados em uma pesquisa qualitativa não são as entrevistas, os grupos focais ou a observação; o principal instrumento de coleta de dados é você mesmo: o pesquisador.

Flick (2009, p. 110) alerta que “as informações a que o pesquisador terá acesso e das quais permanecerá excluído dependem essencialmente da adoção bem-sucedida de um papel ou postura apropriada”. Uma das dificuldades para o pesquisador que adota a abordagem qualitativa é o acesso: pode ser bastante difícil conseguir acessar certos ambientes e, mesmo que se consiga o acesso “formal”, o pesquisador pode não conseguir interagir de modo efetivo com seus pesquisados. A entrada em uma instituição é muito mais fruto de negociação entre o pesquisador e quem autoriza esse acesso do ponto de vista da instituição do que de um convencimento “racional” sobre a importância da pesquisa para a instituição.

Outro aspecto a ser levado em conta é o pesquisador identificar se está entrevistando os personagens em papéis-chave, ou seja, aqueles que são mais relevantes para seu problema de pesquisa, ou se está em contato apenas as pessoas que possuem tempo disponível para as entrevistas. O pesquisador também não deve induzir a respostas, nem efetuar críticas às respostas dadas pelos participantes.

Na sequência, veremos as principais ferramentas de coleta de dados em pesquisas qualitativas.

4.5.1.1 Observação

A observação na pesquisa qualitativa não é apenas uma ação passiva e despreocupada. Nela, o pesquisador deve usar todos os sentidos para conseguir os dados relevantes para a sua pesquisa. Existem diferentes conceitos de observação na literatura de pesquisa qualitativa. *Flick* (2009) apresenta uma classificação em cinco dimensões para procedimentos observacionais:

- a) observação secreta x observação pública: se os observados foram informados, ou não, de que uma pesquisa com observação está em curso;
- b) observação participante x observação não participante: diz respeito ao aspecto de o observador ser um elemento ativo do ambiente observado;
- c) observação sistemática x observação não sistemática: diz respeito à dimensão da observação, ou seja, se ela segue uma estrutura mais padronizada ou se é mais aberta e dinâmica;
- d) observação em situações naturais x observação em situações artificiais: diz respeito ao ambiente dos observados, ou seja, verifica se eles estão em seu ambiente natural ou se foram “deslocados” para outro ambiente;
- e) auto-observação x observar os outros: diz respeito a se a auto-observação reflexiva do pesquisador é considerada para fortalecer a fundamentação daquilo que é observado.

Diferentemente das entrevistas ou narrativas, na observação, o pesquisador tem a oportunidade de acessar as práticas efetivamente como elas foram realizadas. Nas técnicas de entrevista ou narrativas, as práticas são mediadas pelo discurso do entrevistado/ narrador.

Sampieri, Collado e Lucio (2013) alertam para o espectro de papéis que o observador pode exercer. Eles variam entre: um papel de não participante (por exemplo, observando situações por meio de vídeos), uma participação passiva (sem interação), uma participação moderada, uma participação ativa, até uma participação completa (na qual o observador está totalmente inserido no grupo em observação, sendo parte dele).

Um eventual obstáculo para a técnica da observação é o fato de que as pessoas observadas podem modificar seu comportamento em função da presença do observador. Certamente, à medida que o ambiente de estudo for mais aberto ao público (por exemplo, um ambiente com grande circulação de pessoas, como uma biblioteca pública), menos será percebida a presença do observador (se o propósito da pesquisa é observar os usuários). Em um ambiente mais fechado e privado, a presença do observador será notada mais facilmente, ampliando-se a chance de essa presença modificar a forma de atuação dos observados.



Você deve efetuar o registro de suas observações. Tais registros devem conter: o que foi observado, a data, a hora, o local e quem participou da observação. Você, como pesquisador, ainda deve registrar suas impressões e um resumo do que ocorreu durante o processo de observação. Também deve ser incluída sua hipótese sobre o fenômeno observado. A partir do que foi observado, deve registrar seus próximos passos e o que pode ser necessário modificar no plano da pesquisa.

Você deve realizar suas anotações e impressões conforme a observação ocorre. Se decidir tomar nota posteriormente, pode deixar de anotar pontos relevantes. Não se “engesse” durante a observação. Mantenha-se em alerta para alterar seu foco, se for preciso.

É recomendável que você consulte, na *web*, a diversidade de fichas de acompanhamento de observação. Você encontrará uma variedade grande de fichas para observação e, a partir delas, terá condições de construir a mais apropriada ao seu caso.

4.5.1.2 Entrevistas

Uma entrevista é o encontro entre pessoas para que o entrevistador consiga informações do/ a(s) entrevistado/ a(s).

Em que situações é indicado utilizar a técnica de entrevistas?

Entrevistas são fundamentais quando se precisa/ deseja mapear práticas, crenças, valores e sistemas classificatórios de universos sociais específicos, mais ou menos delimitados, em que os conflitos e contradições não estejam claramente explícitos (DUARTE, 2004).

Por outro lado, se o ambiente no qual a pesquisa será realizada estiver convivendo com violência, abuso de poder, corrupção etc., é pouco provável que as pessoas concordem em ser entrevistadas. As entrevistas são classificadas, segundo seus roteiros, como:

- a) não estruturadas: o entrevistador segue apenas um guia geral para a entrevista, com liberdade para a sua realização;
- b) semiestruturadas: o entrevistador vai para a entrevista com um roteiro de perguntas/ assuntos a serem abordados, mas possui liberdade para se aprofundar nos temas desejados;
- c) estruturadas: um roteiro pré-estabelecido é seguido pelo entrevistador.

Quando utilizar as entrevistas? Em algumas situações, a observação é dificultada. Por exemplo, é muito improvável que um pesquisador consiga observar problemas de violência contra a mulher.

Um aspecto negativo das entrevistas é que o entrevistado apresenta a informação segundo sua ótica. Outro possível problema é ele ser influenciado pelo entrevistador. Ainda: o entrevistado pode apresentar respostas falsas para você (por desconhecimento ou má-fé). Talvez ele não tenha interesse em participar de uma pesquisa como a sua. Outra eventual desvantagem das entrevistas é que elas podem apresentar custos superiores aos da utilização de um questionário, por exemplo.

Diferentemente da atitude que se tem ao responder a um questionário em uma pesquisa quantitativa, não há, exatamente, um tempo pré-determinado para a entrevista.

Flick (2013, p. 143) destaca que “é mais provável que os pontos de vista dos sujeitos entrevistados sejam expressos em uma situação de entrevista

com um planejamento aberto do que em uma entrevista padronizada ou em um questionário”.

É recomendado que o entrevistado seja informado do propósito e da importância da pesquisa à qual a entrevista está relacionada.

Ao realizar uma entrevista, você deve escutar seu entrevistado com atenção, conduzindo-a em um local que reduza as possibilidades de interrupções (telefones, chefes ou subordinados demandando a atenção do entrevistado). Você deve valorizar um ambiente de espontaneidade na condução da entrevista. Um procedimento-padrão é começar com as perguntas mais gerais e encaminhar-se para as questões mais específicas no decorrer da entrevista. Nela, não devem ser formuladas perguntas que possam induzir o entrevistado, a quem o entrevistador não deve interromper. O entrevistador também deve evitar ironias e ter paciência com o entrevistado.

Não se esqueça, ao redigir seu trabalho de pesquisa, de explicitar os motivos pelos quais escolheu a entrevista. Você deve apresentar o critério de escolha dos entrevistados, quantos são eles e quais são os dados gerais (idade, escolaridade, profissão, sexo etc.). Para efetuar o registro da entrevista, você pode escolher entre vídeo, gravador, anotações e/ ou um *tablet* ou computador portátil.

Também é recomendável que você faça um pré-teste do instrumento de coleta de dados. Com esse pré-teste, você verificará se a formulação das questões está clara, se as respostas obtidas permitem alcançar os objetivos estipulados e se o tempo de resposta do entrevistado pode ter interferido na consistência das respostas. Faça um ensaio com um conhecido seu, para estimar o tempo da entrevista, testar os equipamentos que farão o registro dela e verificar se seu roteiro está adequado.

Grinnell, Williams e Unrau (2009) apud Sampieri, Collado e Lucio (2013) dividem as perguntas de uma entrevista em quatro tipos:

- a) perguntas gerais: são perguntas-gatilho, que servem para disparar a entrevista em direção ao assunto de pesquisa;
- b) perguntas para exemplificar: estão associadas a questões de maior profundidade. Espera-se que o entrevistado, ao exemplificar, forneça mais detalhes relacionados com o problema de pesquisa;
- c) perguntas de estrutura: é solicitado ao entrevistado que forneça uma lista de conceitos;
- d) perguntas de contraste: é solicitado ao entrevistado que apresente semelhanças e diferenças sobre algum item relevante para a pesquisa.

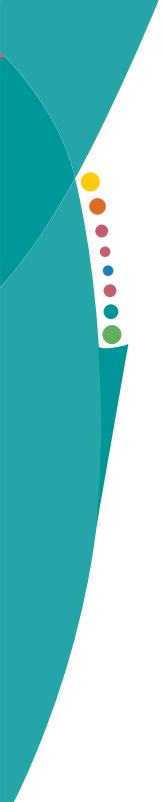
Ao final da entrevista, agradeça ao entrevistado, efetue a transcrição da entrevista o mais rápido possível e proceda a uma reflexão, se for necessário.

Um *software*, de código aberto e gratuito, utilizado em gravação de entrevistas é o *Audacity*: <<http://audacity.sourceforge.net/?lang=pt>>. Acesso em: 10 jan. 2022.

4.5.1.3 Grupos focais (focus groups)

A técnica de grupo focal envolve pequenos grupos de pessoas reunidas para debater um tópico específico. Na literatura de grupos focais, o padrão adotado é o de que os grupos tenham entre seis e dez pessoas (alguns textos falam entre três e dez pessoas).





O tempo de duração de uma sessão de grupo focal oscila entre trinta minutos e duas horas. O grupo focal pressupõe um espaço para uma conversa informal, embora, claro, não seja simplesmente uma “conversa entre amigos”.

O grupo focal exige, ainda, a participação de um moderador. É o moderador quem apresenta as perguntas/ questões (lembremo-nos de que no grupo focal também se pode apresentar um tema para discussão). O moderador gerencia a discussão, mantendo o foco e o ritmo do encontro. Além disso, existe a figura do observador (em alguns casos, até mais de um). *Krueger* (apud LEITÃO, p. 77) denomina o observador de “assistente do moderador.” Certamente, uma eventual desvantagem dos grupos focais é a dificuldade em se conseguir moderadores que sejam experientes na condução desses grupos. O moderador deve ser firme e flexível, para evitar um problema recorrente nos grupos focais: que um ou dois participantes monopolizem a discussão ou a desviem do foco da pesquisa. Com um moderador experiente, por outro lado, temos uma das vantagens dos grupos focais: a flexibilidade, já que o moderador pode aprofundar questões que não estavam inseridas no roteiro planejado, mas que se mostraram relevantes para a pesquisa em curso.

O observador é o responsável pelo registro atento e cuidadoso das informações que serão posteriormente analisadas, sem, contudo, se manifestar ou interferir nas discussões. O observador pode, inclusive, anotar indícios da linguagem corporal dos participantes do grupo focal. Ele também pode auxiliar o moderador na recepção dos participantes e em outras questões operacionais do grupo focal.

Finalmente, por ser uma técnica qualitativa, os grupos focais não permitem uma fácil generalização dos resultados para a população como um todo.

O objetivo dos grupos focais não é o de realizar uma “entrevista em grupo”, mas provocar o debate entre os participantes e estudar essa discussão no grupo.

Certamente, não existe uma regra determinando quantas sessões devem ser realizadas. É o pesquisador quem deverá definir, ao longo do processo de pesquisa, quando os dados para responder ao problema de pesquisa tiverem sido obtidos.

Uma vantagem dos grupos focais é que os próprios participantes tendem a eliminar ideias extremistas ou não verdadeiras. Ao reunirmos um grupo de pessoas para discutir acerca de um tema específico, com um moderador gerenciando o debate, estimulando as manifestações e, de alguma maneira, mantendo a conversa na direção desejada, os indivíduos tendem a se manifestar de forma mais profunda e transparente, deixando de lado aspectos que poderiam ficar ocultos em uma entrevista isolada. Outro aspecto positivo dos grupos focais é o custo: é mais barato entrevistar várias pessoas coletivamente que desenvolver uma série de entrevistas individuais.

Por outro lado, *Flick* (2009, p. 181) alerta para o fato de que um dos pontos fracos do grupo focal é “o número limitado de questões com as quais é possível lidar e os problemas em relação às anotações durante a entrevista”. A literatura recomenda a filmagem da sessão.

A necessidade do moderador é evitar que um ou dois participantes do grupo monopolizem a discussão, impedindo a manifestação de outros participantes e, também, estimular aqueles mais tímidos a emitirem suas

opiniões. O moderador também deve efetuar um direcionamento da discussão visando ao aprofundamento dos itens de interesse da pesquisa, mas sem parecer autoritário ou inflexível. Em algumas circunstâncias, ele deve produzir um debate mais acirrado (quando a discussão estiver “apagada”). Com a participação do moderador, o pesquisador fica liberado para observar o grupo e tomar notas de informações que poderão ajudá-lo nas análises dos dados.

O ambiente para a realização do grupo focal pode ser simplesmente uma sala, com mesa e cadeiras, em um local sem telefones ou interrupções. Em algumas situações, pode ser adequado utilizar-se de um grupo “artificial” (constituído apenas para a pesquisa) ou de um grupo “natural” (um grupo que já existia antes da pesquisa).

Os grupos focais podem ser homogêneos ou heterogêneos. É possível, ao se constituir uma mesma sessão de grupo focal, com participantes com perfis heterogêneos, que alguns deles venham a inibir a participação de outros. Detalhando um pouco mais a questão da homogeneidade/heterogeneidade: considere estudos de usuários. Podemos separar os usuários em diferentes sessões, sendo que, em cada uma delas, teremos um grupo homogêneo. Digamos que nós podemos ter uma sessão com um grupo constituído por aqueles usuários que frequentam a biblioteca uma ou mais vezes por semana, outra sessão com um grupo de usuários que a frequente entre um mínimo de uma até quatro vezes por mês e, por fim, os não usuários da biblioteca. Outro exemplo seria uma pesquisa que abordasse perspectivas para as bibliotecas do século XXI. O pesquisador poderá constituir um grupo focal formado por pesquisadores da área, docentes e bibliotecários. Observe que, nesse caso, temos heterogeneidade na constituição do grupo, contudo, todos os participantes possuem reconhecida experiência na área, o que assegura o mesmo nível de interlocução. O pesquisador, levando em consideração o problema de pesquisa e o trabalho em campo, vai decidir se os grupos devem ser homogêneos ou heterogêneos.

Por exemplo, *Silva* (1998 apud DIAS, 2000), em sua pesquisa sobre a opinião de usuários de bibliotecas, trabalhou com três grupos focais:

- a) usuários frequentes;
- b) usuários pouco frequentes;
- c) não usuários.

O trabalho de *Leitão* (2005) utilizou a técnica de grupos focais. A autora sugere, para o caso de bibliotecas universitárias, que os clientes/usuários externos sejam divididos em:

- a) usuários regulares e frequentes (clientes reais);
- b) usuários que não utilizam a biblioteca com regularidade (clientes reais);
- c) não usuários (clientes potenciais) ou usuários que já a utilizaram, mas não a utilizam mais (clientes perdidos);
- d) usuários que nunca foram contatados.

4.5.1.3.1 O roteiro de questões em um grupo focal

Ao se elaborar o roteiro de perguntas em um grupo focal, deve-se ter em mente o objetivo da pesquisa.

