

ESTIMATIVAS DE PROBABILIDADE DE APROVAÇÃO NO VESTIBULAR PARA ALGUNS CURSOS DA UFU EM FUNÇÃO DO PERFIL SOCIO-ECONÔMICO, UTILIZANDO REGRESSÃO LOGÍSTICA

APARECIDA SILVA SOUZA¹, EDNALDO CARVALHO GUIMARÃES²

RESUMO

O objetivo deste trabalho foi analisar por meio da regressão logística os dados referentes aos alunos aprovados e não aprovados na Universidade Federal de Uberlândia (UFU), nos cursos de Ciências Contábeis noturno, Engenharia Elétrica, Matemática, Medicina e Psicologia, em concursos vestibulares realizados em julho e dezembro de 2009, bem como verificar quais variáveis independentes influenciaram na opção e na aprovação dos candidatos, e também estimar a probabilidade de aprovação do candidato em função do perfil socioeconômico. Os dados relativos ao desempenho e às condições socioeconômicas foram fornecidos pela Diretoria de Processo Seletivo (DIRPS/UFU). Foram examinadas as 42 variáveis socioeconômicas do questionário do manual do candidato respondido no ato da inscrição para o concurso vestibular. O procedimento estatístico realizado foi a análise de regressão logística por Stepwise, bem como a análise do ODDS ($\exp(B)$) adotando como critério de inclusão da variável no modelo a significância de 5%. Os resultados indicaram que 13 (30%) das 42 variáveis socioeconômicas influenciam de algum modo na aprovação ou não do concorrente no vestibular da UFU. Os resultados sugerem ainda que os candidatos com nível socioeconômico menos favorável façam opção pelos cursos menos concorridos, conforme foi indicado na equação linearizada geral obtida nessa pesquisa. Conclui-se que, de modo geral, algumas variáveis do nível socioeconômico podem influenciar a aprovação ou não no vestibular da UFU para os cursos estudados.

Palavras-chave: Regressão logística; Desempenho de aprovados e não aprovados; Perfil socioeconômico.

¹ Acadêmica do curso de Psicologia - INPSI/UFU – Bolsista PIBIC/FAPEMIG (março/10 a fevereiro/11) - Av. João Naves de Ávila, 2160, Bairro Santa Mônica, Uberlândia – MG, CEP: 38400-900 – cidasilvas07@yahoo.com.br

² Prof. Orientador – FAMAT/UFU – Av. João Naves de Ávila, 2160, Bairro Santa Mônica, Uberlândia – MG, CEP: 38400-900 – ecg@ufu.br

ABSTRACT

**ESTIMATES OF PROBABILITY OF PASSING THE ENTRANCE EXAM FOR
SOME COURSES OF THE UFU DEPENDING ON SOCIO-ECONOMIC
PROFILE, USING LOGISTIC REGRESSION**

The goal of this work was to look through logistic regression data for students approved and not approved at the Federal University of Uberlândia, in Accounting Sciences courses nightly, electrical engineering, mathematics, medicine and psychology, vestibular competitions held in July and December 2009, as well as see which independent variables influence on the selection and approval of the candidates, and also to estimate the likelihood of approval of the candidate on the basis of socioeconomic profile performance data and economic and social conditions were provided by the Board of Directors of Selective Process (DIRPS/UFU). The socioeconomic variables were examined 42 of candidate's manual questionnaire answered upon registration for the vestibular contest. The statistical procedure performed was for logistic regression analysis by Stepwise, as well as the review of ODDS ($\exp(B)$) adopting as a criterion for inclusion of the variable in the model of significance 5%. The results indicated that 13 (30%) of socioeconomic variables 42 influence somehow in the adoption or otherwise of competitor in vestibular UFU. The results suggest that candidates with less favourable socioeconomic make option for the less popular courses as indicated in the linearized equation obtained in this survey. It is concluded that, in General, some variables socioeconomic can influence the approval or not of vestibular UFU for courses studied.

Keywords: Logistic regression; Performance of approved and not approved; Socioeconomic profile

INTRODUÇÃO

A revisão da literatura tem demonstrado que o jovem brasileiro para se preparar profissionalmente antes de ingressar no ensino superior esbarra com a fragilidade do sistema educacional, uma vez que os cursos técnicos oferecidos têm como finalidade satisfazer as demandas do mercado. Esses cursos profissionalizantes são ofertados, geralmente, no decorrer do período noturno, direcionados particularmente para as pessoas que estão no mercado de trabalho, organizados com conteúdo específico, apresentando certa fragilidade do ponto de vista de preparação geral e teórica (BONETI, 2008).

