

SER PESSOA COM DEFICIÊNCIA – QUESTÃO DE RISCO

Paulo Tadeu Meira¹², Silvia de Oliveira¹²

Resumo: *Em todo o mundo, as pessoas com deficiência apresentam piores perspectivas de saúde, níveis mais baixos de escolaridade, menor participação econômica, taxa de pobreza mais elevada em termos comparativos às pessoas sem deficiência. Para que as pessoas com deficiência atinjam perspectivas melhores e mais duradoras, devemos capacitar essas pessoas e retirar as barreiras que as impedem de participar da comunidade, de ter acesso à educação de qualidade, de encontrar trabalho decente e ter suas vozes ouvidas. Em, termos estatísticos, uma alternativa bastante útil e que pode servir de suporte e monitoramento das políticas públicas nessa área é propormos para ser utilizado de forma contínua, um índice de risco denominado índice de risco pessoa com deficiência que consiste em avaliar quais fatores estão associados a este risco, assim como a intensidade e a direção de cada um desses fatores, gerando um escore final que pode ser ordenado ou classificado, segundo a probabilidade de pessoas adquirirem uma determinada deficiência. No caso brasileiro, propomos a utilização de técnicas como regressão logística binária e ordinal para seleção de fatores mais significantes utilizando critérios como AIC (do inglês, Akaike Information Criterion), BIC (do inglês, Bayesian Information Criterion), e DIC (do inglês, Deviation Information Criterion), e calcular a probabilidade do risco para as diferentes deficiências (enxergar, ouvir, movimentar e intelectual) para o conjunto de dados da Amostra constituída por 20800804 respondentes do Questionário Completo no Censo 2010 do IBGE por estado, região e país.*

Palavras-chave: *Pessoas com deficiência, regressão logística ordinal estereótipo, risco deficiência, seleção de variáveis, seleção de modelos.*

Abstract: *In this work, we propose the stereotype ordinal logistic model; variable selection using Wald; model selection using AIC, BIC and DIC criteria; and, finally; define the disability risk considering the different degrees of severity for different disabilities visual, hearing, walk and intellectual.*

Keywords: *Disability Risk, Model selection, People with Disabilities, Stereotype Ordinal Logistic Model, Variable Selection.*

¹ MAE - IME/USP. e-mail: poliver@usp.br

² o autor agradece ao IBGE pelo acesso aos microdados dos domicílios selecionados para composição da amostra e a Professora Maria Pavan Soler pela indicação do tema.

1 Introdução

Segundo a Organização Mundial de Saúde (OMS) em 2010, estima-se que um pouco mais de um bilhão de pessoas de todo o mundo, representando por volta de 15% da população mundial e no caso do Brasil, segundo o Instituto Brasileiro de Geografias e Estatística (IBGE) em 2010, estima-se que 45,6 milhões de pessoas, representando aproximadamente 23,9% da população brasileira, convivem com alguma forma de deficiência. Em todo o mundo, as pessoas com deficiência apresentam piores perspectivas de saúde, níveis mais baixos de escolaridade, menor participação econômica e taxa de pobreza mais elevada em termos comparativos às pessoas sem deficiência.

As pessoas com deficiências compõem um grupo de excluídos que sempre despertou diferentes sentimentos, desde a repulsa até a piedade extrema, tendo sido até considerados menos humanos ou desprovidos de humanidade. Atualmente, no âmbito das políticas de inclusão social e educacional, tornaram-se alvo de ações afirmativas, as quais buscam assegurar-lhes seus direitos em vários aspectos da vida em sociedade (Silva, 1987; Oliveira, 2013).

Acredita-se que as baixas condições de trabalho das pessoas com deficiência são devidas a situações como: dificuldade de acesso à educação, infraestrutura inadequada, preconceito, falta de informação e de melhores condições de acessibilidade por parte de escolas e empresas que fazem com que essas pessoas apresentem um menor nível de escolaridade o que dificulta o ingresso delas no mercado formal de trabalho (Garcia, 2010).

Para que as pessoas com deficiência atinjam perspectivas melhores e mais duradoras, devemos capacitar essas pessoas e retirar as barreiras que as impedem de participar da comunidade, de ter acesso à educação de qualidade, de encontrar trabalho decente e ter suas vozes ouvidas (Figueira, 2008).

Para que possa melhor avaliar a necessidade das pessoas com deficiência torna-se necessário, melhor descrever esse conjunto de pessoas para saber respostas a questões como: Quantas são? Onde moram? Como vivem? Quais são as implicações que a deficiência acarreta no acesso dessas pessoas a todos os diferentes serviços humanos de forma autônoma e plena? Em suma, como a deficiência pode influenciar na qualidade de vida destas pessoas?

Em termos estatísticos, mostra a existência de poucos estudos formais, entre os quais destacam-se os dados obtidos através dos censos, que possibilita indagações como: Como as pessoas com deficiência estão distribuídas ao longo do país? Como avaliar o acesso das pessoas com deficiência em termos dos diferentes serviços mencionados anteriormente? Como está a evolução das pessoas com deficiência ao compará-las com as que não apresentam deficiência? Quais seriam as variáveis que mais contribuem para os casos de

deficiências? Como se encontram as pessoas com deficiência em comparação às que não apresentam deficiência? Resposta a essas e outras perguntas possivelmente poderão contribuir para melhor suporte a essas pessoas no sentido de serem melhores assistidas e de recursos serem melhor gerenciados e otimizados pelas ações das políticas públicas nesta área.

No que diz respeito a estudos estatísticos, uma alternativa bastante útil e que pode servir de suporte e monitoramento das políticas públicas nessa área é propormos para ser utilizado de forma contínua, um índice de risco denominado *índice de risco pessoa com deficiência* que consiste em avaliar quais fatores estão mais associados a este risco, assim como a intensidade e a direção de cada um desses fatores, gerando um escore final que pode ser ordenado ou classificado, segundo a probabilidade de pessoas adquirirem uma determinada deficiência. No caso brasileiro, propomos a utilização de técnicas como regressão logística binária e ordinal para seleção de fatores mais significantes utilizando critérios como AIC (do inglês, *Akaike Information Criterion*), BIC (do inglês, *Bayesian Information Criterion*), e DIC (do inglês, *Deviation Information Criterion*), e calcular a probabilidade do risco para as diferentes deficiências (enxergar, ouvir, movimentar e intelectual) para o conjunto de dados da Amostra constituída por 20800804 respondentes do Questionário Completo no Censo 2010 do IBGE por estado, região e país.

Em um trabalho anterior (Oliveira, 2013), consideramos como variáveis respostas, as diferentes deficiências e existência de pelo menos uma deficiência como variáveis binárias, isto é, se um determinado indivíduo é ou não pessoa com deficiência. Nesse trabalho, estamos considerando as diferentes deficiências incorporando seus diferentes graus de severidade e o número de deficiências como variáveis respostas ordinais, o que possibilita uma melhor qualidade em termos de informação e de ajuste no modelo.

Na seção 2 apresentamos uma motivação para o problema, definimos e caracterizamos as variáveis a serem utilizadas, descrevemos e definimos o modelo logístico ordinal estereótipo, seleção de variáveis utilizando teste de Wald, seleção de modelos utilizando os critérios AIC, BIC e DIC e definimos o risco deficiência considerando os diferentes graus de severidade “*não consegue de modo algum*”, “*consegue com muita dificuldade*” e “*consegue com um pouco de dificuldade*” para as deficiências visuais, auditivas e de locomoção, e, no caso de deficiência intelectual propomos a utilização do risco de “*ter*” ou “*não ter*”. Na seção 3 mostramos os resultados e discussões; e, na seção 4 apresentamos as conclusões e sugestões para trabalhos futuros.

2 Material e métodos

a. Motivação

Para que haja inclusão de pessoas com deficiência, é necessário, antes de mais nada, conhecer quais são os fatores que mais influenciam na existência de pessoas com deficiência. Neste trabalho, propomos o ajuste de modelos logísticos ordinal estereótipo para os fatores mais significantes utilizando como seleção critérios como AIC, BIC e DIC; criação e determinação do risco deficiência para o conjunto de dados da Amostra que responderam o Questionário Completo do Censo 2010 do IBGE.

2.2. Descrição dos dados

As variáveis foram obtidas diretamente do questionário aplicado para o conjunto de dados da Amostra que responderam o Questionário Completo e podem ser encontrados no site www.ibge.gov.br em Censo 2010, amostra e micro dados com mais detalhes quanto a sua descrição em Oliveira (2013).

