# MODELOS DE DECISÃO BINÁRIOS:
## UMA REVISÃO[1]

**Francisco Alberto Pino**[2]

***RESUMO:*** *Apresenta-se didaticamente uma revisão sobre modelos de regressão logística, particularmente para modelos de decisão em economia agrícola, como em estudos de adoção tecnológica. Introduzem-se e discutem-se conceitos, estimação de parâmetros e avaliação do modelo.*

***Palavras-chave:*** *modelos de escolha, modelos de resposta binária, modelo de probabilidade, função de conexão logit, dados categóricos.*

## BINARY DECISION MODELS:
### A REVIEW

***ABSTRACT:*** *A review on logistic regression models is didactically presented, particularly for decision models in agricultural economics, such as those applied to technology adoption studies.Such models have also been used in biological, epidemiological (human, animal and vegetable), medical and engineering studies. Concepts, parameter estimation and model evaluation are introduced and discussed.*

***Key-words:*** *choice models, binary response models, probability model, logit link function, categorical data.*

***JEL Classification:*** *C25.*

[1]O autor agradece a colaboração dos Pesquisadores Científicos Mario Antonio Margarido e Vera Lúcia Ferraz dos Santos Francisco. Registrado no CCTC, ASP-17/2006.

[2]Engenheiro Agrônomo, Doutor, Pesquisador Científico do Instituto de Economia Agrícola (e-maill: pino@iea.sp.gov.br).

## 1 - INTRODUÇÃO

Os modelos *logit* constituem um tipo particular de modelo com variáveis dependentes limitadas. Tais modelos têm sido muito utilizados em estudos de biologia, epidemiologia, medicina, economia, engenharia e outros campos. Na verdade, os modelos binomiais *logit* e *probit* são os mais simples dos modelos envolvendo variáveis dependentes qualitativas. Uma dessas aplicações aparece em modelos de decisão em que é necessário escolher entre duas ou mais opções referentes à questão de interesse. Uma questão de particular interesse em economia agrícola que pode ser estudada com esses modelos é a da adoção de tecnologia[3]. Outros exemplos de aplicação aparecem nos estudos de ocorrência de doenças[4] em plantas ou em animais, bem como nos estudos de fatores que afetam a qualidade de uma bebida[5] ou alimento.

O objetivo deste artigo é apresentar uma rápida revisão de caráter didático a respeito dos modelos *logit*, remetendo o leitor interessado em detalhes a outras publicações. Inicia-se apresentando o contexto geral em que esses modelos se inserem. Enfatizam-se os problemas de economia agrícola e fazem-se referências ao SAS® (*Statistical Analysis Software*) como base computacional.

### 1.1 - Contagem e Mensuração

Dados estatísticos podem ser obtidos por dois processos: por contagem ou por mensuração (RODRIGUES, 1970; ISI, 2003). Na contagem, obtêm-se números inteiros e sem erros de aproximação, podendo ocorrer somente erros de sub ou supercontagem. Exemplo: o número de animais ou o número de plantas (pés) na exploração agropecuária são obtidos por contagem. Já na mensuração, estabelece-se uma relação unívoca entre um conjunto de números reais e um conjunto de magnitudes, isto é, obtêm-se números reais, podendo ocorrer também erros de aproximação no procedimento de medição. Há quatro níveis principais de mensuração:

a) Escala nominal: neste primeiro nível, existe um princípio de equivalência segundo o qual a magnitudes equivalentes devem corresponder a números iguais e a magnitudes não equivalentes devem ocorrer números diferentes. Entretanto, não faz sentido ordenar as magnitudes nem realizar operações aritméticas sobre elas. Exemplos: a relação de municípios do Estado, a lista telefônica, as placas de automóvel, os números de camisas de jogadores num esporte coletivo. Cada número da lista telefônica corresponde a uma linha, mas não faz sentido dizer que uma linha é maior do que a outra nem somar números de telefones;
b) Escala ordinal: neste segundo nível, além da equivalência, as magnitudes podem ser ordenadas. Aqui faz sentido ordenar as magnitudes, mas não realizar operações aritméticas sobre elas. Exemplos: uma pessoa pode ser considerada mais bela que outra (embora não se possa dizer o quão mais bela); um Pesquisador Científico de Nível VI tem nível mais alto do que um de Nível III (embora não se possa dizer que o de Nível VI vale duas vezes mais do que o de nível III); em uma pesquisa sobre preferências do consumidor, os artigos em análise podem ser ordenados;
c) Escala intervalar (ou de intervalos): neste terceiro nível, além da equivalência e da ordenação, fazem sentido os intervalos (ou diferenças) entre magnitudes. Portanto, faz sentido ordená-las e realizar operações de soma e subtração, embora não façam sentido as operações de multiplicação e de divisão. Exemplo: escalas de temperatura. Pode-se calcular a média entre duas temperaturas, por exemplo, mas não se pode dizer que 40ºC seja duas vezes mais quente que 20ºC; e
d) Escala racional: neste quarto nível, além da equivalência, da ordenação e dos intervalos, existe um zero absoluto, isto é, uma origem que coincide com a nulidade do atributo de interesse. As quatro operações aritméticas são aplicáveis nesse

[3]Ver, por exemplo, Vicente (1998), Vicente e Vosti (1995), Francisco e Pino (2004), Francisco; Pino; Vegro (2005).

[4]Ver, por exemplo, Athie Júnior et al. (2006).

[5]Ver, por exemplo, Pino e Vegro (2003).

tipo de escala. Exemplo: comprimento linear.

### 1.2 - Respostas Qualitativas e Quantitativas

Em um dado problema estatístico, a resposta pode ser vista de diversas formas, por exemplo, como a reação a um estímulo, a resposta a uma pergunta em um levantamento de dados, bem como o resultado da escolha em um processo de decisão. Respostas desse tipo podem ser qualitativas ou quantitativas.

As respostas quantitativas podem ser expressas por variáveis passíveis de mensuração, geralmente em uma escala racional e com valores reais (eventualmente, não negativos). Exemplos: a produção agropecuária expressa em unidade de massa, bem como o preço de um produto, são respostas quantitativas com valores reais positivos; variações de preços são respostas quantitativas com valores reais.

As respostas qualitativas[6] costumam ser expressas por variáveis categóricas[7], isto é, classificadas em categorias, em uma escala ordinal ou nominal, conforme elas possam ou não ser ordenadas. Exemplos de respostas qualitativas são: sexo, nível de escolaridade, intenção de um consumidor a respeito de comprar ou não determinado produto, intenção de um produtor agrícola a respeito de adotar ou não uma nova tecnologia. As respostas qualitativas podem ser representadas de forma quantitativa, atribuindo-se à resposta em cada categoria um valor numérico arbitrário, procedimento conhecido por codificação.