Questões que apresentem como respostas apenas “sim” ou “não” devem ser evitadas, assim como questões muito complexas, que possam ser de difícil compreensão quando expostas em uma reunião com várias pessoas. Com uma questão complexa, é possível que os participantes utilizem tempo considerável da reunião para compreender o significado da pergunta, em vez de respondê-la.

Ao mesmo tempo, o moderador deve ser flexível o suficiente para perceber em quais situações da discussão é adequado “escapar” do roteiro, que deve ser apenas um guia para a entrevista, e não aprisionar o moderador. O roteiro deve ter as questões ordenadas do geral para o específico, por ordem de importância. As questões não devem ser muito complexas, nem é recomendável que ultrapassem 10 ou 12 perguntas. Dificilmente poderão ser respondidas mais que 12 perguntas por grupos com seis a 12 pessoas, em um encontro de duas horas.



4.5.2 Atividade

Esta atividade atende ao objetivo c: “coletar dados: entrando no ambiente da pesquisa”.

Considere um estudo que tenha como objetivo analisar os problemas no atendimento aos usuários de uma biblioteca em um hospital.

Suponha que você pretende utilizar a técnica de grupos focais.

- a) Quais grupos poderiam integrar o estudo?
- b) Apresente vantagens e desvantagens de se adotar essa técnica.

Resposta comentada

- a) Você poderá ter um grupo constituído por médicos, um segundo grupo constituído por profissionais administrativos e um terceiro grupo constituído por profissionais do hospital que não sejam nem médicos, nem funcionários administrativos.
- b) Algumas vantagens da técnica de grupos focais são o custo relativamente baixo e a flexibilidade (o moderador pode aprofundar a investigação em perguntas que não tenham sido previstas ou estimular a discussão entre os participantes). Também é adequada na avaliação da satisfação dos usuários/ clientes, com respeito a um serviço/ produto oferecido. Como desvantagens do grupo focal, costumam ser citadas: a dependência de um moderador experiente na condução dos grupos, a dificuldade de se generalizar os resultados para a população como um todo e a possibilidade de existir participantes que monopolizem a discussão.



Multimídia

Consulte o *link* <http://portal.mec.gov.br/seb/arquivos/pdf/Avalmat/livro_mec_final_baixa.pdf>.

Nele, você encontra o relatório do Programa Nacional Biblioteca da Escola (PNBE): *Leitura e Biblioteca nas Escolas Públicas Brasileiras*. Nesse estudo de 2008, foi utilizada a técnica de grupos focais.

4.5.3 Análise de documentos e outros materiais

Documentos, registros audiovisuais e artefatos são fontes de dados em pesquisas qualitativas. Organizações (uma escola, uma empresa do governo, as Forças Armadas, uma indústria) produzem grande variedade de documentos e registros escritos, que são fontes de dados para pesquisas qualitativas. Existem documentos oficiais (um registro de contrato de um funcionário novo para a biblioteca) e documentos não oficiais (uma foto do funcionário novo com os antigos funcionários da biblioteca). Os dois documentos em conjunto podem ser utilizados para a investigação das relações humanas no ambiente daquela biblioteca.

Os documentos podem ser jornais, revistas, discursos, normas técnicas internas da organização, relatórios de trabalho, produções técnicas, diários pessoais, arquivos da organização, entre outros. É importante saber a autoria do documento, o objetivo da sua feitura e qual a intencionalidade de sua produção.

Não devemos considerar documentos como uma representação fiel da realidade. Por exemplo, relatórios de treinamentos de funcionários podem conter determinado item. No entanto, é possível que esse item simplesmente não tenha sido apresentado no treinamento. O pesquisador, eventualmente, pode descobrir isso por meio de entrevistas. Assim, podemos efetuar uma “triangulação” entre informações obtidas por uma técnica (análise de documentos) e por outra (entrevistas).

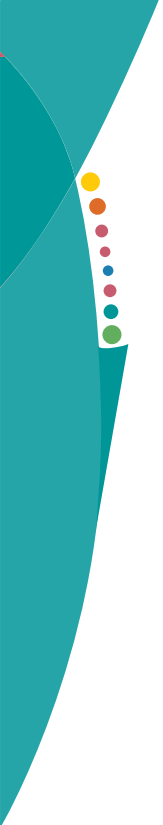
Os documentos podem ser utilizados pelas organizações na construção da versão “oficial” ou como aquilo que elas gostariam que sobrevivesse com o passar do tempo. Em outras palavras: não devemos nos esquecer de quem é o autor do documento e da intenção com que o produziu.

Os materiais audiovisuais (fotografias, filmagens, desenhos, pinturas, mapas, CDs, DVDs) são produzidos por pessoas ou organizações com finalidades diversas. De forma análoga aos registros escritos oficiais, um vídeo institucional busca “vender uma boa imagem” da empresa ou organização. Assim, o pesquisador deve estar atento aos objetivos subjacentes à produção do documento.

4.5.4 Pesquisa qualitativa na internet

A internet também se tornou objeto de pesquisas, tanto quantitativas quanto qualitativas. Podemos pensar nela como um objeto de pesquisa.





Quando diferentes grupos sociais (jovens no Ensino Médio, adultos jovens com Ensino Superior etc.) acessam informações culturais *on-line*, por exemplo, um pesquisador pode investigar, usando entrevistas, grupos focais ou observação, o uso da internet em bibliotecas e/ ou telecentros, ou, ainda, *lan-houses* ou *cyber cafés*).

Muitas pesquisas sobre a internet são quantitativas devido à facilidade de se obter informações de acessos, tempo *on-line* e outros dados quantitativos. Questionários com questões fechadas também são facilmente produzidos em diversas ferramentas disponíveis na internet, algumas gratuitas, como o *Survey Monkey*.

Por outro lado, um pesquisador pode efetuar entrevistas e grupos focais *on-line*, etnografia virtual etc. Entrevistas *on-line* podem ser interessantes para pessoas que estejam geograficamente distantes, que tenham poucos horários para participar, ou possuam algum tipo de constrangimento para responder a uma entrevista pessoalmente. Entrevistas podem ser conduzidas por *e-mail*. Atualmente, a maioria das pessoas não apresenta dificuldades em usar a tecnologia do *e-mail*; a dificuldade maior reside em conseguir endereços de *e-mail* com o perfil desejado para a pesquisa.

Podemos ter entrevistas e grupos focais *on-line* síncronos ou assíncronos. Nos grupos síncronos, é necessário que todos os participantes estejam conectados ao mesmo tempo. Já nas entrevistas e grupos focais assíncronos (que não ocorrem em tempo real), parte dos participantes pode estar *off-line*. Contudo, nem todas as pessoas têm facilidade de acessar a internet em horários ou dias preestabelecidos, apenas para atender a uma pesquisa.

Em relação à análise documental na internet, todos sabemos da enorme disponibilidade de documentos (de várias naturezas) nesse meio. Alguns problemas são: como adotar critérios para selecionar uma página da internet para a pesquisa? Onde está o “início” de uma página da internet, já que as páginas possuem *hiperlinks*?

Uma dificuldade da pesquisa com páginas da internet é que elas podem ser continuamente modificadas. Portanto, é necessário mencionar a data de acesso ao documento.

4.5.5 Netnografia

Com a internet, o computador deixou de ser uma máquina isolada e passou a ser uma ferramenta de informação e comunicação, permitindo o surgimento de comunidades virtuais e redes sociais, tanto as de interesse geral quanto aquelas voltadas ao público acadêmico. Para compreender essa recente e crescente digitalização da sociedade, pesquisadores das mais diversas áreas (Educação, Ciência da Informação, Economia, Marketing, entre outras) vêm investigando os fenômenos sociais virtuais e a forma como as pessoas interagem na *web*.

A palavra netnografia é um neologismo do início do século XXI. *Hobbs* apud *Kozinets* (2014, p. 61) define netnografia como

um coquetel de metodologias que compartilham da suposição de que o engajamento pessoal com o sujeito é fundamental para compreender uma determinada cultura ou ambiente social. A observação participante é o componente mais comum desse coquetel, mas entrevistas, análise de conversação e discurso, análise

documentária, filme e fotografia têm todos o seu espaço no repertório etnográfico. A descrição reside no âmago da etnografia, e independente de como essa descrição seja construída, é o intenso significado da vida social a partir da perspectiva cotidiana dos membros do grupo o que se busca.

De acordo com *Kozinets* (2014, p. 62) a netnografia

[...] adapta os procedimentos etnográficos comuns de observação participante às contingências peculiares de interação social mediada por computador: alteração, acessibilidade, anonimato e arquivamento. Os procedimentos incluem planejamento, entrada, coleta de dados, interpretação e adesão a padrões éticos.

Ainda de acordo com o autor, a netnografia faz uso de “comunicações mediadas por computador como fonte de dados para chegar à compreensão e à representação etnográfica de um fenômeno cultural ou comunal” (KOZINETTS, 2014, p. 62). A netnografia surgiu a partir de pesquisas de marketing e de consumo na internet.

Quanto aos métodos, o pesquisador deve escolher se vai adotar métodos quantitativos, qualitativos ou mistos em sua pesquisa netnográfica. Do ponto de vista qualitativo, o pesquisador pode realizar levantamentos, entrevistas, grupos focais e efetuar análise de redes sociais.

4.5.6 Outras abordagens

Uma forte característica da pesquisa qualitativa é a diversidade de estratégias e abordagens. Além das estratégias mencionadas anteriormente, podemos citar a teoria fundamentada, a metodologia de análise de redes sociais e a do discurso do sujeito coletivo. Você poderá consultar essas abordagens com profundidade em *Valentim* (2005).

4.5.6.1 Triangulação de métodos de coleta de dados

O fato de o pesquisador poder trabalhar com mais de uma fonte e mais de um método de coleta de dados possibilita maior variedade de informações e profundidade. Evidentemente, ele nem sempre terá tempo (ou recursos) para trabalhar com mais de um método ou fonte de informação.

Assim, se uma pesquisa pretende investigar como ocorre o acesso à informação, em *lan-houses* ou *cyber* cafés, por parte de jovens do Ensino Médio, na periferia de uma grande cidade brasileira, ela pode contar com:

- a) entrevistas com jovens nessas condições que acessam e com outros que não acessam regularmente a internet fora de casa;
- b) entrevistas com membros das famílias desses jovens;
- c) observação em um conjunto selecionado de *lan-houses* (ou *cyber* cafés);
- d) discussão com grupos focais de jovens na situação de pesquisa.



Certamente, essa concepção de pesquisa apresentará muito mais informação, com maior variedade e profundidade, do que uma pesquisa que utilize apenas entrevistas com os jovens.



4.5.7 Atividade

Esta atividade atende aos objetivos *a* e *d*: “formular problemas para pesquisas qualitativas” e “analisar os dados em pesquisas qualitativas”.

A partir do que você estudou até aqui, construa um quadro com as limitações das estratégias: grupos focais, observação e entrevistas. Você pode buscar mais elementos nas bases de dados da internet.

Resposta comentada

Observe o Quadro 9:

Quadro 9 – Limitações das estratégias: grupos focais, observação e entrevistas

	Grupo focal	Observação	Entrevistas
Vantagens	Flexibilidade, baixo custo, adequado para a avaliação de satisfação de usuários	Flexibilidade, o pesquisador tem a oportunidade de acessar as práticas efetivamente, como são realizadas	São adequadas à necessidade de mapear práticas, crenças, valores e sistemas classificatórios de universos sociais específicos, nos quais os conflitos e contradições não estejam claramente explicitados
Desvantagens	Dificuldade em conseguir moderador experiente, dificuldade para generalizar o resultado para a população, participantes podem monopolizar ou desviar a discussão do foco da pesquisa	As pessoas que estão sendo observadas podem mudar seu comportamento em função da presença do observador; no caso da observação participante, o pesquisador pode perder em objetividade	Em algumas situações, entrevistas não podem ser conduzidas, o entrevistador pode influenciar o entrevistado. O entrevistado pode fornecer respostas falsas por desinteresse ou má-fé. Entrevistas costumam apresentar custos monetários superiores aos dos questionários

Fonte: Produção do próprio autor (2017).

4.6 A ANÁLISE DE DADOS NA PESQUISA QUALITATIVA

Diferentemente da pesquisa quantitativa, na qual a análise de dados é posterior à coleta, na pesquisa qualitativa, a análise dos dados, em alguns casos, acontece em paralelo à coleta. Contudo, por vezes, isso não ocorre. Por exemplo, podemos gravar a conversa do grupo focal, transcrevê-la e analisá-la *a posteriori*. Recomenda-se, inclusive, que a análise já comece em campo. Conforme você for realizando suas entrevistas, grupos focais, e ao longo da pesquisa de documentos, já deve, com seu diário de campo, ir registrando e analisando os dados colhidos. Esse processo concomitante de coleta e análise de dados é um aspecto crucial da famosa flexibilidade nas pesquisas qualitativas, e é o processo que, possivelmente, irá levá-lo a levantar novas questões para o seu estudo.

Gibbs (2009) também destaca outra diferença entre as pesquisas qualitativas e quantitativas: nestas, a tendência é reduzirmos (ao “resumirmos”) os dados. Já naquelas, muitas vezes, acabamos por ampliar o volume de dados. Por exemplo, a partir de uma entrevista, ao efetuar a transcrição, acabamos por aumentar o volume de informações.

O processo básico da análise qualitativa é uma transformação:

você começa com alguma coleta de dados qualitativos (muitas vezes, volumosa) e depois os processa por meio de procedimentos analíticos, até que se transformem em uma análise clara, compreensível, criteriosa, confiável e até original (GIBBS, 2009, p. 16).

Ainda segundo *Gibbs* (2009) e para outros autores na literatura de análise de pesquisas qualitativas (DENZIN, 1997; GIORGI; GIORGI, 2003, MISHNER, 1986; RIESSMAN, 1993), a análise desse tipo envolve interpretação e recontagem, o que torna o processo mais imaginativo e especulativo.

A seguir, apresentamos, de forma bastante sucinta, as etapas da análise de dados. É claro que, dada a variedade de dados que podem ser obtidos em pesquisas qualitativas, existem especificidades para cada um deles, que não poderão ser detalhadas neste texto, por questões de espaço. Para o leitor que desejar se aprofundar no tema, sugerimos ler *Gibbs* (2009).

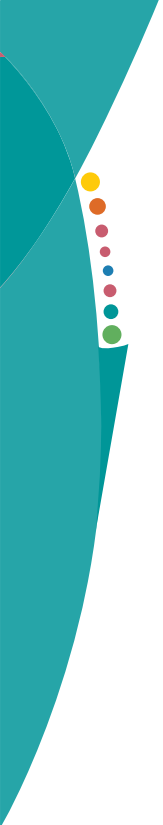
A análise qualitativa de dados pressupõe:

- a) a transcrição (tipicamente de entrevistas);
- b) a produção escrita, por parte do pesquisador, de um diário de pesquisa e de um caderno com notas de campo;
- c) a codificação e categorização de temas;
- d) a análise propriamente dita.

Vamos detalhar um pouco cada um dos itens:

- a transcrição: é a passagem, para o papel ou computador, das gravações de suas entrevistas, discussões com grupos focais, ou mesmo de um único entrevistado. Tipicamente, a transcrição exige





algo como até seis vezes o tempo que foi gasto na coleta. Ou seja: não deixe suas transcrições se acumularem! Além disso, como a transcrição é a redação da linguagem falada, podem ocorrer perdas ou distorções no processo. Esteja sempre atento a esse ponto.

