

APLICAÇÃO DE UM MODELO DE ESCOLHA DISCRETA PARA ANÁLISE DA DIVISÃO MODAL EM CIDADES DE PORTE MÉDIO

SILVA, Thaís. Universidade Federal de Uberlândia. Curso de Pós-Graduação da Faculdade Engenharia Civil. siahtsilva26@yahoo.com.br

MENDES, Flávia Bruno. Universidade Federal de Uberlândia. Curso de Pós-Graduação em Faculdade Engenharia Civil. fbm_udimoc@yahoo.com.br

FARIA, Carlos Alberto. Universidade Federal de Uberlândia. Curso de Pós-Graduação da Faculdade Engenharia Civil. cafarria@ufu.br

ABSTRACT

This research treats on modal split phase that is understood as the third phase of the classic model based on choice probabilities of users in front of the main able alternatives of transport that is provided in medium sized cities. An approach of this phase in the present study was treated by focusing on customer behavior related to the utility concept that constitutes the theoretical foundation of micro economy and discrete choice. Then, from origin-destiny (O/D) research data that were realized on Uberlândia (Minas Gerais/Brazil) city in 2002 it was obtained data bank to realize the calibration of Logit Model on SPSS (*Statistical Package for Social Sciences*) as well as to study the behavior of urban transport demand on modal split phase.

Keywords: urban planning, urban transport, logit model, split modal, demand behavior

INTRODUÇÃO

O crescimento da demanda por transportes urbanos levou ao desenvolvimento de modelos que pudessem ser aplicados ao planejamento de transportes. O modelo clássico ou de quatro etapas tornou-se então, instrumento estratégico empregado por pesquisadores da área de transportes na tomada de decisão. Esse modelo considera o planejamento como um processo sequencial em que interagem diferentes submodelos: produção-atração de viagens, distribuição, divisão modal e alocação do tráfego à rede (Ortúzar e Román, 2003).

)

Essa pesquisa trata da etapa de divisão modal, compreendida como a terceira etapa do modelo de quatro etapas, com base nas probabilidades de escolha dos usuários diante das principais alternativas de transportes disponíveis em cidades de porte médio. A abordagem dessa etapa será tratada a partir do enfoque da teoria do comportamento do consumidor associada ao conceito de utilidade.

O processo de decisão do consumidor na fase de divisão modal dos transportes baseia-se em escolhas discretas, refletindo características importantes dos sistemas de transportes, das viagens e do nível socioeconômico do consumidor. Logo, os indivíduos, usuários dos modos de transportes, diante da necessidade de deslocar-se no espaço e no tempo, são orientados por uma função utilidade e procuram maximizá-la, escolhendo dentre várias alternativas possíveis aquela cujos atributos lhe proporcionem o maior nível de satisfação.

O “juízo de valor” do consumidor sobre as alternativas e atributos dos modos de transportes, passa a ser então, representativo para modelar a demanda por transportes na etapa de divisão modal. Assim, a partir dos dados da pesquisa de origem - destino (O/D) realizada na cidade de Uberlândia (Minas Gerais/Brasil) em 2002, obteve-se o banco de dados para realizar a calibração do modelo de regressão logística multinomial no *software* SPSS (*Statistical Package for Social Sciences*).

OBJETIVOS

Objetivo Geral

O objetivo principal desse trabalho é calibrar modelos de divisão modal baseados na teoria econômica do comportamento do consumidor e na teoria da utilidade, utilizando o *software* SPSS (*Statistical Package for Social Sciences*) e os dados da pesquisa origem/destino (O/D) realizada em 2002 na cidade de Uberlândia (MG).

Objetivos específicos:

- Estimar as probabilidades de escolha dos usuários dos modos de transportes considerando como opções disponíveis ao usuário o ônibus, o automóvel e o caminhar;
- Avaliar o padrão de escolha modal considerando a amostra da população das 65 zonas de tráfego de Uberlândia (MG);
- Mensurar a influência das variáveis sobre o padrão de escolha dos usuários, gerando cenários com base nos atributos das variáveis e definir o perfil do usuário de transporte por cada modal considerado.

JUSTIFICATIVA

A crescente demanda por transportes tem exigido dos órgãos gestores e dos planejadores de transportes esforços para a manutenção do equilíbrio entre a demanda e a oferta desse serviço. Assim, o planejamento de transportes deve ser considerado como instrumento para o desenvolvimento estratégico das cidades brasileiras de médio e grande porte.

Essa integração entre o planejamento em transportes urbanos e o padrão de uso e ocupação do solo pode significar uma saída à estruturação dos centros urbanos, uma vez que, podem facilitar a formulação e aplicação de métodos e modelos baseados em dados desagregados, que refletem melhor o padrão de ocupação, bem como as necessidades e expectativas da demanda por serviços de transportes urbanos.

)

Ferronato (2002) observou que, o uso do solo não somente tem um impacto direto sobre a demanda, mas ele próprio é afetado por mudanças na oferta de transportes. Dessa forma, as aplicações de modelos com base em dados desagregados constituem importante ferramenta para a análise da repartição modal, permitindo a espacialização das novas necessidades e auxiliando de forma mais correta à tomada de decisão.

Considerando então, a aplicação dos modelos com base em dados desagregados, a pesquisa de origem e destino representa a base de dados primária para a calibração dos modelos no enfoque clássico do planejamento de transportes. Apesar da pesquisa OD ser de difícil aplicação e demandar recursos financeiros para sua aplicação é, ainda, a melhor fonte de informações do padrão das viagens do cotidiano (Faria et al, 2004).

Nesse sentido, modelar a etapa de divisão modal com base em dados da pesquisa OD significa buscar a aproximação dos modelos de escolha discreta com a real escolha dos usuários de transportes.

MODELOS PARA ANÁLISE DE DEMANDA E A TEORIA MICROECONÔMICA PARA AS ESCOLHAS DISCRETAS

A análise da demanda em transportes tem adotado tradicionalmente dois tipos de modelos: os modelos agregados ou de primeira geração e os modelos desagregados ou de segunda geração. Os primeiros foram amplamente utilizados nos estudos de transportes até a década de 1970, mas devido a sua pouca flexibilidade, precisão limitada, elevado custo e sua fraca orientação na tomada de decisões políticas (Ortúzar e Willumsen, 1990), esses modelos perderam força no processo de planejamento.

