

EURO DE BARROS COUTO JUNIOR

ABORDAGEM NÃO-PARAMÉTRICA PARA  
CÁLCULO DO TAMANHO DA AMOSTRA COM  
BASE EM QUESTIONÁRIOS OU ESCALAS DE  
AVALIAÇÃO NA ÁREA DE SAÚDE

SÃO PAULO  
2009

EURO DE BARROS COUTO JUNIOR

ABORDAGEM NÃO-PARAMÉTRICA PARA  
CÁLCULO DO TAMANHO DA AMOSTRA COM  
BASE EM QUESTIONÁRIOS OU ESCALAS DE  
AVALIAÇÃO NA ÁREA DE SAÚDE

Tese apresentada à Faculdade de  
Medicina da Universidade de São  
Paulo para obtenção do título de  
Doutor em Ciências

Área de Concentração: Patologia

Orientador: Prof. Dr. Raymundo Soares  
de Azevedo Neto

SÃO PAULO  
2009

**Dados Internacionais de Catalogação na Publicação (CIP)**

Preparada pela Biblioteca da  
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Couto Junior, Euro de Barros

Abordagem não-paramétrica para cálculo do tamanho de amostra com base em questionários ou escalas de avaliação na área de saúde / Euro de Barros Couto Junior. -- São Paulo, 2009.

Tese(doutorado)--Faculdade de Medicina da Universidade de São Paulo.

Departamento de Patologia.

Área de concentração: Patologia.

Orientador: Raymundo Soares de Azevedo Neto.

Descritores: 1.Amostragem 2.Tamanho da amostra 3.Lógica 4.Matemática

5.Coleta de dados 6.Estatística não paramétrica 7.Saúde

USP/FM/SBD-214/09

## DEDICATÓRIA

*A meus filhos, pelo que eles representam para mim: puro orgulho de tê-los como eles são — sensíveis, bondosos, inocentes e inteligentes... hoje, sou o que sou, porque ambos são assim.*

*À minha esposa, que sempre me apoia, para que eu busque progresso a fim de almejar um grau maior de sabedoria.*

*A meu pai e à minha mãe, cujo amor e dedicação, incentivo e atenção me são impagáveis. Sou-lhes eternamente grato por todos os momentos vividos... todos...*

*A meu irmão, que, de modo paciente e solícito, sempre me ajuda, quando dele preciso.*

*A meus sogros, com quem convivo harmoniosamente, e que me viram atingir este momento.*

*A meus avós, in memoriam, pela convivência pacífica e tranquila. Lembro-me continuamente deles.*

## AGRADECIMENTOS

*Ao Prof. Dr. Raymundo Soares de Azevedo Neto, que me orientou neste trabalho, com plenitude e interesse, paciência e respeito, e por sua percepção de que esta proposta tinha algum valor.*

*Ao Prof. Dr. Eduardo Massad, que, pelo tempo em que estive como funcionário do LIM-01, apoiou-me e deu-me a segurança necessária ao desenvolvimento de meu trabalho junto a esse importante laboratório.*

*Aos Profs. Drs. Neli Regina Siqueira Ortega e Jair Minoro Abe, pelos inestimáveis ensinamentos das Lógicas Não-clássicas, que me foram, e por certo, ainda serão extremamente úteis.*

*Ao Prof. Dr. Francisco Antonio Bezerra Coutinho, pela oportunidade que tive em poder observar sua primorosa didática, e pelos ensinamentos tão essenciais recebidos.*

*Ao Prof. Dr. Clóvis de Araújo Perez, pela atenção dispensada a esta proposta, desde sua primeira apresentação pública, no SINAPE-18, em 30 de julho de 2008.*

*Ao Prof. Dr. Paulo Sérgio Panse Silveira, pelo coleguismo e atenção demonstrados durante minha permanência no LIM-01.*

*Ao Sr. James Cox, pela gentileza em ter me fornecido precioso material de estudo e pesquisa.*

*À Sra. Helena Setsuco Ishida Amano e à Sra. Dolores Maria dos Santos, pela compreensão que tiveram, ao permitir meu comissionamento junto à USP, sem o que não teria sido possível desenvolver este trabalho e chegar à sua conclusão.*

*A todos os colegas da PMSP, em especial do DRH, pela compreensão e atenção que tiveram, durante o tempo em que estive ausente, especialmente nas pessoas de Yolanda Muniz Paschoalini, Dalva Aparecida da Silva, Olga Coelho, Vera Lúcia Del Busso Forgioni, Celso Mendes dos Santos; Sidnei Teodoro, Ana Cristina Vieira da Silva, Hubert Richard Voss Giopato; Miriam Siqueira Merchenbacher, Maria Aparecida dos Santos, Silvana Maria de Faria Costa, Rosana Denise Valentim Ruiz; Marina Masae Oda, Cecília Maria de Souza Nascimento, Roberto Moreira Macedo, Lígia Ansaldi da Silva, Luíz Henrique Benigno da Silva, Maria Adelaide Guedes dos Santos, José Santana da Silva, Domingos das Neves Vieira, Marinete Feliciano da Silva Filha; e uma especial gratidão à Ana Lúcia da Conceição Romualdo, que, de modo primoroso, revisou este trabalho.*

*Ao pessoal da Pós-graduação da FM-USP, dos Serviços de Biblioteca da FM-USP e do IME-USP, e ao pessoal do LIM-01, meu agradecimento pela prestimosa ajuda.*

# SUMÁRIO

Resumo  
Summary

1	Introdução	1
1.1	Comentários sobre alguns conceitos usuais .....	8
1.2	Amostragem — conceito e algumas histórias .....	13
1.3	Importância da Amostragem .....	20
1.4	Revisão da literatura .....	24
1.4.1	Lógica Paraconsistente Anotada Evidencial .....	24
1.4.2	Combinatória .....	26
1.4.3	Tamanho da Amostra .....	32
2	Objetivos .....	72
3	Métodos .....	74
3.1	Lógica Paraconsistente Anotada Evidencial .....	74
3.2	Combinatória .....	78
3.3	Tamanho da Amostra .....	80
4	Proposta para o Novo Modelo e Resultados .....	82
4.1	Nova proposta para o cálculo do Tamanho da Amostra .....	82
4.2	Resultados — alguns cálculos demonstrativos .....	90
4.2.1	Cálculo do tamanho da amostra para o SF-36 .....	90
4.2.2	Cálculo do tamanho da amostra para o WHOQOL .....	93
4.2.3	Cálculo do tamanho da amostra para alguns instrumentos da área de Pneumologia .....	96
5	Discussão .....	100
6	Conclusão .....	115
7	Anexos .....	117
8	Referências .....	124

## RESUMO

Couto Junior EB. *Abordagem não-paramétrica para cálculo do tamanho da amostra com base em questionários ou escalas de avaliação na área de saúde* [tese]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2009. 130p.

Este texto sugere sobre como calcular um tamanho de amostra com base no uso de um instrumento de coleta de dados formado por itens categóricos. Os argumentos para esta sugestão estão embasados nas teorias da Combinatória e da Paraconsistência. O propósito é sugerir um procedimento de cálculo simples e prático para obter um tamanho de amostra aceitável para coletar informações, organizá-las e analisar dados de uma aplicação de um instrumento de coleta de dados médicos baseado, exclusivamente, em itens discretos (itens categóricos), ou seja, cada item do instrumento é considerado como uma variável não-paramétrica com um número finito de categorias. Na Área de Saúde, é muito comum usar instrumentos para levantamento com base nesse tipo de itens: protocolos clínicos, registros hospitalares, questionários, escalas e outras ferramentas para inquirição consideram uma sequência organizada de itens categóricos. Uma fórmula para o cálculo do tamanho da amostra foi proposta para tamanhos de população desconhecidos e um ajuste dessa fórmula foi proposto para populações de tamanho conhecido. Pôde-se verificar, com exemplos práticos, a possibilidade de uso de ambas as fórmulas, o que permitiu considerar a praticidade de uso nos casos em que se tem disponível pouca ou nenhuma informação sobre a população de onde a amostra será coletada.

Descritores: 1.Amostragem 2.Tamanho da amostra 3.Lógica  
4.Matemática 5.Coleta de dados 6.Estatística não paramétrica  
7.Saúde

## SUMMARY

Couto Junior EB. *Non-parametric approach for calculation of sample size based on questionnaires or scales of assessment in the health care* [thesis]. São Paulo: "Faculdade de Medicina, Universidade de São Paulo"; 2009. 130p.

This text suggests how to calculate a sample size based on the use of a data collection instrument consisting of categorical items. The arguments for this suggestion are based on theories of Combinatorics and Paraconsistency. The purpose is to suggest a practical and simple calculation procedure to obtain an acceptable sample size to collect information, organize it and analyze data from an application of an instrument for collecting medical data, based exclusively on discrete items (categorical items), i.e., each item of the instrument is considered as a non-parametric variable with finite number of categories. In the health care it is very common to use survey instruments on the basis of such items: clinical protocols, hospital registers, questionnaires, scales and other tools for hearing consider a sequence of items organized categorically. A formula for calculating the sample size was proposed for a population of unknown size, and an adjusted formula has been proposed for population of known size. It was seen, with practical examples, the possibility of using both formulas, allowing to consider the practicality of the use in cases that have little or no information available about the population from which the sample is collected.

Descriptors: 1.Sample studies 2.Sample size 3.Logic  
4.Mathematics 5.Data collection 6.Nonparametric statistics  
7.Health



## 1 Introdução

A História tem demonstrado que, em todas as eras, a Ciência apresenta algum desenvolvimento, que, por sua vez, surge como parte da Cultura Humana, ponto-chave para diferenciar o ser humano dos outros seres vivos.

E a Cultura produziu seus frutos, desde que o ser humano percebeu a si próprio. Os primórdios da Cultura já agregavam saberes que ultrapassavam os instintos. A alma humana foi sendo conformada por itens menos instintivos e mais conscientes, desembocando na construção do pensamento humano e afastando a ignorância. A Natureza proveu o ser humano dessa capacidade, que não era tão-somente a do ato de pensar, mas, sim, a da construção do pensamento. E assim, a raça humana distinguiu-se como tal, passando a dominar seu habitat e tornando-o fértil para o desenvolvimento de pensares mais complexos.

Daí, a observação dos fatos da Natureza levou o ser humano a propor seu registro, por meio de símbolos, e *símbolos vivem mais longamente que seres humanos* (Baitello Jr, 1997). O

registro, por sua vez, propiciou sua descrição; em sendo descritível, cada fato pôde, no geral, ser observado novamente, e essa sequência relativamente simplificada gerou o tênue fio sobre o qual surgiu a base da Ciência, que, segundo René Descartes (1596-1650) é, *em sua totalidade, a verdade e a cognição evidente* (Hutchins, 1971). A totalidade a que se refere não somente Descartes, mas, quase todos os grandes pensadores, diz respeito à unicidade da Ciência, como corpo uno e indivisível, no sentido lato de agregação do conhecimento aquinhoado pela Humanidade.

Assim sendo, em uma visão contemporânea, a integração das ciências, tratadas como frações de um todo é, sem dúvida, a própria natureza da máxima criação humana, que é o Conhecimento. A Ciência e a Arte, como tal se apresentam, são, sob esse aspecto, as faces desse conhecimento, pois consistem em manifestações puras e exclusivas do espírito humano, sob o mesmo fio condutor antes mencionado. Aqui, cabe frisar a capacidade humana de intelecção de si e de compreensão de sua própria existência. Por isso, são reconhecidos, como registros singelos da História, os seres humanos que foram (e são) capazes de realizar arte, tanto quanto ciência, pois elas são, a bem dizer, vertentes de um mesmo núcleo de conhecimento. Se, hoje, a visão que se tem apresenta-se como distorcida, é porque não se teve como deter o resgate dos fatos históricos de um passado glorioso

— e ao mesmo tempo, vil —, que a Humanidade projetou para si mesma no presente. Blaise Pascal (1623-1662), em seu tratado científico sobre o vácuo, já afirma que *a Geometria, a Aritmética, a Música, a Física, a Medicina, a Arquitetura e todas as ciências sujeitas à experimentação e à razão devem ser consideradas, pois tornaram-se perfeitas* (Hutchins, 1971(33):356).

A formal partição da Ciência, sob um certo ângulo, propiciou o aparecimento de especificidades de conhecimento, mas não a desuniu, e não a privou de unicidade: ela continua como una, e as variadas e diversas vertentes do Conhecimento, muito antes, propõem-se a ser um método de organização interna da Ciência, do que propriamente divisões dela. Logo, a Ciência, comprovadamente, continua a ser una, como sempre foi, e como sempre será.

Apesar de as especificidades mostrarem ter vindo para beneficiar o ser humano, a visão totalizadora sobre problemas conjuntos dessas especificidades apresenta-se, hoje, como necessária, inovadora e benéfica. Com isso, pode-se observar, com clareza, a existência de *elementos componentes da Ciência*, que são as conhecidas ciências, por suas naturezas específicas, mas, também, vê-se a integração desses elementos como copartícipes de uma integração sustentável e objetiva. Entre esses elementos

componentes da Ciência, vê-se a inegável presença da Estatística como ferramenta auxiliadora (sem deméritos!) de uma gama extensiva frente a outras ciências, e neste caso, uma ferramenta, já há muito, largamente aplicada na Área de Saúde (Biologia, Medicina, Enfermagem, Fisioterapia, Terapia Ocupacional, Fonoaudiologia, Psicologia, Educação Física, Veterinária etc.), que, por sua vez, formam um vasto e sólido bloco de matérias da própria Ciência.

Antes da abordagem científica na qual aportou a Estatística, após o Renascimento, sua manifestação, dentro das sociedades ditas civilizadas, dizia respeito à contagem bruta de fatos e ocorrências de cunho social e governamental: os dinheiros públicos, as doenças e suas incidências e prevalências, os fatos sociais marcantes (nascimentos, mortes, casamentos, concentrações populacionais e migrações etc.) eram a própria Estatística, mas estava-se longe de uma abordagem amostral, que só surgiu em meados do século XVII, com John Graunt (1620-1674), na Inglaterra.

A Estatística, hoje, pode ser conceituada como a *ciência da contagem*, sendo a Matemática, a ciência que operacionaliza a contagem. Segundo o Professor Aúthos Pagano, emérito estatístico brasileiro, *nada impede seja a Estatística aplicada a outras ciências*

(...), e neste caso, que é o real, ela pode estar na órbita de qualquer outra ciência (Pagano, 1946). Martin Frankel e Benjamin King (1996), em entrevista com Leslie Kish, em 22 e 23 de julho de 1994, receberam estas palavras do grande estatístico:

(...) E hoje, um curso de Estatística é necessário? Penso que sim. HG Wells disse: o pensamento estatístico será, um dia, tão necessário para formar-se um cidadão tanto quanto ter a habilidade de ler e escrever — eu acredito que isso seja necessário, e eu disse, em uma carta escrita em 1994, sobre as investigações médicas que um curso como esse seria muito importante para os médicos, mas um curso não é suficiente para se aprender a construir um experimento ou para fazer levantamentos clínicos. Para qualquer nível de uma educação em Ciência — médica, social, física — um curso de Estatística seria útil para aqueles que pudessem compreender sobre que lêem nos artigos e nas revistas, e essas pessoas deveriam conhecer o suficiente da linguagem estatística para serem capazes de consultar os peritos para ajudá-los a preparar o mapa de seus experimentos. Eles não deveriam, no entanto, preparar esses mapas por si mesmos, pois que não se faz uma cirurgia com apenas um único curso sobre esse assunto. (...)

Com o olhar voltado para a ciência Estatística, pode-se, sucintamente, considerá-la como composta de três grandes partes: Amostragem, Descrição da Amostra, e Análise Estatística propriamente dita.

A Amostragem consiste, em essência, de um processo, cuja origem remonta há milênios e cujos interesses, quase sempre,

redundaram em seu resultado, ou seja, em sua capacidade de contabilizar valores unitários, porém, sob o ponto de vista processual, deve-se considerar que a Amostragem apresenta técnica própria, e muito menor é seu interesse na obtenção do resultado contábil, mas, sim, em conseguir criar o método sob o qual o processo amostral deve ser desenvolvido. Portanto, a Amostragem, em si, alcança o nível processual e tende a ser, no mínimo, uma ferramenta equivalente às demais componentes tanto da Estatística como da própria Ciência. Nesse sentido, modernamente, poder-se-ia até mesmo, vislumbrar a Amostragem como ciência à parte da Estatística, pois que seus eixos correm, em diversas direções, para os campos de atuação de outras ciências.

Quanto à Descrição da Amostra, deve-se considerar, aqui, a necessidade de conjugação de valores amostrados e suas respectivas representações reduzidas, isto é, o uso extensivo de medidas-resumo, caracterizadoras da amostra coletada e que, por sua vez, já são passíveis de uso de teor analítico (ainda, não estatisticamente falando), com a proposta de resumir os dados coletados a pontos de observação relevantes. Surge, neste sentido, o caráter redutor da Estatística, que tem o poder de descrever uma massa de dados coletados por meio de medidas-resumo, medidas estas que bem podem ser valores ou contagens resumidas, arrançados em listas, tabelas, quadros, ou mesmo, desenhos

indicadores de densidades, concentrações e tendências.

Coube a Sir Ronald Aylmer Fisher (1890-1962), afirmar que *a estatística escolhida deveria sumarizar toda a informação relevante fornecida pela amostra.*

E quanto à Análise Estatística propriamente dita, pode-se referenciá-la como necessária ao caráter comprobatório frente a uma hipótese construída. Neste sentido, o dito caráter comprobatório torna-se núcleo essencial para o desenvolvimento das técnicas analíticas. A análise, como questão científica, e neste caso, como questão estatística, pode ser encarada como segregadora, posto ser ela, em termos filosóficos, destinada a ser o divisor de águas para uma tomada de decisão. A Ciência, em si, é vista, hoje, como tendo um caráter decisório mais destacável do que propriamente processual, mas, bem se sabe que o processo científico é que permite a produção e a existência de um teor decisório embasado e seguro.

Ainda, nesta exposição inicial, faz-se necessário destacar a Amostragem como sendo o núcleo desta abordagem, posto que cabe ao processo amostral ser a fase mais importante da aplicação da Estatística sobre os outros elementos componentes da pesquisa científica: *a boa amostra leva ao bom resultado...*

## 1.1 Comentários sobre alguns conceitos usuais

Atualmente, deposita-se sobre a variável aleatória<sup>1</sup> a suprema responsabilidade de ser a base definidora do tamanho da amostra, pois, sendo ela aquele ente estatístico que carrega em sua essência as características necessárias para obter-se o conhecimento prévio sobre o que será abordado na pesquisa científica, quase nada há para ser contestado.

Porém, quer-se propor (antes de contestar!) que o núcleo da variável aleatória, representado por parâmetros previamente fixados, mas, não necessariamente conhecidos *a priori*, trata-se, na verdade, de uma, dentre várias abordagens, o que pode ser comprovado pelo fato de a variável aleatória ser o elemento mais pontual da investigação científica (descontando-se, é claro, seus próprios parâmetros).

Justifica-se a definição de variável aleatória como sendo uma *função* de cunho matemático e que envolve as operações da Teoria dos Conjuntos, porque em sua acepção mais geral, é preciso que exista um rol de critérios matemáticos para chegar-se a essa

---

<sup>1</sup>Em uma breve definição, seja  $\varepsilon$  um experimento e  $S$  um espaço amostral associado ao experimento; uma função  $X$ , que associe a cada elemento  $s \in S$  um número real,  $X(s)$ , é denominada variável aleatória (Meyer, 1983).



definição, que, por vezes, pode não ser suficiente para concretizar um pensamento científico que se sobreponha à ideia de que os itens do *instrumento de coleta de dados* da pesquisa são, verdadeiramente, variáveis aleatórias, conforme sugere essa definição. Por outro lado, a variável aleatória consiste em ninho de parâmetros agregados, pois, sem isso, ela não poderia ser representada por valores, que, efetivamente, existem quando da aplicação da função matemática à qual ela está associada.

O instrumento de coleta pode ser editado em papel, ou, ser uma tela de computador, destinado ao apontamento dos valores observados, durante o processo de coleta dos dados; ele contém itens e cada item pode ou não conter categorias, que caracterizam cada item. Assim, um instrumento de coleta é um rol de itens, cada qual com características próprias. Os principais tipos de instrumentos de coleta são o prontuário, a ficha, o questionário, o protocolo.

Já, os itens formadores do instrumento de coleta também podem ser de alguns tipos, e esses tipos dependem, intrinsecamente, dos *valores* a serem observados (se ordenáveis ou nominais), das expressões textuais aceitas para cada valor (se codificáveis ou não), da sequência de categorias que se apresentam em cada item (se as categorias são mutuamente

exclusivas ou se pode haver múltipla escolha de categorias dentro de um mesmo item).

Se por um lado, tem-se a essência da variável aleatória como produtora da informação suficiente para o cálculo do tamanho da amostra a ser coletada, por outro, é deveras intrigante que essa informação seja, por si só, algo tão importante a ponto de deixar-se o instrumento de coleta à parte de uma decisão dessa monta. Volte-se ao instrumento de coleta de dados: é certo que a palavra instrumento pode não ser a mais feliz, mas, é habitualmente usada para exprimir a ideia de um *conjunto de itens dispostos em um documento adotado para apontamento sistemático de valores observados durante a coleta de dados* (um conceito, neste momento).