Seja $(\Omega, \mathfrak{A}, P)$ um espaço de probabilidade e seja $Y : \Omega \to \mathbb{R}$ uma variável aleatória nele definida e com valores reais.

Diz-se que *Y* é uma variável aleatória *contínua*-se $P[Y^{-1}(y) = 0]$ para todo real $y \in \mathbb{R}$. Por outro lado, dizemos que *Y* é uma variável aleatória *discreta*-se existe uma seqüência enumerável de pontos na reta, $y_i \in \mathbb{R}$, tais que $\sum_i P[Y^{-1}(y_i)] = 1$.

As respostas quantitativas são geralmente representadas por variáveis contínuas, enquanto as respostas qualitativas são representadas por variáveis discretas.

## 2 - MODELOS

Apresenta-se, a seguir, o contexto de modelagem em que se inserem os modelos *logit*.[8]

### Definição 2.1 - Variável de resposta e variáveis explicativas

Suponha-se que um dado fenômeno, ao qual se associa uma variável aleatória *Y*, esteja relacionado a um conjunto de *k* outros fenômenos, a cada um dos quais se associa uma variável aleatória $X_i$, com *i=1,2,...,k*, i.e., que a ocorrência de algum evento dado referente ao primeiro fenômeno dependa, de alguma forma, das ocorrências de eventos referentes a esse conjunto de outros fenômenos. Suponha-se, ainda, que seja possível encontrar alguma função *F* que relacione essas variáveis:

$$Y = F(X_1, ..., X_k) \qquad (2.1)$$

Então, a variável aleatória Y assim definida é chamada variável de resposta ou dependente, enquanto as demais são chamadas variáveis explicativas, ou explanatórias, ou independentes, ou simplesmente regressores.[9]

□

Em muitos casos, é razoável supor que *Y* seja função de uma combinação linear das variáveis independentes, dada por $\beta'\mathbf{X}$, onde $\beta$ é um vetor *k x 1* de parâmetros a serem estimados e $\mathbf{X}$ é o vetor *k x 1* de variáveis independentes:

[6]Usa-se na literatura a sigla QR, de *qualitative response.*

[7]A análise de variáveis categóricas, que representam respostas qualitativas, é conhecida por análise de dados categóricos.

[8]Ver Baltagi (2002); Gourieroux (2000); Greene (1997); Gujarati (1999); Maddala (1992); Pindyck e Rubinfeld (1998); Ramanathan (1998); Wooldridge (1999).

[9]O sinal □ é usado para indicar o final de uma definição.

$$\beta = \begin{pmatrix} \beta_1 \\ \cdots \\ \beta_k \end{pmatrix} \text{ e } \mathbf{X} = \begin{pmatrix} X_1 \\ \cdots \\ X_k \end{pmatrix}$$

Então,

$$Y = F(\beta'\mathbf{X}) + \varepsilon \qquad (2.2)$$

onde $\varepsilon$ é um termo de erro com média nula. Neste caso, a esperança de $Y$ dado X = x é dada por:

$$E\left[Y \middle| \mathbf{X} = \mathbf{x}\right] = F(\beta'\mathbf{x}) \qquad (2.3)$$

**Definição 2.2 - Modelo de Resposta Qualitativa**

É aquele em que a variável de resposta é discreta ao invés de contínua.

□

**Definição 2.3 - Modelo de Resposta Binária (ou quântica)** [10]

É o caso particular de modelo de resposta qualitativa no qual a variável de resposta é binária (ou *dummy*), i.e., ela assume somente dois valores, geralmente 0 para a não-ocorrência do fenômeno e 1 para a ocorrência.

□

**Definição 2.4 - Modelo com Variável Dependente Limitada**[11]

É o modelo em que a variável dependente não pode assumir todos os valores da reta real, em outras palavras, existem restrições, eventualmente substanciais, aos valores que a variável dependente pode assumir.

□

Muitas variáveis econômicas, inclusive em economia agrícola, apresentam tais limitações, como aquelas que podem assumir somente valores positivos (produção, preço, etc.). São exemplos:

a) os modelos com variáveis discretas, em particular binárias;
b) os modelos com variáveis de contagem, que assumem somente valores inteiros não negativos. Exemplo: os modelos de regressão de Poisson;
c) os modelos com variáveis aproximadamente contínuas, com valores positivos, mas que são nulas para uma fração não trivial da população. Exemplo: os modelos tobit (ver mais adiante);
d) os modelos com variáveis truncadas, que ocorrem quando dados amostrais são obtidos de um subconjunto de uma população maior de interesse. Exemplo: estudos sobre renda baseados em rendas acima e abaixo da linha de pobreza; e
e) os modelos com variáveis censuradas. Exemplo: estudos sobre renda baseados em rendas acima e abaixo da linha de pobreza, nos quais ao invés de a renda abaixo da linha de pobreza ser não observada, ela é relatada como se estivesse na linha de pobreza.

Os modelos com dados de contagem (resultantes da enumeração de ocorrências de certos eventos) também trabalham com resultados discretos, mas diferem dos modelos de escolha, porque nos primeiros existe uma ordenação natural da variável dependente e sua magnitude faz sentido (GREENE, 1995).

Embora os modelos LDV possam ser usados em séries temporais e em modelos para dados em painel (*panel data*), eles são usualmente utilizados em modelos para dados *cross-section* (*cross-sectional data*). A questão envolvida é que os métodos de regressão convencionais são inapropriados para esse tipo de modelo.

**Definição 2.5 - Modelo de Probabilidade**

Seja $Y$ uma variável aleatória do tipo Bernoulli ou dicotômica[12], i.e., que ela possa assumir somente dois valores, um para a ocorrência do evento e outro para a não-ocorrência do evento (por conveniência, costuma-se utilizar os valores 0 para a não-ocorrência e 1 para a ocorrência). Então,

[10]Em inglês, *quantal*.

[11]Usa-se na literatura a sigla LDV, *limited dependent variable*.

[12]A variável chama-se politômica se puder assumir três ou mais valores.

$$\begin{aligned} E[Y|\mathbf{X}=\mathbf{x}] &= 1.\Pr[Y=1|\mathbf{X}=\mathbf{x}]+0.\Pr[Y=0|\mathbf{X}=\mathbf{x}]= \\ &= \Pr[Y=1|\mathbf{X}=\mathbf{x}]= \\ &= F(\beta'\mathbf{x}) \end{aligned} \tag{2.4}$$

e este é chamado um modelo de probabilidade, sendo que o lado direito da equação deve ser restrito ao intervalo [0,1].