Outra possibilidade de preparação são os cursos profissionalizantes estruturados por programas e projetos atrelados ao governo estadual ou federal, porém em caráter eventual, não havendo segurança de continuidade dos mesmos, bem como dos cursos proporcionados pelas instituições representativas do comércio e da indústria. No entanto, esses cursos são destinados aos sujeitos que se encontram trabalhando ou desempregados. Desse modo, é possível perceber que as alternativas de preparação profissionalizante das quais o jovem brasileiro têm disponível não possuem o perfil direcionados aos jovens, e se constituem de cursos que não correspondem às necessidades dos mesmos, com exceção do ensino superior (BONETI, 2008).

Dentro desse aspecto tem-se observado que a democratização do sistema de ensino, em nosso país, teve início na última metade da década de 1980, sendo que após 20 anos é possível perceber a descentralização do sistema educacional brasileiro. Entretanto, apesar dos esforços empreendidos nesse período, o ensino superior público ainda não permite acesso a mais de 90% de jovens brasileiros (CORREIA, 2008).

Sabe-se que em uma sociedade do conhecimento a trajetória de formação escolar e profissional está vinculada a um investimento, público e privado, de longa data no percurso do sujeito. No Brasil o sistema de bolsa de estudos não contempla 20% da demanda, e 80% das famílias precisam fornecer recursos para promoção dos estudos de seus jovens (CORREIA, 2008).

De acordo com estudos efetuados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), no Brasil, entre 40% da população mais pobre, somente 4,0% tem ingresso à Universidade, enquanto que entre os 10% da população

com nível socioeconômico mais elevado, 23,4% tem acesso à Universidade. Em relação às regiões menos favorecidas do país os dados indicam que, na Região Nordeste, entre 40% da população mais pobre, apenas 0,9% estão matriculados na Universidade, enquanto que entre 10% da população mais rica, 25,8% frequentam o ensino superior (INEP, 2004).

Outro aspecto importante é que a educação oferecida para a sociedade brasileira, de modo geral, fere a Constituição nacional, uma vez que esse ensino é desigual. Nessa perspectiva, de acordo com estudo desenvolvido pelo Projeto Juventude do Instituto Cidadania, os docentes têm enormes diferenças regionais em termos de remuneração e qualificação. Outro fator relevante é que os recursos das escolas são desiguais, parecendo ser bastante perceptível as diferenças ao se considerar centro e periferia (LASSANCE, 2005 apud MORGADO; VIEIRA, 2009).

Assim, para Morgado e Vieira (2009) um sistema educacional desigual não é capaz de cumprir de modo satisfatório sua função, e colabora de modo direto para uma série de dificuldades socioeconômicas, as quais, por sua vez, contribuem para que os jovens arquem com esses prejuízos no momento de ingressarem no ensino superior.

Diante dessas questões ressalta-se a importância da concretização de investigações científicas e a análise de dados do perfil socioeconômico dos candidatos de concursos vestibulares, visando à obtenção de mais alternativas para os jovens brasileiros no momento de fazer a opção pelo curso de nível superior. Para tanto, é de fundamental relevância a análise e a interpretação de dados utilizando-se os vários métodos estatísticos existentes, de modo que os resultados relativos a essas investigações sejam efetivamente consistentes e confiáveis. As principais metodologias de análises estatísticas podem ser encontradas em livros textos (TRIOLA, 1999, BUSSAB e MORETIN, 2002 e COSTA NETO, 2002).

Nesse sentido, o presente projeto utilizou principalmente estudos mais específicos de Arango (2001) e Vieira (2003), os quais enfatizam que a regressão logística é um procedimento estatístico bastante útil para a observação de eventos, sendo apropriado para a possibilidade de prever ou não a respeito de uma característica ou de uma razão particular, baseada em informações de um conjunto de variáveis independentes. Por meio da conversão da variável dependente em variável de base logarítmica, pode-se estimar a expectativa máxima da ocorrência de um fato

sucedem. Desse modo, o modelo logístico possibilita o conhecimento geral a respeito de um dado evento.

