

## **Identificação das características associadas com a aprovação de candidatos de escolas públicas e privadas, Vestibular-2004, UFMG**

**Identification of the factors associated with the approval of candidates from public and private schools, Vestibular-2004, UFMG**

**Cristiane B. Lopes, Raphael L. F. Ribeiro, Mariana G. Carvalho<sup>1</sup>  
Glaura C. Franco<sup>2</sup>  
Rosângela H. Loschi<sup>3</sup>  
Mauro M. Braga<sup>4</sup>**

### **RESUMO**

Este trabalho visa a conhecer melhor o perfil dos candidatos oriundos de escolas das redes públicas e privadas de ensino que tentaram ingressar na UFMG em 2004. Busca-se identificar quais das características definidas no questionário socioeconômico e cultural aplicado no ato da inscrição do candidato podem estar mais associadas com a aprovação no vestibular. Conclui-se que o local de moradia e o conhecimento de língua estrangeira são as variáveis mais fortemente associadas com a aprovação do candidato de escolas particulares e escolas públicas, respectivamente. Verificou-se que, entre os candidatos que concluíram o ensino médio em escolas públicas, os que estudaram em escolas públicas federais tendem a se concentrar nos grupos com maiores chances de aprovação.

**Palavras-Chave:** CART; Fator Socioeconômico; Curva ROC.

---

<sup>1</sup> Bacharel em Estatística. Departamento de Estatística/ICEX/UFMG.

<sup>2</sup> Doutora em Engenharia Elétrica. Professora do Departamento de Estatística da UFMG.

<sup>3</sup> Doutora em Estatística. Professora do Departamento de Estatística da UFMG.

<sup>4</sup> Doutor em Química. Professor do Departamento de Química e Pró-Reitor de Graduação da UFMG.

**ABSTRACT**

This paper aims to better understand the profile of the candidates for the UFMG entrance examination in 2004, coming from public and private schools. The objective is to identify which of the characteristics defined through a socioeconomic and cultural questionnaire answered by the candidates upon their application for the entrance examination may be associated with their approval at the University exams. It was found that the place where the candidates live and their knowledge of a foreign language are the variants more strongly related to the approval of the candidates of private schools and public schools, respectively. It was also found that, among the candidates who concluded high school at public schools that attained the highest chances of approval, there was a large percentage of candidates that attended high school in federal establishments.

**Keywords:** CART; socioeconomic factor; ROC curve.

**INTRODUÇÃO**

O recorrente tema da democratização do acesso à educação superior no Brasil, em particular no caso do ensino público, ganhou notável alento nos anos recentes. Em parte, isso decorre de transformações que extrapolam a sociedade brasileira, mas, certamente, as questões internas do país desempenharam papel mais relevante para a ênfase que caracterizou a retomada deste debate. De um lado, os mecanismos de globalização das relações econômicas acentuaram, nas mais diferentes sociedades, o importante papel da educação superior, o que ocorreu concomitantemente com a perda da hegemonia das universidades na produção do conhecimento. Santos (1996) menciona, com pertinência, a crise de legitimidade do meio acadêmico, decorrente, entre outros fatores, do confronto entre os saberes erudito e popular.

Por outro lado, o acesso ao ensino médio registrou expressivo crescimento ao longo da última década do século passado, crescimento este concentrado na educação pública. A título de exemplo, cita-se o fato de o número de concluintes do ensino médio em Minas Gerais, nos sete primeiros anos da década de 1990, ter triplicado, aumento este que se verificou, quase todo ele, na rede pública. Em contrapartida, o ensino universitário público cresceu em proporções bem menores, de tal sorte que, em 1990, mais de 40% das matrículas na educação superior brasileira ocorriam em instituições públicas. Ressalta-se que tal percentual, em 2004, foi inferior a 30%. A combinação de todos esses fatores resultou em crescente pressão para que houvesse ampliação do acesso ao ensino superior público, acesso este que, cada vez mais, é matizado pela interferência de fatores socioeconômicos.

Essa pressão, como seria natural, derivou para o debate de mecanismos alternativos de acesso ao ensino superior público. Conquanto esses mecanismos alternativos tenham contemplado nuances diversas, entre as quais se incluem a criação da USP zona leste, a pontuação adicional para egressos da escola pública e para estudantes que se autodeclarem negros, adotada pela UNICAMP, e a expansão do turno noturno, priorizada pela UFMG, o centro do debate das ações das universidades voltou-se para a adoção de quotas, de reserva de vagas, para egressos da escola pública e para estudantes negros e indígenas. Esse caminho, escolhido, inicialmente, em decorrência de lei estadual, trilhado pelas universidades estaduais do Rio de Janeiro, vem sendo, com desenho diversificado, adotado por muitas instituições federais, entre as quais se incluem a UFBA, a UnB e a UFPR.

Essa alternativa foi também adotada pelo Governo Federal, que incluiu a reserva de vagas como um dos pilares de um projeto de reforma universitária que encaminhou ao Congresso. O realce dado à temática, em determinado momento, resultante, possivelmente, da pressão de grupos

organizados da sociedade civil, em especial os movimentos negros, levou as autoridades do Ministério da Educação (MEC) a priorizarem esse aspecto, no contexto da reforma proposta, destacando o projeto de lei específico para tal fim, que tramita em separado do projeto da reforma universitária propriamente dita.

A questão é, contudo, bastante polêmica, tendo em vista a estruturação do sistema de ensino superior voltado para as elites e que, em razão de sua herança liberal, tem o mérito como princípio determinante para o ingresso. Assim sendo, as cotas, sejam elas para escola pública ou para os que se declaram negros ou indígenas, têm mobilizado o debate e agregado segmentos da sociedade favoráveis e contrários a ela. Em consequência, diversas abordagens sobre o tema vêm ocorrendo. Azevedo (2004) discute o ressurgimento da noção de raça entre acadêmicos, políticos e militantes anti-racismo e avalia até que ponto o estabelecimento de cotas raciais deveria se impor como única alternativa para aqueles que gostariam de ver o racismo abolido da sociedade brasileira. Pena e Bortolini (2004) abordam o potencial de contribuição da genética para a determinação de quem pode se beneficiar de ações afirmativas. Maggie e Fry (2004) analisam os caminhos que a mudança de paradigma racial introduzida pela política do estado do Rio de Janeiro tem de estabelecer cotas para negros nas universidades públicas do estado. Rosemberg (2004) considera os desafios do uso da classificação de cor ou raça nas ações afirmativas. Pereira (2003) aborda o pensamento social brasileiro, visando a elucidar o choque social provocado pelas propostas governamentais de cotas e ações afirmativas. Guimarães (2003) analisa o movimento por ações afirmativas no sistema de educação superior brasileiro e as resistências que ele tem despertado na sociedade civil. Peixoto e Braga (2004) analisam o processo seletivo da UFMG, apontando que as chances de aprovação dos candidatos que se declararam brancos quando da inscrição no vestibular, tanto em 2003 quanto em