□

Nos modelos com resposta binária tem-se, então,

$$\begin{aligned} \Pr[Y=1|\mathbf{X}=\mathbf{x}] &= F(\beta'\mathbf{x}) \text{ e} \\ \Pr[Y=0|\mathbf{X}=\mathbf{x}] &= 1-F(\beta'\mathbf{x}) \end{aligned} \tag{2.5}$$

onde os parâmetros do vetor $\beta$ refletem o impacto de mudanças em $X$ sobre a probabilidade.

O problema, neste ponto, está em encontrar uma função $F$ adequada para definir o modelo em (2.4).

**Definição 2.6 - Modelo Linear de Probabilidade**[13]

A opção mais simples é o modelo linear de probabilidade, dado por:

$$F(\beta'\mathbf{x}) = \beta'\mathbf{x} \tag{2.6}$$

□

Entretanto, embora o modelo de regressão linear não imponha restrições sobre as variáveis independentes (que podem ser contínuas, somente maiores ou iguais a zero, dicotômicas, e outras), a variável dependente deve ser contínua, podendo assumir qualquer valor real. Pode-se mostrar que, se $Y$ for dicotômica, a estimação de mínimos quadrados ordinários do modelo linear de probabilidade será viesada (Aldrich e Nelson, 1984) e o termo de erro será heterocedástico, i.e., sua variância dependerá do vetor de parâmetros $\beta$ (Greene, 1997). Por isso, o modelo linear de probabilidade é utilizado somente como termo de comparação para outros modelos mais apropriados e modelos não-lineares de probabilidade devem ser especificados. Segundo Greene (1997), para que o modelo forneça previsões consistentes espera-se que:

$$\lim_{\beta'\mathbf{x}\to+\infty} \Pr[Y=1|\mathbf{X}=\mathbf{x}] = 1$$

e

$$\lim_{\beta'\mathbf{x}\to-\infty} \Pr[Y=1|\mathbf{X}=\mathbf{x}] = 0 \tag{2.7}$$

Essas condições serão satisfeitas se $F$ for qualquer função de distribuição de probabilidade contínua definida na linha reta.

**Definição 2.7 - Função de Ligação**[14]

Em modelos lineares generalizados, a função de ligação é aquela que especifica uma transformação não-linear utilizada para modelar respostas em que a variável dependente relaciona-se com as variáveis explicativas de forma não-linear.

□

Dessa forma, a função de ligação faz com que a distribuição dos valores previstos seja membro da família exponencial de distribuições de probabilidade (gama, Poisson, binomial, normal, logística, etc.).

**Definição 2.8 - Modelo *Probit***

A expressão modelo *probit*[15] serve para designar todos esses modelos não-lineares de probabilidade em que $F$ é uma distribuição de probabilidade.

□

**Definição 2.9 - Modelo *Normit***

É o nome dado ao modelo *probit* quando a distribuição normal é utilizada:

[13]Usa-se na literatura a sigla LPM, *linear probability model.*

[14]Em inglês, *link function.*

[15]Probit é abreviação de *probability unit*, sendo atribuída a C. R. Bliss, segundo Aldrich; Nelson (1984).

$$\Pr[Y=1|\mathbf{X}=\mathbf{x}]=\phi(\beta'\mathbf{x})=\int_{-\infty}^{\beta'\mathbf{x}}\phi(t)dt \quad (2.8)$$

onde $\phi$ representa a densidade da normal padrão:

$$\phi(\beta'\mathbf{x})=\frac{1}{\sqrt{2\pi}}e^{-\frac{(\beta'\mathbf{x})^2}{2}} \quad (2.9)$$

□

A distribuição normal foi a primeira a ser utilizada como função de ligação (ALDRICH e NELSON, 1984). Embora a expressão modelo *probit* designe de forma genérica todos esses modelos não-lineares de probabilidade em que *F* é uma distribuição de probabilidade, a maioria dos autores a tem utilizado somente para o caso da distribuição normal, em lugar de *normit*, atribuindo nomes analogamente construídos nos casos de outras distribuições, como *logit*[16] (distribuição logística), *gompit*[17] (distribuição de Gompertz), *burrit* (distribuição de Burr), *tobit*[18] (um caso de modelo censurado) e assim por diante.

**Definição 2.10 - Modelo *logit***

É o nome dado ao modelo *probit* quando a distribuição logística é utilizada como função de ligação:

$$p=\Pr[Y=1|\mathbf{X}=\mathbf{x}]=\Lambda(\beta'\mathbf{x})=\frac{e^{\beta'\mathbf{x}}}{1+e^{\beta'\mathbf{x}}}=\frac{1}{1+e^{-\beta'\mathbf{x}}} \quad (2.10)$$

□

A função densidade da distribuição logística é dada por:

$$\gamma(\beta'\mathbf{x})=\Lambda(\beta'\mathbf{x})\left[1-\Lambda(\beta'\mathbf{x})\right] \quad (2.11)$$

O modelo *logit* pode ser desenvolvido a partir de 2.6 como resultado da tentativa de remover as restrições em (2.7). Para remover o limite superior, tome-se a razão $p/(1-p)$, onde $p=\Pr[Y=1|\mathbf{X}=\mathbf{x}]$. Essa razão é sempre positiva, uma vez que 0<*p*<1, mas não é limitada superiormente, uma vez que:

$$\lim_{p\to 1}\frac{p}{1-p}=\infty \quad (2.12)$$

Tomando-se o logaritmo natural elimina-se também o limite inferior, uma vez que :

$$\beta'\mathbf{x}=\log\left(\frac{p}{1-p}\right)=\text{logit(p)} \quad (2.13)$$

pode ser qualquer número real. Tomando-se o antilogaritmo e resolvendo para *p* obtém-se (2.10), que é a função logística que caracteriza a distribuição de mesmo nome (ALDRICH e NELSON, 1984).

O modelo de regressão logística é uma extensão da análise de tabelas de múltipla entrada para a estrutura de análise de regressão, na qual se modelam os resultados de probabilidades binomiais[19]. Eles podem ser usados para modelar tanto variáveis de resposta verdadeiramente binomiais (que assumem os valores 0 e 1), quanto dados de proporções, isto é, variáveis que são contínuas no intervalo [0;1]. No primeiro caso, têm-se observações individuais de um resultado binomial. No segundo, os dados podem ter sido obtidos a partir de dados agrupados (unidades experimentais múltiplas observadas quanto à variável em questão) ou a partir de dados de painel (observações múltiplas sobre uma mesma unidade experimental ao longo do tempo)[20]. De modo geral, a diferença entre as regressões logística e linear é que na primeira a variável dependente é disposta em categorias e a resposta é expressa como

[16]Atribuída a Berkson (1944).

