Estimativa de proporções em questões politômicas

Ângelo Henrique Lopes da Silva é servidor do Tribunal de Contas da União, graduado em Engenharia Mecânica (ITA), Mestre e Doutor em Economia (UnB).

1. INTRODUÇÃO

Amostragem é uma das técnicas quantitativas mais presentes nos trabalhos de auditoria não só privada como governamental. Ela permite que, investigando-se uma parte, possa-se concluir sobre o todo. Entre as formulações amostrais, sobressai-se a estimativa de proporções, pela utilidade na identificação e mensuração das inconformidades em bancos de dados, tais como: balanços, folhas de pagamento ou arquivo de contratos. Por meio de tais estimativas, mensura-se a proporção de impropriedades no conjunto de observações auditadas e decide-se se a magnitude dos desvios é significativa para macular a administração.

A estimação de proporções ambienta-se em questões de dois tipos: dicotômicas e politômicas. As questões dicotômicas são aquelas que contêm dois itens de resposta, geralmente representadas pelo binômio sim/não. Esse tipo de questão e as respectivas estimativas possuem teoria bem desenvolvida na literatura sobre amostragem, dispensando assim comentários mais longos sobre o tema neste ar-

tigo. Já as questões politômicas, compostas de questões de mais duas categorias ou classes de resposta, não contam com literatura tão acessível para melhor compreensão das diferenças em relação às questões dicotômicas. Este artigo tem o objetivo de esclarecer e demonstrar a formulação diferenciada que acompanha a amostragem simultânea de várias proporções contidas em questões politômicas.

A primeira seção traz, como contraponto inicial ao tema principal, um resumo teórico da amostragem de uma proporção inserida em questões dicotômicas. Na seção seguinte, apresentamos o desenvolvimento téorico da amostragem de várias proporções contidas em questões politômicas e pontos coincidentes ao modelo dicotômico. Na terceira seção, expomos as correções de Bonferroni e de Sidak, que impõem a diferenciação entre as amostragens dos casos dicotômico e politômico. Na quarta seção, analisamos o impacto dessas correções no dimensionamento amostral. Finalmente, traçamos orientações para aplicação das citadas correções nos trabalhos de amostragem voltados a questões politômicas.

2. QUESTÕES DICOTÔMICAS

A amostragem de proporções em questões dicotômicas fundamenta-se na distribuição binomial, pois ela descreve e opõe duas classes de resultados, tais como: sucesso/fracasso, sim/não, conforme/inconforme etc. Uma distribuição binomial é representada por:

$$P(x) = \binom{n}{x} p^{x}(1 - p)^{n-x}$$

em que p é probabilidade de um evento determinado em cada tentativa e n é o número de tentativas. Para efeito de teste inferenciais, inclusive tendo como objetivo amostragem, o valor esperado de tal evento é $\boldsymbol{n \cdot p}$ e sua variância é de $\boldsymbol{n \cdot p \cdot (1 - p)}$.

Um exemplo proporciona esclarecimento. Caso se deseje realizar uma amostragem para verificar a proporção de contratos afetados por inconformidades, o cálculo do tamanho da amostra $\boldsymbol{n}$ seria:

$$n = \frac{N \cdot z^2 \cdot p \cdot q}{(N - 1) \cdot e^2 + z^2 \cdot p \cdot q}$$

Em que é **N** é o tamanho da população, **q** é igual a **(1-p)**, **e** é a margem de erro, e **z** é o fator da distribuição normal padronizada correspondente ao nível de significância **α**. Geralmente, o produto **p·q** é obtido do histórico de trabalhos anteriores ou, quando totalmente desconhecido, substituído por 0,25, valor máximo que proporcionará um cálculo conservador do tamanho da amostra.

3. QUESTÕES POLITÔMICAS

No caso de se terem questões politômicas, a distribuição que fundamenta o processo de amostragem é a multinomial, que é uma generalização da distribuição binomial para mais de duas proporções. Situações que podem ser modeladas pela probabilidade acima são questões de múltipla escolha, sejam de resposta única ou múltipla, escala de Likert, escala numérica, etc.