2004, foram duas vezes maiores do que as dos que se declararam pretos, índios e amarelos e uma vez e meia superior às dos pardos. Esses autores sublinham, entretanto, que, no universo dos egressos da escola média privada, as diferenças entre brancos e negros são pouco expressivas. Gomes (2004) discute a proposta de cotas raciais como condição que permite ao estudante negro concorrer com candidatos que possuem trajetórias sociais e escolares semelhantes.

É quase um consenso na sociedade que o acesso ao ensino superior tende a melhorar as possibilidades de ascensão social do indivíduo, ampliando sua capacidade de receber mais renda e, por isso, deveria constar no planejamento estratégico de qualquer governo interessado no desenvolvimento da sociedade. Embora para muitos brasileiros as universidades federais sejam elitistas, Moraes (2004) mostra que o percentual de pobres nas universidades públicas desenha uma pirâmide que é similar à observada para as escolas particulares, e é menos aguda do que a observada para a totalidade da população, significando que existem evidências de que, se a educação superior não elimina a perversidade da concentração de renda, também não a aprofunda. Borges e Carnielli (2005) concluem que, embora o mérito individual seja relevante para o ingresso em cursos de maior prestígio social na UnB, o esforço individual não é o bastante para este fim, uma vez que os fatores socioeconômicos contribuem decisivamente para alcançar o mérito de ser aprovado (ver também Ribeiro (1988)). Segundo tais autores, indo em direção contrária à ideologia liberal de que a escola deveria dar oportunidades iguais a todos, a estratificação social é reproduzida na seleção dos candidatos à universidade para cursos socialmente prestigiados. Lage *et al.* (2006) estuda o perfil de candidatos ao Vestibular-2004 da UFMG pertencentes a grupos considerados homogêneos do ponto de vista socioeconômico. Estes autores concluem que, para grupos menos favorecidos do ponto de vista socioeconômico, a realização de

atividades que complementam a formação básica do candidato, como, por exemplo, freqüentar cursos pré-vestibular, escolas de idiomas e saber usar computador tendem a aumentar a chance de ingresso na universidade. Também concluem que a variável raça está associada com a aprovação apenas no grupo de candidatos menos favorecidos do ponto de vista socioeconômico e, ao contrário ao que é esperado, observam que, se comparados a candidatos da raça branca, candidatos das raças preta ou parda, neste grupo, têm mais chance de aprovação. Outra evidência importante encontrada por Lage *et al.* (2006) é que a condição socioeconômica, se determinante, não é a única característica associada com a maior chance de aprovação do candidato, uma vez que o percentual de candidatos aprovados vindos das escolas públicas federais (que, em geral, não são favorecidos do ponto de vista socioeconômico) é substancialmente maior do que o percentual de inscritos vindos destas escolas. Esse efeito positivo das escolas públicas federais em Minas Gerais não é observado para os demais tipos de escola, nem mesmo para as escolas particulares, e poderia ser explicado, entre outras coisas, pelo rigoroso processo de seleção utilizado por tais escolas.

No âmbito da UFMG, a democratização do acesso ao ensino superior é uma questão que vem sendo amplamente discutida. Parte deste debate pode ser encontrada em Peixoto (2004), coletânea que reúne artigos e opiniões de pesquisadores sobre a educação no Brasil e sobre mecanismos para a ampliação do acesso ao ensino superior, como a criação de cursos noturnos e o ensino à distância. Soares (2004) sugere dois caminhos para aumentar o número de egressos de escolas públicas na UFMG: o primeiro é a criação ou a transformação das escolas públicas de ensino médio em ambientes favoráveis ao desempenho acadêmico; e o outro é implementar uma política de cotas em que a universidade, a cada ano, defina um mínimo aceitável de desempenho para cada curso. A idéia é escolher candidatos vindos de escolas públicas que tenham mérito e

foram prejudicados pelas condições encontradas no ensino médio. Soares (2004) salienta que, se em 2000 a reserva de 50% das vagas para candidatos de escolas públicas estivesse em vigor, haveria uma queda considerável nas notas mínimas de ingressos na USP, podendo levar a situações preocupantes com a classificação de um candidato do curso de Medicina que tira zero em uma prova específica.

O primeiro caminho apontado por Soares (2004) motiva o estudo que se pretende realizar aqui. Buscando conhecer os candidatos ao Vestibular-2004 da UFMG, provenientes tanto de escolas públicas quanto de escolas particulares, traça-se o perfil de candidatos de ambos os grupos e identificam-se as características (aqui denominadas fatores), definidas no questionário socioeconômico e cultural aplicado no ato da inscrição do candidato, que podem estar mais associadas com suas aprovações. A meta definitiva é identificar quais características dos candidatos, provenientes tanto de escolas públicas quanto privadas, levam ou à ampliação ou à redução da chance de aprovação, se comparadas à chance de um candidato qualquer ter sido aprovado (que foi de 7% em 2004). Utiliza-se a CART (*Classification and Regression Trees*) para este fim. Com isso, pretende-se fornecer subsídios para uma discussão mais ampla sobre a democratização do acesso à universidade pública brasileira, assim como encontrar evidências empíricas que permitam a definição de políticas públicas mais adequadas para que as universidades, mais especificamente a UFMG, alcancem tal propósito. Embora haja evidência de que candidatos a cursos diferentes sejam heterogêneos, isso não será levado em conta neste estudo, uma vez que se pretende apenas traçar um perfil da totalidade dos candidatos ao Vestibular-2004 da UFMG.

