

# O desafio da exclusão digital: uma análise multivariada com base na Pnad 2005

## Rommel Cysne

Engenheiro civil pela Universidade Federal do Rio de Janeiro (1985), analista de sistemas pela PUC (1988), mestre em estudos populacionais e pesquisas sociais pela Ence/IBGE (2007). Atualmente é analista de planejamento, gestão e infraestrutura em informações geográficas e estatísticas do Instituto Brasileiro de Geografia e Estatística (IBGE).

E-mail: rommelcysne@ibge.gov.br

## José Eustáquio Diniz Alves

Graduado em ciências sociais pela Universidade Federal de Minas Gerais (UFMG, 1980), mestre em economia (1983) e doutor em demografia pelo Cedeplar-UFMG (1994) e pós-doutorado pelo Nepo/Unicamp.

E-mail: jose.diniz@ibge.gov.br

## Sérgio da Costa Côrtes

Especialista pela Faculdades Católicas (1984), mestre em informática pela Pontifícia Universidade Católica do Rio de Janeiro (1999) e doutor em Informática pela Faculdades Católicas (2004). Professor bacharelado da Faculdades Católicas. Atualmente ocupa o cargo de diretor executivo do IBGE.

E-mail: sergio.cortes@ibge.gov.br

## Resumo

O propósito deste artigo é estudar o perfil da população brasileira que tem acesso à Internet. Por meio de uma análise multivariada, verificou-se que em geral os determinantes da exclusão social estão presentes na exclusão digital. O acesso à Internet tem uma relação diretamente proporcional aos níveis de educação e renda. As menores percentagens de presença na rede são encontradas entre as pessoas que estão fora da escola e do mercado de trabalho. Em termos de idade, são os jovens entre 15 e 24 anos que mais utilizam a Internet, especialmente aqueles das escolas privadas. Embora o acesso à Internet tenha crescido mais de 150% no Brasil, entre os anos 2000 e 2005, o país ainda possui baixa parcela da população que usufrui os benefícios da Web. É necessário que se formulem políticas públicas capazes de acelerar o processo de inclusão digital, visando a avançar rumo a uma plena cidadania em todos os seus aspectos e para todos os brasileiros.

## Palavras-chave

Exclusão digital. Cidadania. Acesso à Internet. Modelo de regressão logística.

## The challenge of digital divide: a multivariate analysis based on the Pnad 2005

## Abstract

*The purpose of this paper is to study the profile of the Brazilian population that has Internet access. By means of a multivariate analysis, it was verified that in general the responsible variables by the social exclusion are present in the digital exclusion (digital divide). The access to the Internet has a relationship directly proportional to the levels of education and income tax. The lower percentages of presence on the Web are from people that are out of school and of the labor market. In terms of ages, the young people, between 15 and 24 years old, are those which more frequently use the Internet, specially those of private schools. In spite of the fact that the access to the Internet have grown up more than 150% in Brazil, between the years 2000 and 2005, the country still have a small proportion of the population that use the benefits of the Web. It is necessary to have public politics capable to increase the process of digital inclusion, seeking to move forward to a full citizenship, in all his aspects, and for all Brazilians.*

## Keywords

*Digital divide. Citizenship. Internet access. Logistic regression model.*

## INTRODUÇÃO

Com o advento do fenômeno da globalização, as sociedades contemporâneas vêm passando, nas últimas décadas, por transformações econômicas, políticas e culturais profundas, impulsionadas pelo desenvolvimento de sistemas de comunicação e pelo surgimento da sociedade em rede. Esse fenômeno da globalização, potencializado pela Internet, trouxe em seu bojo, além de maior acesso à informação, a possibilidade de maior debate intelectual e de troca de idéias e pensamentos, criando um ambiente mais propício à democracia.

As novas tecnologias revolucionaram a maneira como lidamos com a informação e o conhecimento, principalmente no que diz respeito a dois de seus aspectos: o espaço e o tempo. Nenhuma outra tecnologia, na história, proveu o mundo de tanta informação e de forma tão acessível. Em termos de sistema de informação, a Internet provê acesso imediato à quantidade gigantesca de informações científicas, culturais, artísticas, em tempo real, de forma direta para o usuário, abrindo possibilidades antes inimagináveis. Sites que oferecem conteúdo para o público jovem vêm se proliferando no ciberespaço, como comunidades, páginas para *downloads* de músicas e de vídeos, *fotologs* e páginas com conteúdo relativo à educação. A produção, o acesso e a disseminação da informação em escala global adquiriram importância crescente nas sociedades atuais, de tal forma que a Internet se transformou em ferramenta estratégica para o desenvolvimento dos mais diversos setores da vida humana.

Na visão de Castells (2003), a Internet é o tecido de nossas vidas, tendo a capacidade de distribuir a força da informação por todo o domínio da atividade humana. Este poderoso mecanismo constitui a espinha dorsal das sociedades contemporâneas e da nova economia mundial, e passou a ser a base tecnológica para a forma organizacional da era da informação: a rede.

Castells teve o grande mérito de abordar o tema da mudança estrutural das sociedades contemporâneas sob a ótica das dimensões das redes e a centralidade da informação. Dois dos seus principais tópicos são a

revolução das tecnologias da informação e a reestruturação do capitalismo à escala global, aspectos que propiciaram a emergência de novo modelo nas economias ocidentais: o modo de desenvolvimento informacional. Distanciando-se um pouco do euforismo da visão de Castells no tratamento teórico da era da informação, cabe mencionar que uma das mais proveitosas análises sobre os processos de ajustamento estrutural da sociedade em suas várias dimensões ao longo do tempo tem sido trazida pela economia com base e inspiração nos princípios de Schumpeter e Kondratieff, que tem sido interpretada por diversos autores como, por exemplo, Christopher Freeman. (ALVES, 2004).