Sejam, então, as probabilidades **p1,...,p2**, satisfazendo a $\mathbf{0 \leq pi \leq 1}$, para **i=1,...,n**, e $\sum_{i=1}^{k} p_1 = 1$. Então, a probabilidade conjunta de se obterem as quantidades $(\boldsymbol{x_1,...,x_n})$, a partir de uma amostra de tamanho $\boldsymbol{m}$, é dada por:

$$P(x_1,...,x_n) = \frac{m!}{x_1!...x_2!}\, p_1^{x_1}...p_2^{x_2}$$

Pode-se demonstrar que a distribuição marginal de qualquer componente $\mathbf{X_i}$ da distribuição multinomial corresponde a uma Binomial$(\mathbf{m, p_i})$, em que **m** é o número de tentativas e $\mathbf{p_i}$ é a probabilidade de obtenção de $\mathbf{X_i}$ em cada tentativa (ver Apêndice A).

Assim, os equacionamentos dos parâmetros de amostragem (esperança e variância) que utilizamos para amostragem de questões dicotômicas são válidos para a amostragem de questões politômicas. Em função da equivalência demonstrada acima, a esperança de $\mathbf{X_i}$ é $\mathbf{m \cdot p_i}$ e a sua variância é $\mathbf{m \cdot p_i \cdot (1\text{-}p_i)}$, que são equivalentes ao caso binomial.

Para benefício de melhor esclarecimento, seja uma questão de auditoria operacional em que se busque identificar a proporção de impropriedades em uma folha de pagamento. Há quatro possibilidades: inconformidade A, inconformidade B, ambas inconformidades ou nenhuma delas. Essa situação seria modelada por uma distribuição multinomial de quatro proporções (**m**=4) e teria tamanho de amostra igual a:

$$n = \max_{p.q} \frac{N \,.\, z^2 \,.\, p \,.\, q}{(N\text{-}1) \,.\, e^2 + z^2 \,.\, p \,.\, q}$$

Essa formulação é equivalente à (1), a não ser pela escolha do produto **p·q** que maximiza o resultado do cálculo, bem com o valor de **z**, que merecerá atenção na seção seguinte. A escolha do valor máximo entre os resultados da fórmula decorre de conservadorismo, no sentido de garantir a obtenção de erros amostrais desejados para todas as proporções. Com o desconhecimento de uma estimativa inicial do produto entre **p** e **q**, a solução, mais uma vez, seria de fazê-lo igual 0,25.

Não obstante à equivalência dos parâmetros estatísticos dos modelos de amostragem binomial e multinomial, há uma diferença na inferência estatística para amostragem multinomial, quanto ao nível de significância do teste, que necessita de um ajuste no valor de **z** da distribuição.

4. AS CORREÇÕES DE BONFERRONI E SIDAK

Segundo Köhl e Magnussen (2010, p. 90), a estimação de intervalos de confiança para **k** classes precisa considerar que as estimativas de precisão são simultaneamente dadas para as **k** classes. Significaria distribuir o nível de significância global **α** pelos **k** intervalos de estimação. O método de Bonferroni distribui o nível de significância global, de modo igualitário, pelas classes individuais. Se temos 4 classes a serem testadas em conjunto com um nível de confiança global de 0,95 (α=0,05), o nível de significância corrigido por Bonferroni para cada classe será de $\alpha_4=\alpha/k$=0,05/4=0,0125. Assim, com a correção de Bonferroni, é possível realizar testes de hipótese conjuntos, mas controlando-se o nível de confiança global que cobre os níveis de confiança de cada um desses testes. Esse raciocício pode ser encontrado também em outros trabalhos, tais como Angers (1989, 1974). A demonstração da correção de Bonferroni pode ser obtida das relações booleanas da estatística (ver Apêndice B).