Este trabalho está assim organizado. Na Seção 2, apresenta-se brevemente a metodologia estatística utilizada. Na Seção 3, descreve-se o banco de dados considerados no estudo. Na Seção 4, apresenta-se a análise dos resultados para os dois grupos considerados no estudo, a

saber, os candidatos que concluíram o ensino médio em escolas públicas ou privadas. A Seção 5 finaliza o artigo, apresentando algumas conclusões mais relevantes.

## **2. Modelo estatístico utilizado:**

### **CART (Classification and Regression Trees)**

O CART pode ser considerado como um modelo de regressão não-paramétrico que tem por objetivo estabelecer a relação entre um conjunto de variáveis explicativas (fatores) e uma única variável resposta, que, neste caso, é “O candidato foi aprovado?”. O modelo é ajustado mediante sucessivas divisões binárias no conjunto de dados, de modo a tornar os grupos resultantes cada vez mais homogêneos em relação à variável resposta, a qual pode ser contínua ou categorizada. O uso do CART não pressupõe que a suposição de normalidade seja verificada para os dados. Uma vantagem do CART é que as interações entre as variáveis explicativas são automaticamente captadas pelo modelo.

A construção de uma árvore, via CART, consiste em determinar, a partir do vetor de variáveis explicativas, aquelas que melhor dividem o conjunto de dados. Para esta escolha, todas as combinações possíveis entre as variáveis são testadas, sendo escolhidas para fazerem parte do modelo aquelas que mais reduzem a heterogeneidade dos subconjuntos criados e tenham maior grau de associação com a variável resposta. Essa divisão resulta em dois subconjuntos mais homogêneos do que o conjunto original. O ideal é que todos os elementos dos subconjuntos resultantes possuam o mesmo valor para a variável resposta, o que implicaria grupos (ou nós) completamente homogêneos. Os nós de uma árvore estão associados aos subconjuntos resultantes da aplicação de uma regra de divisão a determinado conjunto de dados. Os nós são, normalmente, representados por retângulos. O primeiro nó de uma árvore é chamado de nó raiz e equivale a um conjunto de dados completo.

Os nós gerados pela divisão de um nó já existente recebem o nome de nó filho e o nó que os originou é chamado de nó pai. Um conjunto de observações, com o mesmo grau de parentesco do nó pai, recebe o nome de geração.

Quando a variável resposta é uma variável dicotômica, que é o caso estudado neste trabalho, para se medir a homogeneidade ou, equivalentemente, a impureza de determinado nó, é utilizado o índice *Gini*, denotado aqui por  $i(t)$ , que é definido por:  $i(t) = 2p(1/t)(1-p(1/t))$ , em que  $t$  é um nó arbitrário e  $p(1/t)$  é a probabilidade de a variável resposta ser classificada na categoria 1 para o nó  $t$ . Calculado o índice *Gini*, escolhe-se, para fazer parte do modelo, aquelas variáveis com maior grau de associação, isto é, maior valor para  $i(t)$ . Na prática, as probabilidades são estimadas pela proporção de indivíduos de determinada categoria no nó considerado. O processo de divisão é repetido recursivamente até que alguma das regras de parada seja alcançada. Essas regras de parada são definidas pelo usuário e, em geral, consistem em uma limitação no tamanho da impureza do nó, do número de gerações e do número de nós da árvore.

Para variáveis resposta dicotômicas, além das variáveis explicativas associadas à variável resposta, também se pode obter a probabilidade condicional de a variável resposta resultar em sucesso. Conseqüentemente, para o caso em estudo, uma análise via CART fornece a probabilidade de aprovação dos candidatos que possuem as características apresentadas em cada grupo resultante da divisão, além das variáveis que melhor explicam a aprovação do candidato. Para maiores detalhes, ver, por exemplo, Bell (1996) e Diniz e Louzada-Neto (2000).

Uma vez construído o modelo estatístico, faz-se necessário avaliar a sua adequação ou sua capacidade preditiva. Com esta finalidade, usamos, neste trabalho, a curva ROC (*Receiver Operating Characteristic*), que é uma curva contínua obtida a partir das probabilidades de aprovação de cada candidato fornecidas pelo modelo da seguinte forma: para todo ponto de

corde no intervalo (0,1), determinam-se a sensibilidade,  $Se$ , e a especificidade,  $Es$ , do modelo, que, no caso apresentado aqui, são dadas, respectivamente, por:

$$Se = P(\text{classificar como aprovado} / \text{realmente foi aprovado})$$

$$Es = P(\text{classificar como não aprovado} / \text{realmente não foi aprovado})$$

A curva ROC é formada pelos pontos obtidos para cada ponto de corte. Um modelo ideal é aquele cuja área sob a curva ROC seja igual a 1. Um modelo é considerado como tendo muito boa capacidade de classificação se a área sob a curva ROC é superior a 0,8. Para maiores detalhes sobre a Curva ROC, ver Hanley and McNeil (1982), Zweig and Campbell (1993) e Martinez *et al.* (2003).

### 3. Conhecendo os dados

Os dados utilizados neste trabalho são oriundos do questionário socioeconômico aplicado pela UFMG no Vestibular de 2004 e foram cedidos pela COPEVE-UFMG. Considera-se como variável resposta “O candidato foi aprovado?” e as variáveis preditoras são as contidas no questionário socioeconômico e cultural do vestibular da UFMG de 2004, a saber:

#### **Variáveis socioeconômicas:**

- “Trabalha em atividade remunerada”;
- “Por quantos anos teve atividade remunerada”;
- “Renda familiar”;
- “Participação na vida econômica do grupo familiar”;
- “Número de pessoas que vivem da renda mensal”;
- “Classificação social ABIPEME”;
- “Classificação social Fator Socioeconômico (FSE)”.

#### **Variáveis relativas à formação do candidato no ensino médio:**

- “Ano de conclusão do ensino médio”;
- “Onde conclui o ensino médio”;
- “Tipo de ensino médio”;
- “Tipo de escola de ensino médio”;
- “Curso de ensino médio”;

“Quantos anos levou para concluir o ensino médio”;  
“Turno do ensino médio”.

**Variáveis relativas ao histórico do candidato em vestibulares:**

“Há quanto tempo tenta ingressar em curso superior”;  
“Já prestou algum vestibular”;  
“Quantos vestibulares já prestou”;  
“Freqüentou cursinho pré-vestibular”;  
“Motivo pela escolha do curso”.