[17]É dada por $gompit(p)=\ln[-\ln(1-p)]$, sendo também chamada função de ligação log-log complementar.

[18]Semelhante a um modelo proposto por Tobin em 1958, segundo Vicente; Vosti (1995).

[19]Ver, por exemplo, Demaris (1992).

[20]Ver Georgia (2006).

a probabilidade de ocorrência, enquanto na segunda a variável dependente é contínua e a resposta é um valor numérico.

Ainda que outras opções tenham aparecido, tem-se consagrado em aplicações econométricas o uso da distribuição normal e da distribuição logística. Os modelos com essas duas distribuições são semelhantes, exceto nas caudas, que são mais pesadas na logística. Para valores intermediários de $\beta'\mathbf{x}$, como entre -1,2 e +1,2, as duas distribuições fornecem probabilidades semelhantes (Greene, 1997). A distribuição logística tende a fornecer probabilidades maiores do que a normal para *y=0* quando $\beta'\mathbf{x}$ é muito pequeno (e menores quando é muito grande). De modo geral, esperar-se-ão previsões diferentes entre os modelos quando: a) existirem poucas respostas ou poucas não-respostas; e b ) uma variável independente importante tiver grande variabilidade, principalmente quando (a) também for verdadeira. Também existem algumas conveniências matemáticas para escolher um ou outro modelo. Finalmente, o uso de distribuições com caudas mais pesadas tem sido preconizado em alguns modelos quando as variáveis distanciam-se da normalidade, como é o caso de muitas variáveis econômicas (Pino e Morettin, 1993). Entretanto, segundo Aldrich e Nelson (1997), a escolha entre essas duas alternativas permanece arbitrária.

## 3 - APLICAÇÃO: modelos de decisão

Uma teoria da escolha, usada em ciências do comportamento, foi desenvolvida por Luce e Suppes[21] e traduzida econometricamente por McFadden[22] (Aldrich e Nelson, 1984). Suponha-se que um agente (indivíduo ou grupo) deva tomar uma decisão, escolhendo entre duas alternativas. Uma vez que o agente tenha preferências a respeito dessas alternativas, ele irá escolher a sua favorita. Seja $W_{i1}$ o grau de preferência do i-ésimo agente pela primeira alternativa e $W_{i2}$ seu grau de preferência pela segunda alternativa. Logo, a suposição central é que o i-ésimo agente escolherá a primeira alternativa se $W_{i1}>W_{i2}$ e a segunda se $W_{i1}<W_{i2}$.

A fim de modelar esse processo, suponha-se que as preferências sejam funções de variáveis exógenas:

$$W_{ij} = F^*_j(\beta'\mathbf{X}_i) \tag{3.1}$$

*j=1,2*, e *i=1,2,...,n*, análogas à apresentada em 2.2. Considere-se a diferença entre as preferências[23] dada por:

$$\begin{aligned} Z_i &= W_{i1} - W_{i2} = \\ &= F^*_1(\beta'\mathbf{X}_i) - F^*_2(\beta'\mathbf{X}_i) = \\ &= F^*(\beta'\mathbf{X}_i) = \\ &= \beta'\mathbf{X}_i + \varepsilon \end{aligned} \tag{3.2}$$

Freqüentemente, os termos $W_{ij}$ não são discutidos, iniciando-se a discussão diretamente com a equação (3.2) como a descrição primitiva da preferência. Além disso, a forma usual da função $F^*$ é aquela apresentada em (3.2). Contudo, também $Z_i$ não é observável. Seja, então, $Y_i$ a escolha observada feita pelo i-ésimo agente e definida por:

$$Y_i = \begin{cases} 1, Z_i > 0 \\ 0, Z_i < 0 \end{cases} \tag{3.3}$$

i.e., $Y_i$ será igual a um se a primeira alternativa for escolhida e será igual a zero se a segunda for escolhida. Logo,

$$Y_i = \begin{cases} 1, F^*(\beta'\mathbf{X}_i) > 0 \\ 0, F^*(\beta'\mathbf{X}_i) < 0 \end{cases} \tag{3.4}$$

i.e., $Y_i$ é também função das variáveis exógenas, de tal modo que:

[21]Luce, R. D.; Suppes, P. Preference, utility, and subjective probability. In: _____; Busch, R.; Galanter, E. (eds.) **Handbook of mathematical psychology**. New York: John Wiley, 1965. v.3.

[22]McFadden, D. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.) **Frontiers in econometrics**. New York: Academic Press, 1973.

[23]Em termos econômicos, essa pode significar a diferença entre benefício e custo da decisão (Greene, 1997).

$$\begin{aligned}\Pr[Y_i = 1 | \mathbf{X} = x] &= \Pr[Z_i > 0 | \mathbf{X} = \mathbf{x}] = \\ &= \Pr[W_{i1} > W_{i2} | \mathbf{X} = \mathbf{x}] = \\ &= \Pr[\beta' \mathbf{X}_i + \varepsilon > 0 | \mathbf{X} = \mathbf{x}] = \\ &= \Pr[\varepsilon > -\beta' \mathbf{X}_i | \mathbf{X} = \mathbf{x}]\end{aligned} \quad (3.5)$$

e tem-se um modelo de probabilidade semelhante aos discutidos na seção 2. No caso de uma distribuição simétrica, como a normal e a logística, tem-se:

$$\Pr[Y=1|\mathbf{X}=\mathbf{x}]=\Pr[\varepsilon<\beta'\mathbf{X}_i|\mathbf{X}=\mathbf{x}]=F(\beta'\mathbf{X}_i) \quad (3.6)$$

### 3.1 - Modelos de Adoção de Tecnologia

A questão da adoção de uma nova tecnologia por parte de produtores rurais (e, por extensão, a adoção de novas técnicas administrativas, a decisão a respeito de fazer ou não parte de algum tipo de organização) pode ser vista como um problema de escolha entre duas alternativas. Analogamente ao exposto anteriormente, seja $W_{i1}$ o grau de preferência do i-ésimo produtor rural pela alternativa de adoção e $W_{i2}$ seu grau de preferência pela alternativa de não adoção. Da mesma forma, define-se a variável latente ou subjacente $Z_i$, como a diferença entre as preferências, que não pode ser observada empiricamente, que pode variar entre um mínimo e um máximo, mas ao atingir certo valor limite (definido aqui como zero) determina a mudança de estado, no caso, a adoção de uma dada tecnologia ou outra característica. Em outras palavras, ela representa a propensão dos produtores à adoção. A seguir, passa-se para a variável de resposta dicotômica $Y_i$, observável, que representa a decisão do i-ésimo produtor em adotar a nova tecnologia ou característica, valendo as equações (3.1) a (3.6). Da mesma forma, a probabilidade de adoção será função de variáveis exógenas que representam atributos do produtor rural e de sua unidade de produção agropecuária. Uma revisão conceitual sobre a questão da adoção tecnológica pode ser encontrada em Vicente (1998) e em Lima (1996).