Por outro lado, com o tempo e a experiência, o pesquisador percebe que não é necessário transcrever integralmente todas as gravações. Contudo, como alerta *Gibbs* (2009), isso é adequado para alguns tipos de análise, mas para outras, como a análise do discurso ou conversação, uma transcrição detalhada é fundamental;

- o diário de pesquisa e as notas de campo: aproveite todos os momentos que tiver para escrever. Não deixe para começar a escrever apenas no final do processo de coleta e análise de dados. Os três processos podem (e devem) ocorrer em paralelo: coleta, análise dos dados e redação. Esforce-se desde o começo para que sua redação esteja o mais próximo possível daquela pretendida no trabalho final. Se você fizer apenas pequenas notas, quando retomá-las, algum tempo depois, pode ser que já nem se recorde do que elas pretendiam registrar.

Ao fazer esse esforço para registrar seu texto, você estará produzindo, forçosamente, uma reflexão acerca do seu trabalho. Ou seja: mesmo que de maneira menos estruturada, já estará efetuando uma “análise dos dados”.

O que é um diário de pesquisa?

O diário de pesquisa é usado para o pesquisador registrar todos os aspectos interessantes do que fez: datas, suas leituras, contatos telefônicos e *e-mails* das pessoas relacionadas com sua pesquisa, o andamento de seu trabalho (dados coletados, dados processados), dificuldades, impasses, suas reflexões sobre o andamento da pesquisa e ações futuras relevantes para ela (*GIBBS*, 2009).

Já as notas de campo são registros sem maior formalização, rigor na escrita ou planejamento. Servem como lembretes e anotações rápidas, mas imprescindíveis no dia a dia corrido do pesquisador. Contudo, muitas vezes, elas podem ajudá-lo na análise qualitativa dos dados;

- a codificação e a categorização: neste ponto, você vai perceber a importância da transcrição. É possível codificar seus dados sem transcrevê-los, mas é muito mais simples fazê-lo a partir do texto transcrito. De acordo com *Gibbs* (2009), atribuir códigos e categorias é uma forma de indexar e organizar o texto para estabelecer uma estrutura de ideias temáticas. Ao codificar, você poderá estabelecer relações hierárquicas para investigar seus dados ou definir relações entre trechos do texto para algum fenômeno ou ideia específicos. Existem muitos procedimentos de codificação, tais como a codificação teórica, a codificação aberta, a codificação linha a linha e a codificação axial. Para mais detalhes, favor consultar *Gibbs* (2009) ou *Flick* (2009);
- análise de dados qualitativos: como nos outros pontos das pesquisas qualitativas, existe grande variedade e flexibilidade no momento de se efetuar a análise de dados. Existe a análise de conversação, a análise do discurso, a análise de gênero, a análise de narrativa, a hermenêutica, a análise de biografias e narrativas

etc. Faremos alguns apontamentos básicos sobre análise em pesquisa qualitativa. Para mais detalhes, recomendamos que você leia *Flick* (2009), *Gibbs* (2009) e *Sampieri, Collado e Lucio* (2013).

Ao fazer a codificação, você identificou trechos do seu texto por significado, categorizou-os e atribuiu códigos a essas categorias. Neste momento, você irá se fazer perguntas como: qual o significado deste trecho? Quais as relações que podemos estabelecer entre este trecho e outros trechos do texto? Você efetua comparações entre vários momentos textuais, buscando semelhanças e diferenças entre eles. No caso de semelhança entre os dois trechos, é possível imaginar uma categoria para eles. Com essa codificação, vamos estruturando o texto em termos de relações tanto “paralelas” quanto hierárquicas. Assim, segundo *Sampieri, Collado e Lucio* (2013, p. 456),

usamos a codificação para começar a mostrar significados potenciais e desenvolver ideias, conceitos e hipóteses; vamos compreendendo o que acontece com os dados (começamos a gerar um sentido de entendimento sobre a formulação do problema).

Certamente, o pesquisador possui liberdade para modificar várias vezes as categorias que ele mesmo constrói, até identificar aquelas mais adequadas para sua análise. Esse processo continua até que o pesquisador atinja a “saturação de categorias”. *Sampieri, Collado e Lucio* (2013) orientam que, se nossos dados se tornarem repetitivos ou redundantes, é sinal de que os fundamentos estão confirmados, e isso é o que ocorre quando encerramos a primeira fase de categorização, chamada categorização de primeiro nível. Depois, começamos a efetuar a chamada codificação em segundo nível ou central, que é um patamar mais abstrato de categorização. *Berg* apud *Sampieri, Collado e Lucio* (2013, p. 466) recomenda que se utilizem “pelo menos três exemplos de unidades para apoiar cada categoria”, que é avaliada quanto ao seu significado, natureza, relação com as outras categorias e referenciais. Você vai constatar semelhanças e diferenças entre as categorias investigando o significado delas. “O foco da análise se desloca do contexto do dado para o contexto da categoria.” (*BERG* apud *SAMPIERI; COLLADO; LUCIO*, 2013, p. 467).

Deve ser dado destaque, na análise qualitativa, às descrições e significados de cada categoria, bem como ao número de vezes em que cada uma é observada no material coletado. Em *Sampieri, Collado e Lucio* (2013) encontramos ferramentas para auxiliar na análise qualitativa: matrizes (para estabelecer as relações entre as categorias), mapas conceituais, diagramas de hierarquias e calendários (com as datas-chave ou dias críticos).

Na sequência, apresentaremos algumas metodologias utilizadas na Biblioteconomia/ Ciência da Informação.

4.6.1 Análise do discurso

A Análise do discurso surgiu na França, nos anos 1960, adotando como objeto específico o discurso e apresentando várias tendências. Contrariamente a uma visão mais tradicional, que compreende a língua como um sistema de signos ou de regras formais, a Análise do discurso percebe a linguagem como mediação entre o ser humano e sua realidade social, situando-o historicamente. Tal mediação é desenvolvida pelo discurso, que é onde a interpretação se desdobra. Do ponto de vista discursivo, os aspectos social e histórico não podem ser dissociados.



Entre as principais conexões teóricas da Análise do discurso, temos a Linguística, o Marxismo e a Psicanálise. Da relação com a Linguística, temos que a língua não é transparente; da relação com o Marxismo, que a história também não é transparente e, com a Psicanálise, temos que o próprio sujeito também é “opaco” (ORLANDI, 2006).

Para a Análise do discurso, o “dizer” não é propriedade de um autor, não é nossa propriedade. Nessa concepção, o discurso deixa de ser um projeto autoral, de forma que nossa fala, nosso discurso, nosso dizer é função da língua e da história na qual nós, autores, estamos inseridos: o sentido do discurso não está contido nas palavras, mas na mediação entre elas e o mundo exterior. Além disso, a ideologia se apresenta e se manifesta na língua. Enfim, a linguagem vai além do código e não permite a ingenuidade de nossa parte em acreditar que ela seria um instrumento de comunicação perfeito e neutro.

De acordo com Gray (2012, p. 417),

a Análise do discurso tem seu foco em como a linguagem escrita e falada é utilizada em contextos sociais, prestando atenção à estrutura e à organização da linguagem, com ênfase em como as versões dos participantes dos eventos são construídas.

A Análise do discurso se opõe à Análise do conteúdo no sentido de negar o entendimento de que a linguagem é “transparente” e que refletiria o “mundo real”.

De acordo com Bardin (1994), a Análise de discurso busca estabelecer conexões entre as condições de produção do sujeito e a forma como as dimensões semânticas e sintáticas do discurso se apresentam.

Não existe uma Análise do discurso única. Os especialistas apresentam, ao menos, a divisão entre as escolas francesa e norte-americana. Nesta, está mais marcada uma inclinação linguístico-pragmática da Análise de discurso (ORLANDI, 2006). Já na francesa (ou europeia, para alguns estudiosos), há uma inclinação materialista, com o efeito de desterritorializar a língua (e o sujeito).

Entre os teóricos considerados referência na Análise do discurso, citamos Michel Pêcheux (*Análise automática do discurso* ou *Analyse automatique du discours*, em francês, 1969), Michel Foucault (*Arqueologia do saber*, ou *Archeologie du savoir*, em francês) e Eni Orlandi, pesquisadora brasileira da Universidade Estadual de Campinas (UNICAMP), autora de *Introdução às Ciências da Linguagem: discurso e textualidade*.

Outros autores considerados referências na área da Análise do discurso são: o francês Dominique Maingueneau (*Nouvelle tendances dans l'Analyse du discours*, *Novas tendências em Análise do discurso*, de 1997), o russo Mikhail Bakhtin, (*Esthétique de la création verbale*, *Estética da criação verbal*) e o inglês Norman Fairclough (*Language and Power*, *Linguagem e poder*, de 1989, e *Critical Discourse Analysis*, *Análise crítica do discurso*, de 1992).

Michel Pêcheux efetua uma nova organização no campo da Análise do discurso: com seu trabalho, ele questiona a Linguística sobre a historicidade ausente nela. Ao mesmo tempo, também questiona as Ciências Sociais quanto à sua relação de transparência com a linguagem.

Em Foucault (2008, p. 134, apud GASPARI; REIS, 2010), a Análise do discurso expõe que os diferentes textos “remetem uns aos outros, organizam-se em uma figura única, entram em convergência com

instituições e práticas, e carregam significações que podem ser comuns a toda uma época”.

A conclusão extremamente importante a que se chega é a de que a Análise do discurso possibilita um entendimento menos ingênuo da linguagem, e sua contribuição tem grande destaque para a Linguística. A Análise do discurso possibilita ao sujeito identificar que não há neutralidade na linguagem e que não é possível não estar sujeito a ela (com todas as questões problemáticas que tal sujeição impõe).

Em *Freitas* (2010), é problematizada a questão da crescente e significativa utilização da Análise do discurso foucaltiana nas áreas de Ciência da Informação, Biblioteconomia e Arquivologia, em paralelo com a ausência de sistematizações específicas.

Para *Orlandi* (2003), a contribuição da Análise de discurso em sua vertente francesa é a reflexão sobre a impossibilidade de sermos conscientes de tudo, abrindo a perspectiva de menor ingenuidade nas relações com a linguagem.

Na Análise do discurso temos a noção de *formação discursiva*: são temas (honra, liberdade, opressão etc.) e termos (avião, cadeira, cachorro) que tornam concreta uma maneira determinada de visão de mundo.

Por formações discursivas, *Orlandi* (2003) parte da premissa de que as palavras não têm um sentido em si mesmas. Então, esse sentido é resultante dos sentidos em formações discursivas nas quais estejam inseridos.

O Quadro 10, a seguir, produzido por *Orlandi* (UNICAMP, 2006a), apresenta comparações entre a Análise do conteúdo e Análise do discurso:

Quadro 10 – Comparação entre a Análise do conteúdo e a Análise do discurso

Análise do conteúdo	Análise do discurso
O texto é um documento restrito, limitado ao seu próprio conteúdo	Busca mais compreensão do processo produtivo do discurso
Parte do discurso para a enunciação	Parte da enunciação para o discurso
Visa o que o texto quer dizer	Visa o funcionamento do texto diante de determinado contexto social e histórico
Busca o sentido do texto expresso em sua estrutura e o interpreta a partir daí	Não interpreta os textos que analisa, mas os resultados da análise
Não há relação direta com a historicidade	Caráter de historicidade – a maneira como ela dá significado e delimita a natureza

Fonte: Adaptado de *Orlandi* (1987 apud MENDONÇA, 2007, p. 155).

Não há espaço, nesta disciplina, para explorar com profundidade os dispositivos de análise da Análise do discurso. Destacamos, para o leitor interessado, a obra de *Orlandi* (2001).

4.6.2 Protocolo verbal

A técnica do Protocolo verbal ou “pensar alto” (*think aloud*) é uma abordagem na qual os indivíduos manifestam em voz alta seus

pensamentos enquanto executam uma tarefa. Enfatiza-se que esta é uma técnica introspectiva, no sentido de que torna possível o estudo dos processos mentais do sujeito de pesquisa. Ocorre a gravação das falas, bem como do comportamento desse indivíduo. Posteriormente, as gravações são transcritas de forma literal. Como destacado por *Fujita, Nardi e Fagundes* (2003), com o Protocolo verbal, o pesquisador coleta os dados em um ambiente natural, a partir da observação do comportamento dos sujeitos envolvidos na pesquisa. Busca-se a menor interação possível entre o pesquisador e o sujeito da pesquisa. Entre suas aplicações, constam a avaliação de processos mentais em tarefas que exijam a representação da informação e/ ou o uso de estratégias (FUJITA, 2009). Um dos objetivos da técnica é avaliar a usabilidade de um sistema ou objeto.

Com a técnica de Protocolo verbal, o pesquisador pode avaliar cada passo que o sujeito de pesquisa efetua (o pesquisador analisa tanto os pensamentos exteriorizados quanto o gestual do sujeito de pesquisa, inclusive o movimento dos seus olhos). Dessa forma, o pesquisador pode investigar, a partir do que o sujeito está expressando, a usabilidade do objeto avaliado. Os próprios usuários do sistema contribuem com a localização e compreensão das dificuldades ou críticas na recuperação da informação, bem como na utilização do sistema.

Ericson e Simon (1980, 1993) são considerados os primeiros a sistematizar o Protocolo verbal (a partir dos anos 1980). Estes mesmos autores classificam os protocolos em verbalizações concorrentes (os sujeitos verbalizam seus pensamentos em concomitância com a realização da tarefa) e verbalizações retrospectivas (a verbalização é posterior à realização da tarefa).

Fujita (2003, p. 53) apresenta trabalhos na literatura estrangeira com a utilização do “pensar alto” (Quadro 11):

Quadro 11 – Literatura estrangeira com uso do “pensar alto”

Autor, data	Tema
(Ingwersen, 1982)	Processo de recuperação da informação
(Gotoh, 1983)	Processo de indexação
(Endres-Niggemeyer; Neugebauer, 1998)	Processo de elaboração de resumos
(Sauperl, 2002)	Processo de catalogação de assuntos com bibliotecários

Fonte: *Fujita* (2003, p. 53).

Em *Bocato e Ferreira* (2014) são expostas as modalidades do Protocolo verbal individual e do Protocolo verbal em grupo. Vejamos a seguir:

- Protocolo verbal individual: subdivide-se nas situações com interação e sem interação;
- Protocolo verbal individual com interação: é desenvolvido com interação. Neste caso, o sujeito da pesquisa, ao realizar a tarefa, pode receber algum tipo de auxílio por parte do pesquisador;
- Protocolo verbal individual sem interação: não há interação entre o pesquisador e o sujeito ao longo da realização da tarefa;

- d) Protocolo verbal em grupo: um grupo de pessoas desenvolve a leitura e discussão de um texto específico. Ainda em *Boccatto e Ferreira* (2014), são apresentadas vantagens e desvantagens do Protocolo verbal em grupo.

Destaca-se, no Brasil, a aplicação do Protocolo verbal nas pesquisas de *Boccatto e Ferreira* (2014), *Boccatto e Fujita* (2010), *Fujita* (1999, 2003, 2009), *Fujita, Nardi e Fagundes* (2003) e *Rubi* (2008).

4.6.3 *Soft Systems Methodology*

A metodologia sistêmica *soft* ou *Soft Systems Methodology* (SSM) foi desenvolvida a partir da abordagem da Engenharia de sistemas. As primeiras contribuições na área se devem a *Peter Checkland*, da *Universidade Lancaster* (Reino Unido), no final dos anos 1960. Essa metodologia é adequada para a investigação de problemas que ocorram em um sistema. O termo “*soft*” está associado a certa dificuldade na caracterização precisa do problema (ausência de consenso sobre ele). Por exemplo: “Como melhorar serviços da educação pública?” ou “Como gerenciar crises no transporte público em uma metrópole?”. Tais problemas estão colocados de forma não precisa. Segundo *Peter Checkland*, a SSM não é uma ferramenta ou técnica para ser utilizada ocasionalmente, mas uma forma de pensar e agir todos os dias.