**Variáveis relativas ao conhecimento de língua estrangeira:**

“Freqüentou cursinho de idiomas”;  
“Situação com respeito à língua estrangeira”;  
“Opção de língua estrangeira”.

**Variáveis relativas aos pais:**

“Escolaridade do pai”;  
“Escolaridade da mãe”;  
“Situação familiar”;  
“Ocupação do pai”;  
“Ocupação da mãe”.

**Variáveis relativas às características pessoais do candidato:**

“Sexo”;  
“Estado civil”;  
“É canhoto”;  
“Onde nasceu”;  
“Raça”;  
“Onde reside”;  
“Sabe usar o computador”;  
“Tipo de residência”.

Dois índices são utilizados para medir as características socioeconômicas do candidato: o FSE e o critério ABIPEME, desenvolvido pela Associação Brasileira de Institutos de Pesquisa de Mercado ([www.ufrn.br/sites/ fonaprace/perlanexo3.doc](http://www.ufrn.br/sites/ fonaprace/perlanexo3.doc)). O FSE é um índice que combina aspectos da trajetória escolar do candidato com o padrão de renda familiar, profissão e escolaridade dos pais. Ver detalhes em Braga *et al.* (2001) e em [www.ufmg.br/inclusaosocial/ cursos\\_noturnos.doc](http://www.ufmg.br/inclusaosocial/ cursos_noturnos.doc).

Ressalta-se que “Situação com respeito à língua estrangeira” denota o conhecimento do candidato em línguas e “Opção de língua estrangeira” é a escolha de idioma feita pelo candidato por meio da qual ele quer ser avaliado no Vestibular.

A variável “Raça” foi incorporada no questionário socioeconômico da UFMG apenas em 2003 e é informada pelo candidato, portanto, passível de ter sido medida com erro, uma vez que não é uma variável de fácil mensuração. Estudos descritivos preliminares considerando candidatos que fizeram mais de um vestibular mostram que, mesmo após o início da discussão sobre a criação de cotas raciais para as universidades, o percentual de migração entre as diversas opções de raça é bastante similar, indicando uma possível dificuldade do candidato em classificar-se em uma das categorias de Raça. Embora haja evidências biológicas de que os termos “Raça Branca”, “Raça Preta”, etc. não sejam os mais adequados para se referir a coletivos de seres humanos, neste trabalho, define-se esta variável seguindo os padrões utilizados em estudos realizados pelo IBGE (Instituto Brasileiro de Geografia e Estatística). Ver URL: <http://www.ibge.gov.br/>.

Em 2004, dos 60.616 candidatos ao vestibular da UFMG considerados para a análise, 53,3% haviam concluído o ensino médio em escolas públicas (sendo 40,5% em escolas públicas estaduais, 8,2% em escolas públicas municipais e 4,6% em escolas públicas federais) e 46,2% eram provenientes de escolas particulares. As Tabelas 1 e 2 apresentam o percentual de inscritos e aprovados em cada uma das redes de ensino médio, segundo o FSE e a Raça.

**TABELA 1**

Percentual de inscritos e aprovados no Vestibular 2004 UFMG, segundo o Fator socioeconômico (FSE) para cada tipo de escola do ensino médio

FSE	Pública						Particular			
	Federal		Estadual		Municipal		Particular		Curso livre	
	Insc.	Apr.	Insc.	Apr.	Insc.	Apr.	Insc.	Apr.	Insc.	Apr.
0	2,5	0,6	3,3	2,8	7,1	9,9	0,0	0,0	0,0	0,0
1	12,7	10,6	15,5	7,8	20,6	15,9	0,8	0,1	0,9	0,0
2	18,3	15,8	24,3	18,6	23,1	17,2	2,0	0,6	35,3	17,0
3	19,8	18,3	29,5	31,2	23,8	23,3	4,0	2,7	29,3	17,0
4	14,1	14,7	13,4	19,1	12,0	9,0	10,2	6,7	13,0	17,0
5	10,7	12,9	7,2	8,3	6,4	9,5	13,6	9,5	8,4	33,0
6	9,4	12,2	4,0	7,3	4,1	8,2	14,8	11,5	6,5	17,0
7	7,2	10,0	1,9	3,3	2,0	4,7	15,6	14,6	3,3	0,0
8	4,3	3,9	0,6	1,4	0,8	1,3	16,8	19,1	2,3	0,0
9	1,1	1,0	0,1	0,1	0,1	0,9	13,7	20,7	0,5	0,0
10	0,0	0,0	0,0	0,0	0,0	0,0	8,6	14,5	0,5	0,0

Observou-se (TABELA 1) que os candidatos inscritos e aprovados provenientes de escolas particulares tendem a concentrar-se em uma classificação social, determinada segundo o fator socioeconômico (FSE), superior a 4 (apenas 6,8% dos candidatos inscritos e 3,4% dos aprovados possuem FSE inferior a este valor). Situação oposta à observada para os candidatos de escolas públicas, que possuem, em geral, FSE menor ou igual a 6. Ao analisar apenas os candidatos provenientes de escola pública, observa-se que a distribuição dos candidatos das escolas públicas estaduais e municipais são similares, sendo que apenas 2,6% dos candidatos inscritos (4,8% dos aprovados) provenientes de escolas públicas estaduais e 2,9% dos candidatos inscritos (6,9% dos aprovados) provenientes de escolas públicas municipais possuem FSE superior a 6. Embora os candidatos das escolas públicas federais tendam a ter FSE inferiores aos dos candidatos das escolas particulares (67,4% dos inscritos e 60% dos aprovados possuem FSE menor ou igual a 4), nota-se uma concentração um pouco maior destes candidatos em níveis mais altos de FSE (cerca de 12,6% dos candidatos

inscritos e 14,9% dos aprovados possuem FSE superior a 6) do que o observado para os candidatos das escolas públicas estaduais e municipais. Os candidatos provenientes de cursos livres apresentam comportamento referente ao FSE similar ao observado para os candidatos provenientes das escolas públicas estaduais e municipais.

É importante ressaltar que, embora o FSE dos candidatos aprovados seja levemente superior ao dos inscritos para qualquer tipo de escola, pode-se observar que, em geral, os candidatos inscritos e aprovados provenientes de escolas públicas concentram-se na faixa de FSE entre 1 e 6, enquanto os candidatos inscritos e aprovados provenientes de escola particular concentram-se na faixa de FSE entre 5 e 10. Logo, a faixa de concentração de candidatos inscritos e aprovados, segundo o FSE, é a mesma para cada tipo de escola. Ou seja, num olhar puramente descritivo, parece que a definição de cotas para escolas públicas poderia cumprir o papel de aumentar a inclusão social.