Pelo que foi exposto, sugere-se que os instrumentos devam apresentar *tipos* diversos, cuja base classificatória teria, *a priori*, o tipo de variável envolvida na pesquisa. Com isso, a proposta deste trabalho torna-se mais instigante, porquanto não irá aprofundar-se até o atingimento do núcleo da variável aleatória, mas, sim, com base em informações mais imediatamente presumíveis e identificáveis, querer calcular o tamanho da amostra necessária, observando a composição do instrumento, como corpo uno da pesquisa, e que pode fornecer meios para esse cálculo, por

sua forma e conteúdo.

Mas, retorne-se às variáveis aleatórias e seus parâmetros: nas últimas décadas, a Estatística ganhou reforços consideráveis e oportunos, por conta da inclusão em seus métodos das chamadas Estatísticas Não-paramétricas, campo de atuação e desenvolvimento de novas técnicas de teor analítico, mas, que, por sua vez, provocaram sensíveis abalos nas partes *descritiva* e *amostral* da Estatística. Por quê? Porque as técnicas não-paramétricas de análise propunham-se, e ainda propõem-se, a ser uma nova senda a ser percorrida pelos analisadores de dados, quando pouco se conhece sobre a essência da variável aleatória abordada. Ora, a parte analítica da questão parece, então, até mesmo, bem resolvida, porque as análises não-paramétricas encontram seu espaço quase-definitivo junto às técnicas de análises paramétricas, que, por exigirem mais da variável aleatória, ainda pleiteiam características pontuais junto a ela. Assim sendo, a descrição, de cunho não-paramétrico, passa a destacar, com naturalidade, medidas que, antes, tinham menor visibilidade, ou, que tinham usos específicos, como, por exemplo, a mediana e os demais percentis (decis, quartis, tercis etc.), e valores como totais, máximos e mínimos, cujas presenças, em trabalhos de teor analítico-paramétrico (em termos estatísticos), eram menores, visto não haver a necessidade de análise sobre

esses entes.

Do mesmo modo que visto para a Descrição da Amostra, quer-se sugerir que a abordagem não-paramétrica, frente à Amostragem, também encontra-se por ser explorada, desenvolvida e proposta como processo. Assim sendo, chega-se ao cerne desta proposta, que vai tentar direcionar seus esforços para uma visão *menos paramétrica* da Amostragem, voltando-se para uma espécie nova, que é, por fim, o instrumento propositor da pesquisa, em seu conjunto, e não, apenas, a variável aleatória, como exclusivo núcleo de direcionamento para o cálculo do tamanho da amostra.

Frise-se que o referido instrumento de coleta nada mais é do que um conjunto de itens e suas respectivas categorias, selecionados para a boa condução de uma pesquisa. O uso da palavra instrumento, por si só, merece alguns comentários adicionais: a palavra tem uso generalista, isto é, em muitas áreas do conhecimento, fala-se de instrumentos e de expressões derivadas dessa palavra. Logo, deve-se uniformizar seu uso, aqui, destacando, sobremaneira, que a ideia do instrumento — visto como uma coleção de itens pertencentes aos interesses da pesquisa que será conduzida — leva a crer que o estatístico terá livre participação na confecção do instrumento: ledo engano! Essa

confeção é mote originário do pesquisador, e em última instância, poder-se convocar o estatístico para opinar sobre a *forma* dos itens que deverão compor o dito instrumento, não sobre o *conteúdo*. O que seria aceitável considerar, em termos participativos, quanto à presença do estatístico, seria o conhecimento, por parte deste, das intenções da pesquisa e das possibilidades analíticas quando do término da mesma, dadas as dificuldades inerentes ao processo investigatório a ser proposto, porém, repise-se que a participação do estatístico, nesse momento, tem caráter orientador, e não determinante sobre o próprio interesse do que será investigado.

## 1.2 Amostragem — conceito e algumas histórias

Parte da Ciência Estatística destinada ao estudo da amostra, que consiste em subconjunto da população. A população é o ente estatístico de onde provêm todos os elementos que poderão ser submetidos à investigação científica, de interesse do pesquisador responsável pela condução da pesquisa. Um elemento deve, *a priori*, poder ser selecionado para compor a amostra, com a mesma probabilidade de todos os demais. Esse conceito, de equiprobabilidade, obriga o investigador a compor sua própria crítica sobre o processo de coleta de elementos amostrais e por fim, sobre a formação da própria amostra. Bem se sabe, hoje, que

não se consegue atribuir, em muitos casos, probabilidades realmente iguais para a conformação de uma amostra; basta considerar-se os experimentos em que subpopulações de uma população maior são abarcadas para efetivar a coleta da amostra: se nem todos os elementos populacionais estiverem disponíveis, então a probabilidade de seleção não será a mesma para cada ente populacional. Isso ocorre com muita frequência, em pesquisas na área da Área de Saúde, em que a condução de um estudo impõe a seleção de elementos presentes nos serviços clínicos, e portanto, sem a preocupação de que uma probabilidade de seleção seja considerada como efetivamente importante. Aliás, essa probabilidade nem mesmo chega a ser estimada; ela consiste, apenas, em um valor teórico, distante de ser avaliado pelo pesquisador, que, desde logo, está preocupado em ter elementos amostrais suficientes para prover seus resultados, levando-os a uma boa conclusão. Por vezes, essa conclusão nem mesmo poderia ser considerada como de teor estatístico, pois que a amostra de onde os resultados geraram as ditas conclusões também não poderia ser considerada como uma amostra de teor estatístico, dada a não-equiprobabilidade de seleção de seus elementos.

Essa crítica inicial chama a atenção de quem, com certo conhecimento científico, busca, em sua pesquisa, ater-se ao mais elevado grau de confiança que seus dados devem embasar seu

estudo. Mas, em termos reais, isso não ocorre: como se disse, a seleção de elementos provindos da população é, em geral, desprovida dessa equiprobabilidade. A Ciência, que tem em sua base, a observação, ou seja, *a mensuração das manifestações naturais dos fenômenos* (Câmara, 1965), apresenta, por si só, aspectos limitantes bastante visíveis, em termos da seleção de elementos componentes para avaliações em experimentos seriamente conduzidos. Essas limitações, em geral, não consistem em fatores impeditivos para a realização de um experimento, porque existe uma aceitação, de bom senso, de que há uma suficiência inerente ao próprio experimento, e que está assentada sobre os elementos que compõem a amostra coletada, depositando-se, sobre esses elementos, o grau de confiança que será aceito para concluir-se algo a respeito da investigação proposta.

A ideia de amostra existe há milênios: talvez, não com essa denominação, mas, embasada em algo mais primitivo, como a própria repetição da observação. Um dos registros mais antigos encontra-se em Tucídides (460-395 aC), historiador das Guerras do Peloponeso, que escreve que *para se determinar, de maneira acurada, a medida de uma grandeza, faz-se mister observá-la em repetidas vezes* (Câmara, 1965). Logo, essa ideia, a de repetição da observação, está contida na alma da Amostragem e na própria

alma da Estatística. Assim, à característica de repetição agrega-se, desde logo, a característica da contagem, que permite organizar o processo repetitivo. Contar é um ato relativamente simples, e que permite criar um entorno de organização ou de ordenação de objetos selecionados. Conta-se cientificamente, ou seja, organiza-se de modo a reconhecer, numericamente, cada objeto observado, e é essa a razão de a Estatística ser a Ciência da Contagem: neste sentido, ela possui, em sua essência, não o *ato* de contar, mas a *forma* de contar; contagem, aqui, não significa somente a sequência numérica de inteiros positivos, mas, a identificação ordenada de elementos que, necessariamente, têm de existir, para qualificar os objetos contados, que, por sua vez, pertencem ao núcleo da condução de uma pesquisa de caráter observacional.

Matematicamente, a contagem é definida como uma função bijetora que liga cada elemento de um conjunto  $A$  definido a um número inteiro (Griffiths e Hilton, 1975). Assim, pode-se escrever:  $f: A \rightarrow N_n$ , onde  $n$  é o número de elementos do conjunto 'A', e 'N' é o conjunto dos números naturais. Em Amostragem, sempre quer-se que o conjunto 'A' contenha elementos, ditos *elementos amostrais*, que sejam úteis, para as avaliações estatísticas de uma pesquisa.

Historicamente, a *ideia do uso* de uma amostra, apesar



de ser mencionada em textos da Antiguidade e da Bíblia, surgiu, de modo a preencher uma lacuna científica, com Pierre Simon de Laplace, em 1786, quando esse prestigiado matemático francês estimou a população de seu país, usando um cálculo que mostrava proporções de nascimentos em um período, em trinta regiões francesas, e fazendo a extensão dessa quantidade proporcionalmente calculada, para toda a área geográfica abrangida pelas fronteiras francesas (Stigler, 2003).

A primeira discussão pública registrada sobre a ideia de Amostragem ocorreu em 1895, na reunião do ISI (International Statistical Institute), em 1895, em Berna (Suíça), na qual Anders Nicolai Kiaer (1838-1919), que era diretor da Agência Norueguesa de Estatística, relatou sua experiência com amostras e levantamentos amostrais feitos para o governo de seu país (O'Muirheartaigh, 2005). Kiaer definiu que as investigações poderiam ser feitas com base em inquéritos parciais, com base em observações recolhidas em diversas regiões do território investigado, conquanto se obedecesse a uma distribuição proporcional das unidades a serem investigadas em cada região. Essas unidades não seriam escolhidas aleatoriamente, mas, de acordo com um esquema racional baseado em resultados de teor estatístico previamente alcançados.

O geógrafo francês Pierre Émile Levasseur (1828-1911) mencionou que existiam três métodos de levantamento: (a) uma completa enumeração das unidades; (b) a descrição detalhada de um único fenômeno; e (c) a *exploração estatística*, termo este que permitiu esclarecer melhor a proposta de Kiaer. Assim, foi decidido que o terceiro método seria pauta da reunião seguinte do ISI, quatro anos depois, em São Petersburgo.

Kiaer defendeu seu método, no encontro de 1899, insistindo que tais investigações por amostragem eram suficientemente representativas, e que podiam se consideradas como *uma fotografia que reproduz os detalhes do original em suas verdadeiras proporções relativas*. Ele, ainda, afirmou que haveria muitas maneiras de contornar o problema da representatividade, chegando a afirmar que, se as comparações com os resultados do censo mostram-se equivalentes aos do levantamento estatístico para uma determinada característica estudada, então isso deve valer para outras informações populacionais, que podem, então, ser levantadas em uma investigação com partes da população.

Na sessão seguinte do ISI, em 1903, já havia sido constituído um comitê para avaliar o valor dos levantamentos por amostragem, que, de modo corajoso, recomendou que esses levantamentos passassem a ser feitos, por questões de economia

de tempo de coleta e de custo, e que fossem comparados aos resultados dos censos regionais. Ainda em 1903, a noção estatística de *amostra* e seu uso efetivo tomaram corpo com Karl Pearson (1857-1936), que, imaginando sobre as tão discutidas limitações acarretadas por uma investigação de toda a população, afirmou que, *se toda a população fosse tomada, deveriam ser obtidos valores precisos para suas constantes estatísticas, porém, na prática, o pesquisador é apenas capaz de ter uma amostra.*<sup>2</sup>

Somente na reunião do ISI de 1925, seus membros aceitaram e oficializaram os levantamentos estatísticos por amostragem, recomendando-os aos governos sobre os bons resultados alcançados e os benefícios econômicos desses levantamentos, além do elevado grau de credibilidade dos resultados apurados, com base nos estudos feitos nas primeiras décadas do século XX.

---

<sup>2</sup> A palavra *amostra* aparece em uma publicação de Pearson, em 1903, na revista *Biometrika* II, p273. Um colega de Karl Pearson, o zoologista Walter Frank Raphael Weldon (1860-1906) já usava essa palavra, desde 1892, para designar *uma coleção de observações*.

### 1.3 Importância da Amostragem

Em um procedimento tipicamente inferencial (que permite chegar a conclusões sobre uma população a partir do estudo de uma amostra), a técnica de amostragem torna-se essencial (Szwarcwald e Castilho, 1992). Esses mesmos autores ponderam, ainda, sobre o surgimento do problema de selecionar uma amostra, *o mais representativa da população total, diante das limitações de custos e das possibilidades de perda de precisão na estimativa dos parâmetros*. Aqui, é preciso fazer algumas ressalvas:

- (a) a expressão *o mais representativa* deve ser entendida como especificamente *representativa*. A representatividade consiste em busca intrínseca ao ato de conformação de uma amostra, que, por sua vez, tem de ser, obrigatoriamente, representativa, visto que *o mais representativo possível* pode, ainda, consistir em fator insuficiente para alcançar o grau de precisão buscado no processo de avaliação requerido;
- (b) quanto aos custos da amostragem, os mesmos devem ser sempre considerados, porém, não somente os custos de teor financeiro, intrínsecos aos procedimentos de coleta (contratação de pessoal, investimento e manutenção de equipamentos, impressão de cópias de instrumentos de coleta,

gastos indiretos, como transporte e alimentação), mas, também, os custos de teor temporal, que determinam marcadores de tempo, fixando prazos de cumprimento das tarefas agendadas, e para os quais, pode haver itens quanto às possibilidades de término do processo de coleta; em ambos os casos, pode haver gastos e desgastes a serem evidenciados, se possível, antes do início da coleta;

- (c) quanto à perda de precisão na estimativa de parâmetros, a mesma é inerente ao fator redutor de um procedimento de amostragem; uma amostra carrega, por certo, uma massa de imprecisões, tanto menos critérios e regras forem previamente incluídos como membros inseparáveis do processo amostral.

As técnicas de amostragem estão indispensavelmente vinculadas ao nome de William Cochran, que as sistematizou em 1953 (Cochran, 1985). Embora de frequente emprego em investigações populacionais, nem sempre o tratamento analítico dos dados é adequado ao tipo de procedimento usado para a seleção das unidades experimentais, resultando em sérios vieses de interpretação. Com esta perspectiva, um seguro objeto de estudo da Estatística Aplicada, nestes e nos próximos anos, tem sido e será o desenvolvimento de métodos de estimação e

inferência compatíveis com as diferentes técnicas de amostragem. Vale insistir que esta questão não vem recebendo a devida consideração e são inúmeros os exemplos de inferências incorretas, consequentes ao corriqueiro tratamento de que sempre se está diante de amostras confiáveis.

Até as primeiras décadas do século XX, a ideia de uma *amostra aleatória*, cujos elementos deviam ser selecionados equiprobabilisticamente, e conforme é considerada hoje, não era universalmente aceita, visto que vários estatísticos acreditavam que os controles de seleção a serem impostos deveriam ser obrigatórios, e isso seria (mas, de fato, não era) um indicador de maior precisão no grau de representatividade da amostra coletada, frente à população de origem.

Coube, então, ainda no início do século XX, a William Sealy Gossett (conhecido pela alcunha de Student), propor uma série de ideias, que tornaram mais rigorosa as assertivas dos estudos sobre questões estatísticas e também, sobre Amostragem. Seu núcleo de investigações desenvolveu-se em função de seu interesse por amostras consideradas como *pequenas*. Naquela época, já era de conhecimento dos estudiosos da Estatística, que amostras deviam ser grandes o suficiente, para serem representativas da população de onde tinham sido coletadas.

Assim, era consenso, entre os mais afamados estatísticos, que amostras pequenas não poderiam ser usadas como bons fornecedores de informações, e era comum que os trabalhos com amostras ditas pequenas não passassem de meras especulações sobre as investigações populacionais (Crouch e McKenzie, 2006). Após os sérios estudos de Gossett, pequenas amostras passaram a ser o alvo extensivo de muitas pesquisas, principalmente, na Área de Saúde, onde o número de espécimes era (e ainda é!), em geral, limitado, não só por questões éticas, mas, pela própria dificuldade de coleta.

Wiley (2003) afirma que, apesar de, na maior parte dos estudos, o tamanho da amostra não influenciar em demasia os resultados apurados, sabe-se que amostras consideradas pequenas levam a erros de conclusão, e portanto, à perda da pesquisa, devido ao elevado grau dos vieses dos parâmetros estimados. Já, amostras ditas grandes minimizam esses vieses, e quase sempre, permitem que conclusões satisfatórias valorizem o processo de amostragem, bem como os próprios resultados alcançados.

## 1.4 Revisão da Literatura

Esta parte do trabalho engloba, de modo resumido, as referências básicas da Lógica Paraconsistente Anotada Evidencial e da Análise Combinatória, e de modo mais extensivo, os textos sobre as propostas de cálculos de tamanho de amostra, destacando os tipos mais usuais.

### 1.4.1 Lógica Paraconsistente Anotada Evidencial

A Lógica é uma característica humana. Como seres humanos, tentamos, desde há muito, explicar as coisas da Natureza que nos cerca, por meio de teorias e modelos em cujas bases assentam-se pensamentos ditos lógicos. E ainda como seres humanos, talvez, deve-se considerar que temos uma lógica própria e individual, que pode se abstrair (afastar-se ou aproximar-se) da Lógica Matemática ou Lógica Clássica, e que, por certo, pertence à natureza humana. Mas, a Lógica Clássica, apesar de querer mostrar sua característica, nada pode fazer contra os desígnios que a subvertem frente à vontade humana. Um exemplo simples: em um item cujas possibilidades consistem de duas categorias, do tipo sim-não, a Lógica Clássica gostaria de impor que apenas duas respostas são possíveis, e nenhuma ação humana deveria ser



cogitada para contrapor-se a essa imposição, porém, bem sabe-se que tanto a não-resposta (não-escolha do 'sim' e não-escolha do 'não') como a resposta dupla sim-não são, efetivamente, respostas do Raciocínio Humano puro, pois, não menos efetivamente, cabem no bojo das respostas plausíveis e aceitáveis da consciência humana. Assim, a Lógica Clássica, apesar de ser impositora de sua presença, está contida e é abarcada pelo Raciocínio Humano.

Os princípios que norteiam a Lógica Paraconsistente Anotada Evidencial permitem sua aproximação mais aquiescente do Raciocínio Humano (Abe, 1992), pois levam em conta possibilidades que vão além do binarismo proposto em um item do tipo sim-não, mencionado no parágrafo anterior. Assim, um item formado por duas categorias pode ter quatro (e não duas respostas) aceitáveis: *sim*, *não*, *sem resposta* e *sim-não*; esses quatro estados são classificados em quatro diferentes núcleos lógicos, sendo eles o *verdadeiro*, o *falso*, o *paracompleto* e o *inconsistente*, respectivamente, e permitem uma avaliação mais próxima do que exprime a prática e a realidade de uma pesquisa (Abe, 1997). Assim, respostas não previstas pela Lógica Clássica são previstas pela Lógica Paraconsistente Anotada Evidencial.

#### 1.4.2 Combinatória

Apesar de a solução do problema proposto parecer exigir método relativamente sofisticado, o princípio a ser usado nos cálculos que envolverão a Análise Combinatória é muito simples. Por certo, não haveria como relacionar os autores que já publicaram esse princípio, e o assunto é largamente exposto em livros do ensino médio, não necessitando de nenhum complemento para o uso que será feito, neste trabalho. Assim, pode-se citar o trabalho de Grimaldi (1986), que menciona, como um dos Fundamentos da Técnica de Contagem, o uso das combinações para a solução de diversos problemas em que uma contagem ponto a ponto seria desgastante.

Aqui, vale a pena lembrar que a Combinatória trata dos cálculos referentes às combinações e permutações, isto é, de um conjunto de informações científicas cujo escopo responde pela formação da chamada *Teoria Combinatória*, desenvolvida para abarcar os problemas concernentes às contagens e enumerações.

A característica principal da combinação consiste no fato de a ordem de seleção dos elementos disponíveis não ser necessária, diferentemente da permutação, em que essa ordem é estritamente considerada (Grimaldi, 1986). A fórmula da combinação de 'x' elementos 'k a k' é dada por:

$$C = \frac{x!}{k!(x-k)!}$$

De modo resumido e equivalente, a fórmula anterior pode ser reescrita assim:

$$C = \binom{x}{k}$$

Neste caso, tanto os valores de 'x' como os de 'k' são inteiros positivos, e  $0 \leq k \leq x$ , e 'C' é o número de combinações a ser calculado.

Não será desenvolvida nenhuma teoria sobre a combinação, neste momento, visto que sua definição é universalmente conhecida e a aplicação que será feita tem caráter superficial, sendo de simples compreensão.