Segundo Freeman (1996), estamos vivenciando um novo paradigma técnico-econômico, em virtude da emergência da quinta onda longa de desenvolvimento, centrada na utilização intensiva das tecnologias da informação e da comunicação e na topologia de rede como modelo organizacional da produção. Freeman não é muito otimista quanto ao potencial democrático e civilizatório da Internet.

Já para Castells (2003), a rede é um ambiente onde a informação flui em ritmo bastante acelerado e sob pressão da necessidade de ampliação de novos conhecimentos. Reforçadas pela expansão dos meios de comunicação, as sociedades estão caminhando para novo estágio que constitui um marco histórico, a sociedade da informação, no qual a tecnologia e a informação passam a ser seu paradigma. O advento da era do conhecimento, ou da informação, criou nova realidade complexa, uma vez que a modernização provocou assimetrias no mercado de trabalho e na vida cotidiana das pessoas, decorrentes de maior seletividade, acarretando, inclusive, formas de segregações.

O fosso digital definido, de forma simplificada, pela diferença entre os que têm acesso às informações pela Internet e aqueles que não têm, principalmente as parcelas pobres da população residentes nas regiões menos desenvolvidas do nosso país, constitui um desafio a ser enfrentado pelas políticas públicas brasileiras.

A inclusão digital e o acesso às tecnologias da informação e comunicação, em particular a Internet, podem permitir maior compromisso democrático entre governo e cidadão, ampliando a cidadania e motivando as comunidades e os indivíduos a exigirem seus direitos básicos. O Brasil não pode menosprezar as políticas públicas voltadas ao processo de inclusão digital. Contudo, não será possível evoluir nessa área sem o avanço da inclusão social. Por exemplo, ter garantia a boas condições de educação e renda é condição essencial para a expansão da inclusão digital.

Neste artigo serão explorados os condicionantes da inclusão digital no Brasil por meio de uma análise unitemporal *cross-section*, tomando-se por base a Pesquisa Nacional por Amostra de Domicílios de 2005 (Pnad, 2005). O objetivo é verificar as chances de acesso à Internet a partir do ajuste de um modelo de regressão logística. Pretende-se saber, controlando diversas variáveis socioeconômicas, o que acontece com a probabilidade de um indivíduo de determinado nível educacional, inserido em certa faixa de renda, e que pertença à determinada região geográfica, ter acesso à Internet. Outro objetivo, ligado a este, é identificar as variáveis que possuem maior poder de explicação, isto é, que variáveis mais contribuem para explicar esse modelo logístico.

O conceito de acesso à Internet adotado neste trabalho – o mesmo usado pelo Instituto Brasileiro de Geografia e Estatística (IBGE) na elaboração do Suplemento da Pesquisa Nacional por Amostra de Domicílios (Pnad, 2005), refere-se à utilização da Internet pela própria pessoa, por meio de computador de mesa ou portátil (*laptop*, *notebook*, *palmtop*, *pocket pc*, *handheld*) em algum local (domicílio, local de trabalho, escola, centro de acesso gratuito ou pago, domicílio de outras pessoas ou qualquer outro local), pelo menos uma vez, nos últimos três meses anteriores à entrevista.

## ANÁLISE MULTIVARIADA DO ACESSO À INTERNET

A análise multivariada é necessária quando estamos tratando com mais de uma variável explicativa (independente) e queremos investigar os efeitos simultâneos destas variáveis sobre a variável resposta

(dependente). A metodologia adotada no sentido de captar o papel desempenhado pelos vários atributos da Pnad relacionados à divisão digital está baseada no modelo de regressão logística (MRL), uma vez que podemos considerar o acesso à Internet como um evento dicotômico.

Para obter um modelo de regressão logística que se ajustasse adequadamente aos microdados da Pnad 2005, foi utilizado o *software* estatístico *Statistical Analysis System* (SAS) (versão 9.1), tendo que ser levada em conta a incorporação da estrutura do plano amostral, uma vez que a Pnad reúne todos os aspectos que definem um plano (desenho) amostral complexo: estratificação das unidades de amostragem, conglomeração (seleção da amostra em vários estágios, com unidades compostas de amostragem), probabilidades desiguais de seleção em um ou mais estágios, e ajustes dos pesos amostrais para calibração com totais populacionais conhecidos.

O modelo de regressão logística pode ser escrito em termos de logaritmo das chances, chamado de logit. No modelo univariado, a função de ligação logit é definida como:

$$g(x) = \log_e \left( \frac{\text{Prob (evento)}}{\text{Prob (não evento)}} \right) = \log_e \left( \frac{\pi(x)}{1-\pi(x)} \right)$$

Fazendo a transformação, temos que:

$$g(x) = \beta_0 + \beta_1 x$$

Essa é a chamada transformação logit da probabilidade  $\pi$ . A razão  $\frac{\pi}{1-\pi}$  o logit é chamada de odds (chance ou vantagem).

Para o modelo de regressão logística multivariado, o logaritmo das chances é expresso por:

$$\log_e \left( \frac{\text{Prob (evento)}}{\text{Prob (não evento)}} \right) = \log_e \left( \frac{\pi(x)}{1-\pi(x)} \right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p, \text{ onde } x_i$$

variáveis de interesse (independentes) e  $\beta$  é o vetor de parâmetros desconhecidos.

Sendo assim, no modelo de regressão logística, o logaritmo das chances é linearmente relacionado com as variáveis explicativas, e tem-se, portanto, um modelo linear na escala logit, ou seja, o modelo de regressão logística relaciona a vantagem a favor da ocorrência do evento com variáveis explicativas, através da função de ligação *logit*. No nosso caso, o modelo logístico define a probabilidade de uma pessoa acessar a Internet como função linear de um conjunto de variáveis explicativas estudadas.