**TABELA 2**

**Percentual de inscritos e aprovados no Vestibular 2004 UFMG, segundo a raça para cada tipo de escola do ensino médio**

Raça	Pública						Particular		Curso livre	
	Federal		Estadual		Municipal		Insc.	Apr.	Insc.	Apr.
Branca	59,1	65,5	49,7	53,3	49,4	53,3	71,6	73,8	45,2	57,1
Preta	6,7	4,7	9,6	6,2	11,2	9,2	2,9	2,0	10,9	0,0
Parda	25,8	22,4	31,7	32,3	29,5	27,5	18,6	17,0	32,2	0,0
Amarela	2,2	1,4	3,6	2,5	3,1	2,5	2,0	1,1	1,7	0,0
Indígena	0,5	0,0	0,7	0,6	1,0	1,3	0,4	0,2	0,8	14,3
Não deseja declarar	5,7	6,1	4,7	5,1	5,9	6,3	4,6	5,8	9,2	28,6

Com respeito à raça declarada pelo candidato, nota-se na Tabela 2 que, entre os aprovados, a maioria é da raça branca em todos os tipos de escola de ensino médio. Entre os inscritos, os candidatos da raça branca

são a maioria para as escolas públicas federais e particulares e, para os demais tipos de escolas, a proporção de candidatos da raça preta somada à proporção de candidatos da raça parda é próxima da proporção de candidatos da raça branca.

Deve-se ressaltar ainda que os candidatos da raça branca tendem a ter uma condição socioeconômica mais favorável – apenas 43,45% dos candidatos da raça branca inscritos no vestibular têm FSE inferior ou igual a 4, enquanto que 81,23% e 66,55% dos candidatos das raça preta e parda, respectivamente, possuem FSE nesta faixa (resultados não-apresentados).

#### **4. Resultados obtidos**

Nesta seção, apresenta-se a análise para os candidatos que fizeram o ensino médio em escolas particulares e para candidatos provenientes de escolas públicas. Neste último caso, para completar o estudo, apresenta-se também uma análise descritiva para os nós terminais onde foram observadas a mais alta e a mais baixa probabilidades de aprovação, com o intuito de identificar, entre os candidatos inscritos de cada rede de ensino (estadual, federal e municipal), o percentual observado em cada nó. Este tipo de medida pode fornecer alguma evidência sobre a qualidade de ensino praticada em cada uma dessas redes.

Usando a Curva ROC, conclui-se que os modelos ajustados pelo CART para candidatos de escolas particulares e públicas (ver Figuras 2 e 3, respectivamente) podem ser considerados adequados para descrever o comportamento dos dados, pois a área abaixo das curvas é significativamente maior que 0,5.

Para a análise dos resultados fornecidos pelas árvores de classificação do CART, deve-se proceder da seguinte forma: parte-se de cada nó terminal (o qual oferece a probabilidade condicional de aprovação do candidato) e identificam-se as características do candidato,

olhando-se todos os nós precedentes no ramo ao qual o nó terminal pertence, até chegar ao nó raiz. Por exemplo, na Figura 2, o nó 7 é um nó terminal. Assim, o valor 0,0567 (ou 5,67%) é a probabilidade condicional de um candidato ter sido aprovado, uma vez que freqüentou curso de inglês, concluiu o ensino médio ou na Grande BH, no interior de Minas Gerais ou em outro estado e não mora em Belo Horizonte.

#### **4.1. Modelo para candidatos de escolas particulares**

Para candidatos que cursaram o ensino médio em escolas particulares, a variável mais fortemente associada com a aprovação do candidato no vestibular da UFMG de 2004 é o local de moradia (ver modelo ajustado na Figura 1), sendo que um candidato de Belo Horizonte tende a ter maior chance de aprovação. Ressalta-se, no entanto, que a chance de aprovação de um candidato depende de outros fatores além deste. A situação com respeito à língua estrangeira também afeta a chance de aprovação. Por exemplo, um candidato que mora em Belo Horizonte e que não lê o inglês, ou duas ou mais línguas estrangeiras, tende a ter baixa chance de aprovação no vestibular.



probabilidades de aprovação. Observou-se que, da totalidade de candidatos originários de escolas particulares, 29% pertenciam aos três nós com mais baixas chances de aprovação (sendo que 23% pertenciam ao nó 08, que é o nó cuja chance de aprovação é a terceira mais baixa) e somente 6,6% pertenciam aos três grupos com mais altas chances de aprovação. Percebeu-se ainda que tanto os nós com baixa chance de aprovação, quanto aqueles com chances mais altas, eram constituídos por uma maioria de candidatos da raça branca (68% ou mais). Notou-se que 100% dos candidatos dos nós 38 e 43 (probabilidade alta de aprovação) e do nó 42 possuíam FSE maior ou igual a 9. Ressalta-se que apenas uma fatia ínfima (0,22%) dos candidatos originários de escolas particulares pertencem ao nó 42. Para o nó 36, observou-se que 100% dos candidatos têm FSE entre 3 e 8 e, para os nós 13 e 08, notou-se baixa proporção de candidatos com FSE inferior a 4 (10,8% para o nó 13 e 6,7% para o nó 08). Assim, nota-se que, mesmo para nós com baixa chance de aprovação, os candidatos de escola particular tendem a se concentrar em níveis mais elevados de FSE.