Desse modo, se uma questão de auditoria possuir **k** classes, dado que $\sum_{i=1}^{k} p_1 = 1$, apenas **k-1** classes precisarão ser testadas, o que resultará, pela correção de Bonferroni, que o nível de significância será igual ao nível de significância global dividido por **k-1**. Percebe-se que, para o caso de **k=2**, a correção de Bonferroni simplifica-se, resultando no caso dicotômico.

Seja o exemplo anterior, de 2 tipos de inconformidade gerando 4 itens de resposta e considere um nível de significância global de 0,05. O fator **z** a ser inserido na fórmula (2) deverá corresponder ao nível de significância de 0,01667 (=0,05/3), o que redundará em um **z**=2,39398. Logo, considerando **N**=20000, uma margem de erro de **e**=0,03 e o produto **p·q** de 0,25, a amostra será de:

$$n = \frac{20000 \,.\, 2{,}39398^2 \,.\, 0{,}25}{(20000 - 1) \,.\, 0{,}03^2 + 2{,}39398^2 \,.\, 0{,}25} = 1474{,}67 \cong 1475$$

Essa também é a solução apresentada por Oliveira (2004, p. 101-103), ao adotar um nível de significância individual de $\alpha/2(k-1)$, em que **k** é número total de classes. Ao perceber que, em Oliveira (2004), a divisão por dois resulta do teste bicaudal, podemos concluir que se trata de cálculos idênticos.

Aqui, vale um esclarecimento importante a ser feito quanto às afirmações de independência proferidas por tal publicação. Não se trata de independência estatística, mas tão somente de independência algébrica. Não faria sentido afirmar independência estatística entre variáveis aleatórias multinomiais, pois a condição construtiva $\sum_{i=1}^{k} p_1 = 1$ impõe uma característica de dependência estatística entre as variáveis aleatórias $\mathbf{X_1,\ldots,X_n}$, uma vez que a covariância entre $\mathbf{X_i}$ e $\mathbf{X_j}$ quaisquer é não nula e igual $\mathbf{m \cdot p_i \cdot p_j}$ (Casella e Berger, 2002, p. 182). A independência defendida por Oliveira (2004), pelo contrário, diz respeito à mera independência algébrica entre os itens de resposta, apenas afastada quanto ao último item de resposta que completa a soma unitária das proporções como comentado em parágrafo anterior. Deve-se lembrar também, finalmente, de que a formulação de Bonferroni aplica-se ao caso geral de dependência estatística, o que abrange também o caso particular de independência, esta que pode ser definida como dependência nula.

Uma alternativa à Bonferroni seria considerar que, ao se testar um conjunto de proporções, a um dado nível de significância, mais provável será incorrer-se em erros do tipo I –

erro de rejeitar uma hipótese nula verdadeira, denominado nível de significância (**α -**) do que se tivesse testando apenas uma proporção, ou seja, quanto mais testes são realizados, mais provável de que sejam encontrados eventos raros (Abdi, 2012).

Dado um nível de significância **α**, a probabilidade de não se incorrer no erro tipo I em um teste será de 1 - **α**. Caso se façam dois testes, e considerando independência (estatística) entre eles, a probabilidade de não se incorrer no erro tipo I será de $(1 - \alpha)^2$. Assim, para **k** testes, tem-se $(1 - \alpha)^k$. Para obter o nível de significância global, basta tomar o complemento do resultado anterior: $\alpha_{global} = 1 - (1 - \alpha_{individual})^k$, o que resulta na correção de Sidak:

$$\beta = 1 - (1 - \alpha)^{1/k}$$

Na qual se passa a denominar de **α**, o nível de significância global, e de **β**, o individual.