Assim, ressalta-se a relevância desses estudos para corroborar a pesquisa ora realizada, bem como no sentido de permitir à Universidade Federal de Uberlândia (UFU) a obtenção de dados palpáveis a respeito da constituição de estimativas probabilísticas relativas à aprovação nos cursos de Medicina, Psicologia, Engenharia Elétrica, Ciências Contábeis e Matemática em função do perfil socioeconômico dos candidatos.

O objetivo deste trabalho foi analisar, por meio da regressão logística múltipla, dados de candidatos ao concurso vestibular da UFU nos cursos supracitados, visando selecionar variáveis socioeconômicas que influenciem na opção e aprovação no concurso vestibular da UFU.

MATERIAL E MÉTODOS

Os dados utilizados na presente pesquisa foram disponibilizados pela Diretoria de Processo Seletivo (DIRPS/UFU), os quais foram apresentados sob o aspecto numérico e em códigos, impossibilitando, desse modo, a identificação do candidato. Essas informações relativas ao desempenho vestibular e ao perfil socioeconômico referem-se aos aprovados e não aprovados nos concursos vestibulares de julho e dezembro de 2009, e dizem respeito aos concorrentes dos cursos de Ciências Contábeis do período noturno, Engenharia Elétrica, Matemática, Medicina e Psicologia.

A amostra foi composta de 586 concorrentes, sendo 292 aprovados e 294 não aprovados. Desse modo sabe-se que, desses, 103 inscreveram-se para o curso de Ciências Contábeis do período noturno, com aprovação de 60 concorrentes; 109 para o curso de Engenharia Elétrica, tendo 60 aprovações; 94 inscritos para o curso de Matemática, com 52 aprovados; 158 para Medicina, tendo 60 aprovações; e, 122 inscritos para Psicologia, com aprovação de 60 concorrentes.

Aos concorrentes aprovados foi conferido o valor um e aos não aprovados conferiu-se o valor zero, sendo esta considerada a variável dependente no modelo de regressão. As variáveis independentes assumidas são aquelas relativas ao questionário socioeconômico, as quais também receberam codificação pertinente, conforme o Manual do Candidato ao Concurso Vestibular da UFU. Vale ressaltar que o questionário

socioeconômico foi preenchido pelos concorrentes quando da época de inscrição no concurso vestibular.

A regressão logística é um método estatístico utilizado para estimativa de probabilidades de ocorrências em variáveis dependentes do tipo binário. Arango (2001) alega que a regressão logística é um instrumento estatístico vantajoso em acontecimentos nos quais se objetiva prever a presença ou ausência de determinado resultado ou caractere, com base em valores de um conjunto de variáveis independentes. Assim, o modelo logístico permite avaliar a probabilidade máxima depois de converter a variável dependente de base logarítmica, calculando a probabilidade de uma ocorrência acontecer.

Em Arango (2001) e Ayres et al (2005) encontra-se uma descrição geral do modelo logístico simples. A equação geral da logística é dada por (Equação 1):

$$E(Y) = [1 + \exp(-\beta_0 - \beta_i X_i)]^{-1} \quad (1)$$

Em que: Y é a variável binária situação do candidato (1 = aprovado e 0 = não aprovado), X_i são as variáveis preditoras e β_0 e β_i são os parâmetros linearizados da função logística.

Uma qualidade relevante é que a função logística pode ser linearizada. Denotando-se $E(Y)$ por π , pois a resposta média é a probabilidade quando a variável resposta é binária (Arango, 2002). Fazendo-se a transformação (Equação 2 e 3):

$$\pi' = \ln(\pi / 1 - \pi) \quad (2)$$

$$\pi' = \beta_0 + \beta_1 X \quad (3)$$

Esta transformação é denominada de transformação logit da probabilidade π .

A razão $\pi / (1 - \pi)$ na transformação logit é chamada de Odds (chance). A função resposta transformada é denominada como função resposta logit, e π' é denominada de resposta média logit.

Para o cálculo de probabilidade de ocorrência da variável dependente binária, dada uma condição da variável preditora X, Ayres et al (2005) indicam a seguinte função (Equação 4):

$$p = 1 / (1 + \exp(-(\beta_0 + \beta_1 X))) \quad (4)$$

Ao se considerar o concurso vestibular da UFU para os cursos de Ciências Contábeis do período noturno, Engenharia Elétrica, Matemática, Medicina e Psicologia, essa pesquisa visou utilizar o método de regressão logística para observar a influência das variáveis socioeconômicas (preditoras) na aprovação ou não dos candidatos.