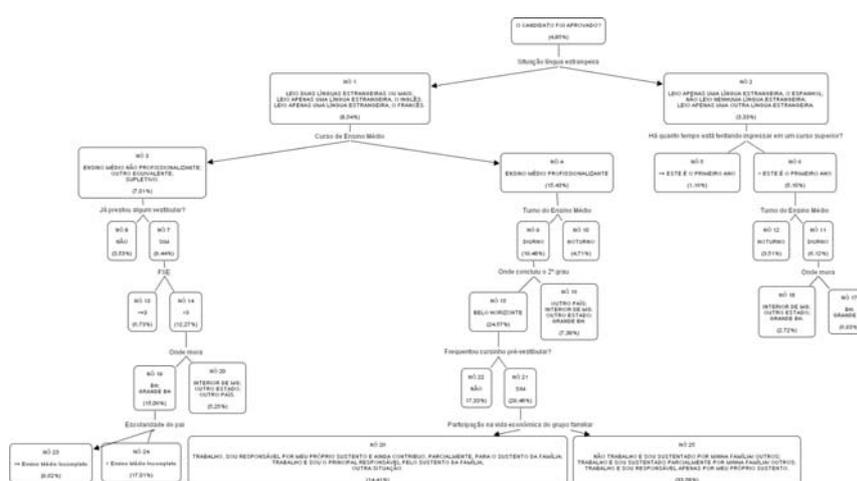
#### **4.2. Modelo para candidatos de escolas públicas**

Contrário ao que ocorreu para os candidatos provenientes de escolas particulares, para aqueles que concluíram o ensino médio em escolas públicas (Figura 2), a variável mais fortemente associada com a aprovação é a “Situação com respeito à língua estrangeira”. Nota-se que o candidato que lê apenas o inglês, ou apenas o francês ou duas ou mais línguas estrangeiras, tende a ter chance mais alta de aprovação no vestibular. Um candidato fora deste grupo poderia ter sua chance de aprovação aumentada se, por exemplo, morasse em Belo Horizonte ou na Grande BH, tivesse feito o ensino médio em curso diurno e estivesse tentando ingressar na universidade pela primeira vez. Percebe-se que, para candidatos vindos de escolas públicas, apesar de o local de moradia não

ser a variável mais fortemente associada, ela também está correlacionada com a chance de aprovação.

Figura 2

Árvore de Decisão para candidatas de escola pública



Analisando aqui também os nós terminais com a mais baixa (nó 5) e a mais alta (nó 25) chances de aprovação, observamos que:

- Uma vez que o candidato não lia nenhuma língua estrangeira ou somente o espanhol, ou outra língua estrangeira que não o inglês nem o francês, e estava tentando ingressar em um curso superior há menos de um ano (ver nó 5), sua chance de aprovação no Vestibular 2004 era de 1,16%.
- Se o candidato que lia o inglês ou o francês, ou duas ou mais línguas estrangeiras, havia cursado o ensino médio profissionalizante em turno diurno em Belo Horizonte, havia freqüentado cursinho pré-vestibular e não trabalhava ou, se trabalhava, não era responsável pelo sustento da família (ver nó 25), sua chance de aprovação era de 33,58%.

Um outro nó que não é terminal, mas que apresenta, em seus segmentos, alta chance de aprovação, é o nó 4, que compreende os candidatos de escolas públicas que, além de falarem inglês, ou francês, ou mais de 2 línguas estrangeiras, eram também provenientes do ensino médio profissionalizante. Exceto para os candidatos que estudaram no período noturno, todos os nós decorrentes desta classe possuem chances de aprovação maiores do que as de um candidato qualquer da escola pública (que era de 4,85%). Ressalta-se que isso não foi observado para escolas particulares (nós 19, 20, 23 e 24, Figura 1). Isso pode ser um indício de que o ensino médio profissionalizante, ao menos o ensino praticado nas escolas públicas federais de Belo Horizonte (COLTEC e CEFET, essencialmente, que é de onde vêm os candidatos provenientes desta rede ao Vestibular da UFMG), pode estar fazendo a diferença na formação do candidato, já que as chances de aprovação aumentam ainda mais neste caso. Como já mencionado, o diferencial apresentado pelos cursos públicos profissionalizantes poderia ser explicado para o vestibular da UFMG, entre outras coisas, pelo rigoroso processo de seleção utilizado por tais escolas.

Far-se-á agora uma análise mais detalhada da frequência de candidatos oriundos de cada tipo de escola pública de ensino médio nos nós terminais com mais altas (nós 22, 24 e 25) e mais baixas (nós 5, 12 e 18) probabilidades de aprovação. Ressalta-se que os percentuais apresentados na Tabela 3 foram calculados com base no total de candidatos inscritos vindos de cada tipo de escola pública. Por exemplo, no nó 12, tem-se que, do total de candidatos inscritos no vestibular oriundos de escolas públicas federais, apenas 5% pertencem a este nó, enquanto que este percentual para candidatos provenientes de escolas municipais e estaduais é bem mais alto, 19,7% e 15,3%, respectivamente. Observa-se da Tabela 3 que os três grupos que apresentam menor probabilidade de aprovação (nós 5, 12 e 18) correspondem a 50,8% dos

alunos originários das escolas públicas inscritos no Vestibular 2004 da UFMG, sendo que 31,2% concentram-se no nó com mais baixa chance de aprovação. O fato que deve ser ressaltado aqui é que, quando comparado com candidatos de escolas estaduais e municipais, exceto para o nó 18, que é o que tem a menor quantidade de alunos, percebe-se presença bem menor de candidatos provenientes da escola pública federal, proporcionalmente. Essas proporções chegam a ser cerca de duas vezes menores (para o nó 5) e três vezes menores (nó 12) para a escola pública federal.

Já os três grupos de candidatos cujas chances de aprovação são as maiores (nós 22, 24 e 25) contêm somente 5,8% do total de candidatos vindos das escolas públicas (Tabela 4) inscritos no vestibular de 2004. O inverso do que ocorreu para os grupos com baixa chance de aprovação é observado aqui, pois é notável a maior proporção de candidatos vindos da escola pública federal em todos os três nós, principalmente para os nós 22 e 25, nos quais se percebe a presença de 10,8% e 11,4%, respectivamente, do total de candidatos provenientes de escolas públicas federais.