Os parâmetros que devemos estimar,  $\beta_0, \beta_1, \beta_2 \dots \beta_p$ , são calculados usando-se o estimador de máxima verossimilhança. Na avaliação desses parâmetros, devemos estar atentos primeiramente ao seu sinal, que nos indicará uma associação positiva ou negativa com relação à categoria-base, além de sua ordem de grandeza. Se  $\beta_i$  é positivo, o fator  $\text{Exp}(\beta_i)$  será maior que 1, significando que as chances (odds) estão aumentando, e em caso contrário, se  $\beta_i$  é negativo, o fator  $\text{Exp}(\beta_i)$  será menor que 1, indicando que as chances (odds) estão diminuindo.

A interpretação do modelo logístico possui a facilidade de permitir analisar as chances estimadas da ocorrência de um evento de interesse em relação à sua não-ocorrência. A vantagem ou chance é definida pela probabilidade de ocorrência do evento sobre a probabilidade de não-ocorrência do mesmo (ou seja, no nosso caso, a probabilidade de a pessoa ter acessado a Internet sobre a probabilidade de não tê-la acessado)<sup>1</sup>.

Inicialmente foram utilizados modelos univariados de acesso à Internet para a determinação do poder de explicação de cada variável, ou melhor, desejou-se saber que proporção das variações registradas na variável dependente (internet) conseguia ser

explicada pela variável independente considerada em cada modelo univariado. Usando modelos univariados devemos supor que, em cada um deles, não existem outras variáveis interferindo no acesso à Internet (além da variável explicativa que está sendo considerada no modelo). No início de nossa análise, esta suposição foi considerada como verdadeira somente para detectar o poder de explicação de cada variável independente (explicativa), mesmo sabendo que, de fato, isso não ocorre.

Todas as variáveis foram consideradas estatisticamente significantes nos modelos univariados (ao nível de significância de 0,01%), constatado através dos relatórios de saída do *Statistical Analysis System – SAS (surveylogistic procedure)*. O poder de explicação de cada variável considerada isoladamente pode ser fornecido mediante simples estatística (*pseudo R<sup>2</sup>*) baseada na comparação do *log-likelihood* do modelo apenas com a constante ( $LL_0$ ) e do modelo de interesse – neste caso, cada um dos modelos univariados – ( $LL_m$ ) usando a fórmula:

$$pseudo \ R^2 = \frac{LL_0 - LL_m}{LL_0} = \frac{(-2\text{Log}L_0) - (-2\text{Log}L_m)}{(-2\text{Log}L_0)}$$

A tabela 1, a seguir, que faz uso desta fórmula, mostra por intermédio de modelos univariados o poder de explicação de cada variável em ordem decrescente. Cabe mencionar que, neste caso, como estamos considerando, de cada vez, apenas uma variável independente para explicar as variações registradas na variável dependente (internet), estamos captando o “efeito bruto” desta variável independente que compõe o modelo univariado.

Podemos constatar que classes de rendimento domiciliar *per capita* é a variável que possui o maior nível de explicação, conseguindo explicar sozinha 23,116% (0,23116) das variações registradas na variável dependente, considerando que não existem outras variáveis interferindo no acesso à Internet, estando, inclusive, à frente da variável referente à existência de microcomputador no domicílio, que consegue explicar sozinha 23,072% das variações na variável dependente. Outra variável que também possui alto

<sup>1</sup> Cabe ainda acrescentar alguns esclarecimentos relativos à execução dos procedimentos de regressão logística no pacote estatístico SAS - *Statistical Analysis System* (versão 9.1). Visando incorporar a estrutura do plano amostral da Pnad, foi utilizada a *procedure surveylogistic*, adotando-se, para os *statements strata*, a variável de estratificação da amostra da pnad v4617, *cluster*, a variável relativa à unidade primária de amostragem da pnad v4618, e *weight*, a variável que representa o peso da pessoa v4729.

poder de explicação é a referente aos grupos de anos de estudo, que explica sozinha 19,752% das variações registradas na variável dependente *acesso à internet*.

Como todas as variáveis explicativas foram consideradas estatisticamente significantes nos modelos univariados, e como cada variável está correlacionada às demais, ao incluir todas essas variáveis explicativas, ao mesmo tempo, no modelo de acesso à Internet, estaremos medindo o “efeito líquido” de cada variável, que no conjunto fornecerá o poder explicativo do modelo logístico multivariado, conforme veremos a seguir.

**PODER EXPLICATIVO DO MODELO LOGÍSTICO**

O modelo multivariado de acesso à Internet proposto pode ser escrito da seguinte forma:

$$\log_e \left( \frac{P_{ijkhlmnopqrstvwxyz}}{1 - P_{ijkhlmnopqrstvwxyz}} \right) = \beta_0 + \beta_j^{\text{sexo}} x_j + \beta_j^{\text{gr\_idade}} x_j + \beta_k^{\text{gr\_anoest}} x_k + \beta_l^{\text{raça}} x_l + \beta_m^{\text{rede\_ens}} x_m + \beta_n^{\text{cond\_ativ}} x_n + \beta_o^{\text{horas\_trab}} x_o + \beta_p^{\text{grup\_ativ}} x_p + \beta_q^{\text{pos\_ocup}} x_q + \beta_r^{\text{class\_rend}} x_r + \beta_s^{\text{celular}} x_s + \beta_t^{\text{tel\_fixo}} x_t + \beta_u^{\text{micro}} x_u + \beta_v^{\text{esg\_sanit}} x_v + \beta_w^{\text{sit\_domic}} x_w + \beta_x^{\text{reg\_metrop}} x_x + \beta_y^{\text{reg\_geog}} x_y$$