Em geral, a Combinatória propõe-se a solucionar os problemas relacionados à contagem, porém, existem problemas de contagem cujo cerne está na classificação previamente adotada e usada para caracterizar as categorias do que se deseja examinar; daí, os problemas de categorização e de seleção do conjunto a ser examinado. Vejam-se, por exemplo, os casos historicamente registrados de duas eleições norte-americanas, em que um candidato dava como certa sua vitória:

- (a) em 1936, Roosevelt e Landon disputaram a eleição, respectivamente, pelos partidos Democrata e Republicano (Stephan, 1948). Um instituto de pesquisa previu que Landon seria eleito e que Roosevelt obteria 43% dos votos, mas, Roosevelt ganhou, com 62%. Apesar de amostra desse instituto de pesquisa ter sido composta por aproximadamente 2,4 milhões de pessoas, elas haviam sido selecionadas de um modo viesado, pois aquele instituto de pesquisa tomou, como base, para seleção dos eleitores, apenas os que tinham telefone e que eram associados de clubes. A maioria da população votante era pobre, e portanto, não tinha telefone e não era associada de clubes; logo, a seleção feita não contemplava a distribuição das intenções de voto de toda a população de eleitores. Outro instituto de pesquisa coletou uma amostra de três mil eleitores, porém tomando o cuidado para conformar uma amostra dita representativa: assim, esse instituto previu que Roosevelt ganharia com 56%; a diferença de seis pontos percentuais deveu-se à técnica de seleção dos elementos que compoariam a amostra (seleção por quotas), e que, a partir da década de 1960 (Kish, 1965), seria alvo de sérias críticas, tendo, inclusive, caído em desuso.
- (b) ainda, nos Estados Unidos, mas, agora, nas eleições de 1948, Truman (Democrata) e Dewey (Republicano) disputaram a

contenda política (Kish, 1965). Três institutos de pesquisa, incluindo aquele que havia acertado o resultado de 1936, afirmaram que Dewey ganharia com 50%, 50% e 53%, respectivamente; Truman ganhou com 50% e Dewey recebeu 45% dos votos válidos. As amostras coletadas foram embasadas pela técnica de quotas, e por isso, o fiasco dos institutos foi geral.

Em ambos os casos, acima, processos planejados com deficiências para as classificações foram responsáveis por erros grosseiros na apuração das prévias eleitorais daqueles pleitos. Em ambos os casos, estratégias para incluir e excluir sujeitos com base nas categorias aos quais eles pertenciam consistiram na derrocada dos resultados das pesquisas. Classificar um sujeito nesta ou naquela categoria é, por certo, algo difícil, quando as incertezas sobrepõem-se e combinam-se a outros obstáculos.

O famoso escritor Isaac Asimov, em um artigo de 1955, intitulado *Franchise*, menciona que, em um futuro distante (para aquela época), que seria o ano de 2008, os Estados Unidos teriam sido convertidos em uma democracia eletrônica, onde um computador denominado Multivac conseguiria selecionar uma única pessoa para responder a uma série de questões, e assim, com as

respostas e algumas características pessoais do escolhido, conseguiria determinar quem seria o próximo presidente do país, sem a necessidade de efetivar-se um real processo eleitoral. Norman Muller, então, é o escolhido como 'eleitor do ano'; de início, ele fica com medo, mas, depois de votar, orgulha-se pelo fato de *os eleitores norte-americanos terem exercido, mais uma vez, sua livre e desimpedida franquia* (Asimov, 1986). Essa pessoa representaria todas as categorias caracterizadoras dos eleitores!

Na Área de Saúde, é comum querer-se adjetivar as categorias formadoras de um item investigado e propor, de modo subjetivo, uma regra de classificação para os elementos que serão os componentes da amostra a ser estudada. Não são poucos os casos em que reclassificações devem ser efetivadas, pois os resultados de teor estatístico ficam à mercê daquela subjetividade que previamente propôs uma primeira classificação. Um exemplo comum consiste na classificação de grau de gravidade em que paciente se encontra: 'leve', 'moderado' e 'grave' constituem em categorias usuais, bem como a classificação baseada nas quatro cruces, ou mesmo, para um estado binário ('estar bem' e 'estar mal'). A dificuldade apresenta-se para qualquer uma dessas possibilidades.

Além do problema de haver um número suficiente de

elementos amostrais em cada categoria criada, pode-se incorrer no erro de agruparem-se elementos díspares sob uma mesma denominação (Katz et al, 1977). A contagem desses elementos em nada conseguirá satisfazer as necessidades de obtenção de um resultado confiável por meio da pesquisa proposta. Por isso, a Amostragem tem seu momento e torna-se presente para trabalhar em prol da Ciência, e hoje, muito mais profundamente, em prol da Área de Saúde, em conjunto com a Matemática, a Física e a Lógica, formando um grupo eficaz de apoio às pesquisas e aos trabalhos científicos.

Enquanto a Estatística é a *ciência* da contagem, a Combinatória é a *arte* da contagem. De início, essa arte fazia-se representar por jogos recreativos e brincadeiras com números, mas, não demorou a que seu escopo se tornasse sério o suficiente, para ser considerado como assunto de evidência ímpar na solução de complexos problemas de contagem. Outro exemplo disso consiste em que, nas últimas décadas, a Ciência Genética tem usado a Combinatória para solucionar os problemas de enumeração de combinações de genes, e com a participação do Cálculo de Probabilidades, tem proposto soluções para identificar os genes provocadores de problemas em seres vivos.

### 1.4.3 Tamanho da Amostra

Um dos artigos fundamentais para a revisão de literatura sobre o cálculo do tamanho da amostra foi publicado em 1991 por Mitchell H. Gail: é um artigo de referências bibliográficas recomendadas, e nele, esta revisão foi feita, em conjunto com as referências mais recentes, publicadas após aquele referido ano.

Os cálculos de tamanho de amostra demonstram-se eficazes, para vários tipos de planejamento de experimentos, previamente conhecidos e usualmente empregados em pesquisas clínicas, bem como em outras áreas do conhecimento (Grunkemeier e Jin, 2007). Os principais métodos de cálculo de tamanho de amostra são de uso comum, e proporcionam o cálculo do tamanho da amostra, com base em características que, previamente, devem ser conhecidas e adotadas.

Em todas as condições de interesse estatístico de uma investigação em que um assunto da Área de Saúde seja o eixo condutor da pesquisa, deve-se estudar e compreender o mapa do experimento a ser executado, e com base nesse mapa, calcular o tamanho da amostra. Esse cálculo deve fornecer um valor que representará o mínimo a ser coletado, dado o desenho do experimento sobre o qual deposita-se parte do interesse da pesquisa. Assim, os tamanhos de amostras a serem calculados, no



geral, apontam para valores ditos mínimos, porém, é preciso, desde já, ficar claro que valores mínimos são desejáveis para consistirem em um determinador do que deverá ser coletado, mas, o valor mínimo, em si, não é um limitador específico e único (Bégaud et al, 2003). Por certo, não há um máximo a ser calculado, mas, tão-somente, adotado, ou, quando for o caso, estimado. O máximo tamanho da amostra a ser considerado é o próprio tamanho da população, posto que qualquer amostra oriunda de uma população, além de ser finita, para a pesquisa em questão, também é menor do que o valor que representa o tamanho da população investigada.

Além disso, algumas questões são comumente referidas, quando se deseja calcular o tamanho da amostra (Scales e Rubinfeld, 2005): (a) a pergunta principal da pesquisa; (b) a variável primária, ou seja, a variável sobre a qual deposita-se o foco do estudo; (c) a análise estatística a ser usada; (d) o grau de confiança ou precisão a ser adotado; e (e) estudo-piloto e suas características. Em cada um desses itens, há limitações que se evidenciam, e para as quais, não se tem solução única e aplicável, de imediato:

(a) a pesquisa pode conter mais de uma pergunta a ser respondida, por meio do processo investigatório ao qual os

próprios itens componentes da investigação estão sendo submetidos. Assim, pode-se verificar que, no geral, pesquisas respondem a várias questões (que podem não ser muitas), porquanto tais questões já existem, desde o planejamento do que vai ser investigado, ou, também, acabam surgindo com o desenrolar do processo investigatório.

- (b) o fato de que, geralmente, existe uma variável eleita como mais importante, chamada, então, de variável primária ou variável principal, e com base em suas características, deposita-se total atenção para o cálculo do tamanho da amostra, é, por certo, algo bastante discutível, visto que, as outras variáveis de interesse, pertencentes ao mesmo conjunto que está sob investigação, podem apresentar-se com ordens de importância bastante elevadas, mesmo que menores, mas, o suficiente para poder interferir nos resultados a serem apurados após a coleta dos dados; percebemos, com certa evidência, que outras variáveis, muitas vezes, não permitem o atingir de um grau de conhecimento sobre elas, impedindo que conclusões correlatas (e próximas) ao foco da pesquisa sejam tiradas com o devido rigor científico.
- (c) apesar de, *a priori*, poder-se conhecer quais serão as análises estatísticas a serem usadas durante o processo de avaliação

da amostra coletada, isso, nem sempre, ocorre nos termos, porque, depois de ocorrido o processo de amostragem, e tendo-se a amostra disponível para as respectivas avaliações de teor estatístico, podem ocorrer alterações na escolha das análises que, previamente, haviam sido prognosticadas, fazendo com que o uso de outras técnicas torne-se necessário. É comum que análises não-paramétricas recebam apelo, para substituir condições que, no momento do planejamento para coleta da amostra, eram, certamente, paramétricas. Essa alteração pode afetar a obtenção de resultados de cunho decisório para a pesquisa, posto que testes semelhantes, mas, diferentes, nem sempre produzem resultados iguais.

- (d) uma adoção sempre tem caráter subjetivo, mesmo que possa advir do uso de técnicas auxiliares; considerando-se sua subjetividade, pode-se afirmar que os resultados oriundos da aplicação dos testes estatísticos escolhidos dependerão (e muito) da precisão que previamente foi adotada; o tamanho da amostra, bem como seus elementos, em si, serão determinados por esse grau de precisão, e por óbvio, tudo o que decorrer do processo de amostragem também ficará sujeito a essa informação, fixada, subjetivamente, no início da pesquisa.

(e) quando alguns valores para o cálculo do tamanho da amostra são necessários e não estão disponíveis, pode-se construir um estudo-piloto, para que tais valores sejam estimados. Uma amostra preliminar deve ser coletada, e algumas estatísticas iniciais devem tornar-se fornecedoras de informações para o cálculo efetivo do tamanho da amostra. Os elementos amostrais coletados durante o estudo-piloto podem, nos casos em que isso puder ser considerado, fazer parte da amostra final; nesses casos, deve-se subtrair do tamanho da amostra calculado, o número de elementos já coletados no estudo-piloto.

Resumidamente: sob a hipótese de que dispõe-se de amostras-piloto, ou, de resultados previamente alcançados, por meio de estudo semelhante, anteriormente conduzido, pode-se fazer uso de valores indicativos, com a intenção de calcular-se o tamanho amostral de um novo estudo, por aplicação direta de fórmulas específicas a cada situação; em outros casos, nenhuma informação é conseguida previamente, e então outros meios para o cálculo do tamanho da amostra podem ser empregados (Schulz e Grimes, 2005). A seguir, consta a exposição de alguns métodos consagrados, e suas respectivas formulações.

Pode-se dividir as fórmulas para cálculo de tamanho de amostra em três blocos, a saber: (a) fórmulas com base em medidas-resumo de teor descritivo — algumas medidas de teor descritivo podem ser, de modo exclusivo, fornecedoras de informações; (b) fórmulas com base em comparações — os testes que promovem a comparação entre parâmetros consistem em subsídio para efetivar-se este cálculo; e (c) fórmulas com base em relacionamentos — os coeficientes de correlação são os entes usados para este tipo.

### 1.4.3.1 Medidas-resumo de teor descritivo

#### (1) Erro-padrão

Considere-se que a variável de interesse seja binária; o erro-padrão da proporção 'P', baseado em observações aleatórias é dado por (Deming, 1966):

$$EP = \sqrt{\frac{P \times (1 - P)}{n}}$$

onde:

P — proporção de ocorrência do evento estudado

n — tamanho da amostra

O erro-padrão é máximo, quando  $P = 0,5$ . Logo, o tamanho da amostra a ser calculado, também é máximo, quando  $P = 0,5$ . Tem-se que:

$$n = \frac{P \times (1 - P)}{[EP]^2}$$

onde:

P — proporção de ocorrência do evento estudado

n — tamanho da amostra

Alguns resultados, quando  $P \times (1-P)$  vale 0,25: se  $EP = 1\%$ , então  $n = 2.500$ ; se  $EP = 2\%$ , então  $n = 625$ ; se  $EP = 3\%$ , então  $n = 278$ ; se  $EP = 4\%$ , então  $n = 156$ , e se

EP = 5%, então n = 100.

## (2) Prevalência

Em pesquisas da Clínica Médica, é bastante corriqueiro ter-se disponível o valor da prevalência do problema estudado, que pode advir de resultado de pesquisa anteriormente conduzida, de literatura já publicada, ou mesmo, de uma amostra-piloto. Essa prevalência consiste em uma estimativa da proporção de ocorrência do evento considerado na população estudada.

Em Cochran (1985), encontra-se a fórmula para o cálculo do tamanho da amostra, em que se considera a população de onde ela será sorteada como finita. Eis a proposta:

$$n = \frac{\frac{z_{\alpha}^2 \cdot P \cdot (1 - P)}{\varepsilon^2}}{1 + \frac{1}{N} \cdot \left( \frac{z_{\alpha}^2 \cdot P \cdot (1 - P)}{\varepsilon^2} - 1 \right)}$$

onde:

n — tamanho da amostra

$z_{\alpha}$  — valor da estatística z, para um nível de confiança alfa

P — prevalência (valor percentual)

N — tamanho da população

$\varepsilon$  — valor do erro admissível

Se o tamanho da população é desconhecido, ou, se é,

consideravelmente, muito grande, pode-se fazer uma simplificação, que imediatamente leva à seguinte fórmula:

$$n = \frac{z_{\alpha}^2 \cdot P \cdot (1 - P)}{\varepsilon^2}$$

onde:

$n$  — tamanho da amostra

$z_{\alpha}$  — valor da estatística  $z$ , para um nível de confiança alfa

$P$  — prevalência (valor percentual)

$\varepsilon$  — valor do erro admissível

Em 1998, Rahme e Joseph tomaram a fórmula acima e desenvolveram uma nova proposta, incluindo os valores de sensibilidade (*sen*) e especificidade (*esp*). Desse modo, a proposta foi construída nestes termos:

$$n = \frac{z_{\alpha}^2 \cdot P \cdot (1 - P)}{\varepsilon^2 \times (\text{sen} + \text{esp} - 1)^2}$$

### (3) Erro Relativo ou Precisão Relativa e o Coeficiente de Variação

O erro relativo ou a precisão relativa é um percentual aceitável para o erro amostral. O erro amostral é o quanto pode ser aceito, em termos percentuais, de desvio do valor populacional desconhecido. O coeficiente de variação (CV) tem de ser estimado: ele é o resultado da divisão do desvio-padrão pela média, que,



também, devem ser estimados. Assim, a fórmula para o cálculo do tamanho da amostra, neste caso, é dada por (McHugh, 1961; Volatier et al, 2002):

$$n = \frac{z_{\alpha}^2 \cdot CV^2}{d^2}$$

onde:

n — tamanho da amostra

$z_{\alpha}$  — valor da estatística z, para um nível de confiança alfa

CV — coeficiente de variação

d — valor do erro relativo admissível

Observação: outro estimador para o coeficiente de variação consiste na relação da diferença entre o terceiro e o primeiro quartil, e, a soma desses quartis, quando eles são conhecidos (Bonett, 2006).

#### (4) Erro ou Precisão em relação à média

Este é um dos métodos mais simples e mais usuais para a determinação do tamanho da amostra. Basta ter-se disponível uma estimativa do desvio-padrão da média da variável de interesse e adotar os demais valores (Guenther, 1981; Volatier et al, 2002):

$$n = \frac{z_{\alpha}^2 \cdot s^2}{d^2}$$

onde:

n — tamanho da amostra

$z_{\alpha}$  — valor da estatística z, para um nível de confiança alfa

s — desvio-padrão estimado

d — valor do erro relativo admissível

O mesmo autor (Guenther, 1981) indica que o valor da média da variável de interesse a ser testado deve satisfazer à seguinte fórmula:

$$n = \left( \frac{z_{\alpha} + z_{\beta}}{\mu_1 - \mu_0} \times s \right)^2 + 0,5 \times z_{\alpha}^2$$

Quando se dispõe de um estudo-piloto e seus resultados numéricos, pode-se estimar um tamanho de amostra mais preciso, usando a fórmula sugerida por Sethuraman et al (2007):

$$n = \frac{s^2}{\sum_{i=1}^k (x_i - \bar{x})^2} \times \left[ \frac{(z_{\alpha} + z_{\beta})}{d} \right]^2$$

O tamanho da amostra calculado tende a ser menor, por conta do ajuste em função da diferença entre os valores observados e sua média, que está no denominador, porém, os outros parâmetros também contribuem decisivamente para o valor

final a ser apurado. Se se tem um estudo-piloto, pode-se estimar, com maior precisão, o valor do desvio-padrão a ser usado, e o erro admissível pode, inclusive, ser adotado como um valor menor do que no caso anterior.

Com certa frequência, pode-se ter dificuldade para montar e conduzir um estudo-piloto cuja amostra seja representativa. Assim, uma amostra relativamente pequena, com poucas unidades, pode ajudar para determinar o estimador do desvio-padrão. Noether (1955) sugere que um estimador interessante para o desvio-padrão seja a amplitude entre o maior e o menor valor encontrados na amostra do pequeno estudo-piloto. Com isso, quando possível, valores máximos e mínimos, e principalmente sua amplitude (diferença entre ambos esses valores) podem ser usados como um auxílio na estimação do desvio-padrão.

Skalski (1992) sugere que um grau de perda de observações pode ser incrementado: seja 'p' a proporção de perdas esperada:

$$n = \frac{z_{\alpha}^2 \cdot S^2}{d^2 \times (1 - p)}$$

Esse ajuste é comum para todos os cálculos de tamanho de amostra, quando a perda apresenta-se como elemento

a ser considerado. Assim, em qualquer uma das formulações, pode-se levar em conta um grau de perda, que propiciará um aumento do tamanho da amostra, tanto maior o valor desse grau venha a ser estimado.

#### (5) Erro Absoluto ou Precisão Absoluta em relação à proporção

Basicamente, a fórmula, a seguir, é a mesma do item 3, porém, aqui, destaque-se que 'P' pode ser uma proporção qualquer, e não um valor de prevalência específico:

$$n = \frac{z_{\alpha}^2 \times P \times (1 - P)}{d^2}$$

onde:

n — tamanho da amostra

$z_{\alpha}$  — valor da estatística z, para um nível de confiança alfa

P — proporção estimada

d — valor do erro relativo admissível

Vollandt e Horn (2004) sugerem uma variação da fórmula acima, para a qual consideram um estimador não-paramétrico para  $P \times (1-P)$ , com base na probabilidade de obter-se uma amostra provinda de uma população cuja variável principal tem distribuição não-normal. Assim:

$$n = t_{\beta}^2 \times \frac{0,5 + \sqrt{(2\Delta)^3 - 6\Delta^2}}{3\Delta^2}$$

onde:

$n$  — tamanho da amostra

$t_{\beta}$  — valor da estatística  $t$ , para um determinado poder do teste

$\Delta$  — valor do erro relativo admissível, com  $0 < \Delta < 0,5$

### (6) Fórmula de Sturges

Em 1926, Sturges considera, apenas, o número de classes (estratos) em que a amostra pode ser dividida. A fórmula deste cálculo é dada por:

$$k = 1 + 3,322 \times \log n$$

onde:

$k$  — número de classes (estratos)

$n$  — tamanho da amostra

$\log$  — logaritmo de base 10

Daí, decorre que:

$$\log n = \frac{k-1}{3,322} \Rightarrow n = 10^{\frac{k-1}{3,322}}$$

Como pode-se ver, a fórmula não pleiteia as características essenciais de nenhuma variável aleatória, ou seja, as informações para o cálculo do tamanho da amostra não necessitam de conhecimento prévio sobre quais variáveis aleatórias

podem estar envolvidas no experimento a ser proposto nem sobre os parâmetros caracterizadores dessas variáveis.

O quadro 1, a seguir, resume o cálculo para 'k' variando entre 1 e 20:

Quadro 1: Tamanhos de amostra com base na Fórmula de Sturges.

k	k - 1	constante	(k - 1) / constante	n
1	0	3,322	0,000	1
2	1	3,322	0,301	2
3	2	3,322	0,602	4
4	3	3,322	0,903	8
5	4	3,322	1,204	16
6	5	3,322	1,505	32
7	6	3,322	1,806	64
8	7	3,322	2,107	128
9	8	3,322	2,408	256
10	9	3,322	2,709	512
11	10	3,322	3,010	1.024
12	11	3,322	3,311	2.048
13	12	3,322	3,612	4.095
14	13	3,322	3,913	8.190
15	14	3,322	4,214	16.381
16	15	3,322	4,515	32.761
17	16	3,322	4,816	65.520
18	17	3,322	5,117	131.039
19	18	3,322	5,418	262.073
20	19	3,322	5,719	524.139

Empiricamente, o número de estratos dificilmente ultrapassa um valor de 8, 9 ou 10 unidades, e portanto, a amostra chega ter uma ordem de grandeza não superior a algumas unidades de

centenas. Percebe-se, claramente, a simplificação do cálculo do tamanho da amostra, para uma aproximação de  $n = 2^k$ , com 'k' inteiro maior do que zero.