**Tabela 3**

**Porcentagem de candidatos inscritos oriundos de cada tipo escola pública nos três grupos com menor chance de aprovação**

Nó	Chance de aprovação	Tipo de escola pública	Freqüência	Percentual
5	1,16%	Federal	441	15,8
		Estadual	8290	33,8
		Municipal	1363	27,4
		<b>Total</b>	<b>10094</b>	<b>31,2</b>
18	2,72%	Federal	171	6,1
		Estadual	1241	5,0
		Municipal	51	1,0
		<b>Total</b>	<b>1463</b>	<b>4,5</b>
12	3,51%	Federal	140	5,0
		Estadual	3763	15,3
		Municipal	978	19,7
		<b>Total</b>	<b>4881</b>	<b>15,1</b>

Tabela 4

Porcentagem de candidatos inscritos oriundos de cada tipo escola pública nos três grupos com menor chance de aprovação

Nó	Chance de aprovação	Tipo de escola pública	Freqüência	Percentual
92	17,01%	Federal	179	6,4
		Estadual	683	2,8
		Municipal	235	4,7
		<b>Total</b>	<b>1097</b>	<b>3,4</b>
22	17,33%	Federal	301	10,8
		Estadual	34	0,1
		Municipal	16	0,3
		<b>Total</b>	<b>351</b>	<b>1,1</b>
25	33,58%	Federal	319	11,4
		Estadual	63	0,2
		Municipal	25	0,5
		<b>Total</b>	<b>407</b>	<b>1,3</b>

Estes resultados evidenciam a diferença no perfil dos candidatos provenientes de escola pública federal em relação àqueles que fizeram o ensino médio em escola pública estadual ou municipal. Parece que o candidato proveniente de escola pública federal tende a ter maior chance de aprovação no vestibular da UFMG.

Os candidatos que constituem os nós 22, 24 e 25 (alta chance de aprovação) são, em sua maioria, da raça branca. Já para os nós 5, 12 e 18 (baixa chance de aprovação), os candidatos se distribuem de forma praticamente igualitária entre candidatos das raças preta ou parda e branca. Para todos os nós (exceto para o nó 24, em que 100% dos candidatos têm FSE superior a 3 e inferior a 10), observa-se que 84% ou mais dos candidatos têm FSE menor ou igual a 5. Isso evidencia, mais uma vez, que o FSE influencia apenas indiretamente na aprovação do candidato, uma vez que pode ser o fator determinante na escolha do tipo de escola em que cursar o ensino médio.

### 3. Discussão e conclusões

Neste trabalho, traçou-se o perfil dos candidatos ao Vestibular 2004 da UFMG, provenientes tanto de escolas públicas quanto de escolas particulares, e identificou-se características definidas no questionário socioeconômico e cultural aplicado no ato da inscrição do candidato que estão mais associadas com as aprovações. Como consequência, foram determinados grupos homogêneos segundo todas essas características (algumas das quais são referentes ao coletivo ao qual o aluno pertence) e também as probabilidades de aprovação em cada grupo.

Notou-se que a variável mais fortemente associada com a aprovação de um candidato proveniente de escola particular é o local de moradia e, de um candidato vindo de escola pública, é situação com respeito à língua estrangeira. Nota-se que um candidato da escola particular de Belo Horizonte tende a ter maior chance de aprovação. No entanto, isso não é regra geral, pois a aprovação depende de outras características dos candidatos. Para os provenientes de escolas públicas, nota-se que tendem a ter maiores chances de aprovação os que lêem inglês, francês ou duas ou mais línguas estrangeiras. Um candidato, seja de escola pública, seja de escola particular, que não é de Belo Horizonte e que não lê inglês, ou duas ou mais línguas estrangeiras, tende a ter chance mais baixa de aprovação no vestibular. A importância do conhecimento de língua estrangeira para o bom desempenho no vestibular já havia sido apontada previamente por Borges e Carnielli (2005) e Peixoto e Braga (2004), por exemplo.

O estudo de Borges e Carnielli (2005) retrata o perfil socioeconômico das famílias dos candidatos da UnB e sua influência na aprovação dos candidatos, medidas pelo nível de escolaridade dos pais, jornada de trabalho do ingressante durante o ensino médio, participação do ingressante na vida familiar e renda familiar. Variáveis similares foram consideradas neste estudo e percebeu-se que a escolaridade do pai e a

participação do candidato na vida econômica do grupo familiar mostraram-se significativamente associadas com a aprovação de alguns grupos de candidatos que cursaram o ensino médio em escolas públicas. A escolaridade da mãe mostrou-se associada apenas com a aprovação de alguns grupos de candidatos de escolas particulares. No entanto, tal associação mostra-se menos forte do que a associação observada entre variáveis relativas à formação básica e variáveis socioeconômicas (Renda e FSE, por exemplo). A raça, apontada por Lage *et al.* (2006) como uma variável associada com a aprovação de candidatos menos favorecidos economicamente, não aparece associada com a aprovação neste estudo. Ou seja, no geral, as evidências sugerem que os aspectos que mais fortemente definem a chance de aprovação do candidato são a condição socioeconômica e a formação básica, e não a raça, como também concluído por Borges e Carnielli (2005).

Para os candidatos que concluíram o ensino médio em escolas públicas, nota-se ainda que, para grupos com menor e maior chances de aprovação, há tendência de concentração de candidatos de escolas públicas federais nos nós com alta chance de aprovação e poucos candidatos provenientes de escola pública federal nos nós com pequena chance de aprovação. Deve-se ressaltar que, em geral, os candidatos de escolas públicas federais de Belo Horizonte (principalmente COLTEC e CEFET) passam por rigoroso processo de seleção e, uma vez que do ponto de vista socioeconômico os candidatos oriundos de escolas públicas federais são tão pouco favorecidos quanto os candidatos de redes públicas estaduais e municipais, há evidência de uma diferença no perfil dos candidatos dessas escolas públicas. Além disso, comparando-se o perfil dos candidatos de escolas públicas e particulares, percebeu-se que a maioria dos candidatos das escolas públicas concentrava-se nos três nós com mais baixa probabilidade de aprovação (chance de aprovação menor ou igual a 3,51%), sendo que, destes, a maioria teve chance de 1,16% de

ser aprovado. Em contrapartida, apenas 29% dos candidatos de escolas particulares se concentraram nos três nós com mais baixa chance de aprovação (chance de aprovação entre 3,16% e 3,33%). Observou-se também que a concentração de candidatos, tanto de escolas públicas quanto de particulares, nos três nós com chance mais alta de aprovação é cerca de 6%. Ou seja, esses fatos indicam haver diferença entre o ensino praticado pelas escolas públicas e privadas.