Santos (2003, p. 263) destaca que

há dois tipos de análise: uma cultural, em que se busca a identificação de valores, normas, papéis e mecanismos de poder, interações, relacionamentos etc. A outra lógica, em que se aplicam as sete etapas da metodologia e que permite, a partir de um exercício inicial de estruturação das questões identificadas na situação em estudo, definir seus subsistemas relevantes.

A abordagem SSM é adotada na modelagem de processos organizacionais, sendo utilizada para verificar problemas em sistemas complexos já existentes, que envolvam atividades humanas, ou ainda para desenvolver sistemas novos. Citamos *Moraes* (2010, p. 148): “tal metodologia está consolidada em círculos acadêmicos como a mais desenvolvida abordagem *soft* e um importante veículo para pesquisa qualitativa”. Ainda: a abordagem SSM busca uma compreensão holística da situação.

Existe uma oposição entre as metodologias em *hard* e *soft*. Citamos *Furnival* (1995) quando explica que as metodologias *hard* concentram-se mais nas soluções técnicas, ao passo que as *soft* tentam abordar as questões mais complexas, relacionadas à estrutura de poder da organização, e a seus fatores psicológicos e políticos, ou seja, fatores fundamentalmente humanos.

As metodologias *hard* estão associadas à Física, Química, Biologia e à Matemática. Já as metodologias *soft* estão associadas às ciências comportamentais e às ciências sociais: Antropologia, Ciência Política, Psicologia, Administração, Economia, Educação. Além disso, a aplicação da metodologia SSM envolve a participação do pessoal de todos os níveis organizacionais.

De acordo com *Bellini et al.* (2004), a metodologia SSM apresenta os seguintes benefícios:



- a) empodera o analista a tratar de situações problemáticas (e não de problemas específicos);
- b) favorece uma compreensão mais profunda sobre os pontos fracos organizacionais;
- c) descola-se de soluções tecnológicas e modismos;
- d) exige a participação das pessoas envolvidas na situação problemática, de modo que a possibilidade de propostas alienadas da realidade da organização serem desenvolvidas é reduzida.

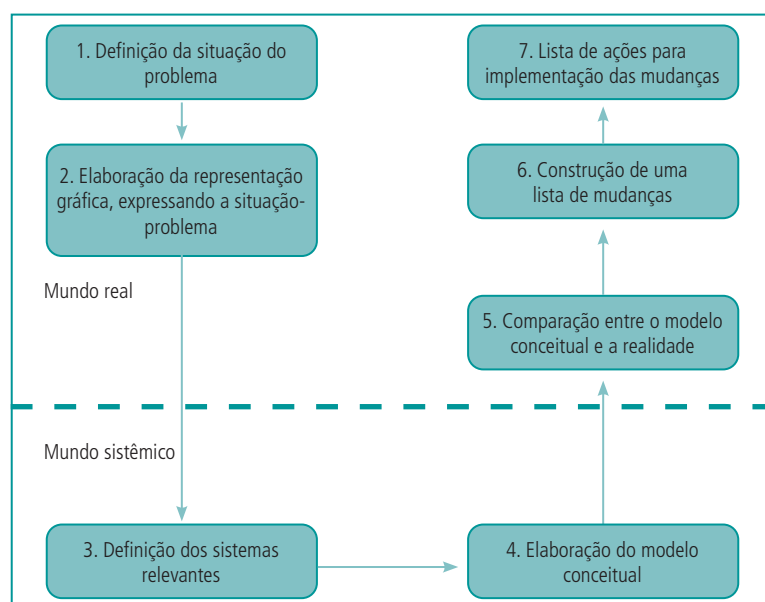
A SSM apresenta duas grandes contribuições para a reflexão sobre sistemas. Em primeiro lugar, ao enxergar o sistema como “sistêmico” e não “sistemático” e, depois, ao tratar de sistemas com atividade humana. Sistemas com atividade humana têm como característica que elementos diferentes da situação-problema interagem com os outros elementos do sistema e não podem ser simplesmente isolados. Os componentes interativos são influenciados por humanos, com toda uma gama de diferentes visões de mundo, interesses, valores e opiniões.

De acordo com *Checkland* (1981), são necessárias sete etapas para a aplicação da SSM:

- a) definição da situação do problema;
- b) elaboração da representação gráfica, expressando a situação-problema;
- c) definição dos sistemas relevantes;
- d) elaboração do modelo conceitual;
- e) comparação entre modelo conceitual e a realidade;
- f) construção de uma lista de mudanças;
- g) lista de ações para implementação das mudanças.

A Figura 34, a seguir, apresenta os estágios da SSM:

Figura 34 – Estágios da SSM



Fonte: *Checkland* (1985).

Para conhecer detalhes de cada uma das etapas da *SSM*, consulte *Moraes* (2010). A seguir, faremos uma descrição sucinta de cada uma dessas etapas.

- Etapa 1: como objetivos da exploração da situação não estruturada, o pesquisador espera obter um quadro mental da situação. Espera-se que ele consiga explicitar as regras, os processos e responsabilidades estabelecidos na situação-problema.
- Etapa 2: diagrama/figura (*rich picture*): por meio do uso de símbolos específicos (frequentemente desenhados à mão), o pesquisador irá representar a estrutura da organização, seus aspectos físicos, os processos nela envolvidos e a forma como tais símbolos se relacionam entre si. Espera-se que o pesquisador consiga representar os sentimentos e críticas que as pessoas apresentam da situação-problema.

Para observar alguns diagramas produzidos no contexto da *SSM*, sugerimos que você consulte a apresentação em *Power Point* “User Centred Approaches to Information Services Development (Soft Systems Methodology): Rich Pictures, The Peter Checkland Model”, disponível no *link*: <<http://pt.slideshare.net/geoffreyawalker/soft-systems-methodology>>.

- Etapa 3: causas (identificação do que é o essencial no sistema). Nesta etapa, *Checkland* (1985) utiliza o termo mnemônico CATWOE. É uma sequência de itens a serem checados. Vejamos: os componentes CATWOE são desenvolvidos para testar uma definição raiz e garantir que ela esteja bem construída, ou seja, que ela contenha o que é importante (nos exemplos, uma situação de sistemas de informação). Veja:
 - **c**liente: clientes, beneficiários (ou ainda, vítimas) da atividade. Exemplo: estudantes;
 - **a**tores: são aqueles que efetuam a transformação. Exemplo: bibliotecários;
 - **t**ransformação: são as mudanças que são realizadas. Exemplo: acesso aos documentos da biblioteca (artigos científicos/ livros) para auxiliar os estudos;
 - **w**eltanschauung (visão de mundo): premissas aceitas do sistema. Exemplo: centro de aprendizagem com apoio atualizado aos estudantes dos serviços da biblioteca;
 - **o**wner (proprietário): são os proprietários do sistema. Exemplo: bibliotecários;
 - **e**nvironment (ambiente): é o que envolve ou engloba o sistema. São influências externas ao sistema, aspectos jurídicos/ legais, restrições diversas ao sistema. Exemplos: leis de direito autoral.
- Etapa 4: nesta etapa, o pesquisador descreve (a partir de uma abordagem sistêmica) modelos das partes relevantes da situação problemática em estudo. O pesquisador concebe os modelos na expectativa de um funcionamento “ideal”. Neste ponto, estamos no “mundo sistêmico”.
- Etapa 5: confronto e comparação dos modelos concebidos na etapa anterior com a realidade. O objetivo desta etapa não é

operacionalizar os modelos concebidos anteriormente, mas, antes de tudo, comparar diferenças e semelhanças entre eles. Perguntas-chave a serem respondidas pelo pesquisador nesta etapa: por exemplo, o pesquisador poderá comparar o histórico do sistema (o que já aconteceu anteriormente) com os modelos conceituais. Também poderá utilizar os modelos conceituais como ferramentas de investigação.

- Etapa 6: nesta etapa, o pesquisador buscará a identificação daquelas mudanças exequíveis e que são adequadas do ponto de vista do que se espera do sistema. *Costa (2003)* e *Longaray, Ensslin e Mackness (2013)* utilizam as expressões “mudanças desejáveis” e “mudanças realizáveis”;
- Etapa 7: finalmente, nesta etapa, são implementadas, na prática, as ações/ transformações identificadas na etapa 6.

Deve-se destacar, como características da *SSM*, que os integrantes do sistema (por exemplo, os bibliotecários frente à instalação de um novo *software* de gestão de bibliotecas) participam do processo. Também é importante ter em mente que qualquer situação humana apresenta uma dimensão política.

4.6.3.1 Limitações ou restrições à abordagem *SSM*

Certamente, a metodologia *SSM* não é adequada para a operacionalização de soluções, ou sua implementação em situações com relações de poder autocráticas, que evitem a discussão coletiva de ideias, ou ainda em casos de alta rotatividade no emprego (MORAES, 2010).

4.6.4 Discurso do sujeito coletivo

Na década de 1990, os professores *Fernando Lefevre* e *Ana Maria Cavalcanti Lefevre*, da *Faculdade de Saúde Pública da USP*, estudaram servidores públicos do município de São Paulo para investigar a opinião deles acerca do Programa de Gerenciamento Integrado da *Secretaria da Saúde de São Paulo* (DUARTE, MAMEDE, ANDRADE, 2009). Os pesquisadores identificaram respostas bastante semelhantes. A partir desse trabalho, surgiu a metodologia do *Discurso do sujeito coletivo (DSC)*. A abordagem do *DSC* busca resolver críticas tanto às pesquisas quantitativas (nas quais temos respostas representativas, mas que nem sempre possibilitam uma compreensão mais profunda do problema) quanto às pesquisas qualitativas (nas quais costumamos encontrar respostas “profundas”, mas nem sempre representativas). O fundamento para o *DSC* é a Teoria da representação social (LEFEVRE, 2015).

O *DSC* é utilizado para se processar depoimentos obtidos em pesquisas de opinião, com características qualitativas e quantitativas. As opiniões de um grupo são coletadas de forma individual. Essas opiniões individuais não são “achatadas”, homogeneizadas ou equalizadas por um procedimento quantitativo que acaba por ignorar diferenças individuais. Por outro lado, opiniões semelhantes em seu sentido são categorizadas e agrupadas. Então, são “somadas”, tanto de forma qualitativa quanto quantitativa. Ou seja, é realizada uma “soma quantitativa”. Assim, podemos encontrar diversos sujeitos individuais expressando, coletivamente, um mesmo discurso.

Não pode ser deixada de lado a fundamentação quantitativa: nas pesquisas de opinião, é relevante que possam emergir todas as opiniões

existentes na coletividade. Deve-se representar, com as ponderações adequadas, a sociedade/ comunidade na pesquisa. Não podemos ignorar a proporção de indivíduos partícipes da elaboração dos vários DSC, com suas características socioeconômicas e demográficas, relevantes para a pesquisa.

Lefevre e Lefevre (2005) apresentam quatro elementos (operações) para constituir o DSC:

- a) expressões-chave (E-Ch): são trechos das entrevistas que descrevem o conteúdo da melhor maneira. É o pesquisador quem deve identificar e destacar as expressões-chave do discurso do depoente;
- b) ideias centrais (IC): cumprem um papel discriminante, pois possibilitam identificar e distinguir os sentidos nos depoimentos. As ideias centrais cumprem um papel classificatório;
- c) ancoragens (AC): as ancoragens identificam as ideologias (ou teorias) apresentadas nos depoimentos. Para que haja uma ancoragem no depoimento, é preciso encontrar, no seu corpo, marcas discursivas explícitas dela (LEFEVRE; LEFEVRE, 2015);
- d) discursos do sujeito coletivo (DSC): ao se alocar, em um discurso-síntese, expressões-chave que apresentem as mesmas ideias centrais ou ancoragens apresentadas em primeira pessoa do singular, teremos um DSC.

Como exemplos de trabalho na área de Biblioteconomia/ Ciência da Informação que faz uso do método DSC, podem ser citados Oliveira (2010), Pizarro (2011), Santos (2014) e Souza (2011). Em Lefevre (2015), você encontrará um banco de trabalhos, na área de Biblioteconomia/ Ciência da Informação, produzidos com o DSC.

Para tratamento dos dados do método DSC, é utilizado um *software* específico: o *Qualiquantisoft*®. O *software* está disponível em três versões: para familiarização, para uso individual e para uso institucional.

4.6.5 Análise do conteúdo

Considera-se H. Lasswell, da Universidade de Chicago, como o primeiro representante da Análise do conteúdo (AC). Já em 1915, ele estudava propaganda e imprensa. No seu trabalho *Propaganda Technique in the World War*, publicado em 1925, ele apresenta análises de imprensa e propaganda do período da guerra (CAMPOS, 2004).

A AC busca compreender, de forma sistemática, um material originado de fontes diversas (pode ser originado de entrevistas, da imprensa, cartazes, filmes, vídeos ou discos). A AC pode ser realizada com procedimentos quantitativos (utilizando-se da determinação da frequência de certos elementos no objeto de estudo) ou qualitativos (fundada na investigação dos valores ou interesses no objeto de estudo). Contudo, ao se adotar uma perspectiva quantitativa, pode-se perder algo relevante apenas por ser de baixa frequência, mas relevante para uma compreensão mais profunda do objeto sob análise. Destacamos de Cordeiro (2004, p. 26):

A Análise de conteúdo não é um simples levantamento quantitativo de determinados tipos de material escrito. A elaboração de índices quantificados dos livros ou artigos que existem numa biblioteca, ou qualquer outro acervo, não é considerada como análise de conteúdo, é simplesmente uma técnica documentária.



A AC pode se dar de duas formas: como verificação (nas situações nas quais pretendemos testar uma hipótese já definida) e como exploração (nas situações nas quais pretendemos explorar temas).

Na obra de *Bardin* (1977), a autora apresentou os conceitos e procedimentos para a AC. Recomenda-se fortemente, ao leitor interessado na AC, que consulte essa referência, onde podem ser encontrados exemplos de aplicação da AC. A autora utilizou a AC para estudar a comunicação de massas, além de as ter usado em investigações psicossociológicas. A AC apresenta a pretensão de neutralidade no que diz respeito ao significado do texto. A AC pretende apresentar um conjunto de técnicas que possibilitem, de forma objetiva e precisa, a investigação sistemática de um corpo textual (que pode ser audiovisual), de modo a quantificar a frequência de palavras, frases ou temas, tendo em vista posteriores avaliações e comparações.

Com a AC, pode ser avaliada uma mensagem para mais de um receptor, ou podem ser avaliados os mesmos receptores em distintas situações. A AC se propõe, também, a analisar sob quais condições foi produzida a mensagem.

A AC pode ser utilizada tanto em pesquisas exploratórias quanto em investigações para verificação. Em procedimentos de verificação, hipóteses previamente estabelecidas podem ser confirmadas ou refutadas.

A AC pode ser classificada, também, em quantitativa ou qualitativa. Na abordagem quantitativa, são utilizadas frequências das ocorrências e/ou quantidades de referências. Já na abordagem qualitativa, objetiva-se uma compreensão mais profunda daquilo que é importante no texto.

A AC ainda pode ser dividida quanto aos métodos adotados. São três classes de métodos: análises temáticas, análises formais e análises estruturais.

4.6.5.1 Categorias na AC

A proposta da AC é a de que as unidades de análises (palavras ou símbolos em geral) podem ser separadas em diferentes categorias, chamadas de categorias conceituais. Estas devem atender ao requisito de serem mutuamente exclusivas, objetivas e exaustivas.

Categorização: quando agrupamos (separamos em conjuntos) itens que possuem aspectos em comum, chamamos esse processo de categorização.