É fácil ver que a variável $Y_i$, em (3.5) é semelhante à apresentada em (2.3) e, portanto, tem-se para a questão da adoção um modelo de probabilidade como o descrito na seção 2.

## 4 - ESTIMAÇÃO

O modelo de escolha binária, apresentado em (3.6), é usualmente estimado pelo método de máxima verossimilhança, sendo cada observação tratada como uma realização simples de uma distribuição de Bernoulli (GREENE, 1997). Para $n$ observações independentes tem-se a seguinte função de verossimilhança:

$$\Pr[Y_1 = y_1,\ldots,Y_n = y_n | X_1 = x_1,\ldots,X_n = x_n] = \prod_{i=1}^{n} [\Pr(Y_i = y_i)]^{y_i} [1 - \Pr(Y_i = y_i)]^{1-y_i} \quad (4.1)$$

Pode-se escrever:

$$L = \prod_{i=1}^{n} [F(\beta' x_i)]^{y_i} [1 - F(\beta' x_i)]^{1-y_i} \quad (4.2)$$

e tomando logaritmos:

$$logL=\sum_{i=1}^{n}[y_i logF(\beta x_i)+(1-y_i)log\,(1-F(\beta x_i))] \quad (4.3)$$

que será maximizada se:

$$\frac{\partial logL}{\partial\beta}=\sum_{i=1}^{n}\left[y_i\frac{f(\beta x_i)}{F(\beta x_i)}+(1-y_i)\frac{-f(\beta x_i)}{1-F(\beta x_i)}\right]x_i=0 \quad (4.4)$$

onde $f$ representa a derivada de $F$.

Nos modelos *probit* e *logit*, a equação (4.3) é não-linear nos parâmetros, o que significa que não se podem obter soluções algébricas, devendo estas ser obtidas por meio de algoritmos iterativos. No caso do modelo *logit*, tem-se, de (2.10), que:

$$F(\beta' x_i) = \frac{\exp(\beta' x_i)}{1+\exp(\beta' x_i)} \quad (4.5)$$

de onde:

$$f(\beta' x_i) = \frac{\exp(\beta' x_i)}{[1+\exp(\beta' x_i)]^2} \quad (4.6)$$

Substituindo-se 4.5 e 4.6 em 4.4, obtém-se:

$$\frac{\partial \log L}{\partial \beta} = \sum_{i=1}^{n} [y_i - F(\beta' x_i)] x_i = 0 \quad (4.7)$$

### 4.1 - Algoritmo de Newton-Raphson

Os algoritmos iterativos iniciam-se com um conjunto inicial de valores *b* para os parâmetros $\beta$. A seguir, determinam a direção correta e o tamanho da mudança em *b*, que aumentará a função objetivo (neste caso, o logaritmo da função de verossimilhança). Repete-se a operação até que mudanças em *b* não consigam aumentar mais a função objetivo segundo algum critério (como a mudança percentual na função objetivo, ou a mudança percentual nas estimativas dos parâmetros). Neste ponto, o processo iterativo termina e diz-se que o algoritmo convergiu.

O algoritmo mais empregado na estimação de modelos *probit* e *logit* tem sido o de Newton-Raphson (Aldrich e Nelson, 1984), que usa a seguinte fórmula para mudar os coeficientes:

$$b^{(i+1)} = b^{(i)} - \mathbf{H}^{-1} g \quad (4.8)$$

onde $b^{(i)}$ indica os valores das estimativas na i-ésima iteração, *g* é o gradiente, i.e., o vetor das derivadas primeiras da função objetivo e H é a matriz hessiana, i.e., das derivadas segundas da função objetivo, dada, no modelo *logit*, por:

$$\mathbf{H} = \frac{\partial^2 \log L}{\partial \beta \partial \beta'} = \sum_{i=1}^{n} F(\beta' x_i)[1 - F(\beta' x_i)] x_i x_i' \quad (4.9)$$

As derivadas segundas não envolvem a variável aleatória $y_i$ e por isso H=*E*(H). Além disso, H é sempre negativa definida, o que significa que o logaritmo da verossimilhança é globalmente côncavo (Greene, 1997). A função de verossimilhança do modelo *logit* é suficientemente suave para garantir a existência de solução única do problema de maximização, sendo que o algoritmo costuma convergir rapidamente mesmo partindo de valores iniciais pobres (Aldrich e Nelson, 1984).

### 4.2 - Matriz de Covariância

A matriz de covariância assintótica do estimador de máxima verossimilhança pode ser estimada usando a inversa de H calculada para os valores das estimativas:

$$AsyVar[\hat{\beta}] = -\mathbf{H}^{-1} \quad (4.10)$$

Outra opção é o estimador de Berndt, Hall, e Hausman, dado, no caso do modelo *logit*, por:

$$B = \sum_{i=1}^{n} [y_i - F(\beta' X_i)] x_i x_i' \quad (4.11)$$

Uma terceira opção baseia-se no valor esperado de H (Greene, 1997).

As probabilidades previstas $\hat{F} = F(\hat{\beta}' x)$, bem como os efeitos marginais estimados $f(\hat{\beta}' x).\beta = \hat{f}\hat{\beta}$ são funções não lineares das estimativas dos parâmetros. Para calcular erros padrões pode-se usar a aproximação linear. Para as probabilidades previstas tem-se a seguinte variância assintótica:

$$AsyVar[\hat{F}] = \left[\frac{\partial \hat{F}}{\partial \hat{\beta}}\right]' V \left[\frac{\partial \hat{F}}{\partial \hat{\beta}}\right] \quad (4.12)$$

Onde:

$$V = AsyVar[\hat{\beta}] \quad (4.13)$$

O vetor de derivadas é:

$$\left[\frac{\partial \hat{F}}{\partial \hat{\beta}} = \left[\frac{d\hat{F}}{d(x'\hat{\beta})}\right]\left[\frac{d(x'\hat{\beta})}{d\hat{\beta}}\right]\right] = \hat{f}x \quad (4.14)$$

de onde:

$$AsyVar[\hat{F}] = \hat{f}^2 x'Vx \quad (4.15)$$

que depende do particular vetor $x$ utilizado. Para o modelo *logit*:

$$\hat{f} = \hat{F}(1-\hat{F}) \quad (4.16)$$

**4.3 - Testes de Hipótese**

Para uma simples restrição pode-se utilizar o teste t de *Student* usual para testar hipóteses sobre os coeficientes. Para um conjunto de restrições $R\beta = q$ pode-se usar o teste de Wald, cuja estatística é dada por:

$$W = (R\hat{\beta} - q)'\{R(EstAsyVar[\hat{\beta}])R'\}^{-1}(R\hat{\beta} - q) \quad (4.17)$$

(GREENE, 1997).