Mais do que fornecer respostas, neste estudo visou-se a fornecer subsídios para uma discussão mais ampla sobre cotas e novas políticas públicas para a universidade. Assim, a constatação de que o conhecimento de língua estrangeira está fortemente associado com a aprovação dos candidatos nos dois grupos pode levar a discussões sobre o nível socioeconômico dos alunos e sobre a qualidade do ensino nas escolas de ensino médio, já que esta é uma variável fortemente associada com estes dois fatores. Outro ponto que vale a pena ser discutido diz respeito ao ensino profissionalizante. Este estudo mostrou que o ensino profissionalizante em escolas particulares não se mostrou tão eficiente quanto o de escolas públicas, que, no caso do vestibular da UFMG, é marcadamente formado por escolas públicas federais. Os resultados encontrados para as escolas públicas federais, que, em relação ao nível socioeconômico, são similares às escolas públicas estaduais e municipais, mas têm nível de aprovação no vestibular próximo ou superior ao das escolas particulares, leva a crer que, mantidos os moldes atuais para ingresso no Ensino Superior, poderia haver maior equidade do acesso ao ensino superior se houvesse melhora na qualidade do ensino ofertado pelas escolas públicas, primordialmente as estaduais e municipais.

**Agradecimentos:** Agradecemos aos professores Antônio Zumpano P. Santos e Maria do Carmo L. Peixoto, da UFMG, e a três revisores anônimos por seus comentários e sugestões, os quais contribuíram para melhorar este artigo substancialmente. Esta pesquisa foi parcialmente financiada pelo CNPq (projetos nos. 300325/2003-7, 472066/2004-8) e pela COPEVE-UFMG.

## REFERÊNCIAS

- AZEVEDO, C. M.M Cota racial e estado: abolição do racismo ou direitos de raça? *Cadernos de Pesquisa*, v. 34, n. 121, janeiro/abril de 2004.
- BRAGA, M.M.; PEIXOTO, M.C.L.; BOGUTCHI, T.F. Tendências da demanda pelo ensino superior: estudo de caso da UFMG. *Cadernos de Pesquisa*, n. 113, p. 129-152, julho 2001.
- BELL, J.F. Application of Classification Trees to the Habit Preference of Upland Birds. *Journal of Applied Statistics*, v. 23, n. 2 e 3, p. 349-359, 1996.
- BORGES, J. L.G.; CARNIELLI, B. L. Educação e estratificação social no acesso à universidade pública. *Cadernos de Pesquisa*, v. 35, n. 124, janeiro/abril de 2005.
- DINIZ, C.A.R.; LOUZADA-NETO, F. *Data Mining*: Uma introdução. Associação Brasileira de Estatística (ABE). 14º Simpósio Nacional de Probabilidade e Estatística, 2000.
- GOMES, N.L. Cotas para a população negra e a democratização da universidade pública. In: PEIXOTO, M.C.L. (Org.). *Universidade e Democracia*: Experiências e alternativas para a ampliação do acesso à Universidade pública brasileira 45-56, 2004.
- GUIMARÃES, A. S. Acesso de negros às universidades públicas. *Cadernos de Pesquisa*, n. 118, p. 247-268, 2003.
- HANLEY, J. A.; MCNEIL, B. J. The Meaning and Use of the Area Under a Receiver Operating Characteristic (ROC) Curve. *Radiology*, n. 143, p. 1, 29-36, 1982.

- LAGE, L.V. et al. Fatores que influenciaram na aprovação de candidatos de diferentes grupos socioeconômicos no Vestibular 2004 da UFMG. *Revista Brasileira de Estatística*, v. 67, n. 226, 2006. (No prelo)
- MARTINEZ, E. Z.; LOUZADA-NETO, F.; PEREIRA, B.B. A curva ROC para testes diagnósticos. *Caderno de Saúde Coletiva*, v. 11, n. 1, p. 7-31, 2003.
- MAGGIE, Y.; FRY, P. A reserva de vagas para negros nas universidades brasileiras. *Estudos Avançados*, v. 18, n. 50, p. 67-80, 2004.
- MORAES, R. C. Ensino superior no Brasil - Balanços e perspectivas a partir de 2003. *Educação e Linguagem*, v. 7, n. 10, p. 68-104, 2004.
- PEIXOTO, M.C.L.; BRAGA, M.M. Demanda pelo ensino superior no Brasil: o caso da UFMG. *Educação e Linguagem*, v. 7, n. 10, p. 124-149, 2004.
- PEIXOTO, M.C.L. (Org.) *Universidade e Democracia: Experiências e alternativas para a ampliação do acesso à Universidade pública brasileira*. Belo Horizonte: Editora UFMG, 2004.
- PENA, S.; BORTOLINI, M. C. Pode a genética definir quem pode se beneficiar das cotas universitárias e demais ações afirmativas? *Estudos Avançados*, v. 18, n. 50, p. 31-50, 2004.
- PEREIRA, A. M. Um raio em céu azul: reflexões sobre a política de cotas e a identidade nacional brasileira. *Estudos Afro-Asiáticos*, v. 25, n. 3, p. 463-482, 2003.
- RIBEIRO, S. C. O Vestibular de 1988: Seleção ou exclusão. *Educação e Seleção*, n. 18, p. 93-109, julho/dezembro 1988.
- ROSEMBERG, F. O branco no IBGE continua branco na ação afirmativa? *Estudos Avançados*, v. 18, n. 50, p. 61-66, 2004.
- SANTOS, B. S. *Pela mão de Alice, o social e o político na pós-modernidade*. São Paulo: Cortez, 1996.
- SOARES, J. F. Implementação de cotas na UFMG para alunos egressos de escolas públicas. In: PEIXOTO, M.C.L. (Org.). *Universidade e Democracia: Experiências e alternativas para a ampliação do acesso à*

Universidade pública brasileira. p. 153-172, 2004.

ZWEIG, M. H.; CAMPBELL, G. Receiver Operating Characteristic (ROC) Plots: A Fundamental Evaluation Tool in Clinical Medicine. *Clinical Chemistry*, n. 39, p. 4, 561-577, 1993.

**Recebimento:** 28/12/2005

**Aprovação:** 24/01/2006

**Contato:**

Rosangela H. Loschi

Av. Antonio Carlos 6627.

Dep. de Estatística ICEX/UFMG

Cidade Universitária - Pampulha

Belo Horizonte/MG

Cep: 31270-901

E-mail: loschi@est.ufmg.br

Mauro Mendes Braga

Av. Antonio Carlos 6627.

Pró-Reitoria de Graduação/UFMG

Cidade Universitária - Pampulha

Belo Horizonte/MG

Cep: 31270-901

E-mail: braga@iufmg.br