A generalização do modelo logístico simples descrito acima para o modelo logístico múltiplo se faz de forma natural. Neste caso, tem-se uma variável dependente (Y), aprovado ou não aprovado, e, diversas variáveis independentes (X), ou seja, obtém-se um coeficiente relacionado a cada variável independente selecionada para o modelo.

Em relação às variáveis independentes utilizadas no modelo, ressalta-se que o total inicial dessas foi de 42, as quais foram: sexo, idade, estado civil, local de residência do candidato, distância da moradia do candidato em relação a Uberlândia, etnia, religião, situação familiar, se o candidato exerce atividade remunerada, idade de início da atividade remunerada, quantidade de pessoas que compõem a família, renda bruta familiar, instrução do pai, instrução da mãe, responsável pela família, existência de computador na residência, utilização de computador, meio de transporte utilizado, intenção de trabalhar durante o curso universitário, situação do imóvel da família, com quem o concorrente reside, participação na economia da família, função profissional do pai, função profissional da mãe, meios de obtenção de informação, preferências culturais, atividade física, hábito de ir ao cinema, atividade cultural desenvolvida, participação em grupo de arte e cultura, férias, tipo de revistas lidas, ocupação do tempo, tipo de ensino fundamental, ensino fundamental, cursa ensino médio, tipo de ensino médio cursado, horário que cursou e/ou cursa ensino médio, cursinho frequentado, expectativa do curso universitário, razão por escolher a UFU, influência na escolha pela UFU.

Devido ao grande volume de variáveis independentes optou-se por utilizar o método de inclusão de variáveis por meio do procedimento Stepwise. Adotando como critério de inclusão no modelo a significância de 5%.

Esta metodologia consiste em incluir variáveis significativas no modelo a cada passo, desta forma o modelo inicia-se com uma única variável independente e passo a passo são incluídas novas variáveis até que a inclusão de outras variáveis não altere, significativamente, os resultados produzidos pelo modelo.

Ao se considerar o concurso vestibular da UFU para os cursos de Ciências Contábeis do período noturno, Engenharia Elétrica, Matemática, Medicina e Psicologia, essa pesquisa objetivou utilizar o método de regressão logística para observar a influência das variáveis socioeconômicas (preditoras) na aprovação ou não dos candidatos.

RESULTADOS E DISCUSSÃO

O procedimento Stepwise foi realizado para os cursos de Ciências Contábeis do período noturno, Engenharia Elétrica, Matemática, Medicina e Psicologia. Desse modo, para o curso Ciências Contábeis, esse procedimento ocorreu em seis passos, sendo selecionadas as variáveis independentes: Influência da Escolha (IE), Hábito de Ir (HI), Preferência (PE), Computador (CO), Responsável (RE) e Estado Civil (EC). Esse resultado especificou que das 42 variáveis socioeconômicas somente 6 contribuíram significativamente para aprovação e/ou reprovação dos candidatos. Assim, a equação linearizada para o curso de Ciências Contábeis foi:

Logit (π) = 1,276+0,550IE+0,515HI-0,354PE-0,914CO+0,401RE-1,285EC. Sendo:

IE= Influência da escolha pela UFU

HI= Hábito de ir ao cinema

PE= Preferência entre cinema/teatro/música/esportes, etc.

CO= Computador na residência

RE= Responsável pela residência

EC= Estado Civil

A análise do ODDS ($\exp(B)$) indicou que o grau de influência de escolha pela UFU, o hábito de ir ao cinema e o responsável pela residência contribuíram positivamente para o aumento da probabilidade de aprovação, enquanto que a preferência, o computador na residência e o estado civil contribuíram negativamente, pois o primeiro grupo apresentou ODDS acima de 1, isto é, 1,733, 1,674 e 1,494 respectivamente, e, o segundo abaixo, ou seja, 0,702; 0,401 e 0,161.

Assim, ao se considerar as codificações adotadas pode-se inferir, por exemplo, que quanto maior o grau de influência de escolha pela UFU maior a chance de aprovação; em contrapartida, o fato de o candidato não possuir computador em casa pode contribuir negativamente com a aprovação do mesmo no concurso vestibular da UFU na área de Ciências Contábeis.