Então, a partir do final da década de 1970, em função das limitações nas abordagens dos modelos agregados, os pesquisadores passaram a empregar os modelos de segunda geração ou desagregados. Esses, por sua vez, utilizam observações individuais, o que permite uma melhor compreensão dos comportamentos de viagem, posto que baseiam-se nas teorias de escolha individual (Espino, 2003). Os modelos de segunda geração fundamentaram-se nos princípios teóricos da economia clássica, no enfoque comportamental e em técnicas de marketing.

Ortúzar e Willumsen (1990) ressaltam que, as vantagens que estes modelos apresentam em relação aos de primeira geração são muitas. Em primeiro lugar, a possibilidade de utilizar dados individuais faz com que seja mais eficiente o uso das informações do que nos modelos agregados. Em segundo lugar, pode-se utilizar toda a variabilidade de que se dispõem com dados individuais. Em terceiro lugar, os resultados dos modelos em nível individual podem ser usados para qualquer nível de agregação. Em quarto lugar, a estimativa dos parâmetros para cada uma das variáveis explicativas consideradas no modelo é explícita, o que proporciona certa flexibilidade para representar as variáveis relacionadas a políticas de planejamento; ademais, os parâmetros estimados têm uma interpretação direta em termos da importância de cada variável explicativa considerada na escolha. E finalmente, trata-se de modelos mais estáveis no tempo e no espaço.

Logo, os modelos desagregados que representam as escolhas discretas permitem que sejam estimadas as probabilidades de escolha a partir da abordagem comportamental. A seguir trataremos as teorias do comportamento do consumidor e da utilidade que são os fundamentos dos modelos de escolha discreta, e dos métodos de obtenção dos dados por Preferência Revelada (PR) e Preferência Declarada (PD) que são a base desses modelos.

Fundamentos Microeconômicos

A teoria econômica clássica propõe um modelo para o comportamento do consumidor onde às preferências que definem esse comportamento pode ser representada por uma função de utilidade.

)

O indivíduo diante de uma necessidade ou desejo em adquirir um bem ou serviço, analisa um conjunto de alternativas disponíveis escolhendo aquela (s) cujos atributos propiciam o maior nível de satisfação. Segundo Ben-Akiva e Lerman (1989), a escolha é resultado de um procedimento realizado pelo usuário de transportes, que compreende os seguintes elementos: o tomador de decisão, as alternativas, os atributos das alternativas e as regras de decisão.

Nestes modelos as escolhas são independentes e baseadas nas respectivas utilidades individuais em relação ao conjunto de escolhas possíveis. Assim, os dados para estimativa dos modelos probabilísticos de escolha discreta são obtidos através das técnicas de preferências declaradas e reveladas.

Preferências Reveladas (PR) e Declaradas (PD)

Os dados oriundos das técnicas de preferências reveladas e declaradas são as fontes de informação empregadas pelos modelos desagregados de escolha discreta. As preferências reveladas baseiam-se nas escolhas efetivamente realizadas pelos indivíduos diante de uma situação existente e contêm informação sobre a importância relativa das diversas variáveis que influenciam a sua decisão (Ortúzar e Roman, 2003).

As preferências declaradas referem-se às escolhas em relação a um conjunto de opções e são apresentados cenários hipotéticos ao consumidor para que ele indique a sua escolha. Essas técnicas começaram a ser aplicadas no princípio dos anos 70 em problemas relacionados com a investigação de mercado, e as primeiras aplicações no campo dos transportes se deve a Louviere et al. (1973) apud Ortúzar e Roman (2003). Em transportes as técnicas são empregadas na realização das pesquisas de origem e destino, as quais fornecem informações sobre o comportamento do consumidor.

Teoria do Comportamento do Consumidor

Uma das premissas do enfoque comportamental é a de que o indivíduo estabelece subjetiva e objetivamente um elenco de opções alternativas na ordem de preferência, e escolhe sempre a mais desejável, dado o conjunto de inclinações (gastos) e dados as condicionantes de ordem econômico-financeira e de oportunidades disponíveis (NOVAES, 1986). Não obstante, o crescimento rápido dos centros urbanos impõe à população uma série de escolhas ao realizarem seus deslocamentos, sendo a escolha do modo de transporte um fator decisivo para a sustentabilidade da qualidade de vida nesses centros.

Dessas escolhas resulta uma série de impactos negativos ou positivos para o desenvolvimento dos centros urbanos. A prioridade por modos de transportes mais sustentáveis pode significar para o futuro de cidades de porte médio, maior mobilidade e acessibilidade, melhor fluidez do tráfego, menos congestionamentos, redução nas emissões de poluentes e tempo em deslocamentos.

Segundo Kotler (1976), os psicólogos clássicos interpretam as necessidades do homem por meio da interação de impulsos, estímulos, sugestões, reações e reforços. O comportamento pode ser entendido como um mecanismo que o indivíduo utiliza para dar resposta a um determinado evento na busca de satisfazer seus desejos e necessidades.

O processo comportamental inicia-se com um estímulo interno que impulsiona o indivíduo a uma ação (Kotler, 1976). O impulso torna-se um motivo quando é dirigido a um objeto específico, a partir do impulso o indivíduo é condicionado a uma reação. A reação de uma pessoa a um determinado estímulo é influenciada por sugestões. As sugestões por sua vez são estímulos menores que determinam quando, onde e como uma pessoa reage.

A reação é, portanto, a resposta do organismo ao estímulo. Se a reação for compensadora, a probabilidade de uma reação similar na próxima vez para a mesma configuração de uma sugestão está reforçada. Se uma reação não for compensadora, a probabilidade de uma reação similar diminui (Kotler, 1976). Assim, a escolha do modo de transporte está

condicionada pelas características do usuário e pela atratividade do modo de transporte e a probabilidade de escolha será tanto maior, quanto mais suprir, os desejos e necessidades individuais dos usuários. Nota-se então, a importância de investimentos em modos de transportes mais sustentáveis e incentivos ao uso do transporte coletivo, a bicicleta e o caminhar.