2.3. Regressão logística ordinal

Muitas das variáveis de estudo nas ciências sociais e humanas são ordinais. Com frequência, a variável dependente toma valores discretos, ou categorias ordenáveis mas cuja a distância entre elas não é conhecida, nem tão pouco constante. Por exemplo, em estudos epidemiológicos e de severidade de deficiência para ver, ouvir ou movimentar conforme estabelecido no Questionário da amostra no Censo Demográfico de 2010 do IBGE que pode ser classificada em “*não consegue de modo algum*”; “*consegue, mas com muita dificuldade*”; “*consegue, mas com um pouco de dificuldade*”; e, por fim, “*não apresenta problemas para ouvir, ver ou se movimentar*”. No caso de deficiência intelectual é dividida em “*ter*” ou “*não ter*”.

Entre os possíveis modelos de regressão logística ordinal é possível destacar: modelo de chances proporcionais que é mais apropriado e de interpretação mais fácil quando a variável resposta a ser considerada é uma variável contínua que foi categorizada; modelo de razão contínua que é utilizado em casos em que existe interesse específico em uma categoria de variável resposta; modelo de chances proporcionais parciais que permite que algumas covariáveis possam ser modeladas com a suposição de chance proporcional, e para as outras variáveis em que este pressuposto não seja satisfeito, são incluídos no modelo parâmetros específicos que variam para as diversas categorias comparadas, e, trata-se de uma extensão do modelo de chances proporcionais; e, por fim; modelo estereótipo, proposto por Anderson (1984) que é utilizado em situações em que a variável resposta é uma variável ordinal que não se trata de uma versão discreta de alguma variável contínua.

Para este trabalho, temos como variável resposta: deficiências, visual, ouvir, locomover e intelectual que tratam-se de variáveis ordinais que não são versões de variáveis contínuas, em vista disso, adotamos neste trabalho o modelo estereótipo.

2.3.1. Especificação do modelo estereótipo

Suponhamos que a variável dependente é constituída por J categorias ($m = 1, \dots, J$) e consideremos K regressores ($J = 1, \dots, K$). O modelo estereótipo define-se numa fase inicial com o modelo de regressão multinomial ao qual se adiciona a condição $\beta_{m|J} \equiv \phi_m \hat{\beta}$, onde J é a categoria de referência, isto é, temos que o modelo de regressão multinomial é dado por

$$\text{Prob}(y = m|x) = \frac{\exp(\beta'_{m|J}x)}{\sum_{j=1}^J \exp(\beta'_{j|J}x)}, \text{ com } m = 1, \dots, J. \quad (1)$$

Substituindo $\beta_{m|J} = \phi_m \tilde{\beta}$ na equação (1) resulta no modelo estereótipo que pode ser escrito matematicamente por:

$$\text{Prob}(y = m|x) = \frac{\exp(\phi_m \tilde{\beta}'x)}{\sum_{j=1}^J \exp(\phi_j \tilde{\beta}'x)} = \frac{\exp(\phi_m \tilde{\beta}_0 + \phi_m \tilde{\beta}_1 x_1 + \dots + \phi_m \tilde{\beta}_k x_k)}{\sum_{j=1}^J \exp(\phi_j \tilde{\beta}_0 + \phi_j \tilde{\beta}_1 x_1 + \dots + \phi_j \tilde{\beta}_k x_k)}, \text{ com } m = 1, \dots, J. \quad (2)$$

Para alguns parâmetros da equação (2) que não são identificáveis, consideramos como restrições $\phi_m \tilde{\beta}_0 \equiv \theta_m$ ($m = 1, \dots, J$), onde $\phi_J = 0$; e $\phi_m \tilde{\beta}_j \equiv -\theta_m \beta_j$ ($m = 1, \dots, J$ e $j = 1, \dots, k$), onde $\phi_J = 0$ e $\phi = 1$. Deste modo, a partir da equação (2), o modelo estereótipo pode ser escrito da seguinte forma:

$$\text{Prob}(y = m|x) = \frac{\exp(\theta_m - \phi_m \beta'x)}{\sum_{j=1}^J \exp(\phi_j \beta'x)}, \quad (3)$$

com $m = 1, \dots, J$ e onde $\theta_J = 0$, $\phi_J = 0$ e $\phi = 1$.

2.3.2. Interpretação dos coeficientes estimados

Aplicando logaritmo na função (3) para duas quaisquer categorias obtemos:

$$\log \left[\frac{p(Y=q/x)}{p(Y=r/x)} \right] = (\theta_q - \theta_r) - (\phi_q - \phi_r) \beta'x. \quad (4)$$

Aplicando a função exponencial a função exponencial à equação (4), segue:

$$\Omega_{q/r} = \frac{p(Y=q/x)}{p(Y=r/x)} = \exp\{(\theta_q - \theta_r) - (\phi_q - \phi_r) \beta'x\} \quad (5)$$

A equação (5) permite avaliar a razão de chances antes e depois de adicionarmos uma unidade à variável x_j , isto é,

$$\frac{\Omega_{q/r}(x, x_k + 1)}{\Omega_{q/r}(x, x_k)} = \exp\{(\phi_r - \phi_q)\beta_x\} \quad (6)$$

O valor obtido na expressão (6) pode ser interpretado como adicionando uma unidade na variável x_k , a razão de chances da categoria r varia $\exp\{(\phi_r - \phi_q)\beta_x\}$ mantendo todas as outras variáveis constantes.

2.3.3. Estimação dos coeficientes estimados

Os parâmetros do modelo estereótipo são estimados pelo método de Máxima Verossimilhança, em que os estimadores são obtidos pelo sistema de equações dadas em (7) a seguir:

$$p_i = \begin{cases} \text{Prob}(y_i = 1|x_i, \phi, \theta) & \text{se } y_i = 1 \\ \vdots & \\ \text{Prob}(y_i = m|x_i, \phi, \theta) & \text{se } y_i = m \\ \vdots & \\ \text{Prob}(y_i = J|x_i, \phi, \theta) & \text{se } y_i = J \end{cases} \quad (7)$$

onde p_i é a probabilidade de observar qualquer valor de y ; e, a $\text{Prob}(y_i = 1|x_i, \phi, \theta)$ foi definida na expressão (3). Ao assumirmos que a amostra é independente e identicamente distribuída, a função de verossimilhança é dada pela expressão (8) a seguir:

$$L(\beta, \phi, \theta|y, x) = \prod_{i=1}^N p_i = \prod_{m=1}^J \prod_{y=m} \text{Prob}(y = m|x, \phi, \theta) \quad (8)$$

em que $\prod_{y=j}$ indica as multiplicações sobre todos os casos *onde* $y = m$ ($m = 1, \dots, J$). Aplicando logaritmo a função de verossimilhança obtida em (8) obtemos o logaritmo da função de verossimilhança dada em (9) logo abaixo:

$$\log(L(\beta, \phi, \theta|y, x)) = \sum_{m=1}^J \sum_{y=m} \log[\text{Prob}(y = m|x, \phi, \theta)] \quad (9)$$

Os parâmetros ϕ 's e θ 's da equação (9) são estimados pelo método de Newton-Raphson.

A razão de chance formada terá uma tendência de crescimento, já que os pesos podem ser construídos pela ordenação. Assim, o efeito das covariáveis na primeira razão de chances é menor que o efeito na segunda e assim sucessivamente.

A determinação destes pesos pode ser feita a priori, sendo estimados por algum estudo piloto ou conjunto de valores apropriadamente escolhidos.

Por exemplo, situação de número de deficiências que um indivíduo possui e que pode variar entre 0 a 4 deficiências, neste caso seriam cinco opções de resposta.

Quanto à qualidade do ajuste para os modelos ordinais, esta pode ser verificada usando os testes de Pearson ou deviance. Esses testes envolvem a criação de uma tabela de contingência na qual as linhas consistem de todas as possíveis configurações das covariáveis do modelo e as colunas são as categorias de resposta ordinal (Abreu, 2007). As contagens esperadas (E_{lj}) dessa tabela são expressas por $E_{lj} = \sum_{l=1}^{N_L} \hat{p}_{lj}$, em que N_L é o número total de indivíduos classificados na linha l e \hat{p}_{lj} representa a probabilidade de um indivíduo na linha l ter a resposta j calculada a partir do modelo adotado.

O teste de Pearson para avaliar a adequação do ajuste compara essas contagens esperadas com as observadas pela fórmula

$$\chi^2 = \sum_{l=1}^L \sum_{j=1}^k \frac{(O_{lj} - E_{lj})^2}{E_{lj}} \quad (10)$$

A estatística deviance também compara contagens observadas (O_{lj}) e esperadas, mas através da fórmula:

$$D^2 = 2 \sum_{l=1}^L \sum_{j=1}^k O_{lj} \log \frac{O_{lj}}{E_{lj}} \quad (11)$$

Os testes para avaliar a qualidade de ajuste do modelo são baseadas nas aproximações das estatísticas (10) e (11) para a distribuição qui-quadrado com $(L - 1)(k - 1)p$ graus de liberdade, em que L e k são conforme definidos anteriormente e p é o número de covariáveis do modelo. Diferenças significativas levam a conclusão de falta de ajuste do modelo aos dados estudados.

Como alternativa, utilizaremos o teste de Wald que é dado por

$$W = (\hat{p} - \hat{p}_0)' \hat{V}_p^{-1} (\hat{p} - \hat{p}_0) \quad (12)$$

em que \hat{V}_p é o estimador consistente da matriz de variância-covariância do estimador \hat{p} do vetor de proporção \hat{p} . Um estimador \hat{V}_p pode ser obtido por método de linearização.

2.3.4. Teste de significância para o modelo

O teste de Wald para os parâmetros considerados individualmente, pode ser obtido comparando a estimativa de Máxima Verossimilhança de determinado coeficiente ($\hat{\beta}_j$) com a estimativa do seu erro padrão (tendo por base a distribuição assintótica dos estimadores de máxima verossimilhança). Assim, a hipótese nula e a hipótese alternativa do teste são respectivamente:

$$H_0: \hat{\beta}_j = \beta_j^* \text{ vs } H_1: \hat{\beta}_j \neq \beta_j^* \quad (j=2, \dots, k),$$

sendo a respectiva estatística sob a hipótese nula:

$$T = \frac{\hat{\beta}_j - \beta_j^*}{\sqrt{\text{var}(\hat{\beta}_j)}} \sim N(0,1)$$

Ao rejeitar H_0 , para um nível de significância α , concluímos que o parâmetro estimado é estatisticamente diferente de β_j^* . Genericamente, utiliza-se $\beta_j^*=0$ o que, perante estas condições, concluímos que o parâmetro é relevante para explicar o comportamento da variável dependente.

2.3.5. Seleção de variáveis

Selecionar variáveis significa escolher um subconjunto que retenha as variáveis predictoras mais importantes excluindo as demais, de tal forma que procura evitar problema como multicolinearidade e que esse subconjunto ajuste tão bem quanto o modelo com todas as variáveis e que contenha as variáveis predictoras mais importantes.

Entre as diferentes estratégias que podem ser utilizadas para selecionar variáveis destacamos forward stepwise e a backward stepwise. A forward stepwise inicia com a constante β_0 , e sequencialmente adiciona ao modelo o preditor X_i mais correlacionado com Y de modo que melhore o ajuste de acordo com a avaliação da estatística F , e a introdução de variáveis para quando não consegue produzir uma estatística F maior de que o percentil 90 ou 95 da distribuição $F_{1, N-k-2}$, em que N é o tamanho amostral e k é o número de variáveis.

Já, a estratégia de seleção backward stepwise inicia com o modelo com todas as variáveis independentes, e sequencialmente, vai excluindo variáveis usando a estatística F para escolher os preditores a eliminar. O preditor que possui a menor estatística F é eliminado e o processo para quando cada preditor eliminado do modelo tem um valor F maior que o percentil de 90 ou 95 da distribuição $F_{1, N-k-2}$. No caso deste trabalho utilizamos o método de forward backward e a estatística de Wald.

Na regressão logística ordinal a significância é assegurada através do TRV (Teste Razão de Verossimilhanças). Assim, em cada passo do procedimento a variável mais importante, em termos estatísticos, é aquela que produz a maior mudança no logaritmo da verossimilhança em relação ao modelo que não contém a variável (Hastie et al., 2009).

Após os parâmetros terem sido estimados, o próximo passo é avaliar se as covariáveis utilizadas e disponíveis para a modelagem são estatisticamente significantes com o evento modelado, como por exemplo, a condição de um indivíduo ser pessoa com deficiência.

Uma forma de testar a significância do coeficiente de uma determinada covariável é comparar os valores observados da variável resposta com os valores preditos obtidos pelos modelos com e sem a variável de interesse (Oliveira, 2008).

A comparação entre os valores observados e preditos é feito a partir do teste de razão de verossimilhanças que é largamente aplicável com a estimação de máxima verossimilhança.

Para testar $H_0 : \theta \in \Theta_0$ versus $H_a : \theta \in \Theta_0^c$ calculamos a estatística (Casella and Berger, 1990):

$$\lambda(x) = \frac{\sup_{\theta_0} L(\theta/x)}{\sup_{\theta} L(\theta/x)}. \text{ Para } n \rightarrow \infty, -2 \ln \lambda(x) \rightarrow \chi_{\nu}^2.$$

onde, ν é obtido através da diferença entre a quantidade de parâmetros existente no modelo testado e a quantidade de parâmetros existente no modelo saturado (Oliveira, 2001).

A qualidade do modelo ajustado pode ser verificado através da comparação entre os valores observados e os valores preditos para a variável resposta (neste caso pode ser uma das diferentes deficiências já mencionadas).

Na escolha de um particular modelo, se por um lado, devemos tentar incluir tantas variáveis independentes quanto possíveis para melhorar a previsão, por outro, desejamos incluir um número mínimo de variáveis por problemas de custo e simplicidade (Oliveira, 2008).

Segundo Draper e Smith (1998) define seleção do melhor modelo como o compromisso de conciliar estes dois objetivos (incorporar uma certa quantidade de variáveis que possa melhorar a previsibilidade do modelo, ao mesmo tempo, descartar variáveis que não são significantes como uma forma de simplificar o modelo e diminuir custos). Esta seleção envolve uma dose de subjetividade o resultado pode ser diferente se alterar o procedimento utilizado para seleção.

2.3.6. Seleção de modelos

Selecionar um modelo significa, após a formulação e ajuste de diferentes modelos plausíveis, selecionar o modelo que "melhor" se ajusta aos dados de um certo experimento de acordo com um determinado critério adotado (Camarinha Filho, 2002).

Em Estatística existe uma vasta literatura pertinente à seleção de modelos (Broman, 1997; Burnham and Anderson, 1998,2002). Uma alternativa para seleção de modelos é a utilização de métodos baseados na função de verossimilhança que fornece diversas medidas estatísticas que auxiliam na comparação entre diferentes modelos. As mais comuns dessas medidas são AIC(do inglês, *Aikake Information Criterion*) proposto por Paulino et al. (2003) e Sakamoto et al. (1986) com penalização dada descontando o valor do dobro da diferença

entre o número de parâmetros entre os dois modelos; BIC (do inglês, *Bayesian Information Criterion*) discutido por Paulino et al. (2003) e tendo como penalização o valor do dobro do número de parâmetros entre os dois modelos multiplicado pelo logaritmo neperiano do tamanho amostral; e, por fim, DIC (do inglês, *Deviation Information Criterion*) discutido também por Paulino et al. (2003) e a penalização é dada pela soma do valor da diferença entre o número de parâmetros entre os dois modelos.

Neste trabalho, utilizamos os critérios AIC, BIC e DIC escolhendo o modelo com menor valor para cada um desses critérios.

2.4. Risco deficiência

Segundo a OMS:

- A prevalência de pessoas com deficiência é alta;
- O número de pessoas com deficiência aumenta devido ao envelhecimento da população e por conta da melhoria em termos globais das condições crônicas de saúde associada à deficiência tais como diabetes, doenças cardiovasculares e doenças mentais;
- Experiências diversas em que a deficiência resultante de interação entre condições de saúde, fatores pessoais e ambientais variando largamente; e, por fim,
- Populações vulneráveis cuja a prevalência varia com as condições de cada país, poder aquisitivo, condições de trabalho e nível de formação. Fatores como esses são considerados como riscos para que as pessoas sejam portadoras de deficiência, que por sua vez pode agravar as situações citadas anteriormente.

Diante deste cenário, tem surgido razões que justifiquem a necessidade de avaliar o bem estar ou qualidade de vida das pessoas com deficiência, propomos a criação do *índice risco pessoa com deficiência* composto pela ponderação das respostas das diferentes variáveis obtidas a partir dos microdados do Censo do IBGE e selecionadas como significativas após aplicar a metodologia stepwise backward em ajuste de regressão logística ordinal do tipo estereótipo para cada deficiência estudada. Essa metodologia surgiu gradualmente a partir de técnicas mais simples a técnicas mais complexas como as multivariadas tal como análise fatorial.

3 Resultados e discussões

Para este trabalho, utilizamos análise de regressão logística ordinal para cada uma das seguintes variáveis respostas:

- Deficiências, que representa o número de deficiências que cada pessoa possui e que pode assumir um valor entre 0 a 4 deficiências;
- Deficiências para enxergar, ouvir e movimentar considerando as categorias: 0, “*para quem não consegue de modo algum*”, 1, “*para quem consegue, mas com muita dificuldade*”, 2, “*para quem consegue, mas com um pouco de dificuldade*”, e, 3, “*para quem não apresenta problema*”; e, por fim;
- Deficiência Intelectual considerando as categorias “*ter*” ou “*não ter*”.

Para cada um dos seguintes blocos: identificação, educação, família e trabalho e para o modelo formado por todas as variáveis significativas, e, para os modelos ajustados em cada bloco foram aplicados:

- a) Métodos de seleção de variáveis stepwise backward e foram excluídas as variáveis que não foram consideradas significativas pelo teste de Wald em cada passo;
- b) Repetição da análise até que não existam mais variáveis para serem excluídas;
- c) Para cada um desses ajustes foi calculado critérios de seleção de modelos AIC, BIC e DIC;
- d) Seleção do melhor modelo entre os diferentes modelos finais para cada uma das diferentes deficiências e número de deficiências para cada um dos seguintes critérios: AIC, BIC e DIC, e, por fim;
- e) Foi calculado o risco de cada indivíduo ser portador de deficiência para os diferentes graus de severidade e número de deficiências.

Nas figuras 1-8 mostram no item a) os gráficos de risco de ser pessoa com uma (representada por p1 em pontos azuis), duas (representada por p2 em pontos vermelhos), três (representada por p3 em pontos verdes), quatro deficiências (representada por p4 em pontos roxos) e pelo menos uma deficiência (representada por pt em pontos pretos) e no item b) de ser pessoa com deficiência visual para cada um dos diferentes graus de severidade: “*não consegue enxergar de modo algum*” (representada por p3 em pontos azuis); “*consegue, mas com muita dificuldade*” (representada por p2 em pontos vermelhos); “*consegue, mas com um pouco de dificuldade*” (representada por p1 em pontos verdes), e, finalmente; “*de ser pessoa com uma deficiência visual*” (representada por pt em pontos roxos) para as variáveis: região na Figura 1, sexo na Figura 2, idade na Figura 3, raça na Figura 4, nível de instrução na Figura 5, trabalho principal na Figura 6, renda categorizada na Figura 7 e número de filhos na Figura 8.

Para a Figura 1, as regiões consideradas foram: 1 – “norte”, 2 – “nordeste”, 3 – “sudeste”, 4 – “sul” e 5 – “centro oeste”.

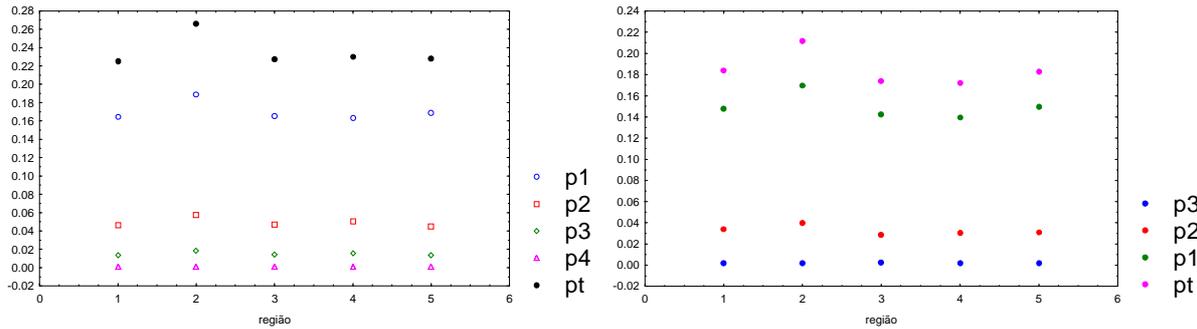


Figura 1. Gráficos de probabilidade de ocorrência: a) de um determinado número de deficiências, b) de ocorrência de deficiência visual de acordo com os seus graus de severidades para variável região.

Nos gráficos da Figura 1 notamos que os maiores riscos de incidência no item a) de uma deficiência e no item b) de deficiência visual encontra-se na região nordeste para todas as diferentes quantidade de deficiência e todos os diferentes graus de severidade. Ao passo que os menores índices de incidência em a) número de deficiências encontra-se na região centro oeste e em b) a menor incidência de risco de deficiência visual encontra-se na região sul.

A Figura 2 apresentam em a) os riscos de ser pessoa com deficiência e em b) o risco de incidência de ser pessoa com deficiência visual considerando os sexos 1 – masculino e 2 – feminino.

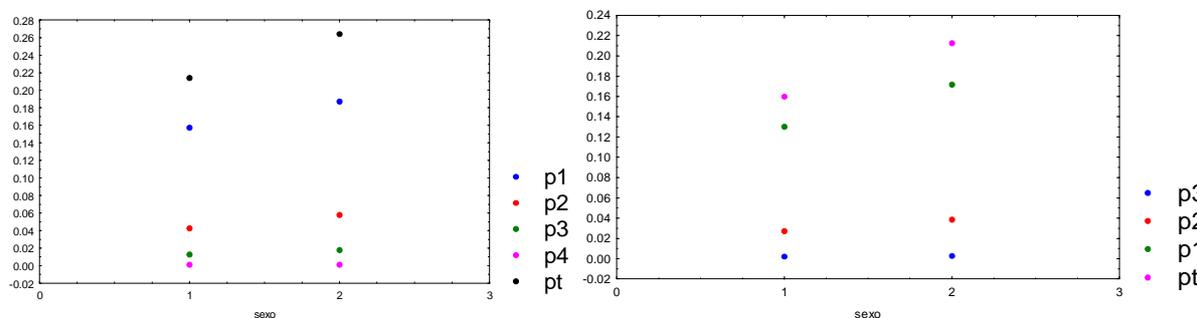


Figura 2. Gráficos de probabilidade de ocorrência: a) de um determinado número de deficiências, b) de ocorrência de deficiência visual de acordo com os seus graus de severidades para variável sexo.

Pelos gráficos da Figura 2 verifica-se que em todos os casos, o maior risco de incidência de a) deficiência e b) deficiência visual são maiores para o sexo feminino.

Já, a Figura 3 apresenta os riscos de incidência de a) deficiência e b) deficiência visual em função da idade.

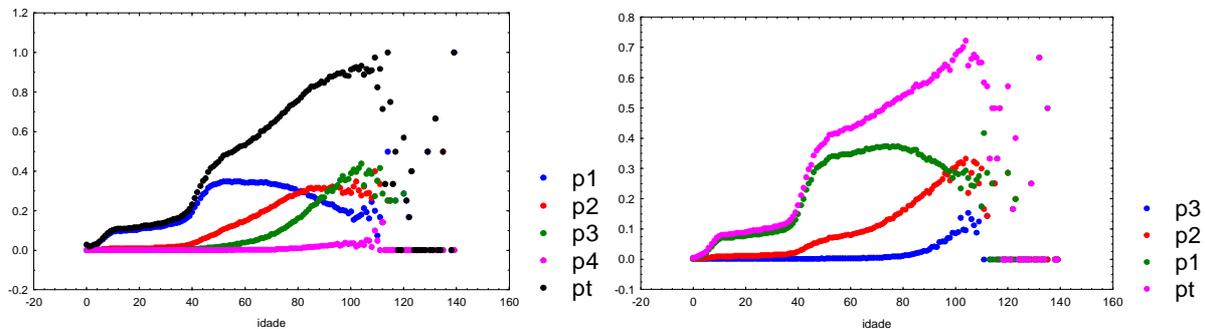


Figura 3. Gráficos de probabilidade de ocorrência a) de um determinado número de deficiências, b) de ocorrência de deficiência visual de acordo com os seus graus de severidades para variável idade.

Na Figura 3 é possível notar que os riscos de deficiência em a) e deficiência visual em b) aumentam na medida em que a idade das pessoas entrevistadas aumentam.

Nota-se também que, após uma determinada idade os pontos começam se aleatorizar, este tipo de ocorrência. Acredita-se que é devido a uma menor quantidade de pessoas entrevistadas em faixas etárias mais altas e isto ocorre principalmente com idades superiores a 80 anos.

Na Figura 4 consideramos as raças 1 - branca, 2 - preta, 3 - amarela, 4 - parda e 5 - indígena.

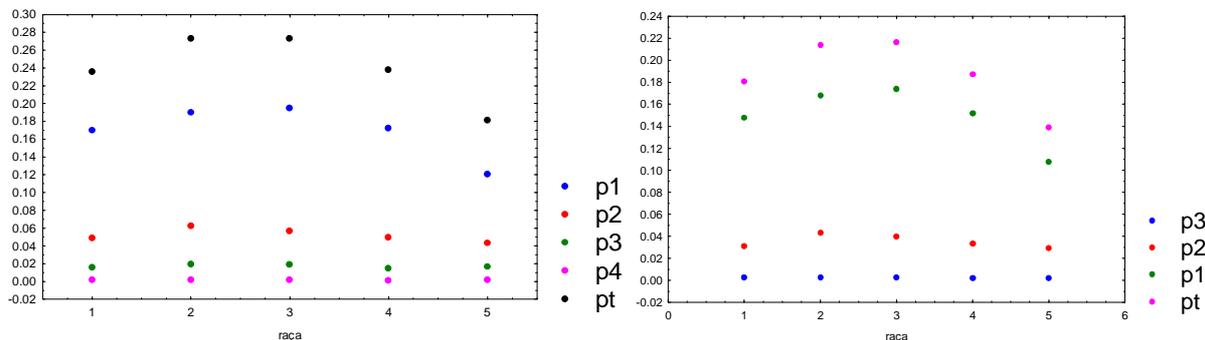


Figura 4. Gráficos de probabilidade de ocorrência a) de um determinado número de deficiências, b) de ocorrência de deficiência visual de acordo com os seus graus de severidades para variável raça.

Dos resultados da Figura 4 notamos que maiores probabilidades de incidência de deficiência e deficiência visual encontra-se em raça amarela e menor incidência em raça indígena.

Já, para a Figura 5, consideramos os seguintes níveis de instrução: 1 - “entre sem instrução e nível fundamental incompleto”, 2 - “entre fundamental completo e nível médio incompleto”, 3 - “entre nível médio completo e superior incompleto”, e, por fim, 4 - “nível superior completo ou mais”.

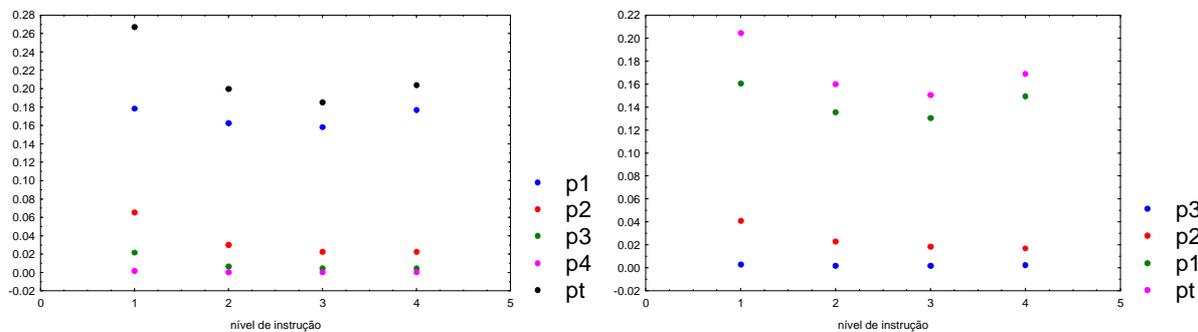


Figura 5. Gráficos de probabilidade de ocorrência a) de um determinado número de deficiências, b) de ocorrência de deficiência visual de acordo com os seus graus de severidades para nível de instrução.

Na Figura 5 observamos que a maior incidência de risco deficiência e de risco deficiência visual é encontrada em 1, “entre sem instrução e nível fundamental incompleto”, enquanto que a menor incidência desses riscos é encontrada em 3, “entre nível médio completo e superior incompleto” em todas as situações.

Para a Figura 6 consideramos os seguintes níveis para trabalho principal: 1 – “empregados com carteira de trabalho assinada”, 2 – “militares e funcionários públicos estatutários”, 3 – “empregados sem carteira de trabalho assinada”, 4 – “conta própria”, 5 – “empregadores”, 6 – “não remunerados”, 7 – “trabalhadores na produção para o próprio consumo”, e, por fim, 8 – “total”.

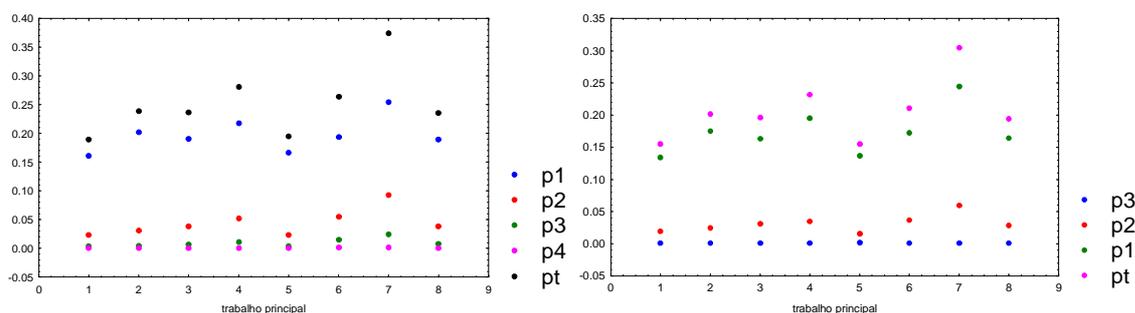


Figura 6. Gráficos de probabilidade de ocorrência a) de um determinado número de deficiências, b) de ocorrência de deficiência visual de acordo com os seus graus de severidades para trabalho principal.

Analisando os gráficos da Figura 7, verificamos que os maiores riscos de incidência de deficiência e de deficiência visual encontram-se em 7, “trabalhadores na produção para o próprio consumo” e o menor risco de incidência em ambos os casos foram encontrados em 2, “empregados com carteira de trabalho assinada”.

Para a situação de renda na Figura 7 estabelecemos: 1 – “entre 0 e 1 salário mínimo”, 2 – “entre 1 e 3 salários mínimos”, 3 – “entre 3 e 7 salários mínimos”, 4 – “entre 7 e 15 salários mínimos”, e, por fim, 5 – “15 salários mínimos ou mais”.

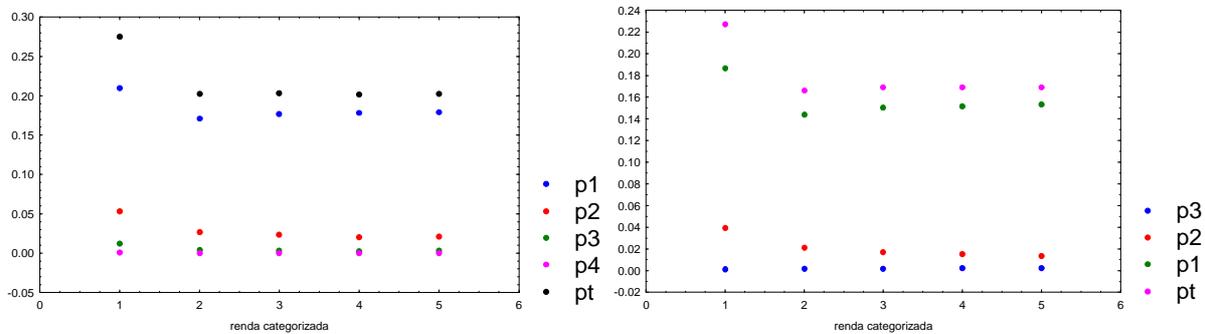


Figura 7. Gráficos de probabilidade de ocorrência a) de um determinado número de deficiências, b) de ocorrência de deficiência visual de acordo com os seus graus de severidades para renda categorizada.

Pelos resultados obtidos nos gráficos da Figura 7 percebemos que o maior risco de incidência de deficiência e de deficiência visual foi encontrada em 1, “entre 0 e 1 salário mínimo” e nota-se que esse risco vai diminuindo na medida em que aumenta a renda da pessoa entrevistada.

Finalmente, na Figura 8 foi feito o gráfico de dispersão para o risco de incidência de deficiência e de deficiência visual em função do número de filhos.

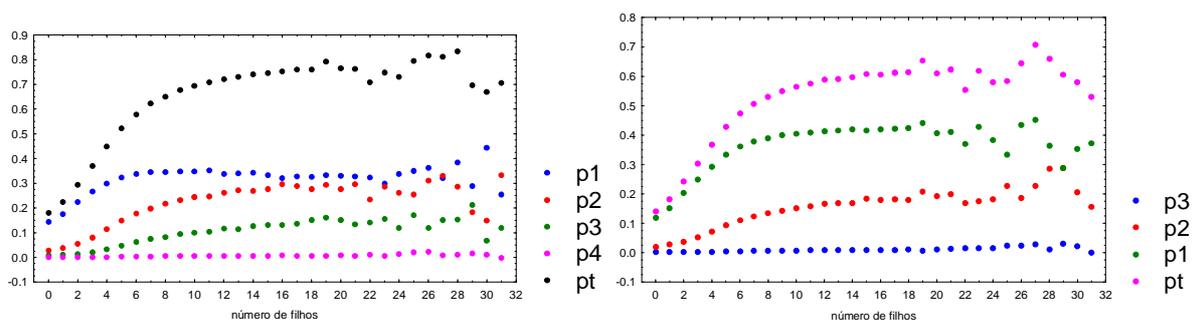


Figura 8. Gráficos de probabilidade de ocorrência a) de um determinado número de deficiências, b) de ocorrência de deficiência visual de acordo com os seus graus de severidades para número de filhos.

Neste caso da Figura 8, em geral nota-se que os riscos de incidência de deficiência e de deficiência visual aumenta na medida que aumenta a quantidade de filhos.

Este resultado pode refletir situações como: maior número de filhos pode significar um maior número de acidentes e menor atenção dos pais para cada filho em termos sociais e econômicos.

As tabelas de 1-5 mostram os resultados pelas análises de regressão logística ordinal estereótipo; critérios de seleção de modelos AIC, BIC e DIC e as estimativas pontuais e intervalares dos parâmetros considerando como variável resposta. O número de diferentes deficiências (Tabela 1), deficiências visual (Tabela 2), auditiva (Tabela 3), caminhar (Tabela 4) e intelectual (Tabela 5) que estão todas marcadas em negrito, bem como as variáveis explicativas incluídas no modelo final para cada uma das diferentes deficiências considerativas significativas segundo o método de stepwise-backward.

Para a variável Número de Deficiências obtemos como significativas as seguintes variáveis predictoras como ajuste para cada um dos diferentes blocos:

Tabela 1. Estimativas pontuais e intervalares dos parâmetros do modelo logístico considerando como variável resposta o número de deficiências (deficiências).

variáveis		Estimativa	Erro padrão	Wald	df	Sig.	Intervalo de confiança 95%	
							Limite inferior	Limite superior
deficiências	0	-.210	.075	7.786	1	.005	-.358	-.063
	1	1.922	.075	650.417	1	.000	1.774	2.070
	2	3.938	.077	2636.890	1	0.000	3.787	4.088
	3	7.034	.104	4562.783	1	0.000	6.830	7.238
região	1	.250	.013	365.872	1	.000	.225	.276
	2	.295	.011	714.216	1	.000	.273	.317
	3	-.071	.010	51.179	1	.000	-.090	-.051
	4	-.181	.011	281.285	1	.000	-.202	-.160
	5	0			0			
naturalidade	1	-.060	.006	97.127	1	.000	-.072	-.048
	2	.076	.014	31.279	1	.000	.049	.103
	3	0			0			
ler e escrever	1	-.428	.015	869.893	1	.000	-.456	-.399
	2	0			0			
creche	1	-.022	.023	.953	1	.329	-.066	.022
	2	-.114	.024	21.574	1	.000	-.162	-.066
	3	-.012	.019	.412	1	.521	-.050	.025
	4	0			0			
condição de ocupação	1	.086	.059	2.130	1	.144	-.029	.201
	2	-.205	.059	12.091	1	.001	-.320	-.089
	3	-.405	.059	47.319	1	.000	-.520	-.289
	4	-.486	.059	67.149	1	.000	-.602	-.370
	5	0 ^a			0			
nível de instrução	1	.037	.014	7.143	1	.008	.010	.063
	2	-.060	.014	17.007	1	.000	-.088	-.031
	3	-.006	.017	.118	1	.731	-.039	.028
	4	0			0			
natureza da união	1	.297	.014	473.400	1	.000	.270	.324
	2	.480	.024	411.506	1	.000	.434	.527
	3	.561	.016	1209.135	1	.000	.529	.592
	4	.819	.023	1244.424	1	.000	.773	.864
	5	0			0			
estado civil	1	-1.055	.015	5009.257	1	0.000	-1.084	-1.026
	2	-.909	.013	5024.643	1	0.000	-.934	-.884
	3	-.509	.013	1621.905	1	0.000	-.534	-.485
	4	0			0			
renda	1	.304	.032	89.934	1	.000	.241	.367
	2	.171	.032	29.101	1	.000	.109	.233
	3	.170	.032	28.047	1	.000	.107	.233
	4	.133	.035	14.378	1	.000	.064	.202
	5	0 ^a			0			
retorno	1	-.221	.017	174.562	1	.000	-.254	-.188
	2	0			0			
trabalho principal	1	-.209	.023	79.698	1	.000	-.255	-.163
	2	-.460	.135	11.602	1	.001	-.724	-.195
	3	.066	.025	7.006	1	.008	.017	.114
	4	-.131	.023	31.478	1	.000	-.176	-.085
	5	-.066	.024	7.620	1	.006	-.113	-.019
	6	-.428	.032	180.577	1	.000	-.491	-.366
	7	0			0			

Identificação: domicílio, idade categorizada, naturalidade, nacionalidade e região; **Educação:** ler e escrever, creche, outra graduação e nível de instrução; **Família:** natureza da união, estado civil, número de filhos categorizados; **Trabalho:** renda, trabalho secundário, trabalho principal, tempo de deslocamento e retorno; e, por fim; **Modelo conjunto:** (Tabela 1) (formado a partir da junção de todas as variáveis preditoras que foram consideradas significativas em cada um dos blocos anteriores): região, naturalidade, ler e escrever, creche condição de ocupação, nível de instrução, natureza da união, estado civil, número de filhos categorizados, renda, retorno e trabalho principal. Para seleção de modelos obtemos: AIC = -7232.953, BIC = -8791.418 e DIC = -6917.953.

Já em Deficiência Visual, foram selecionadas para cada bloco: **Identificação:** região, domicílio, sexo, naturalidade e nacionalidade; **Educação:** ler e escrever, creche, outra graduação e nível de instrução; **Família:** natureza da união, estado civil e número de filhos categorizados; **Trabalho:** renda, tempo, condição, situação e trabalho secundário, e, por fim; **Modelo conjunto:** (Tabela 2) inicializado com todas as variáveis explicativas que foram consideradas significantes para cada um dos diferentes blocos e foram selecionadas: região, naturalidade, ler e escrever, creche, nível de instrução, natureza da união, número de filhos categorizados, retorno, condição e situação. Para seleção de modelos. Para seleção de modelos obtemos: AIC = -2549.708, BIC = -3291.833 e DIC = -2399.707.

Para Deficiência Auditiva foram selecionadas as seguintes variáveis para cada bloco: **Identificação:** região, domicílio, sexo, raça e naturalidade; **Educação:** ler e escrever, creche, outra graduação e nível de instrução; **Família:** natureza da união, estado civil e número de filhos categorizados; **Trabalho:** renda, tempo, condição, situação, trabalho principal e trabalho secundário, e, por fim; **Modelo conjunto:** (Tabela 3) neste caso foram selecionadas as seguintes variáveis: região, naturalidade, ler e escrever, nível de instrução, estado civil, número de filhos categorizados, condição e situação. Para seleção de modelos obtemos: AIC = -2921.348, BIC = -3331.401 e DIC = -2865.348.

Para Deficiência Locomotora obtemos: **Identificação:** região, idade categorizada, e naturalidade; **Educação:** ler e escrever, creche, outra graduação e nível de instrução; **Família:** natureza da união, estado civil e número de filhos categorizados; **Trabalho:** renda, retorno, tempo, condição, situação, trabalho principal e trabalho secundário, e, por fim; **Modelo conjunto:** (Tabela 4) região, naturalidade, ler e escrever, creche, nível de instrução, estado civil, número de filhos, retorno, tempo, condição, situação e trabalho principal. Para seleção de modelos obtemos: AIC = -1258.613, BIC = -2119.480 e DIC = -1084.013.

Tabela 2. Estimativas pontuais e intervalares dos parâmetros do modelo logístico considerando como variável resposta deficiência visual

variáveis	Estimativa	Erro padrão	Wald	df	Sig.	Intervalo de confiança 95%		
						Limite inferior	Limite superior	
deficiência visual	1	-5.190	.072	5147.165	1	0.000	-5.332	-5.048
	2	-2.012	.066	935.571	1	.000	-2.141	-1.883
	3	.177	.066	7.310	1	.007	.049	.306
	4	10.966	.140	6174.537	1	0.000	10.692	11.239
região	1	-.272	.013	421.476	1	.000	-.298	-.246
	2	-.262	.011	544.896	1	.000	-.284	-.240
	3	.141	.010	193.114	1	.000	.121	.161
	4	.275	.011	609.290	1	.000	.253	.297
	5	0			0			
naturalidade	1	.051	.006	67.199	1	.000	.039	.064
	2	-.077	.014	30.191	1	.000	-.105	-.050
	3	0			0			
ler e escrever	1	.369	.014	710.297	1	.000	.342	.396
	2	0			0			
creche	1	-.041	.022	3.940	1	.049	-.084	.002
	2	.043	.024	3.098	1	.078	-.005	.091
	3	-.018	.019	.931	1	.335	-.054	.018
	4	0			0			
nível de instrução	1	-.062	.060	1.064	1	.302	-.181	.056
	2	.226	.060	14.015	1	.000	.108	.345
	3	.425	.060	49.529	1	.000	.307	.544
	4	.460	.061	57.577	1	.000	.341	.579
	5	0			0			
natureza da união	1	-.223	.007	1108.625	1	.000	-.236	-.210
	2	-.135	.008	260.276	1	.000	-.151	-.118
	3	-.067	.016	17.005	1	.000	-.099	-.035
	4	0			0			
número de filhos	1	1.101	.015	5656.802	1	0.000	1.072	1.130
	2	.922	.012	5599.102	1	0.000	.898	.946
	3	.498	.012	1696.917	1	0.000	.474	.522
	4	0			0			
retorno	1	.213	.016	166.427	1	.000	.180	.245
	2	0			0			
condição	1	0			0			
situação	1	0			0			

Fazendo um estudo comparativo entre os modelos contidos nas tabelas 1-5 notamos que o modelo que incluiu uma menor quantidade de variáveis foi o modelo logístico ajustado para deficiência intelectual, enquanto que, o modelo que exigiu a maior quantidade de variáveis independentes foi para número de deficiências.

Tabela 3. Estimativas pontuais e intervalares dos parâmetros do modelo logístico considerando como variável resposta deficiência auditiva

variáveis		Estimativa	Erro padrão	Wald	df	Sig.	Intervalo de confiança 95%	
							Limite inferior	Limite superior
rdeficiência auditiva	1	-6.251	.081	5885.079	1	0.000	-6.410	-6.091
	2	-4.350	.079	3058.665	1	0.000	-4.504	-4.196
	3	-2.495	.078	1017.060	1	.000	-2.648	-2.342
	4	12.767	.288	1962.330	1	0.000	12.202	13.332
região	1	-.098	.018	29.749	1	.000	-.133	-.063
	2	-.299	.015	401.518	1	.000	-.329	-.270
	3	-.015	.014	1.204	1	.273	-.043	.012
	4	.005	.015	.131	1	.718	-.024	.035
	5	0			0			
naturalidade	1	.039	.008	22.219	1	.000	.023	.055
	2	-.110	.018	38.663	1	.000	-.145	-.075
	3	0			0			
ler e escrever	1	.449	.012	1302.286	1	.000	.424	.473
	2	0			0			
nível de instrução	1	-.445	.075	35.178	1	.000	-.592	-.298
	2	-.119	.075	2.475	1	.116	-.266	.029
	3	.123	.075	2.688	1	.101	-.024	.271
	4	.292	.076	14.760	1	.000	.143	.441
	5	0			0			
estado civil	1	-.122	.009	177.545	1	.000	-.139	-.104
	2	-.395	.021	352.262	1	.000	-.436	-.354
	3	-.468	.016	824.993	1	.000	-.500	-.436
	4	-.816	.014	3433.926	1	0.000	-.843	-.789
	5	0			0			
número de filhos condição	1	.826	.016	2645.864	1	0.000	.795	.858
	2	.709	.013	2870.222	1	0.000	.683	.735
	3	.432	.012	1207.655	1	.000	.407	.456
	4	0			0			
condição	1	.042	.014	9.356	1	.002	.015	.068
	2	0			0			
situação	1	0			0			
	2	0			0			

4 Conclusão

O modelo que necessitou de um menor número de variáveis explicativas foi mental ou intelectual com 5, enquanto que, o que necessitou de maior quantidade foi para deficiências com 13.

Tabela 4. Estimativas pontuais e intervalares dos parâmetros do modelo logístico considerando como variável resposta deficiência para caminhar

variáveis		Estimativa	Erro padrão	Wald	df	Sig.	Intervalo de confiança 95%	
							Limite inferior	Limite superior
deficiência para andar	1	-5.591	.125	1987.299	1	0.000	-5.837	-5.345
	2	-2.726	.120	512.061	1	.000	-2.962	-2.490
	3	-1.146	.120	90.832	1	.000	-1.382	-.911
	4	13.027	.234	3110.980	1	0.000	12.569	13.485
região	1	-.182	.022	71.172	1	.000	-.224	-.139
	2	-.370	.018	436.802	1	.000	-.404	-.335
	3	-.086	.016	28.375	1	.000	-.118	-.055
	4	-.047	.018	7.112	1	.008	-.082	-.013
	5	0			0			
naturalidade e ler e escrever	1	.012	.010	1.576	1	.209	-.007	.031
	2	-.157	.020	60.058	1	.000	-.197	-.117
	3	0			0			
creche	1	.333	.029	133.621	1	.000	.277	.390
	2	.586	.039	220.779	1	.000	.509	.664
	3	-.025	.022	1.311	1	.252	-.068	.018
	4	0			0			
nível de instrução	1	-.569	.109	27.157	1	.000	-.783	-.355
	2	-.104	.109	.908	1	.341	-.319	.110
	3	.267	.109	5.963	1	.015	.053	.482
	4	.538	.110	23.881	1	.000	.322	.754
	5	0			0			
estado civil	1	-.125	.010	142.810	1	.000	-.146	-.105
	2	-.530	.022	585.604	1	.000	-.572	-.487
	3	-.621	.017	1326.137	1	.000	-.654	-.588
	4	-.933	.016	3244.635	1	0.000	-.965	-.901
	5	0			0			
número de filhos categorizados	1	1.009	.020	2554.328	1	0.000	.970	1.049
	2	.794	.016	2471.085	1	0.000	.763	.825
	3	.382	.015	639.566	1	.000	.352	.411
	4	0			0			
retorno	1	0			0			
tempo	1	.691	.031	506.529	1	.000	.631	.751
	2	.673	.029	524.780	1	.000	.615	.730
	3	.508	.030	283.211	1	.000	.449	.568
	4	.284	.032	80.361	1	.000	.222	.346
	5	0			0			
condição	1	0			0			
situação	1	0			0			
trabalho principal	1	.542	.033	266.774	1	.000	.477	.607
	2	.497	.241	4.229	1	.040	.023	.970
	3	-.010	.036	.076	1	.783	-.080	.060
	4	.415	.033	158.503	1	.000	.351	.480
	5	.216	.034	40.930	1	.000	.150	.283
	6	.623	.053	139.555	1	.000	.520	.727
	7	0			0			

Tabela 5. Estimativas pontuais e intervalares dos parâmetros do modelo logístico considerando como variável resposta deficiência intelectual

variáveis		Estimativa	Erro padrão	Wald	df	Sig.	Intervalo de confiança 95%	
							Limite inferior	Limite superior
deficiência intelectual	1	-3.795	.087	1923.902	1	0.000	-3.964	-3.625
	2	10.498	.096	11945.597	1	0.000	10.310	10.686
sexo	1	-.103	.006	284.656	1	.000	-.115	-.091
	2	0			0			
idade categorizada	1	.664	.013	2719.203	1	0.000	.639	.689
	2	.073	.008	84.007	1	.000	.057	.089
	3	0			0			
naturalidade	1	-.139	.007	454.576	1	.000	-.152	-.126
	2	-.251	.015	274.087	1	.000	-.281	-.221
	3	0			0			
ler e escrever	1	1.486	.007	46829.311	1	0.000	1.473	1.500
	2	0			0			
nível de instrução	1	-.943	.086	120.203	1	.000	-1.112	-.775
	2	-.303	.087	12.254	1	.000	-.473	-.133
	3	.116	.087	1.802	1	.179	-.054	.286
	4	.362	.089	16.647	1	.000	.188	.535
	5	0			0			

Neste trabalho, ao utilizar modelo logístico ordinal estereótipo com resposta ordinal, foi possível melhorar a qualidade de ajuste, ao ser comparado com o ajuste por modelo logístico proposto em Oliveira, (2013) com resposta binária. Ao utilizar a resposta ordinal, foi incorporado o risco deficiência para os diferentes graus de severidade e quantidade de deficiências.

As diferentes deficiências não são homogêneas quanto as diferentes variáveis predictoras.

Os riscos de incidência de ser pessoa com deficiência e de ser pessoa com deficiência visual são provavelmente maiores em situações como: morar na região nordeste, sexo feminino, idade acima de 80 anos, raça amarela, nível de instrução de no máximo fundamental incompleto, trabalha na produção para o próprio consumo, e possui um número elevado de filhos.

Os menores riscos de incidência são notados em situações com morar na região sul, sexo masculino, idade de no máximo 15 anos, raça indígena, nível de instrução entre nível médio completo e superior incompleto, trabalhador com carteira assinada, e não possui filhos.

A seguir, para as Figuras 1-8 passemos a estabelecer possíveis justificativas e sugestões de trabalhos ou pesquisas que possam aceitar ou não as hipóteses consideradas.

- Figura 1, Estes resultados pode ser justificados pelo baixo investimento efetivo em termos de saúde e infraestrutura que são menores nas regiões nordeste e maiores nas outras regiões como sudeste e sul.

Para que possa, melhor avaliar esta hipótese, pode fazer um levantamento do volume efetivo gasto em cada uma destas regiões em saúde, acessibilidade e infraestrutura que beneficiam pessoas com deficiência entre as diferentes regiões contabilizando a quantidade de pessoas que foram efetivamente beneficiadas e fazer um estudo comparativo entre as diferentes regiões;

- Figura 2: Muito provavelmente, estes resultados refletem maior exposição das mulheres a acidentes domésticos e a dupla jornada da mulher moderna em trabalhar fora e cuidar da família.

Para melhor avaliar este ponto pode ser feito um estudo por amostragem sobre os tempos de trabalho em casa e fora de casa para os homens e para as mulheres e fazer um estudo comparativo;

- Figura 3: Estes resultados mostram que com o decorrer do tempo, a população encontra-se mais envelhecida com maior expectativa de vida e conseqüentemente mais sujeita as doenças da idade avançada e com maior incidência de serem pessoas com deficiência.

Neste caso, poderemos sugerir que comprovam de um lado o aumento de expectativa de vida da população, e por outro lado, aparecimento de doenças que costumam ocorrer em idades mais avançadas e isso pode ser feito considerando os dados da Amostra do Censo IBGE 2010;

- Figura 4: Estes resultados podem ser reflexos das condições culturais e alimentares dos povos orientais e indígenas.

Para um melhor entendimento deste resultado, podemos sugerir um estudo de pesquisa sobre os hábitos de vidas de pessoas que pertençam as raças amarela e indígenas, considerando as possibilidades de se tornarem ou não pessoas com deficiência;

- Figura 5: Acreditam-se que estes resultados possam ser devido a fatores como: baixo nível de instrução pode significar menor conhecimento de informação, baixo poder aquisitivo e maior dependência dos auxílios governamentais.

Neste caso, para comprovar, sugerimos uma pesquisa que possa estabelecer relações entre nível de instrução e renda que muito provavelmente obteremos este tipo de resultado;

- Figura 6: Muito provavelmente, os diferentes tipos de funções refletem o nível de instrução obtido pelos diferentes trabalhadores, pois ser militar ou estatutário

dependem de aprovação em concurso público que necessita um nível de instrução maior, enquanto que trabalho no próprio consumo, em geral, é constituído por pessoas que trabalham no campo e com poder aquisitivo menor.

Para melhor avaliar e compreender esta hipótese sugerimos uma pesquisa que possa estabelecer as médias de remuneração para cada uma das diferentes profissões;

- Figura 7: A justificativa é baseada no fato de que a incidência de risco deficiência ou deficiência visual é menor quanto maior for o poder aquisitivo das pessoas.

Neste caso, sugerimos um estudo em que faça um levantamento de pessoas com deficiência e que não apresentam deficiência e a seguir, com deficiência visual e que não apresentam deficiência visual e que façamos a comparação entre os diferentes níveis de renda; e, finalmente;

- Figura 8: este resultado pode refletir situações como: maior número de filhos pode significar um maior número de acidentes e menor atenção dos pais para cada filho em termos sociais e econômicos.

Neste caso, o que pode ser interessante seria o estabelecimento de uma pesquisa que possa comparar a qualidade de vida entre famílias com diferentes números de filhos e fazer comparações entre os diferentes casos em termos de risco de deficiência.

Os resultados nas figuras 1-8 foram semelhantes para número de deficiências e deficiência visual.

Os resultados deste trabalho mostram, também, a necessidade de outros estudos, pesquisas e análises, pois, quando se fala em risco, nota-se a existência de numerosos métodos para determinar esse risco, seja utilizando os coeficientes de regressão, escores fatoriais ponderação do risco deficiência considerando a ponderação do risco para cada uma das diferentes variáveis explicativas. Por exemplo, sabe-se que o risco deficiência aumenta na medida em que aumenta a idade, o mesmo acontece com o número de filhos e assim por diante.

Entre outras alternativas que valem a pena ser considerada para trabalhos futuros, podemos citar:

1 - Considerar modelo de regressão beta, análise fatorial, modelagem de equações estruturais e o algoritmo BART que em todos esses casos oferecem oportunidades para trabalhos bem interessantes como forma de melhorar a qualidade de ajuste e sua confiabilidade na determinação desse índice.

2 – Repetir análise incluindo variáveis relacionadas a condições de moradia e de posse de outros bens como carro, televisão e microcomputador com internet.

3 - Entre as diversas questões que são necessárias a serem respondidas encontra-se questões a respeito de como as pessoas com deficiência vivem e em que situação elas se encontram ao serem comparadas a pessoas que não apresentam deficiência.

4 - Em situações como está um índice de risco com boa confiabilidade e qualidade de ajuste é de interesse para facilitar o monitoramento dessa situação, da mesma forma que ocorre com o IDH (Índice de Desenvolvimento Humano), muito embora, este último seja um índice mais geral, ainda não leva em conta a questão das pessoas com deficiência.

5 - Outro tópico de interesse seria avaliar a acessibilidade de entorno das moradias das pessoas com deficiência, considerando as localizações destas residências de forma georreferenciadas avaliando as condições da infraestrutura do entorno, para isto, poderemos propor uma modelagem geostatística para que possa incorporar questões como esta.

6 - O problema que existe quando resolve calcular um índice de risco é considerar um método que seja eficiente e confiável e que diminua o risco de discrepância, por motivos como estes, é de interesse que diversos métodos devem ser considerados.

7 - A vantagem em existir um índice como este é que possa ser comparável, ou seja, ele pode ser avaliado e ser estudado para entender por que o valor dele aumentou ou diminuiu, de tal maneira que, quanto maior for o índice, deve refletir uma maior necessidade de intervenção por parte dos poderes públicos para que possa diminuir as barreiras que existem em termos de acesso aos diferentes direitos humanos e de acessibilidade no entorno das moradias das pessoas com deficiência.

8 - Propor melhorias no questionário do censo do IBGE, por exemplo, se um respondente respondeu que é uma pessoa com deficiência perguntar também com que idade ela foi adquirida, pois, segundo a literatura existente (Garcia, 2010), sabe-se que, quanto maior a idade que tornou-se pessoa com deficiência, maiores são as dificuldades dessa pessoa se adaptar.

9 - Em termos estatísticos, necessitamos melhorar as estatísticas nacionais sobre a deficiência, empregando uma abordagem eficiente e de baixo custo para a obtenção de dados mais abrangentes e adicionar perguntas sobre deficiência, conexão estatística entre os diferentes conjuntos de dados, coleta de dados longitudinais, incluindo perguntas sobre a deficiência, para que possa fazer um melhor monitoramento, melhorar a comparabilidade dos dados, desenvolvimento de ferramentas apropriadas e preencher as lacunas entre as investigações e, por fim, fortalecer e apoiar as diferentes investigações considerando a criação de instrumentos que possam medir e monitorar a qualidade de vida e bem estar destas pessoas de forma contínua e periódica.

10 – Considerar a possibilidade de incluir também questões relacionadas a condições de saúde, acessibilidade e lazer.

11 – repetir a análise por região, estado e municípios.

12 - Resultados de trabalhos como este, pode auxiliar na ação de gestores públicos oferecendo um melhor suporte no atendimento das pessoas com deficiência.

5 Referências

- [1] ABREU, M.N.S. **Uso de modelos de regressão logística ordinal em epidemiologia: um exemplo usando a qualidade de vida.** Dissertação de Mestrado. Faculdade de Saúde Pública, UFMG, Belo Horizonte - MG, 2007.
- [2] Anderson JA. Regression and ordered categorical variables. **J R Statist Soc**, 16: 1-30, 1984
- [3] BROMAN, K. W. **Identifying Quantitative Trait Loci in Experimental Cross.** Thesis, University of California, Berkeley, 1997.
- [4] BURNHAM, K.P.; ANDERSON, D.R. **Model selection and multimodel inference.** Springer, New York, USA, 2002.
- [5] BURNHAM, K.P.; ANDERSON, D.R. **Model selection and inference.** Springer, New York, USA, 1998.
- [6] CAMARINHA-FIKHO, J.A. **Modelos lineares mistos: estimativas de matrizes de variância e covariâncias e seleção de modelos.** Tese de Doutorado, ESDALQ-USP, Piracicaba-SP, 2008.
- [7] CASELLA, G.; BERKER, P.L. **Statistical inference.** Wadsworth, Brooks, California, EUA, 1990.
- [8] DRAPER, N.R.; SMITH, H. **Applied regression analysis.** John Wiley, New York, 1998.
- [9] FIGUEIRA, E. **Caminhando em silêncio.** Giz editorial e Livraria Ltda., São Paulo-SP, 2008.
- [10] GARCIA, V.G. **Pessoas com deficiência e o mercado de trabalho.** Tese de Doutorado. Instituto de Economia – UNICAMP, Campinas – SP, 2010.
- [11] HASTIE, T.; TIBSHIRANI, P.; FRIEDMANN, J. **The elements of learning: data mining, inference and population.** Springer, Canada, 2009.
- [12] OLIVEIRA, P.T.M.S., 2013. **Pessoas com deficiência: análise dos resultados do Censo 2010 e a sua evolução.** In: 58 RBRAS/ 15 SEAGRO, no período entre 22 a 26 de julho de 2013, Campina Grande – PB, 2013.
- [13] OLIVEIRA, P.T.M.S. **Aplicação do algoritmo genético no mapeamento de genes epistáticos em cruzamentos controlados.** Tese de Doutorado, IME4-USP, São Paulo – SP, 2008.

- [14] PAULINO, C.D., TURKMAN, A.A., MURTEIRA, B.J.F., **Estatística Bayesiana**.