Um teste de ajustamento, análogo ao teste F usado em análise de regressão para testar a hipótese conjunta de que todos os coeficientes, exceto o intercepto, são nulos, baseado no princípio da razão de verossimilhança é dado pela estatística:

$$c = -2\log(L_0 / L_1] = -2\{\log L_0 - \log L_1) \quad (4.18)$$

onde $L_1$ é o valor da função de verossimilhança calculado com os coeficientes estimados, enquanto $L_0$ é esse mesmo valor calculado para todos os coeficientes iguais a zero, exceto o intercepto. Essa estatística tem distribuição aproximada de qui-quadrado com *k-1* graus de liberdade.

Não existe, nos modelos *logit* e *probit*, uma estatística semelhante ao $R^2$ da análise de regressão, que possa ser interpretado como a proporção da variância da variável dependente que é "explicada" pelas variáveis exógenas. Por isso, algumas medidas de pseudo- $R^2$ têm sido sugeridas (ALDRICH e NELSON, 1984).

**4.4 - Efeitos Marginais e Elasticidade**

Em alguns problemas, os coeficientes estimados não permitem uma interpretação direta. É o que acontece com o modelo *logit* em aplicações econômicas. Neste caso, existe interesse em calcular efeitos marginais e elasticidades, que podem ser mais facilmente interpretados. Entretanto, nem todos os *softwares* calculam-nos diretamente, como é o caso do SAS®.

**Definição 4.1 - Efeito marginal**

Em modelos de probabilidade, o efeito marginal da variável aleatória $X_i$, com *i=1,2,...,k*, é a mudança na probabilidade prevista associada a mudanças nessa variável explanatória $X_i$. Seja o modelo de probabilidade definido em (2.4). Então, o efeito marginal pode ser definido por:

$$\frac{\partial E[Y|\mathbf{X}=\mathbf{x}]}{\partial \mathbf{x}} = \frac{\partial F(\beta'\mathbf{x})}{\partial \mathbf{x}} = \\ = f(\beta'\mathbf{x})\beta \quad (4.19)$$

onde $f(.)$ é a função densidade associada à função de distribuição $F(.)$.

□

Portanto, $f(\beta'\mathbf{x})$ é um fator de mudança de escala nos coeficientes estimados para obter os efeitos marginais e esse fator varia com os valores observados das variáveis explanatórias X. Assim, os efeitos marginais são funções não-lineares das estimativas dos parâmetros e dos níveis de todas as variáveis explanatórias X do modelo.

Por exemplo, para o modelo *probit,* o efeito marginal é dado por:

$$\frac{\partial E[Y|\mathbf{X}=x]}{\partial X_i} = \phi(\beta'\mathbf{X})\beta \quad (4.20)$$

Como $\phi > 0$ sempre, segue-se que, neste caso, a direção da mudança depende do sinal do vetor de parâmetros $\beta$.

Para o modelo *logit*, o efeito marginal é dado

por:

$$\frac{\partial E[Y|\mathbf{X}=x]}{\partial X_i} = \gamma(\beta'\mathbf{X})\beta \qquad (4.21)$$

## 5 - AVALIAÇÃO DOS RESULTADOS

Uma vez obtidas as estimativas dos parâmetros do modelo, é possível avaliar a qualidade dos resultados.

### 5.1 - Ajustamento do Modelo

Duas estatísticas de ajustamento são usualmente calculadas:
a) Critério da Informação de Akaike (AIC); e
b) Critério de Schwarz (SC).

Elas servem para comparar diferentes modelos e, quanto menores seus valores, melhor o ajustamento.

A hipótese $H_0 : \beta = 0$ pode ser testada pelo teste da razão de verossimilhança e pelo teste de Wald, ambos resultando em um valor de qui-quadrado (quanto maior, melhor o ajustamento).

Dois testes de aderência[24] ou ajustamento são usuais, ambos com distribuição de qui-quadrado: a) teste de aderência de Pearson; e b) teste de aderência de resíduos[25]. Entretanto, quando uma ou mais variáveis explicativas são contínuas, os dados costumam ser dispersos demais para o uso desses testes. Para essa situação, utiliza-se o teste de aderência de Hosmer-Lemeshow, que também trabalha com distribuição de qui-quadrado. Valores altos desses testes indicam falta de ajustamento do modelo.

### 5.2 - Razão de Probabilidade[26] (*Odds Ratio*)

Neste contexto, considera-se a probabilidade relativa[27] de um evento acontecer (ou não acontecer), ou de alguma coisa ser classificada em uma de duas categorias (DEMARIS, 1992). Por exemplo, se a probabilidade de os produtores agrícolas adotarem uma dada tecnologia for igual a 0,4 contra a probabilidade de 0,6 de não a adotarem, então, a probabilidade relativa de se adotar tal tecnologia será igual a:

$$\frac{p}{1-p} = \frac{0,4}{0,6} = \frac{2}{3} = 0,67$$

isto é, a probabilidade de se adotar a tecnologia é igual a dois terços da probabilidade de ela não ser adotada. Então, o *logit* dado na equação (2.13) é o logarítmo da probabilidade relativa[28].

**Definição 5.1 - Razão de probabilidade**

A razão de chance ou razão de probabilidade[29] significa qual a probabilidade de um evento ocorrer, se sob as mesmas condições ele não acontecer, sendo calculada por:

$$rc_i = e^{\hat{\beta}_i}, i = 1, 2, ..., k \qquad (5.1)$$

onde $\hat{\beta}_i$ representa a estimativa do parâmetro da i-ésima variável explicativa.

□

A razão de probabilidade é, simplesmente, a razão entre duas probabilidades (DEMARIS, 1992). Por exemplo, se a probabilidade de os produtores agrícolas cooperados adotarem uma dada tecnologia for igual a 0,7 (e, portanto, sua probabilidade relativa for 0,7/0,3=2,33), enquanto a probabilidade de os produtores agrícolas não-cooperados adotarem essa mesma tecnologia for igual a 0,4 (e, portanto, sua probabilidade relativa for 0,4/0,6=0,67), então, a razão de probabilidade de

[24]Em inglês, *goodness-of-fit test.*

[25]Em inglês, *deviance.*

[26]Encontra-se na literatura a tradução pelo galicismo "razão de chance".