#### (7) Amostra-piloto em conjunto com o poder do teste

Quando a pesquisa a ser montada já tem alguns valores disponíveis, por conta de uma amostra-piloto, pode-se considerar o tamanho da amostra-piloto e adotar o poder do teste, conforme sugerem Herrmann e Szatrowski (1982), Francis (1985), e Gillett (1994):

$$n = \frac{n_0 \times (z_\alpha + z_\beta)^2}{z_\alpha}$$

onde:

n — tamanho da amostra

$n_0$  — tamanho da amostra piloto

$z_\alpha$  — valor da estatística z, para um nível de confiança alfa

$z_\beta$  — valor da estatística z, para um valor de poder (1- $\beta$ )

### 1.4.3.2 Comparações

As comparações estudadas podem ser de dois tipos básicos: comparações entre dois grupos, ou, dois momentos (situação em que ocorre pareamento de variáveis); e, comparações entre três ou mais grupos, ou, três ou mais momentos (situação de pareamento).

#### (1) Diferença de duas médias

Quando se deseja estimar o tamanho da amostra, com base nos erros do Tipo I e do Tipo II, em função da diferença de duas médias, tem-se (Lieber, 1990; Edmiston et al, 1993; Hsieh et al, 2003; Karlsson et al, 2003; Willan e Pinto, 2005):

$$n = \frac{2 \cdot (z_{\alpha} + z_{\beta}) \cdot s^2}{\Delta^2}$$

onde:

$n$  — tamanho da amostra

$z_{\alpha}$  — valor da estatística  $z$ , para um nível de confiança alfa

$z_{\beta}$  — valor da estatística  $z$ , para um valor de poder  $(1-\beta)$

$s$  — desvio-padrão estimado da variável de interesse

$\Delta = m_1 - m_0$  — diferença entre as duas médias

Quando se deseja não depender dos erros do Tipo I e do Tipo II e/ou um deles, ao menos, carece de um bom estimador



(principalmente, o Erro Tipo II), pode-se usar as características do *Teste t de Student*, estimando-se a estatística 'z' (Lachin, 1981; McHugh e Le, 1983). Sabe-se que:

$$z = \frac{\mu_1 - \mu_2}{\sigma / \sqrt{n}}$$

onde:

n — tamanho da amostra

z — valor da estatística z, para um determinado nível de confiança alfa

$\sigma$  — desvio-padrão da amostra-piloto

$\mu_1$  — média do grupo 1 da amostra-piloto

$\mu_2$  — média do grupo 2 da amostra-piloto

Assim, pode-se calcular o tamanho amostral, isolando-se a incógnita 'n'. Com isso, obtém-se, simplificadamente, que:

$$n = \frac{z^2 \cdot \sigma^2}{(\mu_1 - \mu_2)^2}$$

Guenther (1981) apresenta uma fórmula semelhante, porém sugerindo que um ajuste seja feito nos seguintes termos:

$$n = \frac{2 \times (z_\alpha + z_\beta)^2 \times s^2}{m_B - m_A} + 0,25 \times z_\alpha^2$$

onde:

n — tamanho da amostra

$z_\alpha$  — valor da estatística z, para um nível de confiança alfa

$z_\beta$  — valor da estatística z, para um valor de poder (1- $\beta$ )

s — desvio-padrão estimado na amostra-piloto da variável de interesse

$m_B - m_A$  — diferença entre as duas médias

Para a comparação entre duas médias de grupos diferentes, como se sabe, pode-se aplicar o *Teste t de Student*, mas, também, a *Análise de Covariância* (ANCOVA), de modo equivalente, pode ser usada. Neste caso, Lazovich et al (2000), e, Borm et al (2007) preconizam, semelhantemente a Guenther, que a fórmula para cálculo do tamanho da amostra, quando a ANCOVA é selecionada, pode ser dada por:

$$n = \frac{2 \times (z_{1-\varepsilon} + z_{1-\beta})^2 \times s^2}{m_B - m_A} + 1$$

Segundo Vickers (2001), o uso da *Análise de Covariância* permite estimar tamanhos de amostra menores do que o *Teste t de Student*, pois é comum que o estimador da diferença entre as duas médias seja maior.

Quando ambos os grupos apresentam as variâncias diferentes, deve-se considerar essa situação, usando os valores das medidas-resumo de cada grupo, para obter o tamanho da amostra a ser coletada, conforme indicam Van Belle e Martin (1993), Dawson (1998), Hayes e Bennett (1999), e, Brasher e Brandt (2007). Assim, de modo similar, considerando-se as hipóteses nula e alternativa como sendo  $H_0: \mu = \mu_0$ , e,  $H_1: \mu = \mu_1$ , pode-se escrever, analogamente que:

$$n = \frac{(z_{\alpha} + z_{\beta})^2 \times (s_0^2 + s_1^2)}{(\mu_0 - \mu_1)^2}$$

onde:

$n$  — tamanho da amostra

$z_{\alpha}$  — valor da estatística  $z$ , para um nível de confiança alfa

$z_{\beta}$  — valor da estatística  $z$ , para um valor de poder  $(1-\beta)$

$s_0$  — desvio-padrão do grupo 0

$s_1$  — desvio-padrão do grupo 1

$\mu_0$  — média do grupo 0

$\mu_1$  — média do grupo 1

O uso do coeficiente de variação (CV), anteriormente visto, e a fórmula de diferença entre duas médias produzem, segundo Van Belle e Martin (1993), uma nova fórmula:

$$n = CV^2 \times (z_{\alpha} + z_{\beta})^2 \times \left[ 1 + 2 \times \left( \frac{\sqrt{\mu_0 \mu_1}}{\mu_0 - \mu_1} \right)^2 \right]$$

O valor do coeficiente de variação, neste caso, pode ser estimado com os valores de desvio-padrão e média de qualquer um dos dois grupos, pois, *a priori*, para este caso, espera-se que ambos os valores de CV sejam estatisticamente semelhantes.

Quando os coeficientes de variação são estimados e considerados como estatisticamente diferentes, pode-se calcular o tamanho da amostra, com base no que propõe Wolfe e Carlin (1999):

$$n = \frac{[\log(1 + CV_2^2) + \log(1 + CV_1^2)] \times (z_\alpha + z_\beta)^2}{(\log d - \frac{1}{2} [\log(1 + CV_2^2) - \log(1 + CV_1^2)])^2}$$

n — tamanho da amostra

CV<sub>1</sub> — coeficiente de variação do grupo 1

CV<sub>2</sub> — coeficiente de variação do grupo 2

z<sub>α</sub> — valor da estatística z, para um nível de confiança alfa

z<sub>β</sub> — valor da estatística z, para um valor de poder (1-β)

d — erro admissível adotado

Um problema comum, na Área de Saúde, diz respeito ao caso em que uma das amostras tem tamanho fixado previamente e deseja-se estimar o tamanho da segunda amostra: normalmente, fixa-se o tamanho da amostra do grupo de estudo (n<sub>e</sub>) e se deseja estimar o número de elementos amostrais do grupo controle (n<sub>c</sub>). Heilbrun e McGee (1985) propõem a seguinte fórmula:

$$n_c = \frac{(z_\alpha + z_\beta)^2}{\frac{d^2}{s_e^2} - \frac{\frac{s_e}{s_c} \times (z_\alpha + z_\beta)^2}{n_e}}$$

Os valores dos desvios-padrão e do número de elementos amostrais do grupo de estudo devem ser, previamente, estimados ou adotados.

Outra abordagem sobre a diferença entre duas médias provém da fórmula em que o estimador da razão de verossimilhança é conhecido (Lindsey, 1997):

$$n = \frac{8 \times L \times s^2}{(\mu_0 - \mu_1)^2}$$

onde:

$n$  — tamanho da amostra

$L$  — estimador de máxima verossimilhança

$\mu_0$  — média do grupo 0

$\mu_1$  — média do grupo 1

## (2) Diferença de duas proporções

Considerem-se as hipóteses nula e alternativa a serem testadas, sendo:  $H_0: p = p_0$ , e,  $H_1: p = p_1$ . Se se quiser calcular o tamanho da amostra, com base nessa proposta, veiculada por Fleiss (1973), Ury e Fleiss (1980), Lachin (1981), McHugh e Le (1983), Hayes e Bennett (1999), Sato (2000), e, em que duas proporções são passíveis de serem testadas, teremos:

$$n = \frac{(z_\alpha + z_\beta)^2 \times [p_0 \cdot (1 - p_0) + p_1 \cdot (1 - p_1)]}{(p_0 - p_1)^2}$$

onde:

$n$  — tamanho da amostra

$z_\alpha$  — valor da estatística  $z$ , para um nível de confiança alfa

$z_\beta$  — valor da estatística  $z$ , para um valor de poder  $(1-\beta)$

$p_0$  — proporção do evento 0

$p_1$  — proporção do evento 1

Já, Edmiston et al (1993) propõem uma fórmula equivalente, mas um pouco diferente:

$$n = \frac{2 \times \left(\frac{p_0 + p_1}{2}\right) \times \left(1 - \frac{p_0 + p_1}{2}\right) \times (z_\alpha + z_\beta)^2}{(p_1 - p_0)^2}$$

Krishnamoorthy e Peng (2007) sugerem uma variante da fórmula acima, como uma aproximação para a estimação do tamanho da amostra:

$$n \cong \frac{(z_\beta \times \sqrt{p_1 \times (1 - p_1)} - z_\alpha \times \sqrt{p_0 \times (1 - p_0)})^2}{(p_1 - p_0)^2}$$

Gordon e Watson (1994 e 1996), Sahai e Khurshid (1996), e, Jung et al (2001) sugerem uma alternativa, para este cálculo, com base na própria Distribuição Normal, quando estimam-se, para ambas as proporções, valores muito pequenos:

$$n = \frac{(c_\alpha + c_\beta)^2}{2 \times \left[ \arcsin \sqrt{p_1 - \frac{1}{2n_0}} - \arcsin \sqrt{p_0 - \frac{1}{2n_0}} \right]^2}$$

onde:

$n$  — tamanho da amostra

$c_\alpha$  — valor do quantil  $c$ , para o erro Tipo I

$c_\beta$  — valor do quantil  $c$ , para o erro Tipo II

$p_0$  — proporção do evento 0

$p_1$  — proporção do evento 1

$n_0$  — tamanho da amostra do estudo-piloto

(3) Diferença de *ranks* (postos) entre 2 grupos (baseado no *Teste de Mann-Whitney*)

Considerem-se duas amostras independentes formadas

pelos elementos  $X_1, \dots, X_m$ , e,  $Y_1, \dots, Y_n$ . Seja o número total de elementos amostrais dado por  $n = a + b$ . A estatística do *Teste de Mann-Whitney* propõe que  $U$  seja o número de elementos amostrais em que  $Y_j$  seja maior do  $X_i$ , com  $i = 1, \dots, a$ , e,  $j = 1, \dots, b$ . Assim,  $\mu(U) = a.b.p$ , onde 'p' é a probabilidade de que  $Y$  seja maior do que  $X$ . Como hipótese nula, pode-se considerar que  $p = 1/2$ , e com isso, tem-se que, sob a hipótese nula,  $\mu_0(U) = a.b.1/2$  e  $\sigma^2_0(U) = a.b.(a+b+1)/12$ . Uma hipótese alternativa aceitável é de que  $p > 1/2$ . Sabendo-se que  $a$  é menor do que  $n$ , pode-se escrever  $a$  como uma parte de 'n'. Assim:  $a = c.n$ , onde 'c' é 'a' fração correspondente de 'a' em 'n'. Logo, a estatística do teste pode ser escrita como (Noether, 1987; Gangnon e Kosorok, 2004; Shieh et al, 2008; Zhao et al, 2008):

$$Q(U) = \frac{12 \cdot c \cdot (1 - c) \cdot n^2 \cdot (p - 1/2)}{(n + 1)}$$

Dáí, pode-se reescrever essa fórmula, isolando  $n$  e chegando a:

$$\frac{n^2}{(n + 1)} = \frac{Q(U)}{12 \cdot c \cdot (1 - c) \cdot (p - 1/2)}$$

Como  $Q(U)$  tem distribuição aproximadamente normal, pode-se escrevê-la como:

$$Q(U) = \frac{(z_\alpha + z_\beta)^2}{(p - \frac{1}{2})}$$

Substituindo-se  $Q(U)$ , e considerando que  $n$  é aproximadamente igual a  $(n+1)$ , temos:

$$n = \frac{(z_\alpha + z_\beta)^2}{12 \cdot c \cdot (1 - c) \cdot (p - \frac{1}{2})^2}$$

Vollandt e Horn (1997) respaldam o resultado alcançado por Noether, uma década antes, e propõem uma simplificação na fórmula, transformando-a em:

$$n = \frac{(z_\alpha + z_\beta)^2}{6 \times (p - \frac{1}{2})^2}$$

#### (4) Diferença entre medianas — *Teste do Sinal*

Seja 'p' a proporção de valores positivos (acima da mediana); e sejam considerados os erros Tipo I e Tipo II (Noether, 1987). Assim, a fórmula de estimação para o tamanho de amostra com base no *Teste do Sinal* pode ser escrita como:

$$n = \frac{(z_\alpha + z_\beta)^2}{4 \times (p - \frac{1}{2})^2}$$

O *Teste do Sinal* é comumente substituído pelo *Teste*



*dos Postos Sinalizados de Wilcoxon*, para a comparação entre dados não-paramétricos pareados. Para esse último teste, Noether (1987) adaptou a fórmula anterior, tendo deduzido a seguinte fórmula, que também é sugerida por Julious e Campbell (1998):

$$n = \frac{(z_{\alpha} + z_{\beta})^2}{3 \times (p - \frac{1}{2})^2}$$

onde:

p — proporção de diferenças entre as duas observações que são positivas

(5) Diferença concomitante de três ou mais grupos, por meio de médias ou de postos (*ranks*)

A comparação entre mais de dois grupos pode, às vezes, ser encarada de forma errônea: é comum pensar-se sobre a possibilidade de fazer-se comparações par a par de todos os grupos, mas, essa técnica deve ser precedida pela comparação concomitante entre todos os grupos. Assim, uma comparação adequada deve pleitear a possibilidade de contraposição de todos os grupos, e depois, quando possível, a comparação par a par entre os grupos, mas, somente, se a diferença apontada na comparação concomitante for estatisticamente significativa (Boos e Brownie, 1995). Os dois testes estatísticos mais usuais, empregados nesta situação, são a *Análise de Variância (ANOVA)* e

o *Teste de Kruskal-Wallis*, respectivamente, teste paramétrico e não-paramétrico.

Boos e Brownie (1995) apresentam uma solução para a estimação do tamanho da amostra, quando a ANOVA é aplicável:

$$n = \frac{F \times (k - 1) \times \sum_{i=1}^k s_i^2}{\sum_{i=1}^k (\bar{X}_{i.} - \bar{X}_{..})^2}$$

onde:

n — tamanho da amostra

F — valor estimado da estatística F

k — número de grupos considerados

$s_i$  — estimador do desvio-padrão o i-ésimo grupo

$\bar{X}_{i.}$  — média do i-ésimo grupo

$\bar{X}_{..}$  — média das médias dos grupos (grande média)

Quando os postos (*ranks*) estão presentes e uma ANOVA não pode ser aplicada, para cobrir um experimento com mais de dois grupos a serem comparados, uma fórmula simplificada, sugerida por Brownie e Boos (1994), Mahoney e Magel (1996), e, Ahnn e Anderson (1998), pode ser usada:

$$n = \frac{k \cdot \sum_{i=1}^k d_i}{\sum_{i=1}^k P_{E_i}}$$

onde:

n — tamanho da amostra

k — número de grupos estudados

$d_i$  — número de eventos esperado em cada um dos k grupos

$P_{E_i}$  — probabilidade de ocorrência do evento, em cada um dos k grupos

Como o *Teste de Kruskal-Wallis* não permite imediata solução para o cálculo do tamanho da amostra; com isso, vale a pena relatar que mais recentemente, Rasch e Šimečková (2007) propuseram um ajuste relativamente complexo para os casos em que muitos empates (postos com mesmos valores) encontrados; a fórmula para essa proposta não será mostrada, aqui, mas repise-se sua complexidade e seu uso específico frente numa situação com um grande número de postos iguais.

(6) Diferença entre dois momentos de observação para uma variável binomial (*Teste de McNemar*)

A estrutura deste teste apresenta-se com a seguinte configuração, em forma de tabela-resumo:

Variável (momento 1)	Variável (momento 2)		Total
	cat 0	cat 1	
cat 0	$p_{00}$	$p_{01}$	$p_{0.}$
cat 1	$p_{10}$	$p_{11}$	$p_{1.}$
Total	$p_{.0}$	$p_{.1}$	$p_{..}$

O *Teste de McNemar* permite a comparação entre dois momentos distintos, de uma variável binária. Feuer e Kessler (1989) e Gönen (2004) propõem a seguinte fórmula, que permite calcular o tamanho da amostra, com base nesse teste:

$$n = \frac{\left( z_{\alpha} \times (p_{01} + p_{10}) + z_{\beta} \times \sqrt{(p_{01} + p_{10})^2 - \frac{1}{4}(p_{01} - p_{10})^2(p_{01} + p_{10} + 3)} \right)^2}{(p_{01} + p_{10}) \times (p_{01} - p_{10})^2 / [1 + (p_{01} + p_{10} - 1) \times (1 + \alpha_{01} + \alpha_{10})^{-1}]}$$

onde:

n — tamanho da amostra

$z_{\alpha}$  — valor da estatística z, para um nível de confiança alfa

$z_{\beta}$  — valor da estatística z, para um valor de poder (1- $\beta$ )

$p_{01}$  — proporção observada

$p_{10}$  — proporção observada

$\alpha_{01}$  — valor absoluto observado

$\alpha_{10}$  — valor absoluto observado

## (7) Processos de Poisson

Muito comuns, em diversas áreas de planejamento e modelagem epidemiológica, os números de espécimes sujeitos à alterações em seus quadros clínicos são usados para a construção de modelos de previsão e controle de endemias / epidemias. Sob esse aspecto, amostras podem ser coletadas e os modelos que levam em conta as alterações clínicas dos sujeitos expostos e não-expostos, por vezes, têm a forma de um processo de Poisson, com taxa conhecida, ou, minimamente, passível de estimação (Van Belle, 2008).

O tamanho da amostra, que permite a comparação de duas populações com distribuição de Poisson, pode ser, simplificada, dado por:

$$n = \frac{(z_{\alpha} + z_{\beta})^2}{(\sqrt{\theta_1} - \sqrt{\theta_2})^2}$$

onde:

n — tamanho da amostra

$z_{\alpha}$  — valor da estatística z, para um nível de confiança alfa

$z_{\beta}$  — valor da estatística z, para um valor de poder (1- $\beta$ )

$\theta_1$  — taxa (média) da amostra 1

$\theta_2$  — taxa (média) da amostra 2

### 1.4.3.3 Relacionamentos

Os estudos de relacionamento, em Estatística, podem ser resumidos com as expressões associação, correlação e regressão; esses estudos permitem avaliar um conjunto de duas ou mais variáveis, formando pares ou blocos de variáveis (com três ou mais variáveis concomitantemente), que estão relacionadas entre si (ou não), mas, que podem ser explicadores, em conjunto, do comportamento de um sistema biológico.

#### (1) Correlação de um par de variáveis

Uma das avaliações estatísticas mais frequentes, mencionada em diversos trabalhos, como o de Bonett e Wright (2000), permite calcular o tamanho da amostra tendo, como base, o coeficiente de correlação de Pearson:

$$n = 2 + \left[ \frac{z_{\alpha} + z_{\beta} \cdot \sqrt{1 - r^2}}{r} \right]^2$$

onde:

$n$  — tamanho da amostra

$z_{\alpha}$  — valor da estatística  $z$ , para um nível de confiança alfa

$z_{\beta}$  — valor da estatística  $z$ , para um valor de poder  $(1-\beta)$

$r$  — valor estimado do coeficiente de correlação, com  $r \neq 0$

## (2) Qui-quadrado

Um dos testes mais usuais apresenta uma forma de representação bastante comum (Woolson et al, 1986; Thomas e Conlon, 1992):

Grupo	Evento		Total
	ocorre	não ocorre	
Estudo	$p_{00}$	$p_{01}$	$p_{0.}$
Controle	$p_{10}$	$p_{11}$	$p_{1.}$
Total	$p_{.0}$	$p_{.1}$	$p_{..}$

Com base nessa forma, onde vê-se as proporções indicadas em cada célula da tabela-resumo, o tamanho da amostra pode ser calculado por meio da seguinte fórmula:

$$n = \left( \frac{z_{\alpha} \times \sqrt{2 \times p_{00}p_{10} \times (1 - p_{00}p_{10})} + z_{\beta} \times \sqrt{p_{00}p_{01} + p_{10}p_{11}}}{p_{10} - p_{00}} \right)^2$$

## (3) Diferença entre curvas de regressão de dois grupos e diferença de correlação (baseado na correlação de Pearson)

Um problema relativamente comum concerne ao cálculo de correlações, em estudos de relacionamento entre pares de variáveis. As correlações permitem concretizar numericamente o estudo do comportamento entre variáveis, e de modo extensivo,

para tantos pares quantos sejam de interesse à pesquisa. As regressões, por sua vez, permitem estudar-se o comportamento concomitante entre várias variáveis.