MODELO DE REGRESSÃO LOGÍSTICA MULTINOMIAL

Neste trabalho a calibração do modelo de escolha discreta deu-se pelo método da regressão logística multinomial. O método é apropriado para análise de experimentos que apresentam variáveis de resposta categórica, cujo interesse principal seja a descrição da relação entre a variável de resposta e o conjunto de variáveis preditoras. Esse tipo de regressão assume uma forma mais geral, já que a variável dependente não se restringe a apenas duas categorias como ocorre no Logit Binomial (Manual SPSS 11.0, 2003). O ajuste por máxima verossimilhança tem por objetivo obter, a partir de uma amostra, estimativas de parâmetros estatísticos, assegurando consistência, eficiência e ajuste dos parâmetros do modelo.

Os dados de entrada para a calibração do modelo devem obedecer às seguintes exigências:

- A variável dependente deve ser categórica e as variáveis independentes podem ser fatores ou covariáveis. Em geral, os fatores devem ser variáveis categóricas (por exemplo: variável sexo: feminino=0 e masculino=1) e as covariáveis devem ser variáveis contínuas (por exemplo: variável tempo de viagem);
- Assume-se que a razão de escolha de qualquer par de categorias é independente das demais categorias de resposta.

Assim, o modelo logit multinomial assume uma coleção de $r+1$ variáveis independentes denotadas por $X = (X_0, X_1, X_2, \dots, X_r)$ onde $x = (x_0, x_1, x_2, \dots, x_r)$ com $x_0 = 1$ e uma variável resposta Y de natureza nominal que pode assumir os níveis $0, 1, 2, \dots, q$. Denota-se as funções logit como sendo:

$$gk \equiv gk(x) = \ln \left[\frac{P(Y=k/x)}{P(Y=0/x)} \right]$$
$$= \beta_{k0} x_{k0} + \beta_{k1} x_{k1} + \dots + \beta_{kr} x_{kr}$$
$$x' \beta_k, \text{ para } k \in \{0, \dots, q\}, \quad (1)$$

onde $\beta_k = (\beta_{k0}, \dots, \beta_{kr})$ e $x_{k0} = 1$

Assumindo-se n observações de Y , denotadas por y_1, \dots, y_n , associadas aos valores de $x_i = (x_{i0}, \dots, x_{ir})$, para $i \in \{1, \dots, n\}$, o logit, dado em (1) apresenta-se como:

$$g_{kn} = g_{kn}(x_n) = \beta_{k0} X_{n0} + \beta_{k1} X_{n1} + \dots + \beta_{kr} X_{nr} + \varepsilon_n \quad (2)$$

Onde $x_{i0} = 1$, para $i \in \{1, \dots, n\}$ e os erros, ε_i , seguem as seguintes suposições, para todo $i, l \in \{1, \dots, n\}$

$$(i) \quad \mathbb{E}(\varepsilon_i/x_i) = 0 \quad (3)$$

$$(ii) \quad \text{Var}(\varepsilon_i/x_i) = \text{Var}(Y_i/x_i)$$

$$(iii) \quad \text{Cov}(\varepsilon_i/\varepsilon_l) = 0, \quad \text{se } i \neq l$$

O exposto acima motiva a seguinte definição: As variáveis Y_1, \dots, Y_n satisfazem um modelo logístico multinomial se uma amostra de tamanho um de cada Y_i pode ser expressa como:

$$Y_k(x_i) = \frac{\exp(g_{ki})}{1 + \exp(g_{ki})} + \varepsilon_i \quad (4)$$

Onde g_{ki} é obtida pela expressão (1), β_{kr} são parâmetros desconhecidos, os erros ε_i possuem as suposições dadas em (3) e $Y_k(x_i)$ representa:

$$\mathbb{P}(Y_i = k/x), \text{ com } i \in \{1, \dots, n\}$$

Para o contexto da regressão logística multinomial, uma das categorias da variável resposta deve ser designada como categoria de referência e as demais serão comparadas com essa referência. Mais detalhes sobre o modelo multinomial podem ser obtidos respectivamente nos trabalhos de Ben-Akiva (1985), Ortúzar e Willumsen (1990) e Train (2002).

Testes de Significância

Nos SPSS 11.0 figuram três estatísticas R^2 . A estatística R^2 de Cox e Snell está baseada na função de verossimilhança e seu valor geralmente é inferior a um, sendo que o valor um indica um ajuste perfeito do modelo. A estatística R^2 de Nagelkerke é uma variação da estatística proposta por Cox e Snell buscando assegurar uma variação entre zero e um (Hosmer e Lemeshow, 2000). Para a estatística R^2 de McFadden valores em torno de 0,4 já indicam um bom ajuste do modelo.

Se o valor de R^2 for próximo ou igual a zero, significa que as variáveis independentes não são importantes na predição da variável dependente, no entanto, se o valor do R^2 for próximo ou igual a um, significa que as variáveis independentes descrevem perfeitamente o modelo em análise.

Outra importante estatística para avaliar a significância do logit multinomial é o teste de Wald. A estatística W uma alternativa comumente utilizada para testar a significância individual dos coeficientes de cada variável independente. O teste de Wald é usado para examinar restrições impostas aos coeficientes da regressão e calcula uma estatística que mede a eficiência das estimativas dos coeficientes da regressão original em satisfazer as restrições da hipótese nula.

Teste da Razão de Verossimilhança: $\{LR = -2\{L(0) - L(b^)\}$*

Tem distribuição de χ^2 com r graus de liberdade, onde r é o número de restrições lineares (parâmetros b). Testa a hipótese de nulidade de todos os parâmetros simultaneamente. Se o valor LR for maior que o valor $\chi^2_{(\alpha, r)}$ então rejeita-se a hipótese de nulidade de todos os parâmetros simultaneamente (Ben-Akiva e Lerman, 1989).

$$Var(\beta_k) = S_k^2 = -\varepsilon \left[\frac{\partial^2 L}{\partial \beta_k} \right] \varepsilon t = \frac{\beta_k}{S_k}$$

Teste da estatística r^2 (Pseudo Coeficiente de Determinação):

$$\rho^2 = 1 - \frac{L(\beta)}{L(0)}$$

Podem ocorrer possíveis comparações com a estatística R^2 (Coeficiente de Determinação) que varia de 0 até 1 (ajuste perfeito). A estatística r^2 tem seu valor teórico também limitado de 0 a 1, mas seu valor acima de 0,2 e próximo a 0,4 indica um ajuste considerado excelente para o modelo logit multinomial (Ortúzar e Willumsen, 1990).

