

6

Um estudo de simulação para avaliação dos estimadores associados aos coeficientes de escalonabilidade da TRIN

Este capítulo apresenta a descrição de dois estudos de simulação adequados ao contexto da amostragem de população finita que foram desenvolvidos com o intuito de investigar empiricamente algumas propriedades estatísticas dos estimadores pontuais e de variância propostos para os *coeficientes de escalonabilidade* H_{ij} , H_i e H no modelo de Homogeneidade Monótona (HM). Este modelo foi empregado para medir o construto de capital econômico dos alunos pertencentes às *populações de referência* de Roraima e do Rio de Janeiro.

Além disso, nestes estudos foram comparados dois planos amostrais. O primeiro é um plano amostral com conglomeração em um único estágio (ACIS), no qual as unidades primárias (escolas) são selecionadas por *amostragem aleatória simples sem reposição*. O segundo plano amostral (AC2-SAEB) apresenta as características de estratificação, conglomeração em dois estágios (escolas e turmas) e seleção das unidades do primeiro estágio com probabilidade proporcional ao tamanho (PPT). Este plano imita aspectos principais do plano amostral empregado em várias edições do Sistema de Avaliação da Educação Básica (SAEB) (Silva et al, 2001, p.5).

6.1. Descrição dos estudos de simulação

Foi usada a simulação para a avaliação da qualidade dos estimadores pontuais e de variância propostos para os *coeficientes de escalonabilidade* H_{ij} , H_i e H nas populações de referência de Roraima e Rio de Janeiro.

O primeiro passo num estudo de simulação consiste em delimitar a *população-alvo* e posteriormente, a *população de referência* e, em seguida, definir os parâmetros e variáveis de interesse.

Para o Estado de Roraima, a *população-alvo* deste estudo é constituída pelo conjunto de alunos que freqüentavam a 8ª série (9º ano) do Ensino Fundamental na rede pública, em área urbana em 2007.

Para o Estado do Rio de Janeiro, a *população-alvo* é formada pelo conjunto de alunos que freqüentavam a 8ª série do Ensino Fundamental na rede pública, em área urbana em 2007.

O *Sistema de referência* (cadastro) é composto pelo conjunto de escolas, turmas e alunos avaliados pelo Sistema de Avaliação do Rendimento Escolar, denominado de Prova Brasil, em 2007 (MEC, 2007).

A escolha deste cadastro como sistema de referência para a realização dos estudos de simulação foi feita porque este contém informações atualizadas sobre o contexto escolar brasileiro e, além disso, disponibiliza um questionário socioeconômico que permite a mensuração do traço latente (capital econômico) através do *modelo de Homogeneidade Monótona* (HM) da TRIN.

Este *sistema de referência*, restrito à área urbana do Estado de Roraima, contém informações sobre 69 escolas públicas, 203 turmas e 4245 alunos na 8ª série do Ensino Fundamental. Deste total de escolas, 41 possuem pelo menos três turmas¹.

A *população de referência* U_1 é formada pelos alunos do Estado de Roraima que freqüentavam a 8ª série do Ensino Fundamental, na rede pública, em área urbana que participaram da Prova Brasil 2007.

Para a área urbana do Estado do Rio de Janeiro, este *sistema de referência* possui informações sobre 1719 escolas públicas, 4370 turmas e 109314 alunos da 8ª série.

A *população de referência* U_2 é formada pelos alunos que freqüentavam a 8ª série do Ensino Fundamental, na rede pública, em área urbana do Estado do Rio de Janeiro que participaram da Prova Brasil 2007.

Em ambas as populações de referência, a distribuição do total de alunos participantes por escola avaliada pela Prova Brasil em 2007 é assimétrica positiva (Figura 5.1).

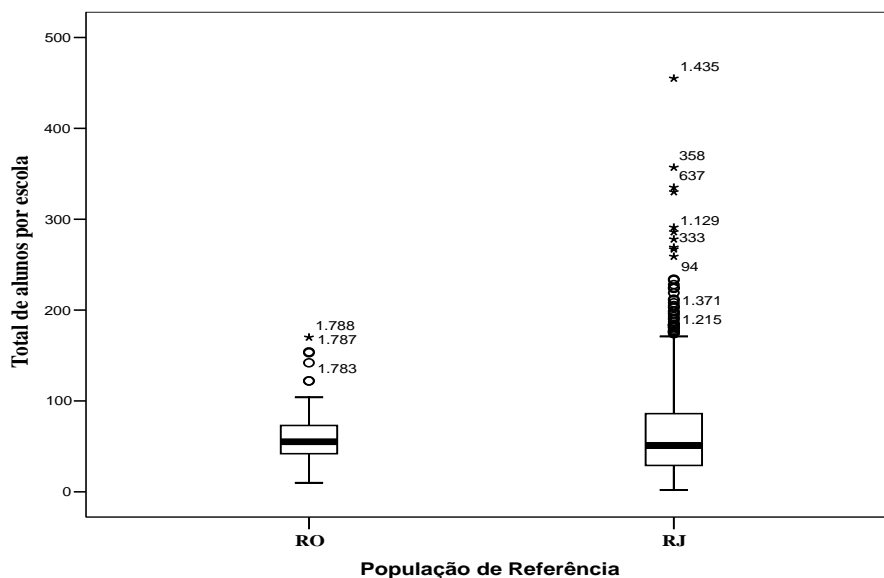
¹ Uma escola é classificada como escola de tamanho grande quando existem pelo menos três turmas. Em caso contrário, uma escola é definida como escola de tamanho pequeno.