[27]Em inglês, *odds.*

[28]Em inglês, *log odds.*

[29]Em inglês, *odds ratio.*

adotar a tecnologia, dado que o produtor é cooperado, em oposição a não ser cooperado, é igual a 2,33/0,67=3,48. Isso significa que os cooperados têm probabilidade de adotar a tecnologia que é 3,48 vezes maior do que a probabilidade dos não-cooperados.

### 5.3 - Respostas Observadas e Probabilidades Previstas

Suponha-se que se disponham de *N* observações, a saber, $Y_1, Y_2, ..., Y_N$, para calcular as estimativas. O número total de pares de observações (isto é, de observações emparelhadas) é dado pelas combinações de *N* duas a duas:

$$C_{N,2} = \binom{N}{2} = \frac{N(N-1)}{2} \qquad (5.2)$$

Seja $n_1$ o número de observações tais que $Y_j = 1$, para algum *j=1,2,...,N*, e seja $n_0$ o número de observações tais que $Y_j = 0$, para algum *j=1,2,...,N*. Então, $n_0 + n_1 = N$. Do número total de pares de observações dado em (5.2), tem-se que $t = n_0.n_1$ são pares de respostas diferentes entre si ($Y_i = 0$ e $Y_j = 1$, para $i \neq j$), enquanto as demais são de respostas iguais ($Y_i = 0$ e $Y_j = 0$, ou $Y_i = 1$ e $Y_j = 1$, para $i \neq j$).

**Definição 5.2 - Par Concordante e Par Discordante**

Um par de observações com respostas diferentes, por exemplo, $Y_i = 0$ e $Y_j = 1$, para $i \neq j$, é chamado:

*a)* concordante, se $P[Y_i = 1] < P[Y_j = 1]$;
*b) discordante*, se $P[Y_i = 1] > P[Y_j = 1]$; e (5.3)
*c) empatado*, se $P[Y_i = 1] = P[Y_j = 1]$.

□

Chamando-se $n_c$ o número de pares concordantes e de $n_d$ o número de pares discordantes, tem-se que o número de pares empatados é igual a $t - n_c - n_d$.

Com essa notação podem-se construir alguns índices de correlação de postos[30], medidas não paramétricas que servem para verificar a capacidade preditiva do modelo estimado e que são definidas a seguir.

**Definição 5.3 - Índice c**

É definido como a relação entre os pares concordantes mais metade dos pares empatados e o total de pares com respostas diferentes:

$$c = \frac{n_c + (t - n_c - n_d)/2}{t}, \text{ com } 0 \leq c \leq 1 \qquad (5.4)$$

□

**Definição 5.4 - Índice D de Sommer**

É definido como a relação entre os pares concordantes menos os pares discordantes e o total de pares com respostas diferentes:

$$D = \frac{n_c - n_d}{t}, \text{ com } -1 \leq D \leq +1 \qquad (5.5)$$

□

É possível verificar que $|D| \leq c$.

**Definição 5.5 - Índice Gama de Goodman-Kruskal**

É definido como a relação entre os pares concordantes menos os pares discordantes e a soma de pares concordantes e pares discordantes:

$$\Gamma = \frac{n_c - n_d}{n_c + n_d}, \text{ com } -1 \leq \Gamma \leq +1 \qquad (5.6)$$

□

É possível verificar que $|D| \leq |\Gamma|$.

[30]Também chamada correlação ordinal ou correlação por posições (em inglês, *rank correlation*).

**Definição 5.6 - Índice Tau-a de Kendall**

É definido como a relação entre os pares concordantes menos os pares discordantes e o número total de pares de observações:

$$\tau_a = \frac{n_c - n_d}{N(N-1)/2}, \text{com } -1 \le \tau_a \le +1 \qquad (5.7)$$

□

É possível verificar que $|\tau_a| \le |D|$.

Em todos esse índices, quanto maior seu valor, melhor a habilidade do modelo em prever as probabilidades da variável de resposta.

## 6 - CONSIDERAÇÕES FINAIS

Quanto aos aspectos computacionais, os modelos *logit*, com as estatísticas e testes referidos neste artigo, podem ser encontrados em diversos *softwares* estatísticos, em particular no SAS® (*Statistical Analysis Software*). Neste caso, utiliza-se o procedimento LOGISTIC do módulo SAS/STAT (SAS, 2002), cuja sintaxe em uma utilização básica é a seguinte:

**proc logistic data=**<nome do arquivo de dados>;
**model** <nome da variável dependente> **=**
<nomes das variáveis explicativas>**;**
<ou>
**model** <número de eventos favoráveis>**/**
<número de experimentos ou tentativas> **=**
nomes das variáveis explicativas>**;**
**run;**

Esse programa permite o procedimento *stepwise* e fornece como resultados, além da estimação dos parâmetros, a maioria das estatísticas de avaliação dos resultados aqui apresentadas. Seguem-se dois exemplos obtidos com a utilização desse programa.

### 6.1 - Exemplo 1

Francisco; Pino; Vegro (2005) estudaram a adoção de tecnologia da informação (TI) pelos cafeicultores do Estado de São Paulo, usando um modelo *logit* com doze variáveis independentes (Tabela 1). A ordem das variáveis na tabela é a mesma de sua entrada no modelo *stepwise*, sendo uma indicação de sua ordem de importância. O valor 18,747 na última coluna (estimativa da razão de probabilidade) indica que a probabilidade de um cafeicultor da região Alta Mogiana adotar TI é cerca de dezenove vezes maior do que nas regiões de plantio marginal ou em declínio, na região Baixa Mogiana é três vezes maior (2,721) e na região centro-sudoeste é duas vezes maior (1,720).

□

### 6.2 - Exemplo 2

Athie Júnior et al. (2006) estudaram a ocorrência de *huanglongbing*, uma doença que afeta as plantas cítricas, na região de Araraquara, Estado de São Paulo, usando um modelo *logit* com três variáveis independentes (Tabela 2). O valor 69,941 na última coluna (estimativa da razão de probabilidade) indica que a probabilidade de uma laranjeira apresentar sintomas da doença e conter a bactéria *Liberibacter americanum* é 70 vezes maior na presença do inseto vetor do que sem sua presença.

□

Em suma, os modelos *logit* e similares têm amplo espectro de aplicações e podem ser úteis em muitos campos científicos, incluindo a economia agrícola. Eles se prestam a modelar determinadas situações em que é necessário usar modelos não-lineares. Embora sua fundamentação matemática exija certa dose de conhecimento e atenção, sua aplicação é relativamente simples, desde que se aprenda a interpretar os resultados.