O poder explicativo desse modelo também pode ser fornecido pela mesma estatística (*pseudo\_R<sup>2</sup>*) baseada na comparação do *log-likelihood* (máxima verossimilhança) do modelo nulo (*LL<sub>0</sub>*) e do modelo com as variáveis de interesse (*LL<sub>m</sub>*). Para este caso, o modelo de interesse é o modelo multivariado proposto, que contém as 17 covariáveis ajustadas para

TABELA 1 Poder de explicação de cada variável, obtido através dos modelos univariados de acesso à Internet

Modelos Univariados	-2 Log L <sub>0</sub> (Intercept Only)*	-2 Log L <sub>m</sub> (Intercept and Covariates)	pseudo_R <sup>2</sup>
Classe de rendimento	132789887	102094469	0,23116
microcomputador	132789887	102152828	0,23072
Grupos de anos de estudo	132789887	106561633	0,19752
celular	132777669	114195122	0,13995
tel fixo	132787616	114383083	0,13860
rede ens	132787519	120252836	0,09440
grup ativ	132789887	120475530	0,09274
esg sanit	132789887	122731663	0,07575
sit domic	132789887	125748477	0,05303
raça	132564428	126763578	0,04376
pos ocup	132789887	127497806	0,03985
reg geog	132789887	128153819	0,03491
reg metrop	132789887	129850526	0,02214
gr idade	132789887	131114494	0,01262
horas trab	132765555	132606596	0,00120
cond ativ	132781690	132716858	0,00049
sexo	132789887	132783608	0,00005

\* O valor de -2Log L<sub>0</sub> não é exatamente igual para todos os modelos univariados em virtude de a *procedure surveylogistic* não considerar os registros cujo conteúdo da variável explicativa do modelo esteja sem informação (*missing values*).  
*pseudo\_R<sup>2</sup>*: poder explicativo das variáveis no modelo

o acesso à internet. A seguir, encontram-se as estatísticas de ajuste do modelo por três critérios (AIC: Akaike Information Criterion; SC: Schwarz Criterion; e -2 Log L: máxima verossimilhança). Utilizando-se a máxima verossimilhança o *pseudo\_R<sup>2</sup>* é de 0,48, ou seja, 48% da variabilidade das covariáveis são explicadas por este ajuste.

Critério	Estatísticas de Ajuste do Modelo	
	Intercepto Somente	Intercepto e Covariáveis
AIC	132515041	68905953
SC	132515051	68906487
-2 Log L	132515039	68905851

Tratando-se de um modelo logístico comportamental, o poder explicativo está bom, já que neste resultado está embutido o desejo das pessoas de acessar (ou não) a Internet. Por exemplo, algumas não querem acessar a Internet, mesmo possuindo bom nível de renda e de escolaridade; algumas delas não gostam; outras não têm tempo, além de diversos motivos, por exemplo, a rede a que elas têm acesso é muito lenta, devido ao tipo de conexão.

Deve ser mencionado que, inicialmente, a variável referente à existência de energia elétrica no domicílio (forma de iluminação – v0219) fazia parte do conjunto de variáveis explicativas (independentes) do modelo. Porém, como a retirada desta do conjunto de variáveis independentes praticamente não alterou o poder explicativo do modelo, além de as mudanças nos coeficientes (e conseqüentemente nas *Odds Ratio*) serem mínimas, optou-se por não considerá-la integrante do conjunto de variáveis explicativas do modelo. Isto pode ser justificado em virtude de o comportamento desta variável (energia elétrica) já ser captado por outras variáveis explicativas, como por exemplo, a existência de microcomputador no domicílio. Presume-se que, na residência onde exista microcomputador\*, a forma de iluminação deva ser elétrica (isto é, de rede elétrica, gerador ou solar).

Os coeficientes  $\beta_i$  do modelo\*\* são apresentados na tabela 2, a seguir, na qual também podem ser

\* Não se levando em consideração o micro portátil que poderia ter sua bateria carregada em outro local diferente do domicílio.

\*\* HOSMER & LEMESHOW (1989) sugerem que se incluam interações no modelo multivariado, uma a uma, e se verifique o ganho conseguido na explicação da variável dependente. Seguindo este raciocínio, foram incluídas no modelo logístico de acesso à Internet algumas interações de primeira ordem. Em virtude do grande número de interações possíveis de primeira ordem ( $172 = 289$ , já que o modelo proposto possui 17 variáveis explicativas), procurou-se verificar o comportamento daquelas mais significantes. (Foram testadas 31 interações). Porém, percebeu-se que o ganho no poder de explicação do modelo foi muito pequeno para todas as interações de primeira ordem selecionadas. Desta forma, não houve prosseguimento nas investigações, e optando-se por não considerar interações no modelo logístico de acesso à Internet proposto.

visualizados outros parâmetros estimados. Os resultados tomaram por base as categorias adotadas como referência para o modelo de regressão logística multivariado, descritas na tabela.