Podem ser adotados vários critérios de categorização. Podemos adotar critérios: semânticos (o agrupamento é realizado em função do tema), léxicos (o agrupamento ocorre em função do sentido das palavras), sintáticos e expressivos.

4.6.5.2 As fases na AC

A AC faz uso de classificação-indexação. De acordo com *Bardin* (1977), a AC apresenta as seguintes fases:

- a) 1ª fase: pré-análise (que se subdivide em leitura flutuante, escolha dos documentos, preparação do material e construção de indicadores);
- b) 2ª fase: fase em que o material será explorado;

- c) 3ª fase: fase em que serão realizados o tratamento dos resultados, a inferência e, por fim, a interpretação.

A seguir, detalhamos cada uma dessas fases:

- 1ª fase – pré-análise: na fase de pré-análise, são definidos os critérios da pesquisa e sua estrutura básica. A primeira fase é um período de intuições (BARDIN, 1977). Primeiramente, são realizadas as denominadas “leituras flutuantes”, que são diversas leituras feitas ainda sem que se prenda a uma sistematização. A finalidade maior das leituras flutuantes é apreender os aspectos mais gerais e as principais ideias do texto (CAMPOS, 2004). Ainda nesta fase, são elencados os documentos, é efetuada a formulação de hipóteses e dos objetivos da pesquisa, bem como a referenciação dos índices e a construção de indicadores.

A 1ª fase pode ser decomposta na escolha dos documentos, formulação de hipóteses, definição dos objetivos e construção dos indicadores a serem utilizados na interpretação (na 3ª fase). É claro que a escolha dos documentos é função do objetivo definido pelo pesquisador.

Após a delimitação do universo de documentos a serem destacados, deve-se constituir um *corpus*. De acordo com Bardin (1977, p. 96), “o *corpus* é o conjunto de documentos tidos em conta para serem submetidos aos procedimentos analíticos”. Os atributos a serem atendidos pelo *corpus* são: a regra da exaustividade (não podemos deixar qualquer dos elementos constitutivos do *corpus* de fora, a menos que se justifique por questões de rigor); regra da representatividade (diz respeito à representatividade da amostragem efetuada); regra da homogeneidade (o conjunto de documentos escolhidos deve ser homogêneo – ou seja, não devem constituir um conjunto com elementos díspares entre si); regra de pertinência (os documentos devem ser pertinentes ao objetivo propostos).

Acerca da construção dos índices, o analista possui grande liberdade. Ele pode tomar como medida da importância de um tema, como transporte individual, a frequência com que as palavras *automóvel* ou *carro* aparecem no texto. O analista, nesse caso, poderá construir tabelas com frequências absolutas e relativas da contagem das palavras-chave no texto. Sob esse aspecto, os computadores podem ser bastante úteis para automatizar parte do processo.

- 2ª fase – exploração do material: na fase de exploração do material, as escolhas realizadas na fase anterior são implementadas e é realizada a codificação e sistematização. Esta fase é a de maior duração. Nela, são extraídos recortes do texto em unidades de contexto e de registro. É a fase de seleção das unidades de análise e nela são fundamentais as unidades de registro (UR) e as unidades de contexto (UC):
 - definição de unidade de registro (UR): são as unidades que visam à categorização e contagem de frequências. É o menor recorte (do ponto de vista semântico) que se retira do texto. Tanto o tamanho quanto a natureza das UR podem variar. Bardin (1977, p. 104) cita como as UR mais utilizadas “a palavra, o tema, o referente (temas em torno dos quais o discurso está organizado), o personagem, o acontecimento e o documento”;



- definição de unidade de contexto (UC): tem o papel de tornar compreensível a unidade de registro. As unidades de contexto podem ser um ou mais parágrafos.

Para a codificação, o pesquisador deve efetuar três escolhas: a escolha das unidades (qual será o recorte adotado); a da forma de contagem (como será desenvolvida a enumeração) e a escolha das categorias (como será realizada a classificação e a agregação) (BARDIN, 1977).

Na fase de categorização, o analista efetuará a classificação dos elementos, diferenciando-os e efetuando um agrupamento por critérios anteriormente definidos. Na categorização, em primeiro lugar, é efetuado o inventário (os elementos são separados); em seguida, procede-se à classificação: os elementos são organizados por critérios. Em Bardin (1977, p. 121), o leitor encontrará exemplos de categorias.

- 3ª fase – resultados: é nesta última etapa que os resultados serão tratados, que será desenvolvida a interpretação e que as inferências serão efetuadas. O investigador, a partir de resultados obtidos na segunda fase, fará o esforço de produzir resultados válidos e significativos. É neste ponto que ele deverá “ultrapassar” o conteúdo manifestado nos documentos e expressar o conteúdo latente. A inferência é definida por vários polos de atenção. O propósito da AC é produzir inferências que permitam ao investigador ir às camadas mais profundas do texto. As inferências são realizadas a partir dos efeitos (os indicadores ou variáveis de inferência) para se analisar as causas (que são as variáveis de inferência).

4.6.6 Método Delphi

A técnica *Delphi* (também conhecida como técnica *Delfos*) não é comumente utilizada em grupos pequenos. Seu uso mais frequente é para previsão de futuro por um grupo de especialistas em determinada área. É bastante comum que sejam efetuadas várias “rodadas”, até que uma medida da convergência (ou consenso) dos pareceres dos especialistas seja atingida. O nome *Delfos* ou *Delphi* remete ao *Oráculo de Delfos*, templo situado na Grécia antiga, no qual as sacerdotisas de *Apolo* (as *pitonisas*) apresentavam profecias e respostas aos consulentes.

A origem dessa técnica data de 1963, quando foi idealizada por Dalkey e Helmer, da *RAND Corporation* (VIRGILLITO, 2010). Embora inicialmente desenvolvida como uma técnica para prognósticos, ao longo dos anos, a técnica *Delphi* vem sendo adaptada para outras aplicações, tal como o entendimento de algum evento que já tenha acontecido (e não como ferramenta de prognóstico). A técnica é utilizada em situação de incerteza e/ ou falta de dados. Por exemplo, para a pesquisa por explicações sobre a sensibilidade do usuário/ consumidor/ cliente com respeito a determinadas questões.

De forma bastante simples, a técnica consiste na aplicação de questões (*survey*), coleta e organização das respostas a um grupo de especialistas. Eventualmente, novos ciclos de questões podem ser realizados junto ao corpo de especialistas. Eles podem modificar suas estimativas, confluindo para o *consenso emergente*. Por outro lado, os especialistas têm liberdade para manter sua estimativa, acrescentando justificativas para tal. Nesse processo *cíclico*, eles podem trocar informações, de forma a arguam

outros especialistas e, até mesmo, convencê-los a alterar suas estimativas. Ao longo das várias rodadas, o consenso tende a aumentar.

As informações discordantes não são simplesmente descartadas: elas são utilizadas para produzir informação mais detalhada. Outro aspecto da técnica é que os participantes permanecem *anônimos* entre si (não há interações “olho no olho” entre os especialistas). Assim, características como cargo, posição social, etnia e gênero acabam por não influenciar na modificação de posicionamento entre os especialistas (VIRGILLITO, 2010). Citando Virgillito (2010, p. 40): “com a técnica *Delphi*, o resultado será uma decisão que levará em conta a opinião da maioria, porém sem os efeitos da interação social”.

A técnica *Delphi* não faz uso de ferramentas matemáticas/ estatísticas sofisticadas. Outro aspecto é que ela dificulta a predominância de uma ou poucas pessoas sobre o restante do grupo. Esta técnica baseia-se na hipótese de que julgamentos intuitivos são uma fonte valiosa de percepção do futuro (KAIRALLA, 1984). O *Delphi* evita a formação de estruturas hierárquicas.

A técnica *Delphi*, inicialmente, foi aplicada à área militar. Ela já foi aplicada à área da Saúde, nas Ciências Contábeis, na Engenharia de *Software*/ Tecnologia da Informação, nos transportes públicos, na Enfermagem, entre outras.

Como exemplo de aplicação da técnica *Delphi* no Brasil, na área de Ciência da Informação/ Biblioteconomia, o leitor pode consultar o artigo de Maira Murrieta Costa, “As bibliotecas brasileiras em 2018: resultados da técnica de Delfos”, publicado no periódico *Perspectivas em Ciência da Informação*, v. 17, n. 1, de janeiro/ março de 2012. Nessa investigação, foram realizadas quatro rodadas de coleta de dados.

4.6.6.1 Etapas para aplicação da técnica *Delphi*

São as etapas:

- a) 1ª etapa: seleção dos respondentes. São selecionados especialistas na área a ser estudada;
- b) 2ª etapa: após a escolha dos respondentes, é realizado um contato com informações sobre a pesquisa que está sendo realizada e um primeiro questionário é enviado a eles. Este é o momento de um pré-teste para avaliar o formato e a adequação do questionário. Cada respondente deve perceber-se como pessoalmente envolvido com o estudo (KAIRALLA, 1984). Podem ser realizadas adaptações e mudanças no questionário, em função de eventuais sugestões e comentários dos respondentes;
- c) 3ª etapa: são enviados os primeiros questionários. Os respondentes são solicitados a incluir críticas e sugestões em suas respostas;
- d) 4ª etapa: organização das respostas. Esta etapa deve ser conduzida por um profissional com profundos conhecimentos do tema;
- e) 5ª etapa: nesta etapa, o especialista é confrontado com suas próprias respostas. A partir das respostas ao primeiro questionário, é desenvolvido o segundo questionário. Os respondentes devem identificar, nele, que foram partícipes e tiveram influência sobre ele. São construídas tabelas indicando as porcentagens de respostas às questões. Solicitações de informações adicionais, por parte



dos respondentes, podem ser atendidas. Caso um respondente apresente respostas que diverjam da maioria, será solicitada a justificativa desse posicionamento. É nesse ponto que ocorre o *feedback* interativo, tão característico da técnica *Delphi*;

- *etapas posteriores*: novos questionários serão enviados, com os respectivos *feedbacks*. Não há um número pré-definido de ciclos a ser atendido.

4.6.6.2 A técnica Delphi eletrônica

Com o advento da internet, os pesquisadores passaram a contar com esta nova plataforma para realizar levantamentos e pesquisas. Nela, continuam valendo as características da técnica *Delphi* “tradicional”: anonimato, busca do consenso, *feedback* interativo e seleção de especialistas. Contudo, existem divergências quanto ao aspecto da qualidade de respostas em questionários via *web*. Respostas que podem ser dadas muito rapidamente podem reduzir um tempo de reflexão necessário, por parte do especialista, para produzir uma resposta mais ponderada. Outro aspecto é que é muito frequente, em pesquisas via *web*, uma taxa de respostas muito baixa.

4.6.6.3 Limites da técnica Delphi

Uma crítica recorrente à técnica *Delphi* é exatamente a busca de consenso: segundo alguns críticos, a técnica permitiria o surgimento de consensos artificiais, dependendo de como é realizado o *feedback* e de como o retorno é dado aos respondentes. Ainda: se os especialistas não possuírem “origens” diversas, será bem mais fácil que ocorram consensos “artificiais”.



4.6.7 Atividade

Esta questão atende ao objetivo *d*: “analisar os dados em pesquisas qualitativas”.

A AC é utilizada para avaliar a frequência de fenômenos na comunicação. A AC pode realizar procedimentos quantitativos (contando quantas vezes determinado elemento aparece em um texto, entrevista, filmes, vídeos etc.) ou qualitativos (investigando valores ou interesses no objeto de estudo). Como vantagens da AC, *Cunha* (1982) apresenta que ela pode ser utilizada para mensurar a legibilidade de um texto ou comunicação e a repetibilidade do método.

Em *Baptista e Cunha* (2007), descreve-se uma pesquisa que utilizou a AC a partir de procedimentos quantitativos:

Spink et al. (2000) analisaram 51.473 transações de busca de 18.113 usuários realizados pelo mecanismo de busca *Excite*. Os autores realizaram a análise por meio dos *logs* (registro de uma operação com arquivo(s) em meio digital), com o objetivo de verificar as falhas nas buscas realizadas. Buscaram similaridades e diferenças nas buscas e compararam com buscas feitas por usuários de sistemas de informação tradicionais, como o *Dialog*.

A partir do trecho citado, aponte uma possível desvantagem do método Análise de conteúdo quando utiliza procedimentos quantitativos.

Resposta comentada

Uma desvantagem da AC, neste caso, é o grande volume de dados que precisa ser analisado. Certamente, podem ser utilizados *softwares* específicos para facilitar essa análise (são os *computer-aided qualitative data analysis software*). Citam-se os *softwares* *Clementine (SPSS)* e o *NUD*IST*.



4.6.8 Atividade

Esta questão atende ao objetivo d: “analisar os dados em pesquisas qualitativas”.

Você acredita que seja viável tratar quantitativamente dados qualitativos, coletados por meio de entrevistas? Se não acredita ser viável, justifique. Se você acredita que seja viável, apresente um exemplo.

Resposta comentada

Sim, é viável. Como exemplo, considere que você efetuou um conjunto de 20 entrevistas. Você poderá agrupar as respostas às suas questões em categorias e, partir delas, criar tabelas de frequência ou gráficos (histogramas, pizzas etc.), tratando quantitativamente um dado que foi coletado, originalmente, em uma perspectiva qualitativa.



4.6.9 Atividade

Esta questão atende ao objetivo c: “coletar os dados entrando no ambiente da pesquisa”.

Explique a noção de saturação na amostragem em pesquisas qualitativas.

Resposta comentada

A noção de saturação indica que a coleta dos dados deve ser realizada até o ponto no qual o pesquisador identifica que deixam de aparecer dados com relevância ou informação nova. Em outras palavras: padrões começam a se repetir.



4.6.10 Atividade

Esta questão atende aos objetivos *c* e *d*: “coletar os dados entrando no ambiente da pesquisa” e “analisar os dados na pesquisa qualitativa”.

O que é a triangulação de métodos na pesquisa qualitativa?

Resposta comentada

A triangulação de métodos é a utilização de mais de uma metodologia para abordar o problema de pesquisa qualitativa. No caso específico da etnografia, a triangulação é considerada um ponto importante para incremento na qualidade na pesquisa qualitativa. A esse respeito, consulte *Flick* (2009, p. 102).

Suponha uma pesquisa qualitativa de estudo de comunidade. Nela, podem ser realizadas observações. Também podem ser feitas entrevistas com pessoas que tenham passado por experiências de alguma forma marcantes ou notáveis. Por fim, podem ser obtidos dados com reuniões de grupos focais com subgrupos homogêneos de maior interesse no problema de pesquisa.



4.6.11 Atividade

Esta atividade atende ao objetivo *b*: “definir a amostra inicial em pesquisas qualitativas”.

Qual são as sugestões para o tamanho da amostra em estudos qualitativos?

Resposta comentada

Sampieri, Collado e Lucio (2013) recomendam que estudos do tipo etnográfico, Teoria fundamentada, entrevistas ou observações devem apresentar amostras entre 30 e 50 casos. Nos estudos que utilizem a técnica de história de vida, o sujeito de estudo deve constar da amostra, bem como o maior número possível de pessoas que tenham tido contato com eles. Para grupos focais, a recomendação da literatura é que contenham entre seis a 12 pessoas por reunião.

Se você pretende descrever um grupo com detalhes, sua amostra deve ser o mais homogênea possível.

Outra recomendação é dada por *Flick* (2009, p.122): “a seleção gradual da amostra”.



4.6.12 Atividade

Esta atividade atende ao objetivo c: “coletar os dados entrando no ambiente da pesquisa”.

Apresente as principais características da técnica *Delphi* nos estudos de usuários.