**Tabela 1** - Adoção de Tecnologia da Informação por Cafeicultores, Parâmetros Estimados do Modelo *Logit*, São Paulo, Novembro de 2004

| Variável | Estimativa do parâmetro | Qui-quadrado[1] | Estimativa da razão de probabilidade (*odds ratio*) |
|---|---|---|---|
| Intercepto | -8,1887 | 9679,3 | - |
| Densidade de cultivo (plantas/ha) | 0,0006 | 2554,8 | 1,001 |
| Área plantada com café (ha) | 0,0500 | 711,8 | 1,051 |
| Região Alta Mogiana | 2,9310 | 2233,7 | 18,747 |
| Nível superior de escolaridade do cafeicultor | 0,2415 | 172,9 | 1,273 |
| Experiência em cafeicultura (anos) | 0,0204 | 264,3 | 1,021 |
| Área de novos plantios/área total de café (%) | 2,0291 | 156,6 | 7,607 |
| Região Baixa Mogiana | 1,1211 | 335,0 | 3,068 |
| Produtor é membro de cooperativa, associação ou sindicato | 0,8593 | 329,9 | 2,362 |
| Produtor absenteísta | 1,2734 | 424,9 | 3,573 |
| Estoques de café mantidos na cooperativa/produção total | 0,6015 | 239,4 | 1,825 |
| Área de café/ área total cultivada | 0,9522 | 63,5 | 2,591 |
| Região Centro-sudoeste | 0,5423 | 54,5 | 1,720 |

[1]Todos significativos para p<0,0001.

Fonte: Adaptada de Francisco; Pino; Vegro (2005).

**Tabela 2** - Incidência de *Liberibacter americanum* em Plantas Cítrícolas Apresentando Sintomas de *Huanglongbing*, Modelo Logit, Região de Araraquara, Estado de São Paulo, Abril/Maio de 2005

| Variável | Estimativa do parâmetro | Qui-quadrado de Wald[1] | Estimativa da razão de probabilidade (*odds ratio*) |
|---|---|---|---|
| Análise de solo | -2,4146 | 607,3009 | 0,089 |
| Presença do vetor | 4,2477 | 37,7457 | 69,941 |
| Muda da região norte | -1,5556 | 14,4352 | 0,211 |

[1]Todos significativos ao nível de 1%.
Índice c: 0,448.
Índice D de Sommer: -0,103.
Índice Gama de Goodman-Kruskal: -0,199.
Índice Tau-a de Kendall: -0,013.
Fonte: Adaptada de Athie Júnior et al. (2006).

## LITERATURA CITADA


ALDRICH, J. H.; NELSON, F. D. **Linear probability, logit, and probit models**. London: Sage, 1984. 94 p.

ATHIE JÚNIOR, J. et al. **Incidência de huanglongbing (greening) em citrus na região de Araraquara.** Submetido para publicação. 2006.

BALTAGI, B. H. **Econometrics.** 3. ed. Berlin: Springer, 2002, 401 p.

BERKSON, J. Application of the logistic function to bioassay. **Journal of the American Statistical Association**, v. 39, p. 357-365, 1944.

DEMARIS, A. **Logit modeling:** practical applications. Newbury Park: Sage, 1992. 86 p.

FRANCISCO, V. L. F. S.; PINO, F. A. Fatores que afetam o uso da Internet no meio rural paulista. **Agricultura em São Paulo,** São Paulo, v. 51, n. 2, p. 27-36, 2004.

_____; _____; VEGRO, C. L. R. Information technology in coffee farms. _____, São Paulo, v. 52, t. 1, p. 77-82, jan./jun. 2005.

GEORGIA Institute of Technology. Discrete dependent variable models. In: **Scientific approaches to transportation research** (NCHRP 20-45). Online documentation, v. 2, cap. 5. Disponível em: <http://onlinepubs.trb.org/onlinepubs/nchrp/cd-22/start.htm>. Acesso em: 20 jun. 2006.

GOURIEROUX, C. **Econometrics of qualitative dependent variables.** Cambridge: Cambridge University Press, 2000. 372 p.

GREENE, W. H. Probit, logit, and other models for binary choice. In: **LIMDEP user's manual**: version 7. Castle Hill, Australia: Econometric Software, 1995. p. 411-485. Disponível em: <http://www.et.bs.ehu.es/docpub/manuales/limdep>. Acesso em: 30 set. 2003.

GREENE, W. H. **Econometric analysis.** 3. ed. Upper Saddle River, NJ: Prentice Hall, 1997, 1075 p.

GUJARATI, D. **Econometrics.** 2. ed. Boston: McGraw-Hill, 1999. 534 p.

ISI. **Glossary of statistical terms**. Disponível em: <http://europa.eu.int/comm/eurostat/research/index.htm?http://europa.eu.int/en/comm/eurostat/research/isi/concepts/>. Acesso em: 1 out. 2003.

LIMA, R. C. Modelos de respostas binárias: especificação, estimação e inferência. **Agricultura em São Paulo**, São Paulo, v. 43, t. 2, p. 19-25, 1996.

MADDALA, G. S. **Introduction to econometrics.** 2. ed. Englewood Cliffs, NJ: Prentice Hall, 1992. 631 p.

PINDYCK, R. S.; RUBINFELD, D. L. **Econometric models and economic forecasts.** 4. ed. Boston: McGraw-Hill, 1998. 634 p.

PINO, F. A.; MORETTIN, P. A. The consistency of the $L_1$-norm estimates in ARMA models. **Communications in Statistics, Theory and Methods**, v. 22, n. 8, p. 2185-2206, 1993.

PINO F. A.; VEGRO, C. L. R. Qualidade do café expresso em condições de campo. **Brazilian Journal of Food Technology**, v. 6, n. 2, p. 345-350, jul./dez. 2003.

RAMANATHAN, R. **Introductory econometrics**: with applications. 4. ed. Fort Worth: The Dryden Press, 1998. 664 p.

RODRIGUES, M. S. **Dicionário brasileiro de estatística.** Rio de Janeiro: Fundação IBGE, 1970. 350 p.

SAS INSTITUTE. **SAS OnlineDoc :** version eight. Disponível em: <www.rz.tu-clausthal.de/sashtml>. Acesso em: 27 mar. 2002.

VICENTE, J. R. Determinantes da adoção de tecnologia na agricultura paulista. **Estudos Econômicos**, São Paulo, v. 28, n. 3, p. 421-451, jul./set. 1998.

_____; VOSTI, S. A. Um teste de dados em nível de imóvel rural do levantamento objetivo IEA/CATI para estudos de adoção de tecnologia. **Agricultura em São Paulo**, São Paulo, v. 42, t. 2, p. 129-148, 1995.

WOOLDRIDGE, J. M. **Introductory econometrics**: a modern approach. USA: South-Western College Publishing, 1999. 824 p.

Recebido em 03/07/06. Liberado para publicação em 01/11/06.