Na população de referência de Roraima, em 96.4% das escolas de tamanho pequeno participaram no máximo 55 alunos, enquanto que em apenas 19.5% das escolas de tamanho grande o total de alunos atingiu este valor. Vale ressaltar que existem apenas 5 escolas que possuem mais de 100 alunos² (Figura 6.1).

Na população de referência do Rio de Janeiro, em 81.1% das escolas de tamanho pequeno participaram no máximo 51 alunos enquanto que em 6.2% das escolas de tamanho grande o total de alunos alcançou este número de alunos. Além disso, existe um número expressivo de escolas que possuem mais de 150 alunos (Figura 6.1).

De modo geral, estes resultados refletem a estrutura das escolas (em termos do total de alunos) em cada população de referência. Estas informações foram importantes no momento da seleção das réplicas de amostras conglomeradas de alunos sob os planos amostrais *ACIS* e *AC2-SAEB* utilizados nos estudos de simulação (Quadro 6.3).

Figura 6.1: Boxplot do total de alunos participantes da 8ª série por escola avaliada na Prova Brasil 2007 Roraima e Rio de Janeiro



A mensuração do capital econômico foi considerada nesta tese em virtude do emprego desta medida como uma *proxy* da condição econômica dos alunos. É comum esta prática, principalmente, em estudos educacionais que visam à

² As observações que estiverem acima do limite superior (cauda superior) estabelecido em cada boxplot e estão representadas pelos asteriscos ou números são escolas destoantes das demais em termos do total de alunos por escola.

avaliação do impacto do nível socioeconômico, por exemplo, na proficiência (Andrade, 2003, 2002; Ferrão e Fernandes, 2000), no desempenho no vestibular da UFMG (Soares, 2003), dentre outros aspectos (Soares; 2005; Soares, 2003; Soares e Manbrini, 2003).

Em cada *população de referência*, a partir do questionário do aluno que participou de pelo menos uma prova (Língua Portuguesa ou Matemática), foram escolhidos dez itens politômicos ordinais, com boas propriedades psicométricas da TRIN, que mensuram o construto de capital econômico, a saber: 14, 5, 6, 12, 15, 8, 13, 7, 10 e 11 conforme consta no Apêndice 1A. Estes foram dicotomizados e rotulados como *item 1*, *item 2*, ..., *item 10*, respectivamente e são apresentados no Quadro 6.1.

Quadro 6.1: Itens dicotomizados do construto capital econômico da 8ª série
Prova Brasil - Roraima e Rio de Janeiro - 2007

Item <i>i</i>	Questionário	Tipo
	Na sua casa tem...?	
1	Banheiro	
2	televisão em cores	
3	Rádio	
4	Carro	
5	alguma empregada doméstica	1: sim
6	Geladeira	0: não
7	Computador	
8	vídeo cassete ou DVD	
9	máquina de lavar roupa	
10	aspirador de pó	

6.2. Geração das réplicas de amostras

Para cada plano amostral descrito a seguir a simulação compreendeu o sorteio de $R^*=2000$ réplicas independentes de amostras de cada *população de referência*.

6.2.1 Plano amostral AC1S

O primeiro plano amostral considerado foi amostragem de conglomerados simples em um estágio (AC1S). Neste plano, cada conglomerado (escola) foi

selecionado segundo a amostragem aleatória simples sem reposição. Na seqüência, todos os alunos nas escolas selecionadas foram incluídos na amostra. Neste estudo, este plano amostral é denominado abreviadamente por *ACIS*.

Para este plano, as *unidades primárias de amostragem (UPAs)* são as “escolas” denominadas como as partes das escolas públicas correspondentes ao conjunto de turmas e alunos da 8ª série do Ensino Fundamental em cada *população de referência*.

Na *população de referência* de Roraima, cada amostra selecionada segundo o plano *ACIS* é composta por 23 escolas sorteadas segundo a *AAS* sem reposição da *população de referência* com 69 escolas.