r^2_{Aj} = valor de r^2 ajustado para r graus de liberdade, é definido como:

$$r^2_{Aj} = 1 - \frac{L(\beta) - r}{L(0)}$$

Teste t e Intervalo de Confiança

O teste t para a significância de um parâmetro b_k , para conjunto de escolha completo (Ben-Akiva e Lerman, 1989):

$$Var(\beta_k) = S_k^2 = -s \left[\frac{\partial^2 L}{\partial \beta_k} \right] e t = \frac{\beta_k}{S_k}$$

IC(b_k) : $b_k \pm taS_k$, onde ta é o valor crítico tabelado, na prática, toma-se $t = 2$ para $n \geq 60$ e indica, com aproximadamente 95% de probabilidade, que o b_k verdadeiro está no intervalo.

Testes da Independência das Alternativas Irrelevantes – IIA

Um propriedade muito importante relacionado ao modelo Logit Multinomial é a suposição da independência das alternativas irrelevantes. Seja C um conjunto de alternativas independentes e C_b um subconjunto de C , a probabilidade condicional de uma alternativa $i \in C_b$ em relação a C , devida a independência, é dada por:

$$P\left(\frac{i}{C_b} \in C\right) = P\left(\frac{i}{C_b}\right) \cdot P\left(\frac{C_b}{C}\right),$$

tomando $i, j \in C_b$, a relação entre as probabilidade condicionais de i e j demonstram a propriedade IIA, como:

$$\frac{P\left(\frac{i}{C_b}\right)}{P\left(\frac{j}{C_b}\right)} = \frac{P\left(\frac{i}{C}\right)}{P\left(\frac{j}{C}\right)}, i, j \in C_b \subseteq C \text{ (Ben-Akiva e Lerman, 1989).}$$

Mostra que, se as alternativas do conjunto C são independentes, então uma alternativa irrelevante pode ser retirada, restando o conjunto C_b , sem que a relação entre as probabilidades não se alteram. Desse conceito, surge a necessidade da ortogonalidade entre as alternativas, tornando-as independentes e não correlacionadas.

METODOLOGIA

A metodologia empregada neste trabalho consta de duas etapas principais. A primeira etapa trata da revisão dos aspectos teóricos sobre o modelo de quatro etapas empregado em

)

planejamento de transportes e dos fundamentos teóricos sobre os quais se assentam os modelos de escolha discreta utilizados para modelar a demanda na etapa de divisão modal. A segunda etapa trata dos aspectos práticos do trabalho, em que a partir dos conceitos estudados e da seleção de uma amostra proporcional do banco de dados da pesquisa origem-destino da cidade de Uberlândia (MG), realizou-se a calibração do modelo de regressão logística multinomial para a previsão da demanda na etapa de divisão modal.

A calibração do modelo de regressão logística para a etapa de divisão modal foi realizada utilizando-se o *software* SPSS – *Statistical Package for Social Sciences*, que através de métodos estatísticos e do algoritmo iterativo de máxima verossimilhança é capaz de estimar os parâmetros para as probabilidades de escolhas dos usuários de transportes. O ajuste por esse método fornece os coeficientes da função utilidade.

O Banco de Dados

A pesquisa OD foi realizada em 2002 pela Faculdade de Engenharia Civil da Universidade Federal de Uberlândia (FECIV) e a Secretaria Municipal de Trânsito e Transporte (SETTRAN). A amostra representativa foi constituída de 11.901 pessoas entrevistadas em 3.126 domicílios.

O banco de dados que será utilizado na calibração dos modelos consta de uma amostra estratificada proporcional por zona de tráfego com margem de erro 3% e nível de confiança de 95%, num total de 1451 indivíduos.

ESTUDO DE CASO E RESULTADOS

A área de estudo compreende a cidade Uberlândia (MG) que é uma cidade de porte médio, e está localizada no Triângulo Mineiro em Minas Gerais/Brasil. A cidade conheceu a partir do final da década de 1960 um crescente processo de urbanização, que se intensificou nas décadas seguintes. Hoje com uma população que segundo dados do Instituto Brasileiro de Geografia e Estatística (IBGE/2008) ultrapassam os seiscentos mil habitantes, e uma frota de aproximadamente 227.876 veículos (Departamento Nacional de Trânsito – Denatran, 2008), assiste-se ao crescimento dos problemas relacionados aos sistemas de transportes urbanos.

O crescimento contínuo e dinâmico da cidade promoveu uma diversificação das atividades (comerciais, econômicas e sociais) no perímetro urbano passando assim, a exigir da população deslocamentos cada vez mais longos entre a origem e destino das viagens. Além disso, os motivos de viagens tornaram-se cada vez mais variados, contribuindo significativamente para o aumento da demanda por transportes.

Devido às características de ocupação, crescimento e desenvolvimento econômico, a cidade foi dividida em setores. O sistema de setorização da área urbana de Uberlândia, criado pela Lei n° 6022 de 24 de maio de 1994, para fins de orientação e direcionamento, dividiu a cidade em cinco setores: Central, Norte, Sul, Leste e Oeste.

Os cinco setores da cidade foram divididos em sessenta e cinco zonas de tráfego conforme dados sócio-econômicos e das viagens realizadas entre o ponto de origem e destino, e os principais modos de transportes disponíveis à população são o ônibus, o táxi, vans, o automóvel, moto, bicicleta e caminhamento.

O zoneamento foi realizado com base na última pesquisa OD de 2002. A Figura 01 mostra a divisão da cidade de Uberlândia em sessenta e cinco zonas de tráfego.