Pela análise da razão de chances (*Odds Ratio*), apresentada na coluna  $\text{Exp}(\hat{\alpha})$  da tabela 2, chegamos a importantes constatações em relação às variáveis predictoras ou explicativas, descritas a seguir:

- sexo: quando comparado ao sexo oposto, a chance de uma pessoa do sexo masculino acessar a Internet é 32,1% maior que a chance de uma pessoa do sexo feminino (categoria de referência);
- grupos de idade: a chance de um indivíduo de 18 a 24 anos de idade de acessar a Internet é 59,1% menor que a chance de um indivíduo que pertença ao grupo de 10 a 17 anos (categoria de referência adotada)\*. As chances de acesso à Internet vão diminuindo com o aumento da faixa etária, de tal modo que uma pessoa que pertença ao grupo de 40 a 49 anos de idade tem chance 87,7% menor que a de um indivíduo do grupo de 10 a 17 anos;
- grupos de anos de estudo: a chance de acesso à Internet cresce com a elevação do nível de instrução. Um pessoa com 4 a 7 anos de estudo tem 2,05 vezes mais chances de acessar a Internet que uma pessoa sem instrução ou com menos de 4 anos de estudo. As pessoas com 11 a 14 anos de estudo possuem 13,86 vezes mais chances de acessar a Internet, e aquelas com 15 anos ou mais têm 41,7 vezes mais chances, quando comparadas com as pessoas da categoria de referência (indivíduos sem instrução ou com menos de 4 anos de estudo);
- cor/raça: o modelo comprova a existência de condições menos favoráveis da população negra e parda no Brasil, em relação à população branca, no que se refere ao acesso à Internet. A chance de um negro ou pardo ter acesso à Internet é 16,4% menor que a chance de um branco (categoria de referência). Já quando se compara o branco com o amarelo, ocorre comportamento inverso: as pessoas

\* Constatou-se que o grupo de 15 a 17 anos de idade apresentou o maior percentual de pessoas que acessaram a rede, no período de referência dos últimos três meses (Pnad, 2005).

TABELA 2

Parâmetros do Modelo de Regressão Logística de Acesso à Internet. Universo: população residente em domicílio particular permanente com 10 anos ou mais de idade e menos de 50 anos de idade

Variáveis	Categorias	Coef (β)	Erro Padrão	Wald Chi-Square	Pr>ChiSq	Exp(β) Odds Ratio
Sexo	Masculino	0.2782	0.0169	271.4696	<.0001	1.321
Grupos de Idade	18 a 24 anos	-0.8936	0.0301	880.3043	<.0001	0.409
	25 a 29 anos	-1.2134	0.0362	1124.2923	<.0001	0.297
	30 a 39 anos	-1.5302	0.0339	2038.9450	<.0001	0.216
	40 a 49 anos	-2.0980	0.0365	3296.8643	<.0001	0.123
Grupos de Anos de estudo	4 a 7 anos	0.7184	0.0359	400.7879	<.0001	2.051
	8 a 10 anos	1.4780	0.0401	1361.7406	<.0001	4.384
	11 a 14 anos	2.6288	0.0416	3993.1367	<.0001	13.858
	15 anos ou mais	3.7305	0.0541	4757.4199	<.0001	41.701
Cor/raça	Negro (pardo + preto)	-0.1793	0.0195	84.5394	<.0001	0.836
	Amarelo	0.1673	0.1355	1.5258	0.2167	1.182
Rede de ensino	Estuda na rede particular	0.7720	0.0360	459.0394	<.0001	2.164
Condição de atividade	Economicamente ativo	0.4483	0.0318	199.3405	<.0001	1.566
Nº. de horas trabalhadas por semana	De 1 a 39 horas	0.1025	0.0266	14.8714	0.0001	1.108
	0 (zero) horas	0.6333	0.2881	4.8326	0.0279	1.884
Grupamentos de atividade do trabalho principal	Indústria	0.7875	0.0729	116.8293	<.0001	2.198
	Construção	0.6329	0.0830	58.1960	<.0001	1.883
	Comércio	0.8715	0.0713	149.4227	<.0001	2.391
	Transp., armaz. e comunic.	0.8022	0.0805	99.2528	<.0001	2.230
	Administração pública	1.5885	0.0813	382.2261	<.0001	4.897
	Educação, saúde e serv. soc.	1.0409	0.0788	174.4468	<.0001	2.832
	Serviços	0.7461	0.0777	92.1152	<.0001	2.109
	Outras atividades	1.6951	0.0770	484.2460	<.0001	5.447
Posição na ocupação no trabalho principal	Empregado sem carteira	-0.2886	0.0297	94.3024	<.0001	0.749
	Func. público estatut./ Militar	-0.1899	0.0478	15.7810	<.0001	0.827
	Conta própria	-0.4481	0.0332	182.0010	<.0001	0.639
	Empregador	-0.0294	0.0525	0.3141	0.5752	0.971
	Outros	0.00021 0	0.1395	0.0000	0.9988	1.000
Classes de rendimento domiciliar per capita	Mais de 1/4 a 1/2 salário mín.	0.2649	0.0483	30.1283	<.0001	1.303
	Mais de 1/2 a 1 salário mín.	0.5390	0.0491	120.4729	<.0001	1.714
	Mais de 1 a 2 salários mín.	0.8946	0.0497	324.4091	<.0001	2.446
	Mais de 2 a 3 salários mín.	1.2569	0.0544	534.0895	<.0001	3.514
	Mais de 3 a 5 salários mín.	1.5700	0.0585	721.4133	<.0001	4.806
	Mais de 5 salários mín.	1.9866	0.0660	907.1055	<.0001	7.291
Telefone móvel celular	Possui	0.7990	0.0193	1715.3132	<.0001	2.223
Telefone fixo convenc.	Possui	0.3384	0.0226	224.5003	<.0001	1.403
Microcomputador	Possui	1.4477	0.0235	3795.3147	<.0001	4.253
Esgotamento sanitário	Adequado	0.2549	0.0305	70.0761	<.0001	1.290
Situação do domicílio	Urbano	0.5954	0.0558	114.0010	<.0001	1.814
Região metropolitana	Reside	0.1224	0.0242	25.6170	<.0001	1.130
Regiões geográficas	Sudeste	0.0281	0.0327	0.7368	0.3907	1.028
	Sul	0.1001	0.0425	5.5524	0.0185	1.105
	Centro-Oeste	0.1160	0.0424	7.4772	0.0062	1.123
	Norte	-0.3194	0.0492	42.1471	<.0001	0.727
Constante	Intercepto	-4.8855	0.1056	2139.3684	<.0001	