Dessa maneira, ao se pensar a situação em que o candidato foi influenciado pela família na escolha pela UFU (1), tenha o hábito de ir ao cinema uma vez por mês (1), tenha preferência por dança (3), não possua computador (3), tenha como responsável pelo sustento da família a mãe (1) e o estado civil solteiro (1), tem-se 8,73% de probabilidade de aprovação.

No que se refere ao curso de Engenharia Elétrica o procedimento Stepwise aconteceu em dois passos selecionando as variáveis independentes: Influência da Escolha (IE) e Curso de Ensino Médio (CEM). Dessa maneira, constatou-se que somente 2 das 42 variáveis socioeconômicas colaboraram para aprovação ou não dos concorrentes para esse curso. A equação linearizada para o curso Engenharia Elétrica foi:

Logit (π) = $-0,311IE+0,239CEM-0,330$. Sendo:

IE= Influência da Escolha pela UFU

CEM= Curso de Ensino Médio que o candidato está cursando

O exame do ODDS ($\exp(B)$) especificou que o grau da variável Curso de Ensino Médio que o candidato está cursando colaborou de modo positivo para o aumento da probabilidade de aprovação, visto que o ODDS apresentado foi 1, 270 acima de 1. Já a Influência da Escolha pela UFU contribuiu de maneira negativa, uma vez que o ODDS especificado foi 0, 733 abaixo de 1.

Esses resultados possibilitam inferir, por exemplo, que quanto maior o grau da variável Curso de Ensino Médio que o concorrente está cursando maior a chance de aprovação; enquanto que o grau de Influência da Escolha pela UFU pode contribuir de modo negativo para a aprovação no vestibular da UFU. Desse modo, ao se considerar a situação em que o concorrente foi influenciado pela família na escolha pela UFU (1) e cursa ensino médio profissionalizante (1), esse tem 48,20% de possibilidade de aprovação. E se o concorrente foi influenciado pela família na escolha pela UFU (1) e cursa ensino médio regular (2), a probabilidade de aprovação é de 54,16%.

Já para o curso de Matemática o procedimento Stepwise foi realizado em cinco passos sendo selecionadas as variáveis independentes: Cursinho (CU), Ensino Fundamental (EF), Responsável (RE), Instrução da mãe (IM), Computador (CO). Este resultado mostrou que das 42 variáveis socioeconômicas apenas 5 contribuíram significativamente para a aprovação e/ou reprovação dos candidatos. Desta forma a equação linearizada para o curso de Matemática foi:

Logit (π) = $-21,331-0,384CU+21,284EF+0,445RE+0,223IM-0,707CO$. Sendo:

CU= cursinho

EF= ensino fundamental

RE= responsável da residência

IM= instrução da mãe

CO= computador na residência

A análise do ODDS ($\exp(B)$) indicou que o grau das variáveis: Instrução da Mãe, Responsável da Residência e tipo de Ensino Fundamental contribuíram positivamente para o aumento da probabilidade de aprovação, já as variáveis Computador na Residência e Cursinho contribuíram negativamente, pois o primeiro grupo apresentou ODDS acima de 1 e o segundo abaixo. Considerando as codificações adotadas pode-se inferir, por exemplo, que quanto maior o grau de instrução da mãe maior a chance de aprovação; em contrapartida, o fato de o candidato não possuir computador em casa pode contribuir negativamente com a aprovação do mesmo no concurso vestibular da UFU na área de Matemática.

Diante desses resultados ao se considerar a situação em que o candidato não fez cursinho (1), cursou o ensino fundamental em escola pública (1), tem como responsável a mãe (2), o grau de instrução da mãe é não alfabetizado (1) e que não possui computador (3), tem-se 13,19% de probabilidade de aprovação. Se o grau de instrução da mãe desse mesmo candidato fosse de nível superior completo (7), a possibilidade de aprovação seria de 36,67%.