Resposta comentada

Não é costumeiro utilizar a técnica *Delphi* (ou *Delfos*) em grupos pequenos. É uma ferramenta utilizada quando o grau de incerteza é elevado.

O objetivo principal desta técnica é identificar tendências.

Na técnica *Delphi* os participantes expressam suas opiniões sobre determinado assunto e são realizadas anotações dessas opiniões. Posteriormente, são realizadas análises estatísticas desses registros. Aquelas com maior votação são selecionadas. Os participantes recebem as mais votadas e desenvolvem uma outra rodada de discussão para escolha.

Essas rodadas são efetuadas até que se obtenha o consenso com respeito às várias opiniões. Ao final do processo, o pesquisador pode determinar variáveis que tenham apresentado maior convergência.

A técnica *Delphi* é considerada uma técnica livre de influência individual. É levada em consideração a opinião do grupo e não a de cada indivíduo, em particular. Ela leva em conta a opinião da maioria, mas sem os efeitos da interação social.

No caso de estudos de usuários, a técnica *Delphi* pode ser utilizada para avaliar a sensibilidade dos usuários com questões relativas ao serviço oferecido. Em particular, como essa técnica permite obter opiniões de especialistas eliminando os efeitos da interação social, eles não estarão influenciados por outras opiniões.



4.6.13 Atividade

Esta atividade atende ao objetivo b: “definir a amostra inicial em pesquisas qualitativas”.

Com o advento da internet, o computador passou a ser uma importante ferramenta de informação e comunicação. A análise de redes sociais adquire cada vez maior importância em várias áreas: Marketing, Sociologia, Psicologia, Ciência política e, claro, diante dos estudiosos da Biblioteconomia e Ciência da Informação.

A palavra netnografia é um neologismo do início do século XXI. Neste livro, na subseção 4.5.5, apresentamos a definição dada por Kozinets (2014, p. 61) para netnografia.

Destaque aspectos positivos e negativos da netnografia.



A netnografia, por ser realizada a partir de comunidades *on-line*, costuma ser mais barata que grupos focais ou a condução de entrevistas. Costuma consumir menos tempo, não exige o deslocamento físico do pesquisador para o campo (já que o campo é a própria internet ou rede social). Também é um método menos invasivo (o pesquisador pode observar as comunicações entre os membros de uma comunidade virtual sem ser percebido).

Como aspecto negativo, pode-se destacar que, nas relações mediadas por computador, perdemos os aspectos da linguagem corporal que o contato presencial oferece. Por outro lado, à medida que a disponibilidade de conexão de maior velocidade e confiabilidade aumenta, os usuários das redes sociais passam a utilizar ferramentas de comunicação que envolvem a imagem (por exemplo, citamos reuniões virtuais via *Skype* ou análogos).

4.7 CONCLUSÃO

Nesta Unidade, você teve a oportunidade de conhecer um pouco da diversidade de estratégias e abordagens da pesquisa qualitativa: as entrevistas, a observação, os grupos focais, a netnografia, entre outras. Uma abordagem não é intrinsecamente melhor que a outra. O que é relevante é observar a adequação da estratégia de coleta e análise de dados à sua questão de pesquisa.

Não deixe de visitar os *links* apresentados para conhecer um pouco sobre os *softwares* específicos utilizados em pesquisa qualitativa.

Quando nosso interesse é a compreensão aprofundada a partir do ponto vista dos participantes da pesquisa sobre o fenômeno em estudo, com o relato de suas experiências pessoais e individuais e com os significados que atribuem às mesmas, recomenda-se a adoção de métodos qualitativos.

Um ponto distintivo dos métodos qualitativos é sua diversidade, seja nas técnicas de coleta de dados, seja nas de amostragem ou de análise de dados.

Os desenhos do processo em pesquisas qualitativas também seguem roteiros bem mais flexíveis que aqueles que devem ser seguidos nos métodos quantitativos.

As hipóteses em estudos qualitativos são amplas e adaptadas aos dados. Também podem ser alteradas ao longo do processo da pesquisa.

A amostragem qualitativa não busca a posterior generalização para a população como um todo. Em algumas circunstâncias, basta uma amostra significativa com um único indivíduo.

A etapa de análise de dados também apresenta grande diversidade.

Cabe ao pesquisador estudar as várias técnicas de coleta e análise de dados para identificar aquelas mais adequadas ao seu problema de pesquisa e os recursos de tempo e dinheiro de que dispõe.

RESUMO

Nesta Unidade, você estudou como é a formulação de problemas em pesquisas qualitativas. A formulação do problema em pesquisas qualitativas deve conter:

- a) os objetivos da pesquisa;
- b) as perguntas da pesquisa;
- c) a justificativa e viabilidade da pesquisa;
- d) a avaliação de deficiências sobre o que se conhece daquele problema;
- e) a caracterização do contexto/ local e do ambiente de pesquisa.

Na sequência, vimos os desenhos do processo em pesquisa qualitativa: Teoria fundamentada, etnografia, estudo de caso e histórias de vida e pesquisa-ação.

Discutimos, então, a questão da amostragem nas pesquisas qualitativas. As amostras, nessas pesquisas, não se preocupam com a generalização dos resultados para a população como um todo.

Entre as técnicas de coleta de dados nas pesquisas qualitativas, abordamos: observação, entrevistas, grupos focais, análise de documentos, pesquisa qualitativa na internet e netnografia.

Com relação à análise de dados nas pesquisas qualitativas, foram estudadas: a Análise do discurso, Protocolo verbal, *Soft Systems Methodology*, Discurso do sujeito coletivo, Análise do conteúdo e método *Delphi*.



Sugestão de Leitura

DICK, B. **Grounded theory**: a thumbnail sketch. [S.l.: s.n.], 2005. Disponível em: <<http://www.aral.com.au/resources/grounded.html>>. Acesso em: 30 jun. 2015.

ENDRES-NIGGEMEYER, B.; NEUGEBAUER, E. Professional summarizing: no cognitive simulation without observation. **Journal of American Society for Information Science**, New York, v. 49, n. 6, p. 486-506, 1998.



GLASER, Barney G.; STRAUSS, Anselm L. **The discovery of grounded theory: strategies for qualitative research**. New York: Aldine de Gruyter, 1967.

GOTOH, T. Cognitive structure in human indexing process. **Library and Information Science**, [S.l.], n. 21, p. 209-26, 1983.

INGWERSEN, P. Search procedures in the library: analysed from the cognitive point of view. **Journal of Documentation**, [S.l.], v. 38, p. 165-191, 1982.

SAUPERL, A. **Subject determination during the cataloging process**. Lanham: Scarecrow Press, 2002.

STRAUSS, Anselm L.; CORBIN, Juliet. **Basics of qualitative research: grounded theory, procedures, and techniques**. Newbury: SAGE, 1990.

REFERÊNCIAS

AMARAL, Renilda Gonçalves do. **Inclusão dos usuários das Unidades de Informação de Águas Lindas de Goiás por meio do acesso à informação**. 2013. Tese (Doutorado em Ciência da Informação) – Faculdade de Ciência da Informação e Documentação, Universidade de Brasília, Brasília, 2013.

ANGROSINO, Michael. **Etnografia e observação participante**. Porto Alegre: Bookman/ Artmed, 2009. (Coleção Pesquisa Qualitativa).

BAPTISTA, Sofia Galvão; CUNHA, Murilo Bastos da. Estudo de usuários: visão global dos métodos de coleta de dados. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 12, n. 2, p. 168-184, maio/ ago. 2007.

BARDIN, Laurence. **Análise de Conteúdo**. [S.l.]: Presses Universitaires de France, 1977. Tradução de: L'analyse de Contenu.

BASTOS, Flávia Maria. **A interação do usuário com catálogos bibliográficos on-line: investigação a partir da Teoria Fundamentada**. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista, Marília, 2013.

BELLINI, C. G. P. et al. *Soft systems methodology*: uma aplicação no “Pão dos Pobres” de Porto Alegre. **ERA-Eletrônica**, São Paulo, v. 3, n. 1, art. 3, jan./ jun. 2004.

BOCCATO, Vera Regina Casari; FERREIRA, Estela Maris. Estudo comparativo entre o grupo focal e o protocolo verbal em grupo no aprimoramento de vocabulário controlado em fisioterapia: uma proposta metodológica qualitativa-cognitiva. **InCID: R. Ci. Inf. E Doc.**, Ribeirão Preto, v. 5, n. 1, p. 47-68, mar./ ago. 2014.

BOCCATO, Vera Regina Casari; FUJITA, M. S. L. O uso de linguagem documentária em catálogos coletivos de bibliotecas universitárias: um estudo de avaliação sociocognitiva com protocolo verbal. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 15, n. 3, p. 23-51, set./ dez. 2010.

CAMPOS, Claudinei José Gomes. Método de Análise de Conteúdo: ferramenta para a análise de dados qualitativos no campo da saúde. **Rev. Brasil. De Enfermagem**, Brasília, v. 57, n. 5, p. 611-614, set./ out. 2004.

CHARMAZ, Kathy. **A construção da teoria fundamentada**: guia prático para análise qualitativa. Porto Alegre: Artmed, 2009.

CHECKLAND, P. **Systems thinking, systems practice**. Chichester: Wiley, 1981.

CORDEIRO, Pedro Aurelio Cerveira. **Análise de Discurso e Ciência da Informação**: ensaio sobre uma possibilidade metodológica. Dissertação (Mestrado) – Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2004.

COSTA, Maira Murrieta. As bibliotecas brasileiras em 2018: resultados da técnica de Delfos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 17, n. 1, p. 74-93, jan./ mar. 2012.

COSTA, Sely M. S. Metodologia de sistemas flexíveis aplicada a estudos em Ciência da Informação: uma experiência pedagógica. **Transinformação**, Campinas, v. 15, n. 2, p. 259-271, maio/ ago. 2003.

CRESWELL, J. W. **Projeto de pesquisa**: métodos qualitativo, quantitativo e misto. 3. ed. Porto Alegre: Artmed, 2010.

CUNHA, Murilo Bastos da. A técnica de Delfos e a pesquisa em Biblioteconomia. **R. Esc. Bibliotecon. UFMG**, Belo Horizonte, v. 13, n. 2, p. 196-206, set. 1984.

DIAS, C. A. Grupo Focal: técnica de coleta de dados em pesquisas qualitativas. **Informação & Sociedade: Estudos**, João Pessoa, v. 10, n. 2, 2000. Disponível em: <<http://periodicos.ufpb.br/ojs/index.php/ies/article/view/330/252>>. Acesso em: 25 fev. 2015.

DRIESSNACK, Martha; SOUSA, Valmi D.; MENDES, Isabel Amélia Costa. Revisão dos desenhos de pesquisa relevantes para enfermagem: parte 2: desenhos de pesquisa qualitativa. **Revista Latino-Americana Enfermagem**, Ribeirão Preto, v. 15, n. 4, jul./ ago. 2007. Disponível em: <<https://www.revistas.usp.br/rlae/article/view/16145>>. Acesso em: 22 jul. 2021.

Semestre

7

DUARTE, R. Entrevistas em pesquisas qualitativas. **Educar**, Curitiba, n. 24, p. 213-225, 2004. Disponível em: <<http://ojs.c3sl.ufpr.br/ojs/index.php/educar/article/view/2216/1859>>. Acesso em: 25 fev. 2015.

DUARTE, Sebastião Junior Henrique; MAMEDE, Marli Villela; ANDRADE, Sônia Maria Oliveira de. Opções teórico-metodológicas em pesquisas qualitativas: representações sociais e discurso do sujeito coletivo. **Saúde Soc.**, São Paulo, v. 18, n. 4, p. 620-626, 2009.

ERICSON, K. A.; SIMON, H. A. Verbal reports on thinking. In: FAERCH, C.; KAPER, G. (Ed.). **Introspection in second language research**. Clevedon: Multilingual Matters, 1987. p. 24-53.

FLICK, Uwe. **Introdução à pesquisa qualitativa**. 3. ed. Porto Alegre: ArtMed/ Bookman, 2009.

FREITAS, Lúcia Silva de. A análise do discurso e o campo informacional: usos atuais e alcance epistemológico: uma atualização. **InCID: R. Ci.Inf. e Doc.**, Ribeirão Preto, v. 1, n. 1, p. 32-55, 2010.

FUJITA, Mariângela Spotti Lopes. A leitura do indexador: estudo de observação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 4, n. 1, p. 101-116, jan./ jun., 1999.

FUJITA, Mariângela Spotti Lopes. **A leitura documentária do indexador**: aspectos cognitivos e linguísticos influentes na formação do leitor profissional. 2003. 210 f. Tese (Livre-Docência em Análise Documentária e Linguagens Documentárias alfabéticas) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2003.

FUJITA, Mariângela Spotti Lopes. **A indexação de livros**: a percepção de catalogadores e usuários de bibliotecas universitárias: um estudo de observação do contexto sociocognitivo com protocolos verbais. São Paulo: Cultura Acadêmica (UNESP), 2009.

FUJITA, Mariângela Spotti Lopes; NARDI, M. I. A.; FAGUNDES, S. A. A observação da leitura documentária por meio de protocolo verbal. In: RODRIGUES, G. M.; LOPES, I. L. (Org.). **Organização e representação do conhecimento na perspectiva da ciência da informação**. Brasília: Thesaurus, 2003. p. 141-178. (Estudos avançados em Ciência da Informação, v. 2).

FURNIVAL, Aridne Chloë. A participação dos usuários no desenvolvimento de sistemas de informação. **Ciência da Informação**, Brasília, v. 25, n. 2, 1995.

GASPAR, Nádea Regina; REIS, Livia de Lima. Um olhar da análise do discurso para a representação temática na Ciência da Informação. **DataGramaZero**, Rio de Janeiro, v. 11, n. 6, dez. 2010.

GASQUE, Kelley Cristine G.D. Teoria Fundamentada: nova perspectiva à pesquisa exploratória. In: MUELLER, Suzana Pinheiro Machado (Org.). **Métodos para a pesquisa em Ciência da Informação**. Brasília: Thesaurus, 2007. p. 83-118.

GIBBS, G. **Análise de dados qualitativos**. Porto Alegre: Artmed/Bookman, 2009.

GRAY, D. **Pesquisa no mundo real**. Porto Alegre: Penso, 2012.

KAIRALLA, Anna Sylvia Silveira. Técnica Delphi para análise de um sistema de informação: estudo de viabilidade. **Ci. Inf.**, Brasília, v. 13, n. 1, p. 11-23, jan./jun. 1984.

KOZINETS, R. **Netnografia**. Porto Alegre: Penso, 2014.

LEFEVRE, Fernando. **Discurso do sujeito coletivo**. São Paulo: IPDSC, 2003. Disponível em: <http://www.fsp.usp.br/quali-saude/Discurso_sujeito_coletivo.htm>. Acesso em: 4 jul. 2015.

LEFEVRE, F; LEFEVRE, Ana Maria Cavalcanti. **O discurso do sujeito coletivo**: um novo enfoque em pesquisa qualitativa. Caxias do Sul: Educs, 2003.

LEFEVRE, F; LEFEVRE, Ana Maria Cavalcanti. **O sujeito coletivo que fala**. **Interface**: Comunic., Saúde, Educ., Botucatu, v. 10, n. 20, p. 517-524, jul./dez. 2006.

LONGARAY, A. A.; ENSSLIN, Leonardo; MACKNESS, John Robert. Uma proposta de integração da *Soft Systems Methodology* à MCDA-C. **Revista Eletrônica Pesquisa Operacional para o Desenvolvimento**, São Carlos, v. 5, n. 3, p. 331-372, set./dez. 2013.

MENDONÇA, Ana Valéria Machado. O uso da análise do discurso do sujeito coletivo em Ciência da Informação. In: MUELLER, Suzana P. M. **Métodos para a pesquisa em Ciência da Informação**. Brasília: Thesaurus, 2007.