Número de observações usadas: 262.245

Categorias de referência em ordem: feminino, 10 a 17 anos, sem instrução e menos de 4 anos, branco, estuda na rede pública, não economicamente ativo, 40 horas ou mais, agrícola, empregado com carteira, sem rendimento a 1/4 do salário mínimo, não possui telefone móvel celular, não possui telefone fixo convencional, não possui microcomputador, inadequado, rural, não reside em região metropolitana, Nordeste

Fonte: microdados da Pnad 2005

que se autodeclararam amarelas possuem condições mais favoráveis de acesso à Internet, quando confrontadas com as brancas, evidenciando a forte ligação dos “orientais brasileiros” com a informática. A chance de um amarelo ter acesso à Internet é 18,2% maior que a de um branco;

- rede de ensino: apesar de existirem algumas iniciativas políticas voltadas para a informatização de escolas públicas, pelo quadro apresentado em relação à rede de ensino, podemos constatar que estas ações ainda são insuficientes. O modelo revela que a chance de um estudante da rede particular de ensino ter acesso à Internet é 116,4% maior que a chance de um estudante da rede pública (categoria de referência), ou de forma equivalente, um estudante da rede particular tem 2,16 vezes mais chances de acessar a Internet que um estudante da rede pública;

- condição de atividade: a chance de uma pessoa economicamente ativa de acessar a Internet é expressivamente maior (56,6%) que a chance de uma pessoa na condição de não-economicamente ativa;

- número de horas trabalhadas por semana: não há tanta diferença nas chances de acesso à Internet entre os indivíduos que possuem jornada de trabalho semanal de 1 a 39 horas e os que têm jornada de 40 horas ou mais, apesar de as chances de acesso à rede serem maiores para as pessoas que possuem jornada com menos horas semanais. O modelo revela que a chance de um indivíduo que trabalha de 1 a 39 horas por semana acessar a Internet é 10,8% maior que a de um indivíduo que possui jornada semanal de 40 horas ou mais. Cabe ressaltar que no grupo que possui jornada de trabalho semanal de 0 (zero) horas, ou melhor, que não trabalha, estão inseridos os estudantes. A chance de acesso à Internet de uma pessoa deste grupo é 88,4% maior que a de uma pessoa do grupo da categoria de referência. A estrutura etária bastante jovem da população estudantil e necessidade dos estudantes de buscar subsídios para o estudo através da *Web* podem ser apontadas como fatores importantes para justificar essa discrepância;

- na análise de acordo com o agrupamento de atividade do trabalho principal, verificamos que as

pessoas que trabalham no setor agrícola são as que possuem as menores chances de acessar a Internet. As pessoas ligadas ao setor da indústria possuem 2,2 vezes mais chances de acessar a Internet que os indivíduos que trabalham no setor agrícola (categoria de referência). Os trabalhadores que pertencem ao agrupamento do comércio (incluindo comércio e reparação e alojamento e alimentação) possuem 2,39 vezes mais chances de acessar a rede que os que pertencem ao setor agrícola. As pessoas que trabalham na administração pública possuem 4,9 vezes mais chances que as do setor agrícola, e as que trabalham em outras atividades (incluindo intermediação financeira, seguros, previdência privada, atividades de informática, pesquisa e desenvolvimento) possuem 5,45 vezes mais chances de acesso à rede que as pessoas que trabalham na categoria de referência (setor agrícola);

- em relação à posição na ocupação no trabalho principal, percebemos que o empregado sem carteira possui 25,1% menos chances de acessar a Internet que os empregados com carteira (categoria de referência), e o que trabalha por conta própria possui chance 36,1% menor de acesso à rede, quando comparado à categoria de referência. A metodologia da Pnad 2005 considera como empregador qualquer pessoa que trabalhe explorando o próprio empreendimento, com pelo menos um empregado, estando inseridos nesta categoria inúmeros donos de pequenos estabelecimentos comerciais (bares, biroskas, armarinhos etc.). Em virtude deste fato, não nos surpreende que a chance de acesso à Internet de uma pessoa que se enquadre na categoria de empregador seja ainda menor que a chance de acesso de um empregado com carteira (chance de acesso 2,9% menor que a de um empregado da categoria de referência). O modelo revela que um indivíduo na categoria outros trabalhadores (trabalhador na produção para o próprio consumo e trabalhador na construção para o próprio uso) possui a mesma chance de acesso à Internet de um empregado com carteira;

- classes de rendimento domiciliar *per capita*: o modelo constatou que a chance de acesso à Internet cresce com a elevação do nível de rendimento domiciliar *per capita*. As pessoas que se enquadram



na faixa de mais de 1/4 a 1/2 salário mínimo possuem 1,3 vez mais chances de acessar a Internet que as pessoas da faixa sem rendimento a 1/4 do salário mínimo (categoria de referência), ou de maneira equivalente, possuem chance 30,3% maior que as pessoas que se enquadram na faixa de rendimento da categoria de referência. Os indivíduos que pertencem à faixa de mais de um a dois salários mínimos possuem 2,45 vezes mais chances de acessar a rede que aqueles que pertencem à faixa de rendimento da categoria de referência. Os que se enquadram na faixa de mais de três a cinco salários mínimos possuem 4,81 vezes mais chances de acessar a Internet que os da faixa sem rendimento a 1/4 do salário mínimo, e os que pertencem à faixa de mais de cinco salários mínimos possuem 7,29 vezes mais chances que aqueles que se situam na categoria de referência (faixa sem rendimento a 1/4 do salário mínimo);