Seja, pois, o exemplo anterior das inconformidades. O fator **z** será correspondente ao nível de significância de 0,01695 ($= 1 - (1 - 0{,}05)^{1/3}$), o que redundará em um **z** = 2,387738. Logo, considerando **N** = 20000, uma margem de erro de **e** = 0,03 e o produto **p·q** de 0,25, a amostra será de:

$$n = \frac{20000 \cdot 2{,}387738^2 \cdot 0{,}25}{(20000 - 1) \cdot 0{,}03^2 + 2{,}387738^2 \cdot 0{,}25} = 1467{,}56 \cong 1468$$

5. IMPACTO DAS CORREÇÕES DE BONFERRONI E SIDAK

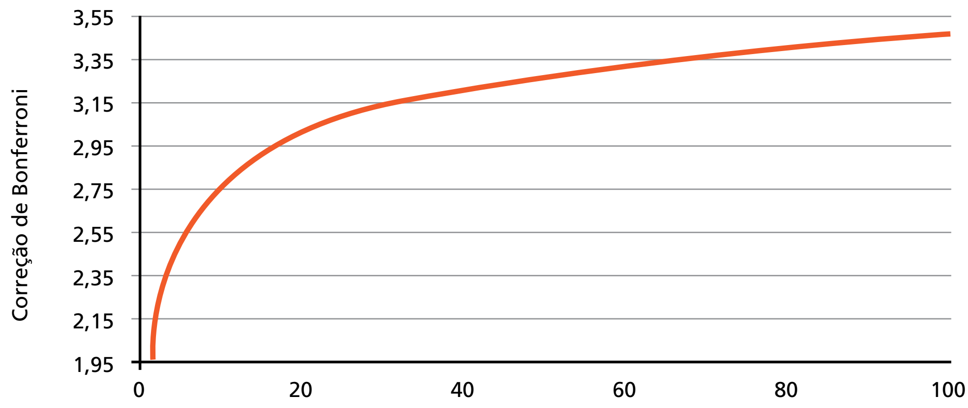
Pode-se demonstrar que a correção de Sidak sempre resulta em valor mais alto do que a de Bonferroni (ver Apêndice C), o que torna esta mais conservadora (pessimista) do que aquela, no sentido de se evitar mais erros do tipo I (Tabela 1).

Uma observação a ser feita é a de que **k**=2 resulta no caso binomial e em uma igualdade entre os níveis de confiança global e individual de cada teste. Apesar de muito próximos, a correção de Sidak possui a vantagem de proporcionar um tamanho de amostra menor do que a de Bonferroni, mas não olvidando que aquela, rigorosamente, aplica-se a testes independentes entre si, caso este que não é o mais frequente nas situações enfrentadas. A situação de dependência estatística na correção de Sidak, ao invés de igualdade, é representada pela desigualdade: $\beta \geq 1 - (1 - \alpha)^{1/k}$. Ao se adotar mais uma vez a solução mais conservadora, faz lógica continuar a se utilizar $\beta = 1 - (1 - \alpha)^{1/k}$. A correção de Sidak para ajuste da estimativa conjunta de proporções é adotada também em outros trabalhos como o manual de amostragem do Tribunal de Contas da União (Brasil, 2004, p. 59-60) e Souza (2000, p. 42-43), ambos trabalhando também com a independência das proporções sob o conceito algébrico, não o

Tabela 1
Correções de α por Bonferroni e Sidak

k	α = 5%		α = 10%	
	Bonferroni	Sidak	Bonferroni	Sidak
2	0,05000	0,05000	0,10000	0,10000
3	0,02500	0,02532	0,05000	0,05132
4	0,01667	0,01695	0,03333	0,03451
5	0,01250	0,01274	0,02500	0,02600
6	0,01000	0,01021	0,02000	0,02085
7	0,00833	0,00851	0,01667	0,01741
8	0,00714	0,00730	0,01429	0,01494
9	0,00625	0,00639	0,01250	0,01308
10	0,00556	0,00568	0,01111	0,01164

Gráfico 1
Variação da correção de Bonferroni em função do número de itens **k**, para **α** = 0,05



estatístico, como antes comentado. Não se deve olvidar outrossim que, em uma questão de auditoria de **k** itens de resposta, com $\sum_{i=1}^{k} p_1 = 1$, o expoente da fórmula de Sidak é $1/(k - 1)$.