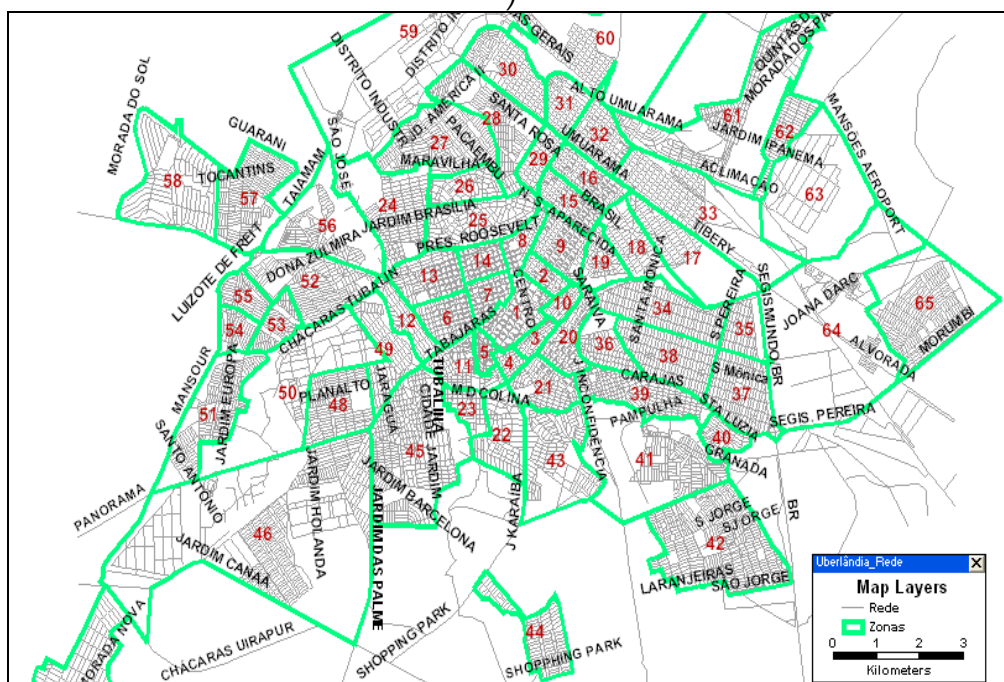


Figura 01 - Zonas de tráfego da cidade de Uberlândia (MG).

ANÁLISE EXPLORATÓRIA DOS DADOS

As informações sobre as viagens na cidade de Uberlândia foram obtidas a partir da pesquisa OD realizada na cidade em 2002. Os dados coletados nessa pesquisa dizem respeito às viagens realizadas e contém informações sobre a origem e o destino de viagem, o tempo de viagem, os modos e motivos de viagem, além de informações socioeconômicas sobre as pessoas que realizaram as viagens.

A análise da amostra selecionada mostra algumas características importantes das pessoas que realizaram as viagens, conforme pode-se observar na Tabela I.

Tabela I - Características gerais da população da amostra.

Situação Familiar:	Nível de Instrução:
Chefe: 452 (31,12%)	Analfabeto/primário incom: 257 (17,69%)
Conjugê: 254 (17,49%)	Primário completo/ginásio incom: 460 (31,68%)
Filho: 581 (40,06%)	Ginásio completo/colegial incom: 223 (15,35%)
Parente: 119 (8,19%)	Colegial completo/superior incom: 332 (22,83%)
Empregado: 10 (0,68%)	Superior completo: 150 (10,33%)
Outros: 01 (0,068%)	Omissos: 29 (1,99%)
Omissos: 34 (2,41%)	
Sexo:	Faixa Etária:
Masculino: 723 (49,79%)	Até 18 anos: 254 (17,50%)
Feminino: 711 (48,96%)	18 à 30 anos: 247 (17,02%)
Omissos: 17 (1,25%)	30 à 40 anos: 181 (12,49%)
	Mais de 40 anos: 346 (23,84%)
	Omissos: 423 (29,15%)

A Figura 02 mostra a porcentagem de viagens realizadas, segmentadas pelo motivo de viagem. Podemos observar que os motivos escola e trabalho representam a maior parcela das viagens num total de 45%, e o motivo residência refere-se às viagens com retorno para casa e representa 43%.

)

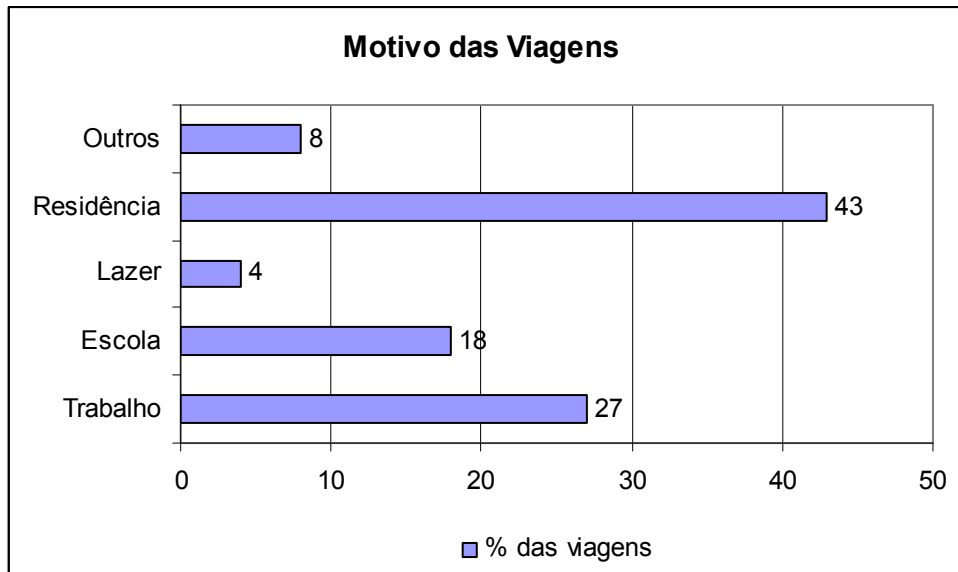


Figura 02 - Distribuição das viagens por motivo.

Com relação à posse de automóvel a Figura 03 mostra que 54% dos indivíduos da amostra possuem automóvel na residência, enquanto que, 39% não possuem. Essas porcentagens confirmam a estatística do IBGE, que mostra um índice de aproximadamente 2,83 veículos por pessoa na cidade.