Para a diferença entre duas correlações, pode-se considerar a seguinte fórmula básica (Zar, 1984):

$$z_{\alpha} = \frac{[\tanh]^{INV} r_1 - [\tanh]^{INV} r_2}{\sqrt{\frac{2}{n-3}}}$$

onde:

$n$  — tamanho da amostra

$z_{\alpha}$  — valor da estatística  $z$ , para um nível de confiança alfa

$[\tanh]^{INV} r_i$  — tangente hiperbólica inversa de  $r_i$ , com  $i = 1, 2$

Assim, tem-se que:

$$n = \frac{2 \cdot z^2}{([\tanh]^{INV} r_1 - [\tanh]^{INV} r_2)^2} + 3$$

Lachin (1981) constrói a seguinte proposta, com base no valor da correlação:

$$n = \frac{(z_{\alpha} + z_{\beta})^2}{\frac{1}{2} \times \ln\left(\frac{1+r}{1-r}\right)} + 3$$

$n$  — tamanho da amostra

$z_{\alpha}$  — valor da estatística  $z$ , para um nível de confiança alfa

$z_{\beta}$  — valor da estatística  $z$ , para um valor de poder  $(1-\beta)$

$r$  — valor estimado da correlação



## (4) Odds-ratio

Considere-se a seguinte tabela-resumo:

Grupo	Evento		Total
	ocorre	não ocorre	
Estudo	$p_{11}$	$p_{12}$	$p_{1.}$
Controle	$p_{21}$	$p_{22}$	$p_{2.}$
Total	$p_{.1}$	$p_{.2}$	$p_{..}$

Sahai e Khurshid (1996) propõem que o tamanho da amostra, em função de um valor de odds-ratio previamente estimado pode ser calculado por:

$$n = (z_{\alpha} + z_{\beta})^2 \times \frac{p_{.1} \times (1 - p_{1.})}{p_{1.} \times (1 - p_{.1})} \times \left( 1 + \frac{1}{(\psi - 1) \times p_{.1}} \right)^2$$

onde:

$n$  — tamanho da amostra

$z_{\alpha}$  — valor da estatística  $z$ , para um nível de confiança alfa

$z_{\beta}$  — valor da estatística  $z$ , para um valor de poder  $(1-\beta)$

$\psi$  — valor estimado do odds-ratio, com  $\psi \neq 1$

Com base na fórmula sugerida por Guenther (1981), Julious e Campbell (1998) desenvolveram uma fórmula para o cálculo do tamanho da amostra, mas, usando o valor estimado do odds-ratio:

$$n = \frac{(z_{\alpha} \times (\psi + 1) + 2 \times z_{\beta} \sqrt{\psi})^2}{(\psi - 1)^2}$$

onde:

$n$  — tamanho da amostra

$z_{\alpha}$  — valor da estatística  $z$ , para um nível de confiança alfa

$z_{\beta}$  — valor da estatística  $z$ , para um valor de poder  $(1-\beta)$

$\psi$  — valor estimado do odds-ratio, com  $\psi \neq 1$

#### 1.4.3.4 Comentários sobre a abordagem bayesiana

Antes de fazer alguns comentários finais sobre este levantamento bibliográfico de vários tipos de cálculo de tamanho de amostra, é preciso mencionar, ainda, o cálculo sob o foco bayesiano. As estatísticas de teor bayesiano são comumente usadas, e como se verá, a seguir, a base de cálculo para o tamanho da amostra contém parâmetros da Teoria Bayesiana, porém, o que se altera especificamente no cálculo do tamanho da amostra são os estimadores a serem usados, mas não a essência (forma) das fórmulas.

Veja-se, um caso mencionado por Adcock (1989 e 1997): são comparados os métodos frequentista e bayesiano, para a estimação do tamanho da amostra, quando se deseja construir um estudo-piloto, onde a amostra está dividida em estratos. O que interessa ser comentado é que, sob ambos os focos (frequentista e bayesiano), as fórmulas do tamanho da amostra dependem, exclusivamente, do número estratos, do número de eventos esperados em cada estrato e da probabilidade de ocorrência do evento em cada estrato; esses valores são estimados de modos diferentes, porém, chega-se a fórmulas cuja *forma* é semelhante. Esse resultado é respaldado por Inoue et al (2005), que afirmam, categoricamente: *[para o cálculo do tamanho da amostra,]*

*bayesianos e frequentistas não são muito diferentes!*

Joseph (1997) comenta sobre os critérios a serem adotados para o cálculo do tamanho da amostra de teor bayesiano, que, no geral, diferem dos critérios frequentistas, mas, que se aproximam (e muito), quando se alcançam as formas das fórmulas de tamanho de amostra. Assim, o autor mostra que, para o cálculo do tamanho da amostra onde o interesse está assentado sobre a diferença entre duas proporções, a fórmula, semelhantemente àquela vista no tem 'B2', é dada por:

$$n = \frac{4 \times z_{\alpha} \times [\pi_1(1 - \pi_1) + \pi_2(1 - \pi_2)]}{(\pi_1 - \pi_2)^2}$$

onde:

$n$  — tamanho da amostra

$z_{\alpha}$  — valor da estatística  $z$ , para um nível de confiança alfa

$\pi_1$  — proporção de ocorrência do evento 1

$\pi_2$  — proporção de ocorrência do evento 2

Os valores de  $\pi_1$  e de  $\pi_2$  são estimados, com base em suas distribuições binomiais *a priori*, fazendo  $\theta = \pi_1 - \pi_2$ , ou seja:

$$f(\theta | x_1, x_2, n_1, n_2) = \int_{\max(0; \theta)}^{\min(\theta+1; 1)} f(\pi_1, \theta | x_1, x_2, n_1, n_2) d\pi_1$$

Adcock (1992) sugere que a estimação dos parâmetros de teor bayesiano seja feita por meio de estudo-piloto, e simplifica uma das fórmulas mais usuais para o cálculo do tamanho de

amostras:

$$n = \frac{0,25 \times \chi_{(1); \theta}^2}{d^2}$$

onde:

$n$  — tamanho da amostra

$\chi^2$  — valor da estatística de qui-quadrado, com um grau de liberdade e o valor estimado de  $\theta$  para a variável principal

$d$  — valor do erro relativo admissível

Uma variante da fórmula (acima) pode ser usada para o cálculo do tamanho da amostra de uma pesquisa em que a Distribuição Multinomial seja considerada (Adcock, 1995 e 1997), onde o parâmetro  $\theta = k$  deve ser o número de classes:

$$n = \frac{\frac{1}{k} \times \left(1 - \frac{1}{k}\right) \times \chi_{(1); \alpha/k}^2}{d^2}$$

O foco deste trabalho não está voltado para uma abordagem sobre as diferenças entre as avaliações frequentista e bayesiana. Sendo assim, o que desejou-se mostrar condiz com o fato de que as *formas* das fórmulas para cálculo de tamanho de amostra de uma corrente e de outra aproximam-se sensivelmente; os estimadores e os processos de estimação de valores a serem usados em cada fórmula diferem-se pelas duas naturezas das correntes estatísticas mencionadas.

Outra questão não abordada, neste trabalho, diz

respeito ao custo do processo de amostragem, que, em geral, traz limitações às pesquisas. Nos casos vistos, foi suposta a não-limitação do número de elementos amostrais a ser coletado devido ao custo. Esse custo, por vezes, precisa ser levado em conta, mas, a intenção deste trabalho não é a de determinar nem comentar sobre fatores limitantes eventualmente presentes no processo de amostragem.

Duas observações finais merecem ser comentadas: a primeira, sobre o poder dos testes, e a segunda sobre a caudalidade das alternativas adotadas em testes estatísticos.

Na prática, é usual que o menor valor de poder do teste a ser adotado valha 80%, para a estimação do tamanho da amostra. O poder de um teste é definido como sendo o complemento da probabilidade do Erro Tipo II ( $1-\beta$ ), e o Erro Tipo II é a *probabilidade de aceitar a hipótese nula, dado que a hipótese alternativa é verdadeira*. Já, o Erro Tipo I é a *probabilidade de rejeitar a hipótese nula, dado que ela é verdadeira* (Rosner, 1986). Como o que se deseja obter em termos do tamanho de uma amostra a ser investigada é que ela seja representativa da população de onde será coletada, então o ideal é que o valor de poder do teste a ser adotado seja o maior possível, ou, o valor do Erro Tipo II seja o menor possível (Adcock, 1997).

Nas fórmulas mostradas, faz-se referência ao nível de significância 'alfa', e não a 'alfa/2'; em todos os casos (sem exceção), os autores demonstram que a unicaudalidade pode ser substituída pela bicaudalidade, dependendo, apenas, das assertivas previamente definidas na aplicação dos testes estatísticos escolhidos e/ou na adoção de critérios de planejamento amostral, quando da escolha do teste estatístico.

Neste trabalho, não se teve a pretensão de coletar todo e qualquer tipo de cálculo de tamanho de amostra: os tipos anteriormente expostos são comuns e usuais; há muitos mais, por conta de que cada planejamento amostral, em uma pesquisa séria, permite avaliações as mais diversas, em função de selecionar-se o mais adequado procedimento para a estimação do tamanho da amostra que lhe será útil.

## 2 Objetivos

Com este trabalho propõe-se a desenvolver um novo método para o cálculo do tamanho de amostra a partir de questionários ou escalas comumente usados em pesquisa na área de saúde.

De modo complementar, propõe-se a testar o novo método, por meio de simulações, e comparar os resultados obtidos com outros métodos já em uso.

Limites deste estudo: tem-se a intenção de não abordar dois aspectos do cálculo do tamanho da amostra: (a) variabilidade do parâmetro escolhido da variável principal; e (b) custo do processo de amostragem, ou, do elemento amostral. Estes dois itens fazem parte, comumente, de estudos sobre cálculos de tamanho de amostra, e são, por vezes, importantes, quando: (a) um único parâmetro de uma variável principal deverá ser usado para o cálculo do tamanho da amostra; e (b) envolvem-se custos de caráter financeiro em uma ou mais partes do processo de coleta dos dados (contratação de pessoal de pesquisa de campo; contratação de pessoal para coleta de material; uso de equipamentos



e/ou kits para coleta; etc.). Neste estudo, não serão abordadas essas duas questões, pois, no caso da primeira, não se tem um parâmetro único, eleito para o cálculo proposto, e no caso da segunda, não se tem interesse, *a priori*, sobre o custo do processo amostral, que consiste em matéria especial e complementar.

Além desses dois aspectos, a reposição e a não-reposição de elementos amostrais, bem como informações referentes aos chamados erros de amostragem (planejamento amostral) não serão abordados neste trabalho.

### 3 Métodos

Os eixos condutores para o desenvolvimento da nova proposta de cálculo de tamanho de amostra estarão assentados sobre dois princípios, a saber: a Lógica Paraconsistente Anotada Evidencial e a Teoria Combinatória.

#### 3.1 Lógica Paraconsistente Anotada Evidencial

Sob a óptica da Lógica Clássica, um item, constante de um instrumento de coleta, *tem de ser* respondido logicamente (classicamente falando), enquanto que, pelo raciocínio humano, isso não é obrigatório, ou seja, um item *pode ser* respondido logicamente (sob os termos da Lógica Clássica).

Ainda, pela Lógica Clássica, para responder ao item, é *preciso* ter conhecimento (ou, espera-se que se tenha); já, com base no raciocínio humano, nem sempre se tem o conhecimento, no momento de responder ao item.

Tanto a Lógica Clássica como o raciocínio humano apresentam dois pontos de toque importantes: (a) para responder

ao item, é preciso *decidir* que se vai responder; e (b) pode-se decidir não responder ao item, por ignorar a resposta ou por falta de vontade — ambos têm caráter decisório, o que não implicaria, *a priori*, uma base lógica constante e única.

O preenchimento indevido (a escolha de mais de uma resposta) e a 'não-resposta' *causam problemas técnicos clássicos*; em contrapartida, por meio do raciocínio humano, tanto o preenchimento dito indevido e a não-resposta não causam, *a priori*, problemas técnicos, visto que um problema técnico surge depois da obtenção da resposta, e não durante o processo de obtenção.

Mesmo toda Lógica sendo binária, a dimensionalidade das lógicas varia, em termos de que elas conseguem traduzir, de modo binário e/ou multidimensional, suas características próprias. Nesse sentido, pode-se dizer que o raciocínio humano apresenta dimensão multifacetada.

Considerando essas informações, podemos, então, perceber que a Lógica Clássica, por sua importância, consiste em base para o raciocínio humano, mas este conseguiu tornar-se mais abstrato, e criar estruturas mais próximas de sua forma, em função das necessidades do próprio raciocínio humano, amoldando-o em uma espécie de padrão lógico-humano, ou seja, um padrão que leva em conta as características multifacetadas desse raciocínio.

A figura 1 evidencia a contrapartida entre a Lógica Clássica, o raciocínio humano e as possibilidades de resposta de um item de um instrumento de coleta, com base na estrutura de conhecimento do ser humano e da tomada de decisão em responder (ou não) a esse item. Logo, cada braço do esquema de árvore dessa figura mostra uma possibilidade possível e esperada, quando da aplicação de um instrumento de coleta.

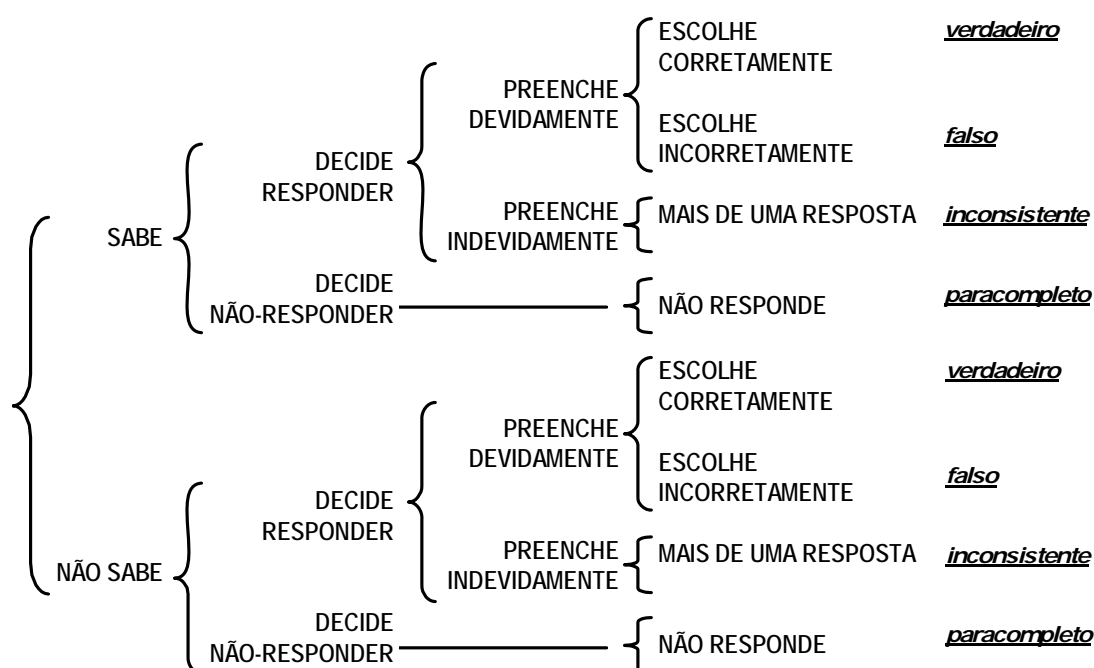


Fig. 1. Esquema didático-lógico, demonstrando a construção do raciocínio humano e sua inter-relação direta frente aos estados lógicos da Lógica Paraconsistente Anotada Evidencial.

Pela Lógica Clássica, pode-se afirmar que somente consideram-se as respostas preenchidas devidamente; e que não é

possível levar-se em consideração as *não-respostas* e os *preenchimentos indevidos*. Já, pelo raciocínio humano, consideram-se as respostas preenchidas e as *intenções* de responder, além de ser possível abordar o problema criado pelas *não-respostas* e pelos *preenchimentos indevidos*.

Por meio da aplicação da Lógica Paraconsistente Anotada Evidencial, pode-se estender as possibilidades de resposta de cada item do instrumento de coleta, pois:

- (a) se o preenchimento for *devido*, pode-se computar as respostas como na Lógica Clássica; e com isso, tem-se os estados lógicos *verdadeiro* e *falso* (respostas ditas *corretas* e *incorretas*).
- (b) se o preenchimento for *indevido*, é possível realizar a computação das *não-respostas* como estado paracompleto – *falta de informação*; e é possível realizar a computação dos erros de preenchimento como estado inconsistente – *excesso de informações*.

Três aspectos da Lógica Paraconsistente, em função das possibilidades de resposta de um item de um instrumento de coleta, podem ser resumidos como se segue, e serão usados, como

ideias básicas, para a proposta do cálculo do tamanho da amostra constante do capítulo de resultados deste trabalho:

- (a) permite a absorção de *situações* não previstas pela Lógica Clássica;
- (b) permite a computação de *todos os tipos* de resposta possíveis (humanamente falando); e
- (c) cria uma estrutura de *estados lógicos*, compatíveis com os tipos de resposta, estrutura essa bem mais próxima das possibilidades de *teor humano*, em relação à tomada de decisões; e

### 3.2 Combinatória

A solução para o problema do cálculo do tamanho da amostra pode ser vislumbrada por meio de técnicas combinatórias, em função do número de elementos populacionais disponíveis, e do tamanho do instrumento de coleta de dados, em função das características de cada variável aleatória que compõe o instrumento. Esquemáticamente, pode-se escrever:

$$n \sim f[N_d, I(v_1, v_2, \dots, v_k)]$$

onde:

n — tamanho da amostra

$f$  — função cujos elementos são usados para o cálculo do tamanho da amostra

$N_d$  — número de elementos populacionais disponíveis

$I$  — função que identifica os tipos e quantidades das ' $k$ ' variáveis aleatórias componentes do instrumento de coleta

$v_i$  — variáveis aleatórias componentes do instrumento de coleta, com  $i = 1, \dots, k$

Os itens formadores do instrumento de coleta e suas respectivas categorias podem ser, então, combinados, de modo a que todas as possibilidades de preenchimento do instrumento sejam abarcadas. As categorias, obedecendo à suas pertinências a cada item, podem ser selecionadas, independentemente da ordem em que isso ocorra. O conjunto de categorias escolhidas sequencialmente forma um *caminho* (maneira de responder ao instrumento), que foi percorrido pelo respondente, dos itens do instrumento, e o número de caminhos possíveis pode ser calculado, com certa facilidade, com o uso da fórmula de combinação. Exemplo simples: o instrumento é formado por dois itens e cada item é binário; há, portanto, quatro categorias disponíveis; a combinação das quatro categorias, duas a duas, menos a combinação entre as categorias de um mesmo item (visto que, dentro de um mesmo item, apenas uma categoria pode ser selecionada), fornece o número de possíveis maneiras para preencher-se esse instrumento; assim, sendo ' $C$ ' o número de combinações possíveis de caminhos (maneiras de responder ao

instrumento), temos, matematicamente que:

$$C = \frac{4!}{2!(4-2)!} - 2 \times \frac{2!}{2!(2-2)!} = \frac{4 \cdot 3 \cdot 2!}{2!2!} - 2 \times \frac{2!}{2!} = 6 - 2 = 4$$

O mesmo princípio será usado para calcular-se o tamanho da amostra, tomando-se, conjuntamente, as regras da Lógica Paraconsistente Anotada Evidencial.

### 3.3 Tamanho da Amostra

Para a efetivação do cálculo do tamanho da amostra proposto neste trabalho, uma primeira restrição deve ser feita: os itens formadores do instrumento de coleta devem ser formados por categorias mutuamente exclusivas. Exemplo: um item pode ser formado por três níveis cujas denominações são 'leve', 'moderado' e 'grave'; ou, por dois níveis, como o 'sim' e o 'não'. Essas categorias são, obrigatoriamente, complementares e exclusivas; quando existem itens em que as categorias não são complementares por suas naturezas, os mesmos devem ser considerados, para efetivação do cálculo do tamanho da amostra, como fornecedores de mais possibilidades de escolha; assim, por exemplo, um item em que as seguintes categorias estejam presentes deve ser transformado: doenças associadas — 'diabetes',



'pressão alta', e 'colesterol'; neste caso, cada uma das categorias deve fornecer a possibilidade de ser ou não ser escolhida; logo, cada categoria deve ser encarada como se fosse um item do tipo sim-não.

As escalas de valores contínuos formam a segunda restrição: essas variáveis devem ser transformadas em categóricas, mesmo que, durante o processo de avaliação dos dados, elas venham a ser usadas como contínuas.

Uma última restrição deve ser considerada: itens que consistem em textos ou que possuem textos complementares não são levados em conta para o cálculo do tamanho da amostra. Eventualmente, algum item textual pode ser transformado em categorias, mas, em geral, não é o que se espera da aplicação de instrumentos em que informações textuais complementares vêm para dar apoio às respostas apontadas em itens categóricos ou categorizados.

## 4 Proposta para o Novo Modelo e Resultados

Neste capítulo, apresenta-se a construção do modelo para a nova proposta de cálculo de tamanho de amostra, com base nos itens componentes do instrumento de coleta, bem como alguns resultados da aplicação desse novo método.