Em relação ao curso de Medicina o procedimento Stepwise realizou-se em dois passos, sendo selecionadas as seguintes variáveis independentes: Quantidade de Pessoas (QP) e Curso de Ensino Médio (CEM). Desse modo, percebeu-se que somente 2 das 42 variáveis socioeconômicas influenciaram na aprovação ou reprovação do candidato. Assim, a equação linearizada para esse curso é:

$\text{Logit}(\pi) = -0,425QP + 0,177CEM + 0,186$. Sendo:

QP= Quantidade de Pessoas que Compõem a Família

CEM= Curso de Ensino Médio que o candidato está cursando

O estudo do ODDS ($\exp(B)$) demonstrou que o grau da variável: Quantidade de Pessoas da Família colaborou de maneira negativa para o aumento da probabilidade de aprovação, visto que o ODDS apresentado está abaixo de 1, ou seja, 0,654. Enquanto que o Curso de Ensino Médio que o concorrente estava cursando contribuiu de maneira positiva, uma vez que o ODDS apresentado foi 1,194 acima de 1. Assim, ao se considerar a situação em que o vestibulando tem sua família composta por três pessoas (3) e cursa ensino médio regular (2), tem-se 32,40% de aprovação. E se a família é composta por cinco pessoas (5) e cursa ensino médio regular (2) a probabilidade de aprovação vai para 17,00%.

No que diz respeito ao curso de Psicologia o procedimento Stepwise foi realizado em dois passos e elencou as seguintes variáveis independentes: Atividade Física desenvolvida pelo candidato (AF) e Local onde o vestibulando reside (UF). Para esse curso apenas 2 das 42 variáveis socioeconômicas influenciaram na aprovação ou reprovação do candidato. Dessa maneira, a equação linearizada foi:

Logit (π) = $-0,249AF - 0,098EF + 2,434$. Sendo:

AF= Atividade Física praticada pelo concorrente

UF= Local onde o concorrente reside

A análise do ODDS ($\exp(B)$) demonstrou que o grau atividade física praticada pelo vestibulando, bem como o local onde o mesmo reside contribuíram negativamente para a aprovação, uma vez que o ODDS apresentado estava abaixo de 1, ou seja, 0,779 e 0,907 respectivamente. Assim, ao se considerar uma situação em que o candidato pratique atividade física uma vez por semana (1) e resida no estado de Minas Gerais (14), o mesmo teria 69,27% de probabilidade de aprovação. Já se o concorrente não praticar atividade física (3) e resida no estado de São Paulo (26), a probabilidade de aprovação do mesmo seria de 29,71%.

Diante do ODDS ($\exp(B)$) apresentado para o curso de Psicologia pode-se inferir, por exemplo, que, quanto maior o grau de atividade física praticada, bem como local onde o concorrente reside menor será a chance de aprovação no candidato no vestibular da UFU.

O processo Stepwise também foi realizado levando-se em consideração todos os cinco cursos em questão, cujo intuito foi alcançar uma equação linearizada geral. Esse procedimento ocorreu em seis passos, escolhendo as seguintes variáveis independentes: Cursinho (CU), Curso de Ensino Médio (CEM), Hábito de Ir (HI), Atividade Física Desenvolvida pelo Candidato (AF) e Computador na Residência (CO). Dessa maneira, ao se considerar todos os cursos analisados, percebe-se que das 42 variáveis apenas 5 influenciaram na aprovação ou não dos concorrentes. Ressalta-se que a variável Código do Curso Escolhido (CC) diz respeito à opção do candidato pelo curso na época da inscrição no vestibular, não sendo, portanto, considerada variável socioeconômica. Assim, equação linearizada foi:

Logit (π) = $-0,498 - 0,148CU + 0,112CEM + 0,217HI - 0,086AF - 0,314CO + 0,178CC$. Sendo:

CU= Cursinho

CEM= Curso de Ensino Médio que o candidato está cursando

HI= Hábito de ir ao cinema

AF= Atividade Física praticada pelo concorrente

CO= Computador na residência

CC= Código do Curso Escolhido

O exame do ODDS ($\exp(B)$) indicou que o grau de opção de curso, o hábito de ir ao cinema e o curso de ensino médio que o concorrente pratica contribuíram de modo positivo para o aumento da probabilidade de aprovação, já que o ODDS demonstrado foi de 1, 199; 1, 242 e 1, 119 respectivamente, ou seja, acima de 1. Já o computador em casa, a atividade física praticada pelo vestibulando e o cursinho frequentado pelo mesmo colaboraram negativamente para a aprovação, uma vez que o ODDS especificado estava abaixo de 1, os quais foram respectivamente 0,730; 0,917 e 0,862.