Na *população de referência* do Rio de Janeiro, cada amostra sorteada de acordo com o plano *ACIS* é composta por 90 escolas selecionadas por *AAS* de um total de 1719 escolas.

6.2.1.1 Ponderação sob o plano amostral ACIS

No processo de amostragem probabilística das unidades populacionais (escolas, turmas, alunos), de acordo com um plano amostral escolhido, cada unidade tem probabilidade positiva e conhecida de ser incluída na amostra. Essas probabilidades de inclusão determinam os pesos amostrais básicos (ou pesos do desenho ou fatores de expansão), que por sua vez permitem a obtenção de estimativas pontuais assintoticamente não viciados dos *coeficientes de escalonabilidade*, bem como as estimativas de variância dos estimadores de H_{ij} , H_i e H . Assim sendo, o peso amostral básico é definido como o inverso da probabilidade de inclusão da unidade na amostra.

Para o plano *ACIS*, como ocorreu apenas a seleção de unidades primárias de amostragem sob *AAS* sem reposição, a probabilidade de inclusão da escola na amostra é definida como a razão entre o tamanho da amostra de escolas e o total de escolas no cadastro. Desta forma, para a população de referência de Roraima, a probabilidade de inclusão de cada escola na amostra assume o valor 23/69. Como conseqüência, o peso de cada aluno proveniente das escolas sorteadas é o valor constante três, que corresponde a 69/23.

Para *população de referência* do Rio de Janeiro, a probabilidade de inclusão de cada escola na amostra assume o valor 90/1719. Conseqüentemente, o peso de cada aluno proveniente das escolas selecionadas é o valor constante 19.1.

Para o plano *ACIS*, na área urbana de Roraima e do Rio de Janeiro, respectivamente, uma vez selecionadas as 2000 réplicas de amostras *ACIS*, foram obtidas as estimativas dos parâmetros H_{ij} , H_i e H e correspondentes estimativas de variância, conforme os estimadores de variância propostos no Capítulo 4.

Para os processos de amostragem e estimação foram desenvolvidos programas na linguagem *R* (versão 2.11), com auxílio das bibliotecas *sampling* e *survey*.

6.2.2 Plano amostral *AC2-SAEB*

Este plano possui as seguintes características: estratificação das unidades primárias de amostragem (segundo o tamanho da escola), conglomeração em dois estágios (sorteio de escolas e turmas) e seleção das unidades do primeiro estágio com probabilidade proporcional a uma medida de tamanho (PPT) da escola, além da ponderação diferente para os diversos alunos na amostra. Neste estudo, este plano amostral é denominado de *AC2-SAEB*.

Na *população de referência* de Roraima, as 69 escolas da 8ª série (conglomerados naturais de alunos) constituem as *unidades primárias de amostragem* enquanto que as 203 turmas formam as *unidades secundárias de amostragem*.

Na *população de referência* do Rio de Janeiro, as 1719 escolas da 8ª série formam as *unidades primárias de amostragem* enquanto que as 4370 turmas compõem as *unidades secundárias de amostragem*.

A medida de tamanho de uma escola é definida pela quantidade de turmas existentes de 8ª série na escola. Dessa forma, uma escola é classificada de tamanho pequeno quando possui no máximo duas turmas, e de tamanho grande quando contém três ou mais turmas da mesma série.

Para este plano *AC2-SAEB* são definidos dois estratos denominados de Estrato 1 e Estrato 2. Na população de referência de Roraima, o primeiro estrato contém 28 escolas de tamanho pequeno, 49 turmas e 991 alunos. No segundo,

existem 41 escolas de tamanho grande nas quais estão distribuídas 154 turmas e 3254 alunos.

Na *população de referência* do Rio de Janeiro, segundo o plano AC2-SAEB, são definidos também dois estratos. No primeiro, existem 1026 escolas de tamanho pequeno, 1562 turmas e 36586 alunos. O segundo estrato contém 693 escolas de tamanho grande nas quais estão distribuídas 2808 turmas e 72728 alunos.

6.2.2.1 Dimensionamento da amostra sob o plano AC2-SAEB

Para especificar quantas unidades de primeiro (escolas) e segundo estágios (turmas) foram selecionadas em cada estrato de interesse em cada *população de referência*; foi empregada a mesma abordagem utilizada por Silva et al (2001, p.5) para a seleção das amostras do SAEB 2001. Além do mais, nos cálculos para o tamanho da amostra de turmas foram mantidas as informações referentes ao total de disciplinas avaliadas (Matemática, Língua Portuguesa e Ciências) e a quantidade de 300 alunos por disciplina selecionados pelo plano amostral do SAEB 2001 na 8ª série, conforme a expressão (6.1).