### 4.1 Nova Proposta para o Cálculo do Tamanho da Amostra

Vai-se, de início, expor a seguinte situação: está-se à procura de um número que represente o tamanho da amostra. Deve-se mostrar que esse número pode ser aceito como um tamanho de amostra. Para tanto, considere-se o número total de caminhos possíveis (formados por todas as possibilidades de respostas a serem consideradas em itens categóricos de um instrumento de coleta). Esse número é, por certo, um candidato a ser um tamanho de amostra, pois permite que todas as possibilidades de resposta sejam observadas. Tome-se este exemplo: se se tem um item formado por duas categorias ditas *originais*, então quatro respondentes, *a priori*, seriam suficientes,

para ter-se uma *amostra inicial*, pois há quatro possibilidades esperadas, aquelas que serão chamadas de *categorias efetivas*: a escolha da primeira categoria efetiva, a escolha da segunda categoria efetiva, a escolha de ambas as categorias efetivas concomitantemente, e a escolha de nenhuma categoria efetiva, conforme foi exposto nas ideias metodológicas. Logo, quatro é um tamanho amostral sugestivo para essa proposta. Extensivamente, se se tiver uma lista de itens com categorias originais previamente definidas, será preciso um número suficientemente grande de elementos para abarcar todas as possibilidades de preenchimento dos itens dessa lista. Assim, esse número de elementos pode ser considerado, também, como o número de elementos amostrais, ou seja, como um tamanho de amostra. O número de categorias efetivas de um item pode ser calculado por:  $2^k$ , onde  $k$  é o número de categorias originais.

Se se puder aceitar que um tamanho de amostra inicial seja considerado como o número de elementos componentes que cobrem todos os caminhos possíveis (combinações possíveis) formados pelos itens e suas respectivas categorias efetivas, então pode-se dar um próximo passo para o cálculo do tamanho da amostra com base na estrutura do instrumento de coleta. Essa suposição é fundamental, neste momento, mas deve receber uma crítica: ela não contempla uma possível diferenciação de

representatividades entre as categorias efetivas de cada item, o que seria esperado, porque, em geral, alguns caminhos são mais frequentes do que outros; além disso, alguns caminhos podem, inclusive, ser impossíveis de serem observados. Porém, a técnica de cálculo de tamanho de amostra que está sendo proposta parece numericamente viável, pois o número de caminhos possíveis pode, efetivamente, ser considerado como um tamanho de amostra, em função do crescimento do número de itens, e por consequência natural, em função do crescimento do número de categorias efetivas.

Esse aceite tem teor subjetivo e é intuitivo. Se se pensar em um estado real, isso pode ser aceito, de imediato, porque os instrumentos de coleta de dados médicos podem ter vários itens compostos por várias categorias originais — em geral, não menos do que algumas dezenas de itens binários (com duas categorias), ternários (com três categorias), quaternários (com quatro categorias) etc. são comuns nesses instrumentos de coleta. Assim, intuitivamente, pode-se aceitar que um tamanho de amostra inicial possa ser considerado como sendo o número de caminhos possíveis a ser observado, e cuja base assenta-se sobre o número de categorias efetivas do instrumento de coleta adotado.

Matematicamente, pode-se representar esse tamanho

de amostra inicial como mostra a fórmula [4.1.1]:

$$n_1 = \binom{c^E}{2} - \sum_{i=1}^k \binom{c_i^E}{2} \quad [4.1.1]$$

onde:

$n_1$  — tamanho da amostra inicial

$c^E$  — número de categorias efetivas do instrumento de coleta

$c_i^E$  — número de categorias efetivas do i-ésimo item

$k$  — número de itens do instrumento de coleta

Ainda, considerando a prática, sabe-se que alguns desses caminhos podem aparecer mais do que outros. Assim, existe um *grau de similaridade* entre os respondentes do instrumento de coleta, que, por sua vez, podem produzir, então, caminhos iguais, permitindo que alguns caminhos sejam mais representativos do que outros (Borg e Groenen, 2005). Considere-se o grau de similaridade como sendo um número entre zero e um, representando, portanto, um percentual: se for um, então a similaridade é total, ou seja, todos respondem aos itens do instrumento de coleta de modo igual; se for zero, então a similaridade é nenhuma, ou seja, todos respondem, de modo diferente uns dos outros, aos itens do instrumento de coleta.

Assim, pode-se escrever que:

$$n_2 = (1 - g^S) \times n_1 \quad [4.1.2]$$

onde:

$n_2$  — tamanho da amostra considerando o grau de similaridade

$g^S$  — grau de similaridade

$n_1$  — tamanho da amostra inicial

Na prática, os graus de similaridade são valores que se afastam de zero e de um, e portanto, estão no entremeio desses dois extremos. Sendo assim,  $n_2$  será, sempre, menor do que  $n_1$ . Com isso, conhecendo-se o grau de similaridade, calcula-se o tamanho da amostra. O problema reside no fato de que esse grau de similaridade é, na prática, desconhecido, apesar de passível de estimação. Logo, deseja-se buscar um estimador para o grau de similaridade.

Intuitivamente, como foi dito, tanto maior seja o instrumento de coleta, e menor será o grau de similaridade esperado, pois, supõe-se que quanto maior o número de possibilidades diferentes de caminhos a serem percorridos, tanto menor será a similaridade entre os elementos componentes da amostra. E o número de possibilidades diferentes de caminhos é tanto maior, quanto maior for o número de itens do instrumento de coleta; e ainda, quanto maior for o número de itens do instrumento de coleta, tanto maior será o número de categorias originais disponíveis desse instrumento. Assim, pode-se afirmar que o grau de similaridade apresenta-se como inversamente proporcional ao

número de categorias originais do instrumento de coleta. Então, pode-se escrever que:

$$1 - g^S \approx \frac{1}{c^0} \quad [4.1.3]$$

onde:

$g^S$  — grau de similaridade

$c^0$  — número total de categorias originais do instrumento de coleta

Como, para um mesmo número de itens de um instrumento de coleta, o fator que permite o aumento de caminhos é o número de categorias originais de cada item, pode-se considerar o número total dessas categorias do instrumento de coleta como sendo o estimador ideal para ser usado na estimação do grau de similaridade. Assim, pode-se reescrever:

$$n_2 = \frac{1}{c^0} \times n_1 \quad [4.1.4]$$

onde:

$n_2$  — tamanho da amostra considerando o grau de similaridade

$c^0$  — número total de categorias originais do instrumento de coleta

$n_1$  — tamanho da amostra inicial

Com isso, o tamanho da amostra poderá ser calculado do seguinte modo:

$$n = \frac{\binom{c^E}{2} - \sum_{i=1}^k \binom{c_i^E}{2}}{c^O} \quad [4.1.5]$$

onde:

$c^E$  — número de categorias efetivas do instrumento de coleta

$c_i^E$  — número de categorias efetivas do i-ésimo item

$k$  — número de itens do instrumento de coleta

$c^O$  — número total de categorias originais do instrumento de coleta

Ou, de outro modo:

$$n = \frac{\binom{c^E}{2} - \sum_{i=1}^k \binom{c_i^E}{2}}{\sum_{i=1}^k c_i^O} \quad [4.1.6]$$

onde:

$c^E$  — número de categorias efetivas do instrumento de coleta

$c_i^E$  — número de categorias efetivas do i-ésimo item

$k$  — número de itens do instrumento de coleta

$c_i^O$  — número de categorias originais do i-ésimo item

Assim, o numerador da fórmula [4.1.6] tem sua base assentada sobre o número de categorias efetivas (de teor paraconsistente), enquanto que o denominador é caracterizado pelo número de categorias originais (de teor clássico).

Para amostras ditas pequenas, e/ou, para cálculos de tamanho de amostra em que seja necessário, por razões técnicas, levar em conta o tamanho da população da qual a amostra será



coletada, pode-se ajustar o cálculo acima, considerando a fórmula

[4.1.7]:

$$n = \frac{\frac{\binom{c^E}{2} - \sum_{i=1}^k \binom{c_i^E}{2}}{\sum_{i=1}^k c_i^O}}{1 + \frac{1}{N} \times \left[ \frac{\binom{c^E}{2} - \sum_{i=1}^k \binom{c_i^E}{2}}{\sum_{i=1}^k c_i^O} - 1 \right]} \quad [4.1.7]$$

onde:

$c^E$  — número de categorias efetivas do instrumento de coleta

$c_i^E$  — número de categorias efetivas do i-ésimo item

$k$  — número de itens do instrumento de coleta

$c_i^O$  — número total de categorias originais do i-ésimo item

$N$  — tamanho da população

Pela fórmula [4.1.7], quando o tamanho da população for relativamente pequeno, haverá um ajuste a ser levado em conta, e tanto maior seja o tamanho da população, o ajuste será cada vez menor, fazendo com que a fórmula [4.1.7] possa ser substituída pela fórmula [4.1.6].

## 4.2 Resultados — alguns cálculos demonstrativos

Três partes constituem este item: cálculo do tamanho da amostra para dois instrumentos muito conhecidos e de largo uso, que são o SF-36 e o WHOQOL, e uma parte final, onde 23 instrumentos da área de Pneumologia foram selecionados, para cálculos de tamanho de amostra, como demonstração geral da proposta feita no item 4.1.

### 4.2.1 Cálculo do tamanho da amostra para o SF-36

Um exemplo de instrumento de coleta bastante comum é o SF-36 (Ciconelli et al, 1999), que investiga a qualidade de vida de pacientes submetidos a quaisquer tratamentos, sendo composto por onze macro-itens subdivididos em 36 itens que, por sua vez, são compostos por sete itens com duas categorias, dez itens com três categorias, nove itens com cinco categorias, e dez itens com seis, totalizando os 36 itens mencionados.

Serão realizados os cálculos do tamanho da amostra, em dois casos: o primeiro, em que a população da qual a amostra será coletada seja muito grande, e o segundo, uma população de tamanho igual a 200.

Os valores básicos a serem calculados para uso nas

fórmulas de cálculo do tamanho da amostra são:

$$c^E \text{ — número de categorias efetivas do instrumento de coleta} \\ 7 \times 2^2 + 10 \times 2^3 + 9 \times 2^5 + 10 \times 2^6 = 1036$$

$$c_i^E \text{ — número de categorias efetivas do } i\text{-ésimo item} \\ 4, 8, 32 \text{ e } 64$$

$$k \text{ — número de itens do instrumento de coleta} \\ 36$$

$$c_i^O \text{ — número de categorias originais do } i\text{-ésimo item} \\ 2, 3, 5 \text{ e } 6$$

$$\sum c_i^O \text{ — número de categorias originais do instrumento de coleta} \\ 7 \times 2 + 10 \times 3 + 9 \times 5 + 10 \times 6 = 139$$

$$N \text{ — tamanho da população} \\ 200$$

Para o primeiro caso, substituindo-se os valores na fórmula [4.1.6], vem que:

$$n = \frac{\binom{c^E}{2} - \sum_{i=1}^k \binom{c_i^E}{2}}{\sum_{i=1}^k c_i^O} = \frac{\binom{1036}{2} - 7 \times \binom{4}{2} - 10 \times \binom{8}{2} - 9 \times \binom{32}{2} - 10 \times \binom{64}{2}}{149}$$

$$n = \frac{536130 - 24946}{149} = 3430,765 < 3431$$

Para o segundo caso, usando a fórmula [4.1.7], tem-se que:

$$n = \frac{\binom{c^E}{2} - \sum_{i=1}^k \binom{c_i^E}{2}}{\sum_{i=1}^k c_i^O}$$

$$n = \frac{\binom{1036}{2} - 7 \times \binom{4}{2} - 10 \times \binom{8}{2} - 9 \times \binom{32}{2} - 10 \times \binom{64}{2}}{149}$$

$$n = \frac{1}{1 + \frac{1}{200} \times \left[ \frac{\binom{1036}{2} - 7 \times \binom{4}{2} - 10 \times \binom{8}{2} - 9 \times \binom{32}{2} - 10 \times \binom{64}{2}}{149} - 1 \right]}$$

$$n = \frac{\frac{536130 - 24946}{149}}{1 + \frac{1}{200} \times \left[ \frac{536130 - 24946}{149} - 1 \right]} = \frac{3429,765}{1 + \frac{1}{200} \times 3428,765}$$

$$n = \frac{3429,755}{18,144} = 189,032 < 190$$

Este exemplo mostra a aplicação da fórmula proposta para o cálculo do tamanho da amostra, em duas situações diferentes, e os valores encontrados podem ser, de modo bastante aceitável, em termos intuitivos, considerados como tamanhos de

amostra plausíveis, cada um dos quais, para as investigações a que se propõem os pesquisadores.

Em um serviço clínico, no qual existe um número limitado de prontuários disponíveis, o tamanho da amostra a ser calculado deve ter, como base, esse número de prontuários disponíveis (ou um número menor, que abarque somente os prontuários bem preenchidos, ou que tenham dados suficientes frente ao que se deseja investigar), e com isso, a fórmula [4.1.7] presta-se, de modo mais realista, ao cálculo do tamanho da amostra a ser usada em uma possível investigação de uma pesquisa científica.

#### 4.2.2 Cálculo do tamanho da amostra para o WHOQOL

Do mesmo modo como feito no item anterior, vai-se calcular o tamanho da amostra para um dos mais conhecidos instrumentos, o WHOQOL, que é usualmente aplicado para coletar-se dados de qualidade de vida, e que foi proposto pela Organização Mundial da Saúde. O primeiro instrumento construído contém cem itens com cinco categorias cada (WHOQOL-100); uma segunda versão desse instrumento (versão abreviada) contém 26 itens com cinco categorias cada (WHOQOL-bref): esta é a versão para a qual

será calculado o tamanho da amostra.

Assim, considerando um tamanho de população de 200 sujeitos, tem-se que os valores básicos a serem calculados para uso nas fórmulas de cálculo do tamanho da amostra são:

$$c^E \text{ — número de categorias efetivas do instrumento de coleta} \\ 26 \times 2^5 = 832$$

$$c_i^E \text{ — número de categorias efetivas do } i\text{-ésimo item} \\ 32$$

$$k \text{ — número de itens do instrumento de coleta} \\ 26$$

$$c_i^O \text{ — número de categorias originais do } i\text{-ésimo item} \\ 5$$

$$\Sigma c_i^O \text{ — número de categorias originais do instrumento de coleta} \\ 26 \times 5 = 130$$

$$N \text{ — tamanho da população} \\ 200$$

Para o primeiro caso, substituindo-se os valores na fórmula [4.1.6], vem que:

$$n = \frac{\binom{c^E}{2} - \sum_{i=1}^k \binom{c_i^E}{2}}{\sum_{i=1}^k c_i^O} = \frac{\binom{832}{2} - 26 \times \binom{32}{2}}{130} = \frac{345696 - 12896}{130} = 2560$$

Para o segundo caso, usando a fórmula [4.1.7], tem-se que:

$$n = \frac{\frac{\binom{C^E}{2} - \sum_{i=1}^k \binom{C_i^E}{2}}{\sum_{i=1}^k C_i^O}}{1 + \frac{1}{N} \times \left[ \frac{\binom{C^E}{2} - \sum_{i=1}^k \binom{C_i^E}{2}}{\sum_{i=1}^k C_i^O} - 1 \right]} = \frac{\frac{\binom{832}{2} - 26 \times \binom{32}{2}}{130}}{1 + \frac{1}{200} \times \left[ \frac{\binom{832}{2} - 26 \times \binom{32}{2}}{130} - 1 \right]}$$

$$n = \frac{\frac{345696 - 12896}{130}}{1 + \frac{1}{200} \times \left[ \frac{345696 - 12896}{130} - 1 \right]} = 185,574 < 186$$

De modo resumido, pode-se esquematizar, no quadro 2, o cálculo do tamanho da amostra, usando as duas fórmulas propostas.

Quadro 2: Resumo dos cálculos de tamanho de amostra dos dois exemplos (SF-36 e WHOQOL).

instrumento	núm. de categs. originais por item	núm. de itens	núm. de categs. originais	núm. de categs. efetivas por item	núm. de categs. efetivas total	combinações do núm. de categs. efetivas total	combinações do núm. de categs. efetivas por item duas a duas	núm. de categs. efetivas a ser descontado	n pela fórmula [4.1.6]
SF-36	2	7	14	4	28	—	6	42	—
	3	10	30	8	80	—	28	280	—
	5	9	45	32	288	—	496	4.464	—
	6	10	60	64	640	—	2.016	20.160	—
Totais	—	—	149	—	1.036	536.130	—	24.946	3.431
WHOQOL	5	26	130	32	832	—	496	12.896	—
Totais	—	—	130	—	832	345.696	—	12.896	2.560

Tendo-se chegado ao resultado proposto pela fórmula [4.1.6], pode-se, com facilidade, aplicar a fórmula [4.1.7], chegando-se a um resultado ajustado pelo tamanho da população disponível.

#### 4.2.3 Cálculo do tamanho da amostra para alguns instrumentos da área de Pneumologia

Com o intuito de exemplificar um pouco mais, 24 instrumentos de coleta foram selecionados, entre os muitos disponíveis, e cujos resultados dos tamanhos de amostra serão mostrados, a seguir, no quadro 3. Esses exemplos foram retirados



do sítio da American Thoracic Society ([http:// www.atsqol.org / sections / instruments](http://www.atsqol.org/sections/instruments)). O sítio da ATS não informa sobre o conteúdo do instrumento: somente o nome, os números de itens e de categorias estão disponíveis para cada instrumento, sendo essas as informações necessárias para calcularem-se os tamanhos de amostra para cada um desses instrumentos.

Quadro 3: Cálculos de tamanho de amostra, para alguns instrumentos de coleta de dados da área de doenças pulmonares e áreas correlatas.

	Instrumento	Núm. de Itens	Número de Categorias Originais	Tamanho da Amostra		
				Fórmula [4.1.6]	Fórmula [4.1.7]	
					N = 100	N = 1000
1	Airways Questionnaire (AQ-20)	20	2	76	44	71
2	Airways Questionnaire (AQ-30)	30	2	116	54	105
3	Chronic Obstructive Pulmonary Disease Activity Rating Scale (CARS)	12	3	118	55	106
4	Functional Status II-R (FS II-R) - forma curta	14	3	139	59	122
5	Nottingham Health Profile (NHP)	38	2	148	60	130
6	Life Satisfaction Inventory-A (LSI-A); Life Satisfaction Inventory-Z (LSI- Z)	20	3	203	68	169
7	Daily Diary Card (Geddes)	8	4	224	70	184
8	Daily Diary Card (MRC)	5	5	410	81	291
9	Functional Status II-R (FS II-R) - forma longa	43	3	448	82	310
10	Functional Severity Index	6	5	512	84	339
11	Living with Asthma Questionnaire (LWAQ)	68	3	715	88	418
12	Integrated Therapeutics Group Child Asthma - forma curta	8	5	717	88	418
13	Dartmouth COOP Charts	9	5	820	90	451
14	Rotterdam Symptom Checklist (RSC)	30	4	928	91	482
15	Human Activity Profile Test (HAP) - bloco de 'atividade'	94	3	992	91	499
16	Functional Outcomes of Sleep Questionnaire	35	4	1088	92	522
17	Therapy Impact Questionnaire	36	4	1120	92	529
18	Symptom Distress Scale (SDS)	13	5	1229	93	552
19	About my Asthma (AMA)	44	4	1376	94	580
20	Integrated Therapeutics Group - Asthma (TG-ASF) - forma curta	15	5	1434	94	590
21	Functional Performance Inventory (FPI)	65	4	2048	96	673
22	Quality of Life - Anti-Cancer Drugs (QOL-ACD)	22	5	2151	96	683
23	Severe Respiratory Insufficiency Questionnaire (SRI)	24	5	2356	96	703
24	Rhinasthma	30	5	2970	97	749

Os cálculos acima recorreram à fórmula [4.1.6] (coluna 'Fórmula [4.1.6]'), e à fórmula [4.1.7] (colunas 'N = 100' e 'N =

1000'), respectivamente, para indicar tamanhos de amostras que podem ser coletadas de populações com tamanhos cem e mil, e que podem ser a realidade de serviços clínicos ambulatoriais e/ou de hospitais.

## 5 Discussão

O foco deste trabalho teve a intenção de mostrar uma abordagem não-paramétrica sobre uma possibilidade de cálculo de tamanho de amostra. *Ser não-paramétrico*, neste caso, significa não haver interferência de parâmetros oriundos de variáveis estatísticas; uma variável tem parâmetros, que são calculados, em várias condições usuais, tanto para descrição como para análise. Nesse sentido, ou seja, no sentido estatístico da palavra parâmetro, o instrumento de coleta não possui parâmetros, mas, sim, itens e suas respectivas categorias; itens podem ser (e são) considerados como variáveis estatísticas que possuem parâmetros, mas, a base da contagem de categorias foge ao que se poderia conceber como um parâmetro estatístico.

Em geral, os cálculos de tamanho de amostra que dependem de parâmetros ligados ao teste estatístico que será aplicado, quando da análise dos dados, podem deixar a desejar, por conta da adoção cega dos valores de poder do teste, da adoção do tipo de caudalidade, e mesmo do valor do Erro Tipo I (nível de significância), o que, para os cálculos que independem dessas

adoções, pode ser evitado. A proposta mostrada neste trabalho vem ao encontro dessa independência, assim como algumas outras formas de cálculo do tamanho da amostra que também não estão vinculadas a parâmetros ligados à variável de interesse.