MORAES, Cássia Regina Bassan de. **Gestão do conhecimento nas organizações**: modelo conceitual centrado na cultura organizacional e nas pessoas. 2010. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

MORAES, Cássia Regina Bassan de; Fadel, Bárbara. Gestão do conhecimento nas organizações: perspectivas de uso da metodologia sistêmica soft. In: VALENTIM, Marta (Org.). **Gestão, mediação e uso da informação**. São Paulo: Ed. UNESP, 2010.

MUELLER, S. P. M. (Org.). **Métodos para pesquisa em Ciências da Informação**. Brasília: Thesaurus, 2007.

OLIVEIRA, Joelma de Souza Passos et al. **Introdução ao Método Delphi**. Curitiba: Mundo Material, 2008.

ORLANDI, Eni. **Análise de discurso, princípios e procedimentos**. Campinas: Pontes, 2001.

ORLANDI, Eni. **Introdução às Ciências da Linguagem**: discurso e textualidade. Campinas: Pontes, 2006.



ORLANDI, Eni. Michel Pêcheux e a Escola Brasileira de Análise de Discurso. In: SEMINÁRIO DE ESTUDOS EM ANÁLISE DO DISCURSO, 1., 2003, Porto Alegre. **Anais...** Porto Alegre: Ed. UFRGS, 2003. Disponível em: <<http://www.ufrgs.br/analisedodiscurso/anaisdosead/1SEAD/Conferencias/EniOrlandi.pdf>>. Acesso em: 6 jul. 2015.

PAVÃO, Caterina Marta Groposo. **Comportamento de busca e recuperação da informação em serviços de descoberta em rede no contexto acadêmico**. 2014. Tese (Doutorado) – Universidade Federal do Rio Grande do Sul, Porto Alegre, 2014.

PINTO, Cândida Martins. A teoria fundamentada como método de pesquisa. In: SEMINÁRIO INTERNACIONAL DE LETRAS, 12., 2012., Santa Maria. **Anais...** Santa Maria: Ed. UNIFRA, 2012.

PIZARRO, Daniela Camara. Ética Profissional do bibliotecário atuante no segmento empresarial de Santa Catarina. **Liinc em Revista**, Rio de Janeiro, v. 7, n. 1, p. 231-247, mar. 2011. Disponível em: <<http://www.ibict.br/liinc>>. Acesso em: 4 jul. 2015.

RUBI, M. P. **Política de indexação para construção de catálogos coletivos em bibliotecas universitárias**. 2008. 196 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, UNESP, Marília, 2008.

SAMPIERI R. H.; COLLADO, C. F.; LUCIO, M. P. B. **Metodologia da pesquisa**. 5. ed. Porto Alegre: McGraw-Hill/ Penso, 2013.

SANTOS, Luciana Mara dos. **O discurso dos bibliotecários com relação às contribuições da sua formação para o desenvolvimento de suas atividades em empresas de tecnologia**. 2014. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Centro de Ciências da Educação; Universidade Federal de Santa Catarina, Florianópolis, 2014.

SOUZA, C. L. M. V. A problemática dos métodos quantitativos e qualitativos em biblioteconomia e documentação: uma revisão de literatura. **Ciência da Informação**, Brasília, v. 8, n. 12, p. 174-182, 1989. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1381/1007>>. Acesso em: 25 fev. 2015.

TAYLOR, S. J.; BOGDAN, R. **Introduction to qualitative research methods: a guidebook and resource**. New York: John Wiley & Sons, 1997. p. 24-43.

TEZZA, Rafael; BONIA, Antonio Cezar. O idoso e a internet: uma etnografia sobre interação e aprendizagem. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 15, n. 1, p. 185-197, jan./abr. 2010.

TRIPP, D. Pesquisa-ação: uma introdução metodológica. **Educação e Pesquisa**, São Paulo, v. 31, n. 3, p. 443-466, set./dez. 2005.

VALENTIM, M. L. P. **Métodos qualitativos de pesquisa em Ciência da Informação**. São Paulo: Polis, 2005.

VERGARA, Sylvia Constant. **Métodos de pesquisa em Administração**. 5. ed. São Paulo: Atlas, 2012. 159 p.

VIRGILLITO, Salvatore Benito (Org.). **Pesquisa de marketing**: uma abordagem quantitativa e qualitativa. São Paulo: Saraiva, 2010.

YIN, R. K. **Estudo de caso**: planejamento e métodos. 4. ed. Porto Alegre: Bookman, 2010.



UNIDADE 5

MÉTODOS MISTOS



5.1 OBJETIVO GERAL

Compreender as características dos métodos mistos: seus fundamentos, processos, vantagens e dificuldades.

5.2 OBJETIVOS ESPECÍFICOS

Esperamos que, ao final desta Unidade, você seja capaz de:

- a) descrever os fundamentos dos métodos mistos de pesquisa;
 - b) descrever os diferentes sequenciamentos dos métodos mistos de pesquisa;
 - c) descrever os principais desenhos de pesquisa mista.
- 

5.3 MÉTODOS MISTOS DE PESQUISA

Durante muito tempo, existiu uma oposição, nas Ciências Sociais, entre as abordagens quantitativas e as qualitativas, por vezes com posições bastante aguerridas entre os seguidores de cada escola. Existia o entendimento de que um pesquisador adotava ou abordagens quantitativas, ou qualitativas, mas nunca ambas. Ao longo do tempo, com a contribuição de pesquisadores de várias áreas (Educação, Ciências Sociais, Marketing, Administração etc.), essa visão vem sendo modificada. Em muitos problemas de pesquisa, as abordagens quantitativas e qualitativas não são opostas, na verdade, elas se *complementam* e possibilitam um aprofundamento e uma compreensão maior da questão investigada.

Desenhos de pesquisa mistos utilizam, ao menos, métodos quantitativos e qualitativos. Como os métodos mistos ainda são relativamente recentes, existe polêmica e controvérsia a respeito desse novo paradigma metodológico, seja nos seus desenhos, seja em suas análises. *Johnson et al.* apud *Grey* (2012, p. 166) escrevem:

Um estudo das definições oferecidas por 19 pesquisadores que usam métodos mistos concluiu que três definições consideravam que a mistura ocorria na etapa de coleta de dados, duas sugeriam que a mistura acontecia tanto na etapa de coleta quanto na análise de dados, enquanto quatro delas pressupunham que a mistura ocorria em todas as etapas do processo de pesquisa.

Do ponto de vista filosófico, considera-se que a pesquisa de métodos mistos segue um sistema pragmático. De acordo com *Sampieri, Collado, Lucio* (2013) e *Creswell* (2010), os métodos mistos recebem vários nomes: pesquisa integrativa, multimétodos, métodos múltiplos, estudos de triangulação ou pesquisa mista.

A combinação e o peso dos aspectos quantitativo e qualitativo podem variar. Para alguns autores da literatura de métodos mistos (*JOHNSON et al.; SAMPIERI; COLLADO; LUCIO, 2013*), a pesquisa de métodos mistos é vista como um contínuo no qual é possível mesclar os enfoques quantitativo e o qualitativo, centrando-se mais em um ou dando a eles o mesmo peso. Veja Figura 35, a seguir:

Figura 35 – *Continuum no desenho de pesquisa*

Q U A N	Q U A N – qual qual – Q U A N	Q U A N – Q U A L Q U A L – Q U A N	Q U A L – quan quan – Q U A L	Q U A L
------------------	----------------------------------	--	----------------------------------	------------------

Fonte: *Tréz* (2012, p. 1138).

Na Figura 35, segundo *Teddie e Tashakkori* (2009) apud *Tréz* (2012), temos uma ilustração do *continuum* da abordagem mista.



5.4 A SEQUÊNCIA DOS MÉTODOS NAS PESQUISAS MISTAS

Os métodos mistos podem ter seus componentes executados no tempo de forma sequencial ou concomitante. No caso sequencial, podemos aplicar um método quantitativo e, em seguida, um qualitativo ou, inversamente, aplicamos primeiro um método quantitativo, seguido de um qualitativo.

Já em uma execução *concomitante*, aplicamos um método quantitativo e um qualitativo simultaneamente. Nesse caso, as coletas e a análise de dados são efetuadas basicamente ao mesmo tempo. Nele, a análise qualitativa dos dados não interfere na análise quantitativa, nem o contrário ocorre. Só depois de concluída a interpretação dos dados, em cada um dos métodos, de forma separada, é que procedemos a uma conclusão, unindo as duas abordagens. Assim, coleta e análise de dados são efetuadas sem que uma influencie a outra. Apenas ao final é realizada uma síntese, a partir das conclusões de cada uma das abordagens.

No caso de uma execução em *sequência*, em um primeiro momento ocorre a coleta e a análise qualitativa de dados, ou a coleta e a análise quantitativa de dados. De forma geral, quando a etapa qualitativa é a primeira a ser executada, pretende-se compreender melhor o problema, com uma amostra inicial menor, para, na etapa quantitativa, o estudo ser realizado com uma amostra maior, que permita a generalização para a população (SAMPIERI; COLLADO; LUCIO, 2013).

Em muitas pesquisas mistas, pretendemos que a análise dos dados seja efetuada por sua “transformação”: dados qualitativos em dados quantitativos, ou o contrário. Nesse caso, temos a pesquisa mista, cujo desenho é denominado de transformação.

5.5 DESENHOS DE PESQUISA MISTA

Neste ponto, vamos seguir a ideia de desenhos de pesquisa mista proposta por Sampieri, Collado e Lucio (2013), respeitando a notação e a tipologia dos autores. Se você consultar outros textos da literatura sobre métodos mistos de pesquisa, por exemplo, Creswell (2010), ou ainda Johnson e Onwuegbuzie (2004), encontrará outros desenhos. No caso dos últimos, é apresentada uma matriz com nove possibilidades de desenhos de pesquisa mista, combinando *status* dominante e mesmo *status* das abordagens qualitativa e quantitativa, com as características de estratégias simultâneas ou sequenciais.

Vejamos uma pequena lista de notações:

- a) é costume abreviar-se o método: “quan” para quantitativo e “qual” para qualitativo;
- b) o símbolo de adição “+” é utilizado para significar uma coleta ou análise de dados que seja concomitante;
- c) o símbolo “→” denota uma coleta ou análise sequencial de dados;
- d) o símbolo “O” nos diz que o desenho pode ser realizado com dois formatos;
- e) nas situações em que há maior peso para um dos métodos, este é escrito em caixa-alta. A representação QUAN/ qual significa que o método qualitativo está inserido (embutido) no método quantitativo.

Na estratégia sequencial, o pesquisador aprofunda o conhecimento obtido na investigação dos dados com uma abordagem, após a outra ter sido realizada. Assim, o investigador pode iniciar sua pesquisa com uma abordagem quantitativa, chegar a algumas conclusões e, em seguida, aprofundá-las com uma abordagem qualitativa. Teríamos o esquema:

QUAN → QUAL: ambas em caixa-alta, significando que o peso adotado para as duas abordagens é o mesmo.

Na estratégia simultânea, o pesquisador mistura uma abordagem quantitativa com uma qualitativa (os pesos podem ser iguais ou não, é uma escolha do pesquisado). Se a estratégia for a de uma pesquisa qualitativa com maior peso que uma pesquisa quantitativa, teremos o esquema: QUAL + quan.

Usamos, ainda, o termo conversão (ALISE; TELIDDLIE, 2010 apud TRÉZ, 2012) para os desenhos em paralelo, nos quais um dado é transformado, podendo ser quantificado ou qualificado, também sendo analisado qualitativa e quantitativamente.

Quando devemos usar métodos mistos de pesquisa? *Creswell* (2010) aponta as seguintes situações:

- a) quando as abordagens (qualitativa e quantitativa), juntas, fornecem uma melhor compreensão do problema de pesquisa, em vez de serem trabalhadas isoladamente;
- b) quando uma das abordagens, tomada isoladamente, não consegue responder ao problema;
- c) quando é necessário incorporar aspectos qualitativos a uma investigação quantitativa;
- d) quando, a partir de um estudo quantitativo, se prossegue com um qualitativo para melhor entendimento do problema de pesquisa.

Sampieri, Collado e Lucio (2013) apresentam os seguintes desenhos:

- a) desenho exploratório sequencial;
- b) desenho explicativo sequencial;
- c) desenho transformador sequencial;
- d) desenho de triangulação concomitante;
- e) desenho incrustado concomitante com o modelo dominante;

- f) desenho incrustado concomitante de vários níveis;
- g) desenho transformador concomitante e desenho de integração múltipla.

5.5.1 Desenho exploratório sequencial

Este tipo de desenho ocorre na primeira fase de coleta e análise de dados qualitativos. Na sequência, vem a coleta e a análise de dados quantitativos. Considera-se esse tipo de desenho como relativamente fácil de ser executado. Uma desvantagem é, justamente, o fato de ele ser sequencial: temos que aguardar a conclusão da primeira etapa antes de começarmos a segunda.

A ilustração desse procedimento está na Figura 36:

Figura 36 – Esquema do desenho exploratório sequencial



Fonte: Sampieri; Collado; Lucio (2013, p. 567).

5.5.2 Desenho explicativo sequencial

Neste caso, a coleta é efetuada em primeiro lugar e, em seguida, vem a análise quantitativa de dados. Pode-se dar mais ênfase ao lado quantitativo, ao qualitativo ou apresentar pesos iguais. Sampieri, Collado e Lucio (2013) explicam que esse tipo de desenho é utilizado quando o estudo quantitativo apresenta resultados “inesperados ou confusos”. A parte qualitativa será utilizada para auxiliar na interpretação e explicação das descobertas quantitativas iniciais, assim como para aprofundá-las. O desenho aparece representado a seguir, na Figura 37:

Figura 37 – Esquema do desenho explicativo sequencial



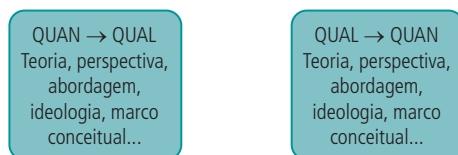
Fonte: Sampieri; Collado; Lucio (2013, p. 569).

5.5.3 Desenho transformador sequencial

Neste desenho, diferentemente dos casos sequenciais anteriores, há uma teoria que orienta o estudo. Como Sampieri, Collado e Lucio (2013, p. 572) apresentam, podem ser “feminismo, ação participativa, o enfoque das inteligências múltiplas etc.”. Esse guia teórico é mais importante para o estudo que o método”. É ele que vai orientar o pesquisador na direção a ser seguida.

A ilustração desse desenho está na Figura 38:

Figura 38 – Esquema do desenho transformador sequencial

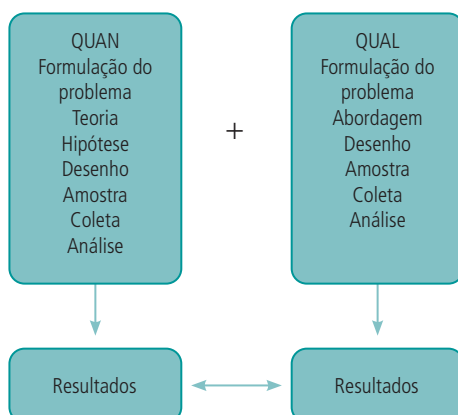


Fonte: Sampieri; Collado; Lucio (2013, p. 573)

5.5.4 Desenho de triangulação concomitante

Este tipo de desenho é utilizado quando se pretende efetuar uma validação denominada “cruzada” entre os dados qualitativos e quantitativos. É recomendável para situações de pesquisa nas quais se pretende efetuar a confirmação de resultados. Nele, faz-se a coleta e a análise dos dados, tanto qualitativos quanto quantitativos, em paralelo (concomitantemente). É um desenho adotado com bastante frequência (SAMPIERI; COLLADO; LUCIO, 2013). A seguir, na Figura 39, temos sua representação:

Figura 39 – Esquema do desenho de triangulação concomitante

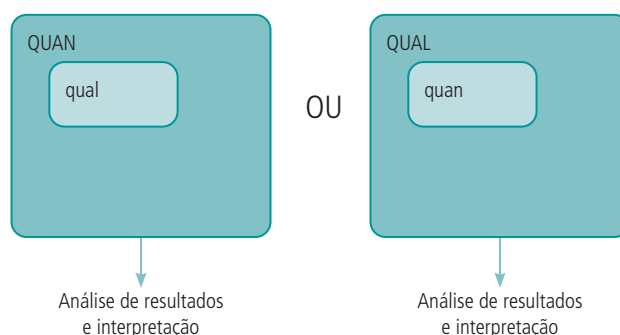


Fonte: Sampieri; Collado; Lucio (2013, p. 573).