O valor **z** da distribuição normal, não obstante aumente bastante, não explode com o quantitativo maior de itens **k**, não comprometendo assim o tamanho da amostra. Na Figura 1, para **α**=5%, plotamos os valores de **z** em função de **k**, pela formulação de Bonferroni, que a opção mais conservadora a resultar em **z** mais elevado. Tomando-se, como parâmetro de comparação o **k**=2, em que convencionalmente **z**=1,9599, podemos verificar quão reduzido é o avanço de **z** com aumento de **k**, a ponto de **z** ser apenas ainda 3,4780, quando **k** atinge 100. Nesse caso, o aumento do tamanho da amostra, considerando a população suficientemente grande, seria de 216% (=(3,4780/1,9566)2-1), mesmo considerando um aumento de **k**=2 para **k**=100.

6. TRATAMENTO DE QUESTÕES POLITÔMICAS

Questões com itens de respostas cujas proporções são complementares, ou seja, cuja soma iguala-se à unidade, possuem uma evidente dependência estatística. Nesses casos, embora ainda seja possível o desmembramento da questão em questões dicotômicas, tal procedimento não é recomendável, uma vez que o que se busca é uma amostragem com nível de significância global da questão como todo.

Um exemplo prático em auditoria pode esclarecer a importância de uma amostragem com nível de significância global. Seja um trabalho de auditoria que busca em contratos os tipos de inconformidade A, B e C, por meio de um questionário que contenha uma questão de múltipla escolha. Há duas formas de se implementar a questão (Tabela 2).

Tabela 2
formatos de questões de múltipla escolha

Forma 1 – resposta única	Forma 2 – resposta mútipla
Marque o item correpondente às inconformidades verificadas no contrato examinado: (a) Inconformidade A (b) Inconformidade B (c) Inconformidade C (d) Inconformidades A e B (e) Inconformidades A e C (f) Inconformidades B e C (g) Inconformidades A, B e C (h) nenhuma delas	Marque, caso haja, as inconformidades encontradas no contrato examinado (pode marcar mais de um item): (a) Inconformidade A (b) Inconformidade B (c) Inconformidade C

Observe que as formas são equivalentes quanto ao objetivo pretendido de mapear todas as situações de inconformidade dos contratos. A título de exemplo, caso um contrato possua as inconformidades A e C, perante a forma 1, a marcação será o item (e), enquanto, perante a forma 2, os itens (a) e (c) serão apontados simultaneamente. Mesmo a inexistência de inconformidades, que, na forma 1, possui o item (h), pode ser registrado na forma 2, pela ausência de marcação.

Outrossim, os procedimentos de amostragem são equivalentes. Uma quantidade de tipos **n** implica a construção de 2^n itens na forma 1 da questão. No caso exemplificado, três tipos de inconformidades redundarão em $2^3=8$ itens na forma 1, ou seja, o número de partes do conjunto das inconformidades {A,B,C}. Então, 8 seria o valor de **k** para dimensionamento da amostra, segundo uma amostragem com o nível de significância global regida pela fórmula (2), em que seriam estimadas 7 ($=2^n-1$) proporções, devido a 7 proporções independentes algebricamente, sendo a oitava dependente das restantes em razão do somatório das proporções igual à unidade. No entanto, como já comentado, a forma 2 representa uma economia de espaço na implementação da questão, sem comprometer os resultados buscados. Do mesmo modo, o número de itens **k** a serem considerados no dimensionamento da amostra deve considerar o número de partes do conjunto 2^n.