A contraposição de *métodos* de cálculo do tamanho da amostra não faz sentido, *a priori*, porque os métodos, em si, são diferentes, únicos e de uso específico, em cada caso. Uma possível, mas, também, não muito interessante comparação entre os cálculos de tamanho de amostra pode ser feita com os *resultados* apurados com cada uma das técnicas de cálculo. Assim sendo, escolheu-se um tipo bastante usual de cálculo do tamanho da amostra, e alguns resultados para algumas condições possíveis foram calculados; esses resultados foram contrapostos aos resultados da proposta desenvolvida neste trabalho.

Escolheu-se o método cuja base de informação é a diferença entre medianas, mencionado na parte 1.4.3.2, item '4' (Noether, 1987):

$$n = \frac{(z_{\alpha} + z_{\beta})^2}{4 \times (p - \frac{1}{2})^2}$$

Fixando  $z_{\alpha} = 1,96$ , e variando  $z_{\beta}$ , com três valores de uso comum (80%, 90% e 99%), pode-se gerar os resultados que são mostrados no Quadro 4.

Quadro 4: Apresentação de alguns valores para a determinação do tamanho da amostra, usando um método usual (diferença de medianas).

Proporção de valores positivos	Valor de $\beta$		
	80%	90%	99%
45,0%	784	1050	1841
42,5%	349	467	818
40,0%	196	263	461
37,5%	126	168	295
35,0%	88	117	205

Com base na nova proposta de cálculo, considerando a estrutura de um instrumento de coleta, propôs-se calcular o tamanho da amostra para cinco instrumentos diferentes, cujos resultados também encontram-se a seguir, no Quadro 5:

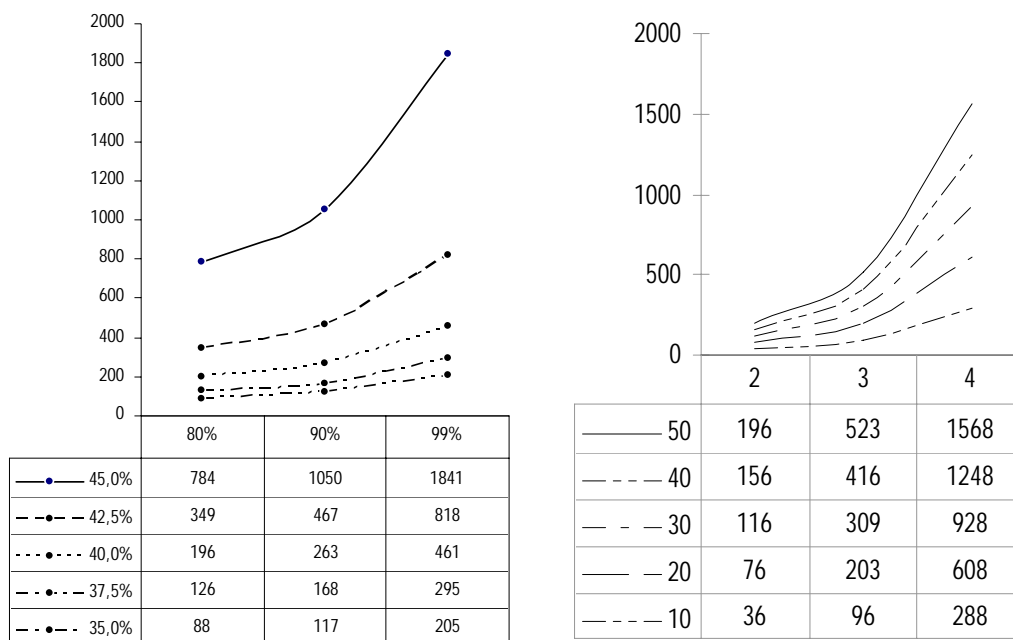
Quadro 5: Apresentação de alguns valores para a determinação do tamanho da amostra, usando um método usual (diferença de medianas).

Tipos de instrumento de coleta				
10 itens com 2 categorias	10 itens com 3 categorias	20 itens com 2 categorias	10 itens com 2 categorias e 10 itens com 3 categorias	20 itens com 3 categorias
36	96	76	136	203

Apesar de ambos os métodos não poderem ser comparados por suas metodologias de construção, os resultados apresentam-se com uma lógica esperada: para o método usual, tanto maior é o tamanho da amostra, quanto maior é a proporção

de valores positivos encontrados para a diferença entre as medianas, e tanto maior ainda é esse tamanho de amostra, quando os valores de beta aumentam; para o método proposto, tanto maior é o tamanho da amostra, quanto mais itens e principalmente, mais categorias são consideradas.

Pode-se representar graficamente (figura 2) o comportamento dos resultados apurados por ambos os métodos; essa representação não tem a intenção de tornar-se uma base de comparação entre os métodos, mas deve ser encarada como uma possível comparação entre resultados apurados.



(a)

(b)

Figura 2: Representação gráfica dos resultados de tamanhos de amostra, considerando os dois métodos (usual e proposto) — Figura 2a: o eixo das abscissas representa o valor da estatística beta; o eixo das ordenadas representa o tamanho da amostra; há cinco curvas representando alguns valores de proporções de valores positivos para a diferença entre as medianas; Figura 2b: o eixo das abscissas representa o número de categorias de cada item; o eixo das ordenadas representa o tamanho da amostra; há cinco curvas representando alguns números de categorias escolhidos.

Observa-se que, apesar de a comparação entre ambos os métodos, em termos de suas construções, não ser plausível, os resultados apurados assemelham-se, em cada uma das condições estudadas, pelos valores exemplificados. Pode-se, então, afirmar que os comportamentos de ambos os tipos de cálculo assemelham-se por seus resultados, mas, é falso afirmar que exista uma equivalência metodológica que leve a resultados semelhantes,



como por exemplo, dizer-se que um instrumento com 50 itens binários, por produzir um tamanho de amostra de '196' elementos amostrais, equivale-se a uma proporção de 40% de valores positivos para a diferença entre medianas, com um poder adotado de 80%, cujo tamanho de amostra também é '196': os resultados coadunam-se, em termos numéricos, mas, pelo fato de ambos os métodos diferirem-se por suas características próprias, não há nenhum vínculo entre ambos os resultados apurados.

Os resultados vistos no quadro 3 apontam para uma *performance* mais adequada da nova proposta de cálculo de tamanho de amostra, nos casos em que os instrumentos têm menos categorias por item, pois esse número (o de categorias) produz, quando acima de quatro e em um instrumento com um número grande de itens, tamanhos de amostra muito grandes, que se coadunam com os tipos vistos no item 1.4.3.1, e que são, na prática, tamanhos de amostra passíveis de alcançar valores maiores do que as propostas usuais constantes dos itens 1.4.3.2 e 1.4.3.3.

Qualquer método para cálculo de tamanho de amostra pode produzir tamanhos de amostra tão grandes ou tão pequenos quantos se deseje (Deming, 1966). O menor tamanho de amostra aceitável é o unitário, e o maior, um número finito, porém,

consideravelmente grande. Na Área de Saúde, *em geral*, os tamanhos de amostra limitam-se por valores relativamente pequenos, pois a coleta de dados médicos, em virtude de questões éticas e da gama de dificuldades nos processos de coleta (desistências, óbitos, exclusões, e perdas de diversos outros tipos etc.), também ficam limitadas dentro do processo de seleção das amostras.

Noether (1955 e 1987) sugeriu algumas formas de estimação para o tamanho da amostra, em que parâmetros como amplitudes e totais são considerados, vindo ao encontro da nova proposta feita neste trabalho, porquanto a base das contagens tem como essência valores que são totalizadores do número de itens e de categorias componentes do instrumento de coleta. Respaldam essa proposta as formas de cálculo de tamanho de amostra de Sturges (1926), Herrmann e Szatrowski (1982), Francis (1985), e Gillett (1994).

A proposta de Skalski (1992) não é única: ela foi selecionada para mostrar que a previsão de perdas de elementos amostrais pode ser levada em conta, para estimação do tamanho da amostra, fazendo um paralelo com a aplicação da Lógica Paraconsistente Anotada Evidencial (Abe, 1992 e 1997). O estado paraconsistente *paracompleto* (falta de informação) equivale-se à

preocupação que se tem, quando as perdas são iminentes, e serve não somente como característica típica de pesquisas nas áreas biológicas (por conta de perdas que, em geral, efetivamente podem ocorrer!), mas, também, como auxiliar importante para a nova proposta de cálculo do tamanho da amostra. Perdas são ocorrências comuns, em estudos em que procedimentos invasivos e tratamentos com elementos cujas enfermidades atingem graus de gravidade maiores; com isso, elas devem ser, quando possível, consideradas, o que, para os cálculos de tamanho amostral usuais, nem sempre são lembradas.

Os testes estatísticos cuja base é a comparação entre valores a serem testados são de aplicação específica e dependem da adoção de parâmetros e de critérios rigorosos, em função da distribuição da variável de interesse principal (variável primária). Todos esses métodos de cálculo não correm em paralelo com a nova proposta feita neste trabalho, pois, além de não existir uma variável primária (pois o instrumento é visto em seu todo), os métodos usuais, com base em diferença de médias servem para cálculos em que, exclusivamente, a variável principal comanda o tamanho da amostra a ser calculado, conforme demonstram Guenther (1981), Lieber (1990), Edmiston et al (1993), Van Belle e Martin (1993), Lindsey (1997), Dawson (1998), Hayes e Bennett (1999), Lazovich et al (2000), Hsieh et al (2003), Karlsson et al

(2003), Willan e Pinto (2005), Borm et al (2007), e, Brasher e Brandt (2007).

Quanto ao uso das proporções, tais valores gerados por números relativos a totais, as mesmas afastam-se da proposta deste trabalho, conquanto os testes estatísticos para proporções, sem exceção, pleiteiam um ou dois valores de proporções a serem usados, como atestam os autores Fleiss (1973), Ury e Fleiss (1980), Lachin (1981), McHugh e Le (1983), Edmiston et al (1993), Hayes e Bennett (1999), Sato (2000), e, Krishnamoorthy e Peng (2007). Os testes relacionados a três ou mais valores que são considerados tendem a usar vários valores disponíveis (estimados previamente, ou, conhecidos), o que parece aproximar esse tipo de cálculo de tamanho de amostra à nova proposta, mas, também limitam-se à necessidade de comparações mais complexas, em que os testes estatísticos já prevêm a necessidade para sua aplicação, conforme foi proposto por Boos e Brownie (1995), Mahoney e Magel (1996), Ahnn e Anderson (1998, e, Rasch e Šimečková (2007). Apesar de os testes não-paramétricos tenderem a aproximar-se do uso de contagens, porque sua base está disposta sobre essa característica, a estrutura dos cálculos de tamanho de amostra por meio desses testes foge da nova proposta, por conta do uso específico de valores eleitos para essa abordagem, como pôde ser visto em propostas de autores consagrados como Noether

(1987) e, Feuer e Kessler (1989), e mais recentemente, com Vollandt e Horn (1997), Julious e Campbell (1998, seguidores da proposta de Noether), Gangnon e Kosorok (2004), Gönen (2004), Shieh et al (2008), e Zhao et al (2008).

Associações, correlações e regressões formam o contexto dos estudos em que os relacionamentos entre variáveis são abordados. Esses estudos tendem a ser mais complexos, devido ao tipo de modelagem que sempre envolve mais de uma variável, concomitantemente. Na Área de Saúde, o relacionamento entre variáveis pode ser um fator de confusão, provocado por, ao menos, uma das variáveis estudadas; assim, é fato que todo tipo de cuidado deve ser almejado, sem o que de nada são válidas as estimativas dos parâmetros amostrais, e muito menos, do aspecto inferencial de extensão sem limites de conclusões para a população. Desse modo, a escolha de variáveis especialmente destinadas ao cálculo do tamanho da amostra, nesse caso, precisa tanto quanto em outros casos, de rigores absolutos e entendimento pleno sobre o experimento a ser planejado e montado. Percebe-se clara preocupação sobre esta questão nos trabalhos de Lachin (1981), Zar (1984), e, Bonett e Wright (2000), que usam as correlações como informação de base para os cálculos propostos; a nova proposta usa o grau de similaridade (Borg e Groenen, 2005), que apresenta certa equivalência frente à correlação, o que permite

aproximar ambas as propostas, sob esse foco. Na nova proposta, a similaridade é usada como fator preponderante como um limitador ao tamanho da amostra a ser calculado.

As tabelas do tipo 2 x 2 são produtoras de muitas estatísticas importantes, e os estudos sobre esse tipo de tabela são largamente divulgados. Os valores de seus totais, de suas marginais e de suas células internas são usados, com frequência, na determinação de estatísticas definidoras, tanto de comparações como de relacionamentos entre variáveis, o que permite afirmar sobre sua importância nos estudos em que essa modelagem é buscada. Woolson et al (1986), Thomas e Conlon (1992), Sahai e Khurshid (1996), e Julious e Campbell (1998, como seguidores de Guenther) são bons exemplos do uso de medidas-resumo como qui-quadrado e odds-ratio, para o cálculo do tamanho da amostra, em concordância com a nova proposta, pelo uso de totalizadores como estimadores das maneiras de preenchimento do instrumento de coleta: cada célula de uma tabela 2 x 2 equivale-se à repetição dos caminhos (maneiras de preenchimento), o que permite uma aproximação entre as formas de cálculo de tamanho de amostra.

Adcock (1989, 1992 e 1997) reserva a seus textos comentários sobre a visão frequentista e a bayesiana: ele demonstra a semelhança, quanto aos *resultados* alcançados, mas

destaca as diferenças entre ambos os *métodos*. Ambas essas visões coadunam-se entre si, em termos de resultados, porém, como já foi comentado em relação aos cálculos anteriormente vistos, de cunho frequentista, também a abordagem bayesiana investe sobre a fixação de características específicas de cada tipo de pesquisa sobre a qual sua ingerência se faz presente. Pelo que foi visto, Inoue et al (2005), que respalda Adcock, e, Joseph (1997) apresentam-se como inovadores, enquanto divulgadores do teor bayesiano de suas propostas para cálculo de tamanhos de amostra. Há uma clara aproximação entre o teor bayesiano, sobre o conhecimento prévio de valores estimados, e a nova proposta (deste trabalho), que tem como base, somente, as contagens dos itens e categorias, que, então, neste sentido, apresentam-se como valores previamente conhecidos, e que não precisam de estimação, mas, sim, de cálculos. A estatística bayesiana prevê a existência antecipada de valores, ou seja, de conhecimento sobre os estimadores, e a nova proposta só se caracteriza sobre contagens factíveis e previamente estabelecidas; apesar das diferenças em ambas as abordagens, há pontos de toque entre ambas, por seus teores previsionários dos valores estimados previamente.

Finalmente, pode-se listar algumas vantagens e desvantagens do método referente à nova proposta para cálculo de tamanho de amostra. Elas cobrem as perguntas que, de praxe, são

feitas, quando do planejamento amostral, ou seja, no momento em que são definidas as características do experimento que será montado, para o cumprimento dos objetivos pleiteados pelo pesquisador.

Vantagens:

- (a) permite calcular o tamanho da amostra em função da composição do instrumento de coleta, possibilitando que o próprio instrumento, nos casos possíveis e/ou necessários, seja readequado em função dos itens que efetivamente são de interesse na investigação.
- (b) leva em consideração *todas* as combinações de respostas possíveis; intuitivamente, isso permite afirmar que qualquer combinação de respostas foi considerada, em função de todas as possibilidades de combinação existentes.
- (c) tendo-se feito o cálculo do tamanho da amostra para um instrumento de coleta, o mesmo não se altera, o que pode facilitar o uso desse número em diversas investigações diferentes.
- (d) não necessita de informações externas; só necessita do conhecimento da estrutura do instrumento de coleta.



- (e) é um cálculo relativamente fácil de ser feito, pois depende, apenas, das contagens dos elementos formadores de sua formulação, e não de um estudo-piloto ou pesquisa prévia para conseguir-se os valores necessários para o cálculo do tamanho da amostra.
- (f) não depende do tamanho da escala de qualquer variável, pois cada escala é transformada em uma estrutura de categorias.
- (g) este cálculo independe da estimação de valores prévios de nível de significância e de poder do teste, bem como da adoção, também prévia, de algum teste estatístico.

Desvantagens:

- (a) no caso de instrumentos de coleta muito grandes, a depender do número de itens e principalmente, do número de categorias originais, o tamanho da amostra a ser calculado pode ser grande o suficiente, para inviabilizar a pesquisa; a sugestão, neste caso, é a de sempre considerar populações limitadas por valores da ordem de grandeza das dezenas, centenas ou algumas unidades de milhar; se isso não puder ser viabilizado, então deve-se recorrer a outros métodos de cálculo de tamanho de amostra.

- (b) no caso de instrumentos de coleta muito pequenos, o problema anterior se inverte, e a amostra pode ser pouco representativa para a condução da pesquisa; sugere-se, então, também, recorrer a outros métodos de cálculo de tamanho de amostra.
- (c) nem sempre, é possível transformar os itens e suas eventuais categorias, para que o instrumento possa consistir em base para o cálculo do tamanho da amostra; com isso, surgem efetivas dificuldades para a criação de categorias em itens cujas respostas são abertas, ou mesmo, em escalas numéricas cujas divisões não se tornam claras para a criação de categorias sugestivas.
- (d) Chen (1989) alerta para o volume de classificações errôneas em categorias oriundas de variáveis que originalmente não são categóricas: esse número pode ser relativamente elevado, e prejudicará o cálculo do tamanho da amostra, quando não se conhece, com precisão, quais serão as categorias a serem determinadas para cada variável em que isso for necessário.

## 6 Conclusão

A proposta de desenvolver um novo método para o cálculo do tamanho de amostra em pesquisa na Área de Saúde foi atingida, por meio do uso das ferramentas expostas e da construção de um modelo que supre esse tipo de cálculo, quando informações sobre o cerne do experimento a ser montado estão ausentes.

Os resultados alcançados demonstram as possibilidades de uso desse novo método, com seus devidos pontos positivos e negativos. Assim sendo, o que foi proposto, neste trabalho, parece ser tão viável como o que já existe, nos mesmos termos de uso e de aplicação das técnicas para o cálculo do tamanho da amostra. Calcular o tamanho da amostra que será coletada, e depois, usada para as investigações biológicas e eventuais avaliações estatísticas, é algo tão poucas vezes lembrado e ao mesmo tempo tão destacável, quanto a própria circunstância da pesquisa que se deseja montar.

Em vários experimentos na Área de Saúde, nenhum cálculo de tamanho de amostra é necessário (estudos de caso,

disponibilidade de espécimes em número reduzidíssimo etc.), e em outros, esse cálculo é crucial, pois faz parte do *teor* da conclusão do trabalho científico em questão. A proposta feita aqui, portanto, serve e não serve: o uso desse método dependerá das circunstâncias da investigação a ser conduzida. Nesse sentido, tudo indica que esta proposta vai se unir a outras já existentes, e será solução que ficará disponível, nos casos em que pouca informação externa pode ser arrebanhada. Talvez, seja essa a característica mais propícia para o uso dessa técnica: a falta de informação e a existência de um instrumento previamente adotado.

## 7 Anexos

### 7.1 SF-36 (adaptado de Ciconelli et al, 1999)

1- Em geral você diria que sua saúde é:

Excelente (5.0); Muito Boa (4.4); Boa (3.4); Ruim (2.0); Muito Ruim (1.0)

2- Comparada há um ano, como você classificaria sua saúde em geral, agora?

Muito Melhor (1); Um Pouco Melhor (2); Quase a Mesma (3); Um Pouco Pior (4); Muito Pior (5)

3- Os seguintes itens são sobre atividades que você poderia fazer atualmente durante um dia comum. De acordo com a sua saúde, você teria dificuldade para fazer estas atividades? Neste caso, quando?

Atividades	Sim, muita dificuldade	Sim, um pouco de dificuldade	Sem dificuldade
a) Atividades vigorosas, que exigem muito esforço, tais como correr, levantar objetos pesados, participar em esportes intensos.	1	2	3
b) Atividades moderadas, tais como mover uma mesa, passar aspirador de pó, jogar bola, varrer a casa.	1	2	3
c) Levantar ou carregar mantimentos	1	2	3
d) Subir vários lances de escada	1	2	3
e) Subir um lance de escada	1	2	3
f) Curvar-se, ajoelhar-se ou dobrar-se	1	2	3
g) Andar mais de 1 km	1	2	3
h) Andar vários quarteirões	1	2	3
i) Andar um quarteirão	1	2	3
j) Tomar banho ou vestir-se	1	2	3

4- Durante as últimas 4 semanas, você teve algum dos seguintes problemas no seu trabalho ou com alguma atividade regular, como consequência de sua saúde física?

	Sim	Não
a) Você diminui a quantidade de tempo que se dedicava ao seu trabalho ou a outras atividades?	1	2
b) Realizou menos tarefas do que você gostaria?	1	2
c) Esteve limitado no seu tipo de trabalho ou a outras atividades?	1	2
d) Teve dificuldade de executar seu trabalho ou outras atividades (p. ex. necessitou de um esforço extra)?	1	2

5- Durante as últimas 4 semanas, você teve algum dos seguintes problemas com seu trabalho ou outra atividade regular diária, como consequência de algum problema emocional (como sentir-se deprimido ou ansioso)?