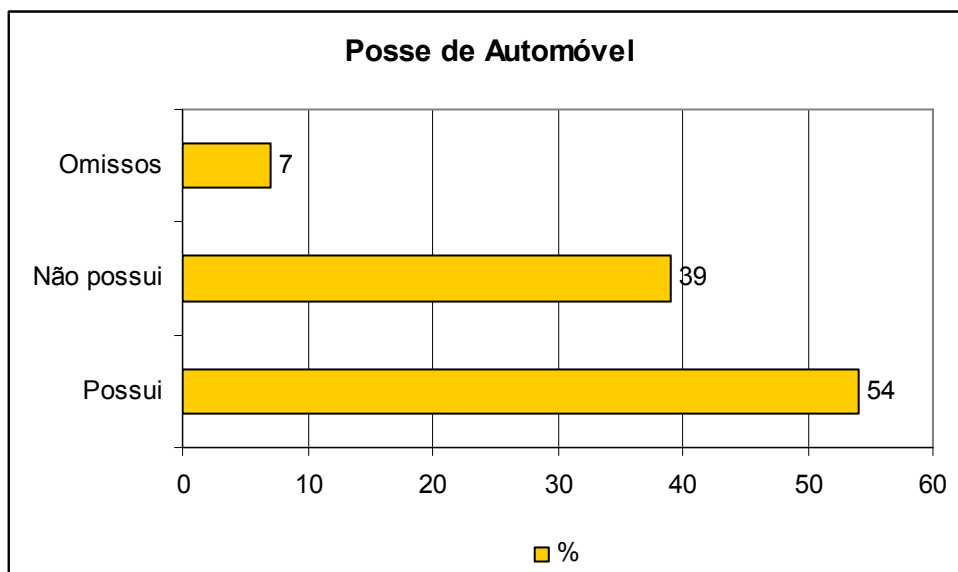


Figura 03 - Porcentagem de indivíduos que possuem ou não automóvel.

A posse ou não de automóvel significa para os serviços de transportes a subdivisão do mercado em viajantes cativos do transporte coletivo e viajantes com escolha. As características dos viajantes que influenciam suas decisões quanto à escolha da modalidade de transporte são aquelas que determinam a disponibilidade de automóveis aos viajantes e, conseqüentemente, o “*estatus*” cativo ou com escolha (Hutchinson, 1974).

A Figura 04 apresenta a distribuição das viagens de acordo com o modo de transporte utilizado. Verifica-se que 73,85% das viagens são realizadas pelos modos não motorizados (caminhamento e bicicleta) e transporte coletivo, considerados importantes para a sustentabilidade do tráfego nas cidades. Apesar de a cidade possuir a segunda frota de veículos do estado, as viagens pelo modo automóvel somam 21,01% das viagens. Já o

modo bicicleta foi utilizado em uma pequena parcela das viagens (6,06%), o que pode ser ocasionado pela falta de espaços adequados (ciclovias e bicicletários) para esse modo de transporte.

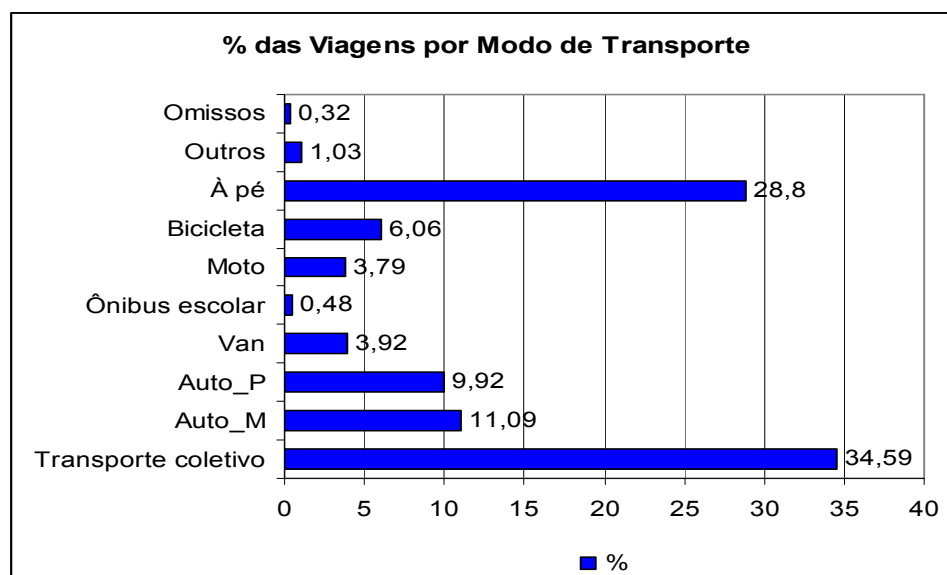


Figura 04 - Distribuição das viagens pelo modo de transporte utilizado.

RESULTADOS E CONSIDERAÇÕES

Na Tabela II são apresentadas as variáveis que foram selecionadas para a análise e calibração do modelo.

Tabela II - Variáveis de análise

Variável Dependente (Y)	Politômica
Modo de Transporte	1 = ônibus 2 = auto 3 = caminhada 0 = outros
Variáveis Independentes (X ₁ , X ₂ ,..., X _n)	Catégoricas
Motivo de viagem	1 = trabalho; 2 = escola; 3 = residência; 4 = outros
Posse de auto	0 = não possui; 1 = possui
Sexo	0 = feminino; 1 = masculino
Tipo de domicílio	0 = coletivo; 1 = particular
Situação familiar	0 = não chefe; 1 = chefe de família
Classe de atividade	1 = setor agrícola; 2 = construção civil; 3 = funcionário público; 4 = setor de serviços; 5 = autônomos
Renda mensal	1 = 0,0 – 2,0 SM; 2 = 2,1 – 4,0 SM; 3 = 4,1 – 6,0 SM; 4 = 6,1 – 8,0 SM; 5 = 8,1 – 10,0 SM; 6 = 10,1 – 20,0 SM; 7 = ≥ 20,1 SM
Nível de instrução	1 = analfabeto/priminc; 2 = primcom/gininc; 3 = gincom/colinc; 4 = colcom/supinc; 5 = supcomp
Condição de moradia	0 = não própria; 1 = própria
Classificação ABA/ABIPEME	1 = E; 2 = D; 3 = C2; 4 = C1; 5 = B2; 6 = B1; 7 = A2; 8 = A1
Idade	1 = até 18 anos; 2 = 18 a 30 anos; 3 = 30 a 40 anos; 4 = mais de 40 anos
Variáveis Independentes	Contínuas
Tempo de viagem; Número de pessoas no domicílio; Número de viagens realizadas; Tempo de residência; Número de famílias no domicílio.	