Tabela 3
fator **z** em função de **n**

n	k	z
1	2	1,959964
2	4	2,393980
3	8	2,690110
4	16	2,935199
5	32	3,153563
6	64	3,355000
7	128	3,544271
8	256	3,724016
9	512	3,895869
10	1024	4,060938

Ao se utilizar Bonferoni, percebe-se que, mesmo com o aumento exponencial de **k** itens decorrente da reduzida elevação da quantidade de tipos n, o **z** não cresce exacerbadamente, como se pode ver na Tabela 3. Como se verifica, **n**=10 gera **k**=1024 itens, mas mal consegue dobrar o valor de **k**, quando comparado com seu valor convencional de **z**=1,9599. Mas, ao se vislumbrar que, dificilmente, uma questão de auditoria conteria 10 tipos de inconformidade e as consequentes combinações, conclui-se que a majoração do

fator **z** devido à correção de Bonferroni certamente não inflará o tamanho da amostra a ponto de inviabilizá-la.

Por fim, é possível o desmembramento de questões politômicas em tantas questões dicotômicas quantos forem os itens de resposta independentes algebricamente. Mas essa deve ser uma solução adotada com parcimônia e apenas quando, definitivamente, fizer sentido uma estimação, em separado, das proporções de cada item de resposta. Melhor explicando, há questionamentos em auditoria que se afastam da condição construtiva de dependência estatística e complementaridade das proporções: $\sum_{i=1}^{k} p_1 = 1$. Suponha o objetivo de identificar contratos que apresentem inconformidades do tipo A e do tipo B. As proporções (ou probabilidades) de ocorrências entre inconformidades de tipo A e B são, a princípio, independentes estatisticamente e devem ser tratadas em questões dicotômicas separadas. Aliás, nesse caso, uma questão politômica não seria necessária, a menos que haja interesse em se mensurar a ocorrência simultânea dos dois tipos de inconformidade (interseção das ocorrências), o que seria incluir um item de resposta dependente estatisticamente aos itens anteriores.

APÊNDICE A

Adotando a demonstração por Casella (2002, p. 181-182), faz-se necessário primeiro apresentar o teorema multinomial.

Teorema multinomial: Seja m e n inteiros positivos. Seja A o conjunto de vetores $x=(x_1,...,x_n)$ tal que cada x_i é um inteiro não negativo e $\sum_{i=1}^{n} x_i = m$. Então, para quaisquer números reais $p_1,...,p_n$,

$$(p_1 + \ldots + p_n)^m = \sum_{x \in A} \frac{m!}{x_1! \ldots x_n!} p_1^{x_1} \ldots p_n^{x_n}$$

Considerando que os pi probabilidades tais que $\sum_{i=1}^{n} p_i = 1$, tem-se que $(p_1 + \ldots + p_n)^m = 1$.

Calcula-se, então, a distribuição marginal de $\mathbf{X}_n$ de uma distribuição multinomial. Para um valor fixo de x_n, com o objetivo de calcular distribuição marginal de $\mathbf{f}(\boldsymbol{x}_n)$, deve-se somar todos os possíveis valores de $(x_1,...,x_{n-1})$, ou seja, a soma sobre todos os $(x_1,...,x_{n-1})$ tal que os x_i sejam todos inteiros não negativos e $\sum_{i=1}^{n-1} x_i = m - x_n$, conjunto o qual chamaremos de B. Então,

$$f(x_n) = \sum_{(x_1,\ldots,x_{n-1}) \in B} \frac{m!}{x_1! \ldots x_n!} p_1^{x_1} \ldots p_n^{x_n}$$

$$f(x_n) = \sum_{(x_1,\ldots,x_{n-1}) \in B} \frac{m!}{x_1! \ldots x_n!} p_1^{x_1} \ldots p_n^{x_n} \frac{(m-x_n)!\,(1-p_n)^{m-x_n}}{(m-x_n)!\,(1-p_n)^{m-x_n}}$$

$$f(x_n) = \frac{m!}{x_1(m-x_n)!} p_n^{x_n} (1-p_n)^{m-x_n} \sum_{(x_1,\ldots,x_{n-1}) \in B} \frac{(m-x_n)!}{x_1! \ldots x_{n-1}!} \left(\frac{p_1}{1-p_n}\right)^{x_1} \ldots \left(\frac{p_{n-1}}{1-p_n}\right)^{x_{n-1}}$$

Usando o fato de que $\boldsymbol{x_1 + ... + x_{n-1} = m - x_n}$ e $\boldsymbol{p_1 + ... + p_{n-1} = 1 - p_n}$, e ainda o teorema multinomial, pode-se ver que a somatório acima é igual a 1. Assim, demonstra-se que a distribuição marginal de $\mathbf{X_n}$ é uma binomial $(\mathbf{m,p_n})$.