5.5.5 Desenho incrustado concomitante de modelo dominante

Por ser concomitante, este desenho coleta os dados em paralelo (quantitativos e qualitativos). Nesse caso, há um método que é dominante, podendo ser um ou outro. O método subordinado é inserido como uma “caixa” menor na “caixa” do método dominante. O subordinado responde a questões distintas daquelas do método dominante. A Figura 40, a seguir, ilustra o método:

Figura 40 – Esquema de desenhos incrustados concomitantes de modelo dominante

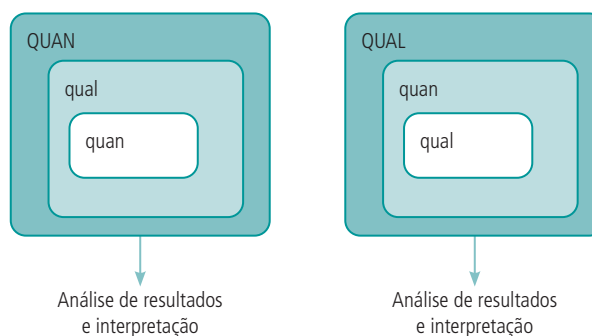


Fonte: Sampieri; Collado; Lucio (2013, p. 575).

5.5.6 Desenho incrustado concomitante de vários níveis

Este é um desenho no qual podemos coletar os dados (qualitativos ou quantitativos) em mais de um nível. Este desenho abre a possibilidade de que dados sejam acessados em mais de um grupo ou níveis de análise, como mostra a Figura 41:

Figura 41 – Esquema do desenho incrustado concomitante de vários níveis



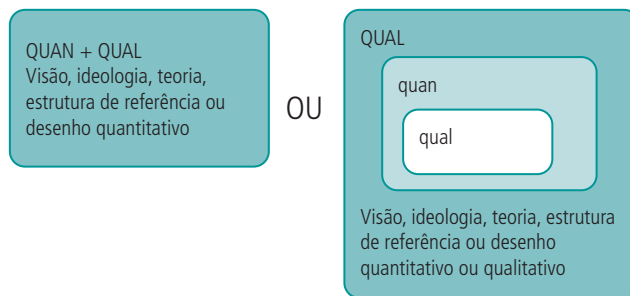
Fonte: Sampieri; Collado; Lucio (2013, p. 579).

5.5.7 Desenho transformador concomitante

O conjunto de desenhos anteriores é considerado e combinado para constituirmos o presente desenho. O concomitante na denominação do desenho está associado ao fato de coletarmos dados quantitativos ou qualitativos em paralelo. O pesquisador escolhe se deseja atribuir maior ênfase à abordagem qualitativa ou quantitativa, ou se prefere atribuir pesos idênticos a elas.

De modo análogo ao desenho transformador sequencial, aqui, a coleta e a análise de dados também são orientadas por teorias, ou ideologias, na visão de Sampieri, Collado e Lucio (2013). O esquema a seguir (Figura 42) é devido aos autores:

Figura 42 – Esquema de desenho transformador concomitante



Fonte: Sampieri; Collado; Lucio (2013, p. 580).



5.5.8 Atividade

Esta atividade atende ao objetivo a: “descrever os fundamentos dos métodos mistos de pesquisa”.

Existem problemas de pesquisa nos quais adotar métodos qualitativos é mais adequado: o pesquisador pode expor a fala e as narrativas do ponto de vista dos sujeitos de pesquisa, ou, ainda, ele pode estar interessado nas diferentes compreensões e na subjetividade dos sujeitos de pesquisa. Em outros casos, uma pesquisa quantitativa é mais adequada: o interesse do pesquisador está no entendimento das relações entre variáveis ou em saber se um grupo ou determinado tratamento produz resultados médios distintos dos de outro grupo.

E agora? Em quais problemas de pesquisa se justifica a adoção de métodos de pesquisa mistos?

Reflita, consulte livros e procure artigos na internet que tenham adotado métodos de pesquisa mistos antes de ler a resposta comentada, a seguir.

Resposta comentada

Creswell e Clark (2013, p. 24) listam indicações para a utilização de métodos de pesquisa mistos. Aqui, apresentamos três delas. Consulte o texto completo para conhecer outras indicações.

- Apenas uma fonte de dados pode não ser o bastante para a sua pesquisa: enquanto pesquisas qualitativas apresentam dificuldade para generalizar os resultados para toda a população, pesquisas quantitativas, de uma forma geral, não têm como característica o aprofundamento das perspectivas e compreensões dos sujeitos de pesquisa. Assim, a adoção de métodos mistos pode permitir um entendimento mais profundo do problema de pesquisa. Novamente recomendamos a leitura de *Leitão* (2005), para que você conheça uma situação de pesquisa na qual o uso exclusivo de dados quantitativos não “conta toda a história”.

Semestre

7

- É necessário explicar os primeiros resultados obtidos na pesquisa: suponha que você tenha realizado uma pesquisa quantitativa. Produziu os questionários, coletou os dados, construiu as tabelas de frequência e gráficos estatísticos. Mas você sente que seu estudo está incompleto. Na verdade, falta um entendimento e explicações mais profundas. É a situação típica do desenho misto sequencial: quantitativa-qualitativa. O pesquisador obtém, a partir dos dados quantitativos, uma primeira leitura do problema de pesquisa. Na etapa qualitativa, por meio de entrevistas, grupos focais, histórias de vida ou narrativa, realizará a busca pelo entendimento que os testes estatísticos não produziram.
- O pesquisador precisa generalizar os resultados qualitativos para a população como um todo: é uma situação inversa da exposta no item anterior, que indica que você precisa generalizar os resultados de uma pesquisa qualitativa para a população como um todo. Agora, temos o caso típico do desenho sequencial qualitativo-quantitativo. Neste caso, o pesquisador tem dificuldades para definir as variáveis que serão mensuradas, por exemplo. Assim, ele explora qualitativamente em uma primeira etapa, para investigar variáveis e questões, para, após isso, desenvolver a etapa quantitativa e buscar estabelecer relações entre as variáveis definidas na etapa anterior.



Multimídia

Consulte o *link* a seguir. Nele, você poderá assistir a uma palestra apresentada pela professora *Maria Cristiane Barbosa Galvão*, da USP de Ribeirão Preto, cujo título é: *Métodos de pesquisa mistos e revisões de literatura mistas: uma janela de oportunidades científicas*. Disponível em: <<http://iearp.blogspot.com.br/2013/02/metodos-de-pesquisa-mistos-e-revisoes.html>>. Acesso em: 22 jul. 2021.

5.6 CONCLUSÃO

Destacamos como o talvez principal aspecto sobre os métodos de pesquisa mistos a afirmação de *Sampieri, Collado e Lucio* (2013, p. 548):

A meta da pesquisa mista não é substituir a pesquisa quantitativa nem a pesquisa qualitativa, mas utilizar os pontos fortes de ambos os tipos combinando-os e tentando minimizar seus potenciais pontos fracos.

Nas últimas décadas, os métodos de pesquisa mistos vêm ganhando cada vez mais adeptos. Cada vez menos se acredita em uma dicotomia simplista (e muitas vezes empobrecedora) opondo quantitativo OU qualitativo.

Os métodos de pesquisa mistos podem apresentar uma preponderância quantitativa, uma preponderância qualitativa, ou mesmo um equilíbrio entre ambas as abordagens. Também podem apresentar desenhos concomitantes, sequenciais, incrustados, transformadores ou combinações deles.

RESUMO

Nesta Unidade, você teve contato com a pesquisa de métodos mistos: uma metodologia relativamente nova. Você foi apresentado a alguns desenhos utilizados nas pesquisas com métodos mistos. Vimos os desenhos seguintes, referentes a métodos de pesquisa mistos:

- a) desenho exploratório sequencial;
- b) desenho explicativo sequencial;
- c) desenho transformador sequencial;
- d) desenho de triangulação concomitante;
- e) desenho incrustado concomitante de modelo dominante;
- f) desenho incrustado concomitante de vários níveis;
- g) desenho transformador concomitante.

Ao redigir sua pesquisa de métodos mistos, é recomendável que você apresente as características principais dessa metodologia, bem como um resumo histórico da evolução dela. Isso deve ser feito para que o leitor do seu trabalho consiga se situar frente a essa nova metodologia.

Você deve indicar se seu desenho de pesquisa é sequencial ou paralelo, bem como os outros aspectos específicos do desenho que você adotou.



REFERÊNCIAS

CRESWELL, J. W. **Projeto de pesquisa**: métodos qualitativo, quantitativo e misto. 3. ed. Porto Alegre: Artmed, 2010.

CRESWELL, J. W.; CLARK, V. L. **Plano de pesquisa de métodos mistos**. Porto Alegre: Penso, 2013.

SAMPIERI R. H.; COLLADO, C. F.; LUCIO, M. P. B. **Metodologia da pesquisa**. 5. ed. Porto Alegre: McGraw-Hill/ Penso, 2013.

TRÉZ, T. A. Caracterizando o método misto de pesquisa na educação: um *continuum* entre a abordagem qualitativa e quantitativa. **Atos de Pesquisa em Educação**, Blumenau, v. 7, n. 4, p. 1132-1157, 2012.

CONCLUSÃO DA DISCIPLINA

O método, em uma pesquisa científica, é o seu coração, ou seja, sua parte vital. Se ele não estiver adequado ou não for desenvolvido com o rigor necessário, os resultados estarão totalmente comprometidos. Não importa se você vai escolher uma abordagem quantitativa, qualitativa ou mista para desenvolver a sua pesquisa. O que vale é você definir seu método e segui-lo com muito rigor. Depois, quando for redigir o seu texto, detalhe essa parte do trabalho com o mesmo rigor com que o desenvolveu. Com esses cuidados, sua pesquisa terá a qualidade buscada nos estudos científicos atualmente.

BIBLIOGRAFIA DA DISCIPLINA

ALMEIDA JÚNIOR, O. F. Sobre os métodos e as técnicas de pesquisa: reflexões. In: VALENTIM, M. L. P. (Org.). **Métodos qualitativos de pesquisa em ciência da informação**. São Paulo: Polis, 2005. p. 161-171.

ARAÚJO, C. A. Á. Um mapa dos estudos de usuários da informação no Brasil. **Em Questão**, Porto Alegre, v. 15, n. 1, p. 11-26, 2009.

CATIVELLI, Adriana Stefani. Marketing Aplicado em bibliotecas: análise de conteúdo dos artigos publicados em periódicos da Ciência da Informação. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 11, n. 3, p. 30-45, set./ dez. 2013.

COSTA, F. J. **Mensuração e desenvolvimento de escalas:** aplicações em Administração. Rio de Janeiro: Ciência Moderna, 2011.

CUNHA, Murilo Bastos da. A técnica de Delfos e a pesquisa em Biblioteconomia. **R. Esc. Bibliotecon. UFMG**, Belo Horizonte, v. 13, n. 2, p. 196-206, set. 1984.

FREITAS, F. H. A.; FREIRE J. R. O. Sobre o uso da *Web of Science*: fonte para a história da Ciência. **Revista da SBHC**, Rio de Janeiro, v. 2, n. 1, p. 129-147, 2003.

LENNART, B.; INGWERSEN, P. Toward a basic framework for webometrics. **Journal of the American Society for Information Science and Technology**, [S.l.], v. 55, n. 44, p. 1216-1227, 2004.

MAGALHÃES, M. N.; LIMA, A. C. P. **Noções de probabilidade e estatística**. 3. ed. São Paulo: Instituto de Matemática e Estatística da Universidade de São Paulo, 2001.

MATTOS, A. M. Desenvolvimento de coleções em bibliotecas universitárias: uma abordagem quantitativa. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 14, n. 3, p. 38-60, 2009.

MEIRELES, Magali Rezende Gouvêa; CENDÓN; Beatriz Valadares. Aplicação Prática dos processos de Análise de Conteúdo e de análise de citações em artigos relacionados às redes neurais artificiais. **Inf. Inf.**, Londrina, v. 15, n. 2, p. 77-93, jul./ dez. 2010.

PAULISTA, Paulo Henrique. **Desenvolvimento de software para apoio à realização de auditoria interna de sistema de gestão da qualidade**. 2009. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Itajubá, Itajubá, 2009.

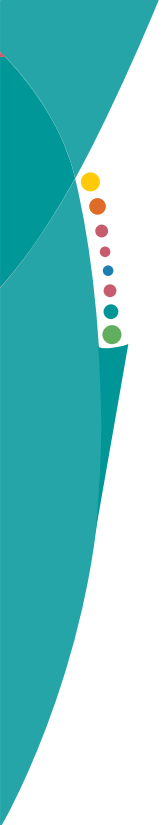
ROZADOS, H. B. F. **Impactos da intranet em unidades de informação de ensino superior da grande Porto Alegre**. 2000. Dissertação (Mestrado em Comunicação e Informação) – Faculdade de Biblioteconomia e Comunicação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2000.

SHARPE, N.; DE VEAUX, R.; VELLEMAN, R. **Estatística aplicada:** Administração, Economia e Negócios. Porto Alegre: Bookman, 2011.

SOUZA, C. L. M. V. A problemática dos métodos quantitativos e qualitativos em biblioteconomia e documentação: uma revisão de literatura. **Ciência da Informação**, Brasília, v. 8, n. 12, p. 174-182, 1989. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1381/1007>>. Acesso em: 25 fev. 2015.

VANZ, S. A. S.; STUMP, I. R. C. Procedimentos e ferramentas aplicados aos estudos bibliométricos. **Informação & Sociedade**, João Pessoa,





v. 20, n. 2, p. 67-75, 2010. Disponível em: <<http://www.ies.ufpb.br/ojs/index.php/ies/article/view/4817/4358>>. Acesso em: 25 fev. 2015.

VOLPATO, G.; BARRETO, R. **Estatística sem dor**. Botucatu: Best Writing, 2011.

WORNELL, I. Informetria: explorando bases de dados como instrumentos de análise. **Ciência da Informação**, Brasília, v. 27, n. 2, p. 210-216, 1998. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/367/328>>. Acesso em: 21 fev. 2015.

VELASQUES, Matheus Trindade et al. Protocolos verbais na pesquisa qualitativa: um estudo de caso. In: CIC, 19., 2010, [S.l.]. **Anais...** [S.l.: s.n.], 2010.



UFRJ
100
ANOS
1920 | 2020
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

Faculdade de Administração
e Ciências Contábeis
Departamento
de Biblioteconomia



CAPES

MINISTÉRIO DA
EDUCAÇÃO



Agência Brasileira do ISBN

ISBN 978-85-85229-30-6



9 788585 229306

Agência Brasileira do ISBN

ISBN 978-85-85229-31-3



9 788585 229313