A variável dependente ou resposta refere-se à escolha do modo de transporte, enquanto as variáveis independentes referem-se às características relativas à viagem e ao usuário.

Modelo de Regressão Logística Multinomial

No SPSS 11.0 foi aplicado o modelo de regressão logística multinomial, utilizando o método *stepwise forward* para definir o modelo final que minimiza o número de variáveis e maximiza a precisão do modelo. A Tabela III mostra um resumo dos casos processados na análise e as variáveis que entraram no modelo final. A categoria de referência adotada para calibrar o modelo refere-se a categoria de menor valor, ou seja, $Y=0$, referindo-se à “outros modos de transporte”.

Tabela III - Casos processados

		N	Marginal Percentage
Modo	ônibus	409	39,8%
	auto	231	22,5%
	caminhamento	275	26,8%
	outros	113	11,0%
Auto	não possui	430	41,8%
	possui auto	598	58,2%
NivlInst	an/priminc	173	16,8%
	primcom/gininc	316	30,7%
	gincom/colinc	169	16,4%
	colcom/supinc	253	24,6%
SitFam	supcom	117	11,4%
	não chefe	662	64,4%
Idade1	chefe de família	366	35,6%
	até 18 anos	254	24,7%
ClasAtiv	18 a 30 anos	247	24,0%
	30 a 40 anos	181	17,6%
	mais de 40 anos	346	33,7%
	setor agrícola	9	,9%
Renda	construção civil	29	2,8%
	funcionário público	101	9,8%
	setor de serviços	291	28,3%
	autônomo	598	58,2%
	0,0 - 2,0 SM	590	57,4%
ClasABA	2,1 - 4,0 SM	233	22,7%
	4,1 - 6,0 SM	87	8,5%
	6,1 - 8,0 SM	41	4,0%
	8,1 - 10,0 SM	25	2,4%
	10,1 - 20,0 SM	44	4,3%
	>= 20,1 SM	8	,8%
	E	4	,4%
Valid	D	181	17,6%
	C2	249	24,2%
	C1	222	21,6%
	B2	303	29,5%
	B1	63	6,1%
	A2	6	,6%
Missing		1028	100,0%
Total		423	
Subpopulation		1451	
		858 ^a	

Aplicação de um Modelo de Escolha Discreta para Análise da Divisão Modal em Cidades de Porte Médio (SILVA, Thaís; MENDES, Flávia Bruno; FARIA, Carlos Alberto)

Na Tabela IV observa-se que a estatística de probabilidade – 2log diminui o que indica um bom ajuste do modelo final, além disso, o modelo final mostra-se significativo a um nível de significância de $\alpha = 0,050\%$.

Tabela IV - Informações de modelo

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	2601,333			
Final	1836,153	765,181	78	,000

Na Tabela V verifica-se o teste qui-quadrado, mais conhecido como, *Hosmer and Lemeshow's goodness of fit test*. Segundo Hosmer e Lemeshows (2000) se a estatística do teste for maior que o nível de significância α adotado, rejeita-se a hipótese de que não há diferença entre os valores observados e preditos implicando, assim, que o modelo descreve bem os dados no nível adotado.

Tabela V - *Hosmer and Lemeshow's goodness of fit test*

	Chi-Square	df	Sig.
Pearson	5815,193	2493	,000
Deviance	1789,170	2493	1,000

Os valores obtidos na tabela VI para os R^2 de Cox e Snell, e Nagelkerke são aceitáveis em torno de 56%, o R^2 de McFadden alcançou aproximadamente 30% o que indica a relação entre as variáveis X e Y. Esses valores são considerados aceitáveis devido ao tamanho da amostra (1451 observações), já que para amostras grandes os valores dessas estatísticas podem ser menores que 60% e 40%. Considerando-se então os valores desses testes pode-se dizer que o modelo possui um bom ajuste.

Tabela VI - Pseudo R - Quadrado

Cox and Snell	,525
Nagelkerke	,567
McFadden	,287

Após as interações, oito das dezesseis variáveis apresentaram significância para compor o modelo final. Pode-se verificar na tabela VII que as variáveis tempo de viagem, posse de auto, situação familiar, renda, nível de instrução, idade, classe de atividade e classe socioeconômica são significantes com $\alpha < 0,050\%$. Foram excluídas as variáveis: motivo de viagem, sexo, tipo de domicílio, condição de moradia, número de pessoas no domicílio, número de viagens realizadas, tempo de residência e número de famílias no domicílio.

Tabela VII - Teste de razão de verossimilhança (ajuste do modelo)

Effect	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	1836,153 ^a	,000	0	,
TTV	2178,888 ^b	342,736	3	,000
AUTO	1875,417 ^b	39,265	3	,000
NIVLINST	1865,347 ^b	29,195	12	,004
SITFAM	1866,425 ^b	30,272	3	,000
IDADE1	1857,807 ^b	21,654	9	,010
CLASATIV	1862,215 ^b	26,063	12	,011
RENDA	1887,223 ^b	51,070	18	,000
CLASABA	1873,407 ^b	37,254	18	,005

A Tabela VIII apresenta a classificação dos estimadores em corretos e incorretos. Segundo Figueira (2006) as tabelas de classificação são tabelas de ordem dois, para o caso da regressão logística dicotômica, e de ordem $2 \times (q + 1)$ para o caso da regressão logística politômica. As colunas apresentam os valores preditos da variável dependente e as linhas os valores observados. Em um modelo perfeito, todos os casos estariam na diagonal principal (destacado em vermelho na tabela) e a porcentagem de acertos seria de 100%.

Na tabela VIII são apresentados os valores de classificação do modelo. Assim, tem-se que nas previsões para a escolha:

1. do ônibus como modo de transporte, ou seja, $Y=1$, foram classificados 409 casos observados, houve 318 acertos (valores predito) e 91 erros, alcançando uma taxa de acerto de 77,8% considerada moderadamente boa;
2. para a escolha do automóvel ($Y=2$), verificou-se 231 valores observados, 134 acertos e 97 erros, ou seja, 58,0% dos casos foram corretamente classificados;
3. para o modo caminhada ($Y=3$), 275 valores observados, 181 corretamente classificados e 94 casos classificados incorretamente, com 65,8% dos casos corretamente classificados;
4. já para a escolha de outros modos de transporte ($Y=4$), 113 corresponde aos valores observados, 14 valores preditos corretamente classificados e 99 classificações incorretas, sendo que, somente 12,4% dos casos foram corretamente classificados.