De acordo com o ODDS ($\exp(B)$) apresentado é possível inferir que quanto maior o grau de opção pelo curso, hábito de ir ao cinema, bem como o curso de ensino médio que o vestibulando cursa maior será a probabilidade de aprovação; em contrapartida, o fato de não ter computador na residência pode colaborar de modo negativo na aprovação do mesmo no concurso vestibular da UFU.

Desse modo, ao se considerar um concorrente que fizesse cursinho (1); cursasse o ensino médio regular (2), não tivesse o hábito de ir ao cinema (5); praticasse atividade física uma vez por semana (1), tivesse computador com acesso à internet (1); e, optasse pelo curso de Ciências Contábeis (3), o mesmo teria 68,93% de ser aprovado. Caso esse mesmo candidato optasse pelo curso de Matemática (4) ou Engenharia Elétrica (5), a probabilidade de aprovação seria respectivamente de 72,61% e 76,00%. Já se a opção de curso fosse para Psicologia (1) ou Medicina (2) a probabilidade de aprovação seria de 60,84 e 64,99% respectivamente.

Ao se observar a equação para curso de Ciências Contábeis e Matemática percebe-se que computador na residência (CO) influencia de modo negativo para a aprovação dos concorrentes, enquanto que o responsável pela residência (RE) exerce influência positiva. No entanto, o tipo de ensino fundamental frequentado pelo vestibulando aparece somente no curso de Matemática e exerce influência positiva para a aprovação nesse curso.

No que se refere aos cursos de Engenharia Elétrica e Medicina, bem como Matemática, considerando a equação geral, observa-se que o grau da variável Curso de Ensino Médio que o candidato está cursando (CEM) exerce influência positiva para a

aprovação dos candidatos nos cursos escolhidos. Esse fato parece ser relevante uma vez que o grau do código do curso escolhido (CC), o qual aparece na equação geral, em conjunto com o grau do curso de ensino médio (CEM) e o hábito de ir ao cinema (HI) contribui de modo relevante para a aprovação no vestibular da UFU.

Assim, pode-se inferir, por exemplo, que o candidato que faz a opção pelos cursos de Matemática (4) e Engenharia Elétrica (5), e cursa ensino médio regular (2) e tem o hábito de ir ao cinema duas vezes por mês (2) terá uma maior probabilidade de aprovação, caso seja comparado com a opção pelos cursos de Psicologia (1) e Medicina (2) mantendo-se as demais condições.

Em relação ao curso de Psicologia as variáveis que mais parecem exercer influência na aprovação ou não do candidato são: Atividade Física praticada pelo concorrente (AF) e Local onde o mesmo reside (UF). Dados como esse nos possibilitam pensar que, provavelmente, a equação geral seria uma melhor opção para traçar o perfil de candidatos aprovados no vestibular UFU, bem como a utilização de somente algumas variáveis socioeconômicas mais significativas.

Com base nesse questionamento, fez-se o cálculo da equação linearizada geral utilizando as seguintes variáveis: Quantidade de Pessoas na família (QP), Renda Bruta Familiar (RB), Instrução Pai (IP), Instrução Mãe (IM), Responsável (RE), Transporte (T), Situação Imóvel (SI), Função Paterna (FP), Função Materna (FM), Ensino Fundamental (EF), Ensino Médio (EM), Cursinho (CU) e Código do Curso (CC), obtendo-se a seguinte equação geral: $\text{Logit}(\pi) = 0,430 - 0,190\text{FP} - 0,150\text{QP} + 0,148\text{CC}$. Esse resultado parece sugerir ser mais adequado para esse tipo de estudo a utilização seletiva das variáveis socioeconômicas e também a equação linearizada geral.

Outro aspecto percebido foi a necessidade de um estudo mais pormenorizado de mais dados, ou seja, pesquisar outros concursos vestibulares da UFU abrangendo uma quantidade maior de cursos.

Esses dados possibilitam ainda a sugestão de novas pesquisas relacionadas à percepção desses alunos em relação ao próprio desenvolvimento ao longo do curso, às possibilidades oferecidas pelo mesmo, bem como quais perspectivas futuras relativas à atividade profissional serão desenvolvidas ao longo do curso optado.

Outra pesquisa sugerida estaria relacionada com o momento da entrada e de saída dos mesmos no curso prestado, de modo que se visualize como se apresentava a autoestima dos mesmos nesses momentos. É válido também que se repita uma pesquisa

similar ao estudo ora empreendido em vestibulares prestados em anos distintos, de maneira que se confirmem ou não os dados obtidos no presente estudo.