APÊNDICE B

É conhecida a relação:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Como $P(A \cup B) \leq 1$, tem-se que $P(A \cap B) \geq P(A) + P(B) - 1$ que é conhecida por desigualdade de Bonferroni.

Essa desigualdade permite que se limite por baixo a probabilidade de eventos simultâneos, a interseção, em termos das probabilidades de eventos individuais (Casella e Berger, 2002, p. 11). Assim, suponha que os eventos **A** e **B** possuam cada um uma probabilidade de 0,95.

Então a probabilidade de ambos ocorrerem é limitada inferiormente por:

$$P(A \cap B) \geq P(A) + P(B) - 1 = 0{,}95 + 0{,}95 - 1 = 0{,}90$$

Pode-se generalizar a desigualdade de Bonferroni (Casella e Berger, 2002, p. 13). Aplicando à desigualdade Booleana o conjunto A^C, tem-se que:

$$P\left(\bigcup_{i=1}^{n} A_i^c\right) \leq \sum_{i=1}^{n} P(A_i^c)$$

Como, $\cup A_i^c = (\cap A_i)^c$ e $P(A_i^c) = 1 - P(A_i)$, temos

$$1 - P\left(\bigcap_{i=1}^{n} A_i^c\right) \leq n - \sum_{i=1}^{n} P(A_i)$$

$$P\left(\bigcap_{i=1}^{n} A_i^c\right) \geq \sum_{i=1}^{n} P(A_i) - (n - 1)$$

que é a versão generalizada da desigualdade de Bonferroni.

Por fim, se $\mathbf{1 - \alpha}$ é o nível de confiança simultâneo sobre um conjunto de **n** níveis de confiança individuais $\mathbf{1 - \beta}$, utilizando a desigualdade de Bonferroni, obtém-se como limite para o nível de confiança $\boldsymbol{\beta}$**:**

$$1 - \alpha \geq \sum_{i=1}^{n} (1 - \beta) - (n - 1)$$

$$\Rightarrow \beta \geq \alpha/n$$

Por conservadorismo, usa-se $\boldsymbol{\beta = \alpha/n}$.

APÊNDICE C

Demonstra-se que: $\frac{\alpha}{2(k-1)} \leq \frac{1-(1-\alpha)^{1/(k-1)}}{2}$, para $k-1 \geq 1$.

Substituindo-se ***k*-1** por ***n*** para simplificar.

Então, quando um inteiro $n \geq 1$,

$[1-(\alpha/n)]^n = 1 - n(\alpha/n) + [(n-1)/2n]\alpha^2 - \{[(n-1)(n-2)]/6n^2\}\ \alpha 3 + \ldots$

Como $\alpha < 1$ e é pequeno, $[(n-1)/2n]\alpha^2 - \{[(n-1)(n-2)]/6n^2\}\ \alpha 3 + \ldots$ é positivo.

Logo, $[1-(\alpha/n)]^n \geq 1-\alpha \Rightarrow 1-(\alpha/n) \geq (1-\alpha)^{1/n} \Rightarrow 1-(1-\alpha)^{1/n} \geq (\alpha/n)$

Logo, voltando com **k-1** no lugar de **n** e dividindo por 2, tem-se:

$$\frac{1-(1-\alpha)^{1/(k-1)}}{2} \geq \frac{\alpha}{2(k-1)}, \ \forall k \geq 2$$