Assim, o tamanho da amostra de escolas é uma função do tamanho da amostra de turmas. Dessa forma, primeiramente é dimensionado o tamanho da amostra de turmas através da seguinte expressão:

$$n_{turmas} = arredondar \frac{3 \times 300 \times N_{turmas}}{N_{alunos}} \quad (6.1)$$

onde: n_{turmas} : número de turmas a selecionar para a amostra; N_{turmas} : total de turmas no cadastro; N_{alunos} : total de alunos no cadastro.

Em seguida, a partir do tamanho total da amostra de turmas (n_{turmas}), obtido pela expressão (6.1), foi empregada a *alocação proporcional* para o dimensionamento da amostra de turmas nos estratos definidos pelo plano AC2-SAEB.

O Quadro 6.2 mostra os tamanhos da amostra de turmas e de escolas em cada estrato, em cada população de referência.

Na *população de referência* de Roraima, para o sorteio da amostra de unidades do primeiro estágio em cada estrato foi empregada *amostragem proporcional à medida de tamanho da escola* (quantidade de turmas). Desta

forma, no primeiro estrato foram selecionadas 11 escolas e no segundo, 16 escolas (Quadro 6.2).

Quadro 6.2: Tamanho da amostra de escolas e turmas nos estratos de interesse das populações de referência: Roraima e Rio de Janeiro sob o Plano Amostral AC2-SAEB - 2007

População de referência	Estrato	Turmas	Escolas
RR	1	11	11
	2	32	16
RJ	1	48	48
	2	84	42

Na seqüência, foram selecionadas turmas, através da *amostragem aleatória simples sem reposição*, dentro das escolas selecionadas na primeira etapa, dentro de cada estrato. Uma vez escolhida uma turma, todos os alunos foram incluídos.

Cabe ressaltar que na escola de tamanho pequeno é sorteada uma turma e na escola de tamanho grande são selecionadas duas turmas. Desta forma, a amostra de 43 turmas está distribuída em 11 turmas e 32 turmas, respectivamente, no primeiro e segundo estratos.

Na *população de referência* do Rio de Janeiro, para a seleção da amostra de escolas em cada estrato foi empregada novamente a *amostragem proporcional à medida de tamanho da escola*. Desta forma, a amostra de escolas é composta por 48 escolas de tamanho pequeno e 42 escolas de tamanho grande, totalizando assim, 90 escolas sorteadas. Isto implica que a *fração amostral* no 1º estágio assume o valor de 5.2%. Além disso, no segundo estágio foram selecionadas 48 turmas e 84 turmas nas escolas selecionadas, respectivamente, para compor a amostra de turmas no primeiro e segundo estratos (Quadro 6.2).

Para o plano AC2-SAEB, uma vez selecionadas as 2000 réplicas de amostras de escolas, turmas e seus alunos, foram obtidas as estimativas dos coeficientes de escalonabilidade H_{ij} , H_i e H e correspondentes estimativas de variância, conforme os estimadores de variância propostos no Capítulo 4.

Para os processos de amostragem e estimação foram desenvolvidos programas na linguagem R (versão 2.11), com auxílio das bibliotecas *sampling* e *survey*.

6.2.2.1 Ponderação sob o plano amostral AC2-SAEB

Como é de interesse neste estudo obter estimativas pontuais dos coeficientes de escalonabilidade bem como avaliar a qualidade dessas estimativas pontuais, é preciso definir os pesos amostrais básicos para as escolas, turmas e alunos, em cada estrato de cada *população de referência*, segundo o plano AC2-SAEB, conforme descrição a seguir.

a) **Peso da Escola**

O peso da escola l no estrato h é definido como:

$$w_{hl} = \frac{1}{n_h} \frac{M_h}{M_{hl}}, \quad h=1,2. \quad (6.2)$$

onde

n_h : número de escolas selecionadas em cada estrato h ;

M_h : total de turmas no estrato h ;

M_{hl} : total de turmas na escola l no estrato h .

b) **Peso da Turma**

O peso final da turma t na escola l no estrato h é dado por:

$$w_{hlt} = \frac{1}{n_h} \frac{M_h}{M_{hl}} \frac{M_{hl}}{m_{hl}}, \quad h=1,2. \quad (6.3)$$

onde

m_{hl} : número de turmas a selecionar na escola l no estrato h .

Convém destacar que na composição da amostra de turmas em cada estrato de cada *população de referência*, o valor m_{hl} assume o valor unitário (uma turma selecionada na escola selecionada l no Estrato 1) ou o valor 2 (duas turmas selecionadas na escola selecionada l no Estrato 2), conforme o plano amostral do SAEB 2001 (Silva et al., 2001). Vale notar que os pesos das turmas assim definidos são constantes por estrato.

c) **Peso do Aluno**

Como o plano AC2-SAEB apresenta apenas dois estágios de seleção (escola e turma), o peso básico do aluno, dentro de cada estrato, coincide com o peso da turma selecionada na escola selecionada.