CONCLUSÕES

A metodologia estatística de seleção de variáveis preditoras que influenciam na opção e aprovação de concurso vestibular da UFU para os cursos de Ciências Contábeis do período noturno, Engenharia Elétrica, Matemática, Medicina e Psicologia mostrou-se eficiente do ponto de vista metodológico, entretanto, faz-se necessário novos estudos com amostras maiores, mais dados de concursos vestibulares e uma análise prática das variáveis envolvidas na equação para apurar a confiabilidade das estimativas.

Neste estudo, verificou-se que, de modo geral, as variáveis: influência da escolha pela UFU; hábito de ir ao cinema; preferência entre cinema/teatro/música/esportes/outros; estado civil; curso de ensino médio que o candidato está cursando; cursinho; tipo de ensino fundamental; responsável pelo sustento da residência; instrução da mãe; computador na residência; quantidade de pessoas que compõem a família; atividade física praticada pelo concorrente; local onde o vestibulando reside, podem contribuir significativamente na aprovação dos candidatos ao concurso vestibular da UFU.

REFERÊNCIAS

ARANGO, H. G. **Bioestatística**: Teórica e Computacional. Rio de Janeiro: Guanabara Koogan, 2001.

AYRES, M.; AYRES Jr, M.; AYRES, D. L.; SANTOS, A. S. dos. **BioEstat 4.0: Aplicações estatísticas nas áreas das ciências biológicas e médicas**. Belém: Sociedade Civil Mamirauá; Brasília: CNPq, 2005. 324 p.

BIOESTAT aplicações estatísticas nas áreas das ciências bio-médicas. Disponível em: <<http://people.ufpr.br/~vayego/pedeefes/manual.pdf>>. Acesso em: 06 ago 2011.

BONETI, L. W. Formação profissional e emprego num contexto de pobreza: jovem pobre, pobre jovem: a condição de acesso ao ensino superior no Brasil. In: **Um olhar sobre o jovem no Brasil**. 1. ed. Brasília: Ministério da Saúde Fundação Oswaldo Cruz, 2008,p.107-116. Disponível em:

<http://www.espm.br/ConhecaAESPM/CAEPM/nucleodeestudosdajuventude/Documents/Banco%20de%20Dados%20Jovens/10.%20SOCIOLOGIA%20DA%20JUVENTUDE/10.24.%20olhar_sobre_jovem_brasil.pdf>. Acesso em: 10 mar 2011.

BUSSAB, W. O, Morettin, P. A. **Estatística básica**. 5. ed. São Paulo: Saraiva, 2002. 526 p.

CORREIA, S. M. S. Brasil: uma sociedade de jovens? In: **Um olhar sobre o jovem no Brasil**. 1 ed. Brasília: Ministério da Saúde Fundação Oswaldo Cruz, 2008, p. 11-27. Disponível em:<

http://www.espm.br/ConhecaAESPM/CAEPM/nucleodeestudosdajuventude/Documents/Banco%20de%20Dados%20Jovens/10.%20SOCIOLOGIA%20DA%20JUVENTUDE/10.24.%20olhar_sobre_jovem_brasil.pdf>. Acesso em: 10 mar 2011.

COSTA NETO, P. L. DE O. **Estatística**. 2. ed. São Paulo: Edgar Blucher, 2002. 266 p. INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). **Acesso ao ensino superior é marcado por desigualdade entre ricos e pobres**. Informativo, Brasília, ano 2, n. 40, p. 1, 25 mai 2004. Disponível em: <<http://www.inep.gov.br/informativo/pdf/informativo40.pdf>>. Acesso em: 24 mar 2011.

MORGADO, M.A; VIEIRA, H.Z. **Interferência da Condição Sócio-Econômica na Escolha dos Jovens pelo Curso Superior na UFMT**. Disponível em:< <http://www.ie.ufmt.br/semiedu2009/gts/gt14/ComunicacaoOral/HELIZE%20DE%20SOUZA%20VIEIRA.pdf>>. Acesso em: 24 mar 2011.

TRIOLA. M. F. **Introdução à Estatística**. 7. ed. Rio de Janeiro. 1999.

VIEIRA, S. **Bioestatística: tópicos avançados**. 5. ed. Rio de Janeiro: Elsevier, 2003