- telefone móvel celular: pelo modelo pode-se constatar que a chance de uma pessoa que possui aparelho telefônico celular de acessar a Internet é 122,3% maior que a chance de uma que não possui celular (categoria de referência);

- telefone fixo convencional: a existência de telefone fixo nas residências aumenta a chance de acesso à Internet, possibilitando, inclusive, o acesso à rede via linha telefônica discada. Quem reside em domicílio com telefone fixo convencional possui 1,4 vez mais chances de acessar a Internet que uma pessoa que não possui telefone fixo (categoria de referência), ou de forma equivalente, chance 40,3% maior;

- microcomputador: obviamente a existência desse equipamento no domicílio aumenta expressivamente a chance de uma pessoa acessar a Internet. Um indivíduo que o possui tem 4,25 vezes mais chances que um indivíduo que reside em domicílio sem microcomputador (categoria de referência), isto é, chance 325,3% maior;

- esgotamento sanitário: as condições de habitabilidade podem ser expressas, *grosso modo*, pelo tipo de esgotamento sanitário presente nos domicílios, refletindo muitas vezes o nível social e educacional das famílias. O modelo revela que a

chance de uma pessoa residente em domicílio que possui esgotamento sanitário adequado acessar a Internet é 29,0% maior que aquela residente em domicílio que possui esgotamento sanitário inadequado (categoria de referência);

- situação do domicílio: as pessoas que vivem em áreas urbanas têm maiores chances de acesso à Internet que os residentes em áreas rurais. O percentual é de 81,4% dos primeiros em relação aos segundos;

- região metropolitana: constatação semelhante à observada em relação à situação do domicílio (rural/urbano) é verificada sobre residir ou não em região metropolitana. A chance de um indivíduo residente em domicílio situado em região metropolitana acessar a Internet é 13,0% maior que a chance de um que não reside em região metropolitana (categoria de referência);

- regiões geográficas: em termos regionais, o modelo revela que as chances de acesso à Internet são bastante diferenciadas. Uma pessoa da região Sudeste tem, na média, as mesmas (1,03 vez) chances que uma pessoa da região Nordeste (categoria de referência), mas possui 2,8% vezes mais chances que uma pessoa da região Nordeste. Uma pessoa da região Sul tem 1,1 vez mais chances de acessar a Internet que uma pessoa da região Nordeste, e uma pessoa da região Centro-Oeste tem 1,12 vezes mais chance que uma pessoa da região tomada como categoria de referência (uma chance 12,3% maior). Já a chance de um morador da região Norte acessar a Internet é 27,3% menor que a de um da região Nordeste (categoria de referência).

## CONCLUSÃO

Esta artigo analisou as características do acesso à Internet na população brasileira, fazendo uso de uma análise multivariada, por meio da qual foi possível definir os contornos da exclusão digital e evidenciar as disparidades existentes em nosso país.

Na exploração das dimensões da divisão digital, ficou evidenciado que os níveis de instrução e renda dos indivíduos constituem fatores demasiadamente importantes, além de mostrar que a Internet precisa ser rapidamente incorporada como uma ferramenta

educacional básica por todo o sistema escolar, notadamente o da rede pública de ensino. No Brasil, percebe-se ainda que as escolas são territorial e institucionalmente (públicas / privadas) diferenciadas por classe social e raça, existindo um fosso substancial entre elas, em termos de tecnologia.

Com a finalização da abordagem empírica, no que concerne à exclusão digital, os seguintes resultados tornam-se visíveis:

1) o aumento do nível educacional diminui a exclusão digital. As pessoas que possuem menos de quatro anos de estudo têm participação nitidamente menor na população dos incluídos digitais do que os que possuem nível de escolaridade mais alto. No caso dos que possuem 15 anos ou mais de estudo, observa-se que a participação deste grupo na população dos incluídos digitais é bastante expressiva, o que nos conduz à constatação que a melhora do nível de escolaridade constitui fator fundamental para a redução da exclusão digital;

2) a melhoria e efetiva informatização da rede pública de ensino pode trazer enormes contribuições para a redução da exclusão digital. Ainda são graves os problemas das escolas públicas brasileiras, existindo grande número de alunos matriculados em escolas que não possuem laboratórios de informática ou qualquer ponto de acesso à Internet. Garantir a melhoria da qualidade do ensino público e, conseqüentemente, da aprendizagem, equipando essas escolas com computadores com acesso à Internet em banda larga e demais ferramentas da tecnologia da informação é indispensável para reduzir a exclusão digital, e, portanto, deve estar presente de forma marcante nas políticas públicas brasileiras;

3) a ampliação da infra-estrutura de comunicação contribui para a inclusão digital. A universalização das telecomunicações, facilitando a obtenção de telefones de linha convencional em larga escala pelos indivíduos de camadas sociais menos abastadas, favorece a inclusão digital, uma vez que possibilita o acesso à Internet através de linha discada (desde que exista um microcomputador, um *modem*, e seja instalado um *software* de conexão). Visando a alcançar o objetivo da inclusão digital para todos e o acesso economicamente viável às tecnologias da

informação e comunicação (TICs), a difusão de uma vasta rede de fibras ópticas nos meios urbanos, a proliferação de antenas que permitam o acesso à Internet via satélite nas regiões mais distantes e a adoção da tecnologia *wireless* devem configurar-se como elementos-chave nas políticas públicas voltadas para a melhoria e ampliação da infra-estrutura de comunicação no território nacional;