	Sim	Não
a) Você diminui a quantidade de tempo que dedicava-se ao seu trabalho ou a outras atividades?	1	2
b) Realizou menos tarefas do que você gostaria?	1	2
c) Não realizou ou fez qualquer das atividades com tanto cuidado como geralmente faz.	1	2

6- Durante as últimas 4 semanas, de que maneira sua saúde física ou problemas emocionais interferiram nas suas atividades sociais normais, em relação à família, amigos ou em grupo?

De forma nenhuma (5); Ligeiramente (4); Moderadamente (3); Bastante (2); Extremamente (1)

7- Quanta dor no corpo você teve durante as últimas 4 semanas?

Nenhuma (6.0); Muito Leve (5.4); Leve (4.2); Moderada (3.1); Grave (2.0); Muito Grave (1.0)

8- Durante as últimas 4 semanas, quanto a dor interferiu com seu trabalho normal (incluindo o trabalho dentro de casa)?

De maneira alguma (1); Um pouco (2); Moderadamente (3); Bastante (4); Extremamente (5)

9- Para cada questão abaixo, por favor, dê uma resposta que mais se aproxime da maneira como você se sente, em relação às últimas 4 semanas.

	Sempre	A maior parte do tempo	Boa parte do tempo	As vezes	Poucas vezes	Nunca
a) Por quanto tempo você se sente cheio de vigor, força, e animado?	6	5	4	4	2	1
b) Por quanto tempo se sente nervosa(o)?	1	2	3	4	5	6
c) Por quanto tempo se sente tão deprimido que nada pode animá-lo?	1	2	3	4	5	6
d) Por quanto tempo se sente calmo ou tranquilo?	6	5	4	4	2	1
e) Por quanto tempo se sente com muita energia?	6	5	4	4	2	1
f) Por quanto tempo se sente desanimado ou abatido?	1	2	3	4	5	6
g) Por quanto tempo se sente esgotado?	1	2	3	4	5	6
h) Por quanto tempo se sente uma pessoa feliz?	6	5	4	4	2	1
i) Por quanto tempo se sente cansado?	1	2	3	4	5	6

10- Durante as últimas 4 semanas, por quanto tempo a sua saúde física ou problemas emocionais interferiram em suas atividades sociais (como visitar amigos, parentes, etc.)?

Sempre (1); A maior parte do tempo (2); Boa parte do tempo (3); Poucas vezes (4); Nunca (5)

11- O quanto verdadeiro ou falso é cada uma das afirmações para você?

	Definitivamente verdadeiro	A maioria das vezes verdadeiro	Não sei	A maioria das vezes falso	Definitivamente falso
a) Eu costumo adoecer um pouco mais facilmente que as outras pessoas	1	2	3	4	5
b) Eu sou tão saudável quanto qualquer pessoa que eu conheça	5	4	3	2	1
c) Eu acho que a minha saúde vai piorar	1	2	3	4	5
d) Minha saúde é excelente	5	4	3	2	1

## 7.2 WHOQOL-bref (adaptado de WHOQOL Group, 1998)

PROGRAMA DE SAÚDE MENTAL - ORGANIZAÇÃO MUNDIAL DA SAÚDE GENEVRA  
GRUPO WHOQOL

### VERSÃO EM PORTUGUÊS DOS INSTRUMENTOS DE AVALIAÇÃO DE QUALIDADE DE VIDA - BREF (WHOQOL) ou Breve 4 - 1998.

#### Instruções:

Este questionário é sobre como você se sente a respeito de sua qualidade de vida, saúde e outras áreas de sua vida. **Por favor, responda a todas as questões.** Se você não tem certeza sobre que resposta dar em uma questão, por favor, escolha entre as alternativas a que lhe parece mais apropriada. Esta, muitas vezes, poderá ser sua primeira escolha.

Por favor, tenha em mente seus valores, aspirações, prazeres e preocupações. Nós estamos perguntando o que você acha de sua vida, tomando como referência as **duas últimas semanas**. Por exemplo, pensando nas últimas duas semanas, uma questão poderia ser:

	nada	muito pouco	médio	muito	completamente
Você recebe dos outros o apoio de que necessita?	1	2	3	4	5

Você deve circular o número que melhor corresponde ao quanto você recebe dos outros o apoio de que necessita nestas últimas duas semanas. Portanto, você deve circular o número 4 se você recebeu "muito" apoio.

Você deve circular o número 1 se você não recebeu "nada" de apoio.

**Por favor, leia cada questão, veja o que você acha e circule no número e lhe parece a melhor resposta.**

		muito ruim	ruim	nem ruim nem boa	boa	muito boa
1	Como você avaliaria sua qualidade de vida?	1	2	3	4	5

		muito insatisfeito	insatisfeito	nem satisfeito nem insatisfeito	satisfeito	muito satisfeito
2	Quão satisfeito(a) você está com a sua saúde?	1	2	3	4	5



As questões seguintes são sobre <b>o quanto</b> você tem sentido algumas coisas nas últimas duas semanas.						
		nada	muito pouco	mais ou menos	bastante	extremamente
3	Em que medida você acha que sua dor (física) impede você de fazer o que você precisa?	1	2	3	4	5
4	O quanto você precisa de algum tratamento médico para levar sua vida diária?	1	2	3	4	5
5	O quanto você aproveita a vida?	1	2	3	4	5
6	Em que medida você acha que a sua vida tem sentido?	1	2	3	4	5
7	O quanto você consegue se concentrar?	1	2	3	4	5
8	Quão seguro (a) você se sente em sua vida diária?	1	2	3	4	5
9	Quão saudável é o seu ambiente físico (clima, barulho, poluição, atrativos)?	1	2	3	4	5

As questões seguintes perguntam sobre <b>quão completamente</b> você tem sentido ou é capaz de fazer certas coisas nestas últimas duas semanas..						
		nada	muito pouco	médio	muito	completamente
10	Você tem energia suficiente para seu dia-a-dia?	1	2	3	4	5
11	Você é capaz de aceitar sua aparência física?	1	2	3	4	5
12	Você tem dinheiro suficiente para satisfazer suas necessidades?	1	2	3	4	5
13	Quão disponíveis para você estão as informações que precisa no seu dia-a-dia?	1	2	3	4	5
14	Em que medida você tem oportunidades de atividade de lazer?	1	2	3	4	5

As questões seguintes perguntam sobre <b>quão bem ou satisfeito</b> você se sentiu a respeito de vários aspectos de sua vida nas últimas duas semanas						
		muito ruim	ruim	nem ruim nem boa	bem	muito bem
15	Quão bem você é capaz de se locomover?	1	2	3	4	5
		muito insatisfeito	Insatisfeito	nem satisfeito nem insatisfeito	satisfeito	muito satisfeito
16	Quão satisfeito (a) você está com o seu sono?	1	2	3	4	5
17	Quão satisfeito (a) você está com sua capacidade de desempenhar as atividades do seu dia-a-dia?	1	2	3	4	5
18	Quão satisfeito (a) você está com sua capacidade para o trabalho?	1	2	3	4	5
19	Quão satisfeito (a) você está consigo mesmo?	1	2	3	4	5
20	Quão satisfeito (a) você está com suas relações pessoais (amigos, parentes, conhecidos, colegas)?	1	2	3	4	5
21	Quão satisfeito (a) você está com sua vida sexual?	1	2	3	4	5
22	Quão satisfeito (a) você está com o apoio que você recebe de seus amigos?	1	2	3	4	5
23	Quão satisfeito (a) você está com as condições do local onde mora?	1	2	3	4	5
24	Quão satisfeito (a) você está com o seu acesso aos serviços de saúde?	1	2	3	4	5
25	Quão satisfeito (a) você está com o seu meio de transporte?	1	2	3	4	5

A questão seguinte refere-se a <b>com que frequência</b> você sentiu ou experimentou certas coisas nas últimas duas semanas						
		nunca	algumas vezes	frequentemente	muito frequentemente	sempre
26	Com que frequência você tem sentimentos negativos tais como mau humor, desespero, ansiedade, depressão?	1	2	3	4	5

## 8 Referências

- Abe JM Fundamentos da Lógica Anotada, Tese de Doutorado, FFLCH - USP, 135 pp, 1992.
- Abe JM Some Aspects of Paraconsistent Systems and Applications, *Logique et Analyse*, 157: 83-96, 1997.
- Adcock CJ Pre-sampling procedures. *The Statistician*, **38**: 107-15, 1989.
- Adcock CJ Sample size determination: a review. *The Statistician*, **46**(2): 261-83, 1997.
- Adcock CJ The Bayesian approach to determination of sample sizes – some comments on the paper by Joseph, Wolfson and du Berger. *The Statistician*, **44**(2): 155-61, 1995.
- Ahnn S e Anderson SJ Sample size determination in complex clinical trials comparing more than two groups for survival endpoints. *Statist Med*, **17**: 2525-34, 1998.
- Asimov I *Earth is room enough*, Glasgow: Grafton Books, 1986.
- Bégaud B; Fourrier A; Moore N; Moride Y The delusion of reducing sample size. *Eur J Clin Pharmacol*, **59**(8-9): 711-2, 2003.
- Björner A e Stanley RP *A combinatorial miscellany*, 1999.
- Bonett DG Confidence interval for a coefficient of quartile variation. *Computat Statist Data Anal*, **50**(11): 2953-7, 2006.
- Bonett DG e Wright TA Sample size requirements for estimating Pearson, Kendall and Spearman correlations. *Psychometrika*, **65**(1): 23-8, 2000.
- Boos DD e Brownie C Anova and Rank-Tests When the Number of Treatments Is Large. *Statist Prob Letters*, **23**(2): 183-91, 1995.
- Borg I e Groenen PJF *Modern multidimensional scaling theory and applications*, 2ed. New York: Springer, 2005.

Borm GF; Fransen J; Lemmens WAJG A simple sample size formula for analysis of covariance in randomized clinical trials. *J Clin Epidemiol*, **60**: 1234-8, 2007.

Brasher PMA e Brant RF Sample size calculations in randomized trials: common pitfalls. *Can J Anaesth - Journal Canadien D'Anesthesie*, **54**(2): 103-6, 2007.

Brownie C e Boos DD Type-I Error Robustness of Anova and Anova on Ranks When the Number of Treatments Is Large. *Biometrics*, **50**(2): 542-9, 1994.

Chen TT A Review of Methods for Misclassified Categorical-Data in Epidemiology. *Statist Med*, **8**(9): 1095-106, 1989.

Ciconelli RM; Ferraz MB; Santos W; Meinão I; Quaresma MR Tradução para a língua portuguesa e validação do questionário genérico de avaliação de qualidade de vida SF-36 (Brasil SF 36). *Rev Bras Reumatol*, **39**(3): 143-150, 1999.

Cochran W *Sampling Techniques*, 2nd ed., Boston: John Wiley & Sons, 1985.

Crouch M e McKenzie H The logic of small samples in interview-based qualitative research. *Social Sci Inform*, **45**(4): 17, 2006.

Dawson JD Sample size calculations based on slopes and other summary statistics. *Biometrics*, **54**(1): 323-30, 1998.

Deming WE *Some theory of sampling*. New York: Dover, 1966.

Edmiston CE; Josephson A; Pottinger J; Ciacco-Tsivitis M; Palenik C The numbers game: sample-size determination. *AJIC*, **21**: 151-4, 1993.

Feuer EJ e Kessler LG Test Statistic and Sample-Size for a 2-Sample McNemar Test. *Biometrics*, **45**(2): 629-36, 1989.

Fisher RA On the Mathematical Foundations of Theoretical Statistics. *Phylos Trans R Soc*, April 19, 1922.

Fleiss JL *Statistical Methods for Rates and Proportions*. New York: John Wiley & Sons, 1973.

Francis K Sample size determination: power in research, *TIPS*: 350-2, 1985.

- Frankel M e King B A conversation with Leslie Kish. *Statist Sci*, **11**(1):65-87, 1996.
- Gail MH A bibliography and comments on the use of statistical models in Epidemiology in the 1980s. *Statist Med*, **10**: 1819-85, 1991.
- Gangnon RE e Kosorok MR Sample-size formula for clustered survival data using weighted log-rank statistics. *Biometrika*, **91**(2): 263-75, 2004.
- Gillett R An average power criterion for sample size estimation. *The Statistician*, **43**(3):389-94, 1994.
- Gönen M Sample size and power for McNemar's test with clustered data. *Statist Med*, **23**(14): 2283-94, 2004.
- Gordon I e Watson R A note on sample size determination for comparison of small probabilities, *Control Clin Trials*, **15**: 77-9, 1994.
- Gordon I e Watson R The myth of continuity-corrected sample size formulae. *Biometrics*, **52**: 71-6, 1996.
- Griffitts HB e Hilton PJ Matemática Clássica — uma interpretação contemporânea, São Paulo: Edgard Blucher, EDUSP, 1975.
- Grimaldi, RP *Discrete and Combinatorial Mathematics — an applied introduction*. Reading (Massachusetts): Addison-Wesley Publishing Company, 1986.
- Grunkemeier GL e Jin R Power and sample size: how many patients do I need? *Ann Thorac Surg*, **83**: 1934-9, 2007.
- Guenther WC Sample-Size formulas for Normal Theory T-Tests. *Am Statist*, **35**(4): 243-4, 1981.
- Hayes RJ e Bennett S Simple sample size calculation for cluster-randomized trials. *Int J Epidemiology*, **28**:319-26, 1999.
- Heilbrun LK e McGee DL Sample size determination for the comparison of normal means when one sample size is fixed. *Computational Statistics & Data Analysis*, **3**: 99-102, 1985.
- Hermann N e Szatrowski TH Asymptotic formulas for expected sample size savings in curtailed binomial tests. *Sequential Analysis*, **1**(3): 221-45, 1982

Hsieh FY, Lavori PW, Cohen HJ; Feussner JR. An overview of variance inflation factors for sample-size calculation. *Eval Health Prof* **26**(3): 239-57, 2003.

Inoue LYT; Berry DA; Parmigiani G Relationship between Bayesian and frequentist sample size determination. *Am Statist*, **59**(1): 79-87, 2005.

Joseph L Bayesian and mixed bayesian / likelihood criteria for sample size determination. *Statist Med*, **16**:769-81, 1997.

Julious SA e Campbell MJ Sample size calculations for paired or matched ordinal data. *Statist Med*, **17**: 1635-42, 1998.

Jung SH; Kang SH; Ahn C Sample size calculations for clustered binary data. *Statist Med*, **20**(13): 1971-82, 2001.

Karlsson J; Engebretsen L; Dainty K Considerations on sample size and power calculations in randomized clinical trials. *Arthroscopy*, **19**(9): 997-9, 2003.

Katz D; Gutek BA; Kahn RL; Barton E *Bureaucratic encounters — a pilot study in the evaluation of government services*. Michigan: Ann Arbor, 1977.

Kish L *Survey sampling*, New York: John Wiley & Sons, 1965.

Krishnamoorthy K e Peng J Some properies of the exact and score methods for Bionmial proportion and sample size calculation. *Communications in Statistics – Simulation and Computation*, **36**: 1171-86, 2007.

Lachin JM Introduction to sample size determination and power analysis for clinical trials. *Control Clin Trials*, **2**: 93-113, 1981.

Lazovich D; Murray DM; Brosseau LM; Parker DL; Milton FT; .ugan SK Sample size considerations for studies of intervention efficacy in the occupational setting. *Ann Occup Hyg*, **46**(2): 219-27, 2002.

Lieber RL Statistical Significance and Statistical Power in Hypothesis-Testing. *J Orthop Res*, **8**(2): 304-9, 1990.

Lindsey JK Exact sample size calculations for exponential family models. *The Statistician*, **46**(2): 231-7, 1997.

Mahoney M e Magel R Estimation of the power of the Kruskal-Wallis test. *Biom J*, **38**(5): 613-30, 1996.

- McHugh RB Confidence interval inference and sample size. *Am Stat*, April, 1961.
- McHugh RB e Le CT Confidence estimation and size of a clinical trial. *Control Clin Trials*, **5**: 157-63, 1984.
- Meyer PL *Probabilidade — aplicações à Estatística*. 2ed. São Paulo: LTC, 1983.
- Noether GE Sample size determination for some common nonparametric tests. *J Amer Stat Assoc*, **82**(398): 645-7, 1987.
- Noether GE Use of the Range Instead of the Standard Deviation. *J Amer Stat Assoc*, **50**(272): 1040-55, 1955.
- O'Muircheartaigh C Balancing Statistical Theory, Sampling Concepts, and Practicality in the Teaching of Survey Sampling, *ISI 55<sup>th</sup>*, 2005.
- Pagano A *Lições de Estatística*, 2ed, São Paulo: Gráfica da Prefeitura de São Paulo, 1946.
- Rahme E e Joseph L Estimating the prevalence of a rare disease: adjusted maximum likelihood. *The Statistician*, **47**: 149-58, 1998.
- Rasch D e Šimečková M Determining the size of experiments for the one-way ANOVA model I for ordered categorical data. In: *Lopez-Fidalgo J; Rodriguez-Dias JM; Torsney B (eds) MODA 8 – Advances in Model-oriented Design and Analysis*. Springer-Verlag, Heidelberg, 175-82, 2007.
- Rosner, B *Fundamentals of Biostatistics*, 2nd ed, Boston: Duxbury Press, 1986.
- Sahai H e Khurshid A Formulae and tables for the determination of sample sizes and power in clinical trials for testing differences in proportions for the two-sample design: A review. *Statist Med*, **15**(1): 1-21, 1996
- Sato T Sample size calculations with compliance information, *Statist Med*, **19**:2689-97, 2000.
- Scales DC e Rubenfeld GD Estimating sample size in critical care clinical trials. *J Crit Care*, **20**(1): 6-11, 2005.
- Schulz KF e Grimes DA Epidemiology 1 — sample size calculations in randomized trials: mandatory and mystical. *Lancet*, **365**: 1348-53, 2005.



- Sethuraman VS; Leonov S; Squassante L; Mitchell TR; Hale MD Sample size calculation for the power model for dose proportionality studies. *Pharmac Stat*, **6**:35-41.
- Shieh G; Show-Li J; Randles RH On power and sample size determinations for the Wilcoxon-Mann-Whitney test. *Nonparametric Statistics*, **18**(1): 33-43, 2006.
- Skalski JR Sample size calculations for Normal variates under Binomial censoring, *Biometrics*, **48**: 877-82, 1992.
- Stephan FF History of the Uses of Modern Sampling Procedures. *J Am Stat Assoc*, **43**(241): 12-39, 1948.
- Stigler SM *The history of Statistics*. Cambridge: The Belknap Press of Harvard University, 1986.
- Sturges HA The Choice of a Class Interval. *J Am Stat Assoc*, **21**: 65-6, 1926.
- Szwarcwald CL e Castilho EA Os caminhos da Estatística e suas incursões pela Epidemiologia. *Cad Saúde Pública*, **8**(1), Rio de Janeiro Jan/Mar, 1992.
- The WHOQOL group Development of the World Health Organization WHOQOL-BREF quality of life assessment. *Psychol Med*; **28**: 551-8, 1998.
- Thomas RG e Conlon M Sample size determination based on Fisher's Exact Test for use in 2 x 2 comparative trials with low event rates. *Control Clin Trials*, **13**: 134-47, 1992.
- Ury HK e Fleiss JL On Approximate Sample Sizes for Comparing 2 Independent Proportions with the Use of Yates Correction. *Biometrics*, **36**(2): 347-35, 1980.
- Van Belle G *Statistical rules for thumb*, 2nd ed, Boston: John Wiley & Sons, 2008.
- Van Belle G e Martin DC Sample-Size as a Function of Coefficient of Variation and Ratio of Means. *Am Statist*, **47**(3): 165-7, 1993.
- Vickers A How to reduce the number of patients needed for randomized trials: a basic introduction. *Complement Ther Med*, **9**: 234-6, 2001.
- Volatier JL; Turrini A; Welten D Some statistical aspects of food intake assessment. *Eur J Clin Nutr*, **56**(2): 546-52, 2002.

Vollandt R e Horn M A nonparametric method for selecting a good population. *Communications in Statistics*, **33**(7):1541-55, 2004.

Vollandt R e Horn M Evaluation of Noether's method of sample size determination for the Wilcoxon-Mann-Whitney test. *Biom J*, **39**(7): 823-9, 1997.

Wiley RH Is there an ideal behavioural experiment? *Anim Behav*, **66**: 585-8, 2003.

Willan AR e Pinto EM The value of information and optimal clinical trial design. *Statist Med*, **24**: 1791-806, 2005.

Wolfe R e Carlin JB Sample-size calculation for a log-transformed outcome measure. *Control Clin Trials*, **20**(6): 547-54, 1999.

Woolson RF; Bean JA; Rojas PB Sample size for case-control studies using Cochran's statistic. *Biometrics*, **42**: 927-32, 1986.

Zar JH *Biostatistical analysis*, 2nd ed, Englewood Cliffs, NJ: Prentice Hall International, 1984.

Zhao YD; Rahardja D; Qu YM Sample size calculation for the Wilcoxon-Mann-Whitney test adjusting for ties. *Statist Med*, **27**(3): 462-8, 2008.