O modelo apresentou, portanto, uma porcentagem de acerto geral de 62,9% que pode ser considerado razoável. O valor de 62,9% é obtido somando-se o total de acertos ($318+134+181+14=647$) e dividindo-se pelo total de observações 1028.

Tabela VIII - Tabela de classificação.

Observed	Predicted				Percent Correct
	ônibus	auto	caminha mento	outros	
ônibus	318	31	47	13	77,8%
auto	32	134	53	12	58,0%
caminhamento	54	34	181	6	65,8%
outros	39	31	29	14	12,4%
Overall Percentage	43,1%	22,4%	30,2%	4,4%	62,9%

O modelo calibrado considera os seguintes modos de transportes:

- Ônibus (incluindo o transporte coletivo, o ônibus escolar e a van);
- Automóvel (incluindo automóvel como motorista e passageiro);
- Caminhamento;
- Outros (incluindo a bicicleta, a moto, o caminhão).

Obedecendo-se a equação (2) para o logit:

$$g_{kn} = g_{kn}(x_n) = \beta_{k0}X_{n0} + \beta_{k1}X_{n1} + \dots + \beta_{kn}X_{nr} + \varepsilon_n \quad (2)$$

Tem-se que:

As equações para as utilidades dos modos de transporte, obtidas no modelo final com as variáveis socioeconômicas e de características da viagem são os seguintes:

$$U_{\text{ônibus}} = -2,357 + 0,066(TTV) - 0,522(\text{Auto}) - 0,864(\text{Situação Familiar}) + 0,222(\text{Idade})$$

$$U_{\text{auto}} = -2,504 - 0,028(TTV) + 1,500(\text{Auto}) + 0,220(\text{Nível Instrução}) + 0,374(\text{Renda})$$

$$U_{\text{caminhamento}} = 2,462 - 0,063(TTV) - 0,708(\text{Auto}) + 0,262(\text{Classe Atividade}) - 0,351(\text{Renda}) - 0,203(\text{Classe ABA})$$

$$U_{\text{Outros}} = -1,062 - 0,009(TTV) - 0,218(\text{Nível Instrução}) + 1,589(\text{Situação Familiar}) - 0,440(\text{Idade})$$

Com base nas utilidades de cada modo de transporte podem ser calculadas as probabilidades de escolha, conforme a expressão (4):

$$\mathbb{P}(Y_i = k/x) = \frac{\exp(g_{ki})}{1 + \exp(g_{ki})} + \varepsilon_i \quad (4)$$

As variáveis empregadas para análise do modelo mostraram-se significativas para explicar as escolhas dos indivíduos com relação ao modo de transporte numa cidade de porte médio. Com base na análise dos parâmetros, pode-se dizer que o modelo final apresentou um bom ajuste sendo plausível sua utilização para orientação do planejamento de transportes.

Pode-se, por exemplo, diante da necessidade de mudanças no transporte coletivo, avaliar as características individuais do usuário e assim as probabilidades de escolha desse modo em relação à outros.

REFERÊNCIAS BIBLIOGRÁFICAS

- Cavalcante, R. A. (2002). Estimativa das Penalidades Associadas com os Transbordos em Sistemas Integrados de Transporte Público. Dissertação de Mestrado em Engenharia de Transportes. COPPE/Universidade Federal do Rio de Janeiro, Rio de Janeiro; RJ. 151 p.
- Ben-Akiva, M. E e Lerman, S. R. (1985). Discrete Choice Analysis: theory and application to travel demand. p. 100-129. Acesso em 2010. <<http://books.google.com.br/books>>
- Brasil. Instituto Brasileiro de Geografia e Estatística. Acesso em: 2009. <<http://www.ibge.gov.br/home/estatistica/população>>
- Espino, R. S. (2003). Análisis y Predicción de La Demanda de Transporte de Pasajeros: Una Aplicación al Estudio de Dos Corredores de Transporte en Gran Canaria. Tese de Doutorado. Universidad de Las Palmas de Gran Canaria. 188 p.
- Faria, C. A.; Jardel, I. M.; Sorratini, J. A.; Macedo, M. H. (2004). Matriz Origem/Destino – Um Instrumento para a Otimização do Transporte Público Urbano por Ônibus. Métodos Computacionais em Engenharia. APMTAC, Lisboa, Portugal. 8 p.

)

- Ferronato, L. G. (2002). Potencial de Medidas de Gerenciamento da Demanda no Transporte Público por Ônibus. Dissertação de Mestrado em Engenharia de Produção. Universidade Federal do Rio Grande do Sul, Porto Alegre; RGS.119 p.
- Hosmer, D. W. e Lemeshow, S. (2000). Applied Logistic Regression. Inglaterra: John Wiley e Sons Ltda., p. 1-28.
- Hutchinson, B. G. (1979). Princípios de Planejamento dos Sistemas de Transporte Urbano. Rio de Janeiro: Guanabara Dois S. A. 416 p.
- Kotler, P. (1994). Marketing Management. 8a. ed. Prentice Hall Ed. Cap. 7: Analyzing Consumer Markets and Buyer Behavior.
- Manual Statistical Package for Social Sciences (2003). Versão 12.0 Copyright© de SPSS Inc.
- Novaes, A. G. (1986). Sistemas de Transportes: Análise da Demanda. Edgard Blücher Ltda. Vol. 1. São Paulo - SP, 151 p.
- Ortúzar, J. de D. e Román, C. (2003). El Problema de Modelación de Demanda Desde Una Perspectiva Desagregada: el caso del transporte. *Eure*, v. 29, n° 88, p. 149-171.
- Ortúzar, J. de D. e Willumsen, L. G. (1990). Modelling Transport. Inglaterra: John Wiley e Sons Ltda., 375 p.
- Train, K. E. Discrete Choice Methods With Simulation. Cambridge: University Press, 2002, 330 p.