4) crianças e adolescentes são grupos bastante excluídos em nosso país, quando comparados com os demais grupos etários, menos no que se refere à inclusão digital. Os jovens estão mais bem representados na tecnologia digital do que os indivíduos de idades mais avançadas, o que suscita algum otimismo quanto ao futuro, à medida que ao longo do ciclo de vida as novas gerações vão difundir seus conhecimentos e práticas;

5) em nosso país, a divisão étnica continua sendo indicativa do fato de que a era da informação não é cega para cores. No Brasil o *apartheid* digital reflete, em certa medida, as condições sociais e educacionais diferentes, quando comparados brancos com afro-brasileiros. Observa-se uma discrepância de chances entre brancos e negros (pardos e pretos) no acesso à Internet. Confrontando grupos raciais diferentes, com características observáveis (por exemplo, anos de estudo e tipo de emprego) semelhantes, constata-se os negros menos incluídos digitalmente que os indivíduos da raça branca;

6) na elaboração das políticas de inclusão digital, o governo federal deve estar convicto da necessidade de generalização da banda larga, de forma a atingir todos os rincões do país. A massificação da banda larga é essencial para o desenvolvimento econômico do Brasil, e de qualquer outro país, uma vez que se configura como o “sistema nervoso” da nova economia globalizada;

7) devido à alta desigualdade de renda existente em nosso país, as ações de mobilização de recursos para financiamento subsidiado de equipamentos de informática a preços populares (baixo custo), com prestações acessíveis e com prazos para aquisição dilatados, voltadas principalmente às populações de baixo poder aquisitivo e marginalizadas digitalmente, devem ocupar lugar de destaque nas orientações de

políticas públicas, visando a impulsionar a inclusão digital de inúmeros cidadãos brasileiros;

8) não adianta falar e discutir as maravilhas da era tecnológica, microcomputadores, tecnologia *wireless*, protocolo *Wap*, redes de fibra óptica, provedores de acesso, tamanho da banda de conexão, esquecendo do grande número de famílias brasileiras que não têm sequer energia elétrica e formas adequadas de esgotamento sanitário em seus domicílios. Para enfrentar os novos desafios da sociedade da informação, o governo brasileiro precisa assumir também essas barreiras, avançar na implementação de políticas estruturais e realizar ações redistributivas de combate à pobreza.

No decorrer das análises realizadas neste artigo, foi possível perceber que o resultado cumulativo de diferentes camadas de desigualdades traduz-se em vastas diferenças nos efeitos do uso da Internet, tendo, portanto, relevantes implicações no desempenho educacional dos indivíduos, e conseqüentemente, no tratamento da exclusão digital.

## REFERÊNCIAS

- AGRESTI, Alan. *An introduction to categorical data analysis*. New York: Wiley Series in Probability and Statistics, 1996.
- ALVES, Nuno de Almeida. *Planos de ação para a sociedade da informação e do conhecimento: mudança tecnológica e ajustamento estrutural*. 2004. Disponível em: <<http://www.scielo.oces.mctes.pt/pdf/spp/n44/n44a06.pdf>>. Acesso em: 24 nov. 2007.
- CASTELLS, Manuel. *A sociedade em rede*. São Paulo: Editora Paz e Terra 2005.
- \_\_\_\_\_. *A galáxia da Internet: reflexões sobre a Internet, os negócios e a sociedade*. Rio de Janeiro: Jorge Zahar Ed., 2003.
- DOBSON, A. J. *An introduction to generalized linear models*. 1<sup>st</sup> ed. [S.l.]: Chapman & Hall/CRC, 1990.
- FREEMAN, Christopher; PEREZ, Carlota. Structural crisis of adjustment: business cycles and investment behavior. In: FREEMAN, Christopher (Org.). *Long wave theory*. Cheltenham: Edward Elgar, 1996. (Elgar Reference Collection).
- GIDDENS, Anthony. *Mundo em descontrol: o que a globalização está fazendo de nós*. Rio de Janeiro: Record, 2005.
- HOSMER, D. W.; LEMESHOW, S. *Applied logistic regression*. New York: A Wiley-Interscience Publication, 1989.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. *Pesquisa nacional por amostra de domicílios 2005: acesso à Internet e posse de telefone móvel celular para uso pessoal*. [S.l.], 2007.
- LÉVY, Pierre. *As tecnologias da inteligência: o futuro do pensamento na Era da Informática*. Rio de Janeiro: Editora 34, 1993.
- NEGROPONTE, Nicholas. *A vida digital*. São Paulo: Cia das Letras, 1995. 231 p.
- NERI, Marcelo Côrtes. *Mapa da exclusão digital*. Rio de Janeiro: Fundação Getúlio Vargas, 2003.
- SILVA, Pedro Luis do Nascimento; PESSOA, Djalma Galvão Carneiro. *Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral*. 2002. Disponível em: <[http://www.ie.ufrj.br/eventos/seminarios/pesquisa/texto06\\_04\\_11.pdf](http://www.ie.ufrj.br/eventos/seminarios/pesquisa/texto06_04_11.pdf)>. Acesso em: 19 jun. 2007.
- SILVEIRA, Sérgio Amadeu da. *Exclusão digital: a miséria na era da informação*. São Paulo: Fundação Perseu Abramo, 2005.
- TAKAHASHI, T. (Org.). *Sociedade da informação no Brasil: livro verde*. Brasília: Ministério da Ciência e Tecnologia, 2000.
- \_\_\_\_\_. *SAS Institute Inc. SAS 9.1 language reference: concepts*. 3<sup>rd</sup> ed. Cary: [s.n.], c2005.
- \_\_\_\_\_. *STAT 2005/3004 multivariate data analysis: part 2: logistic regression*. [S.l.: s.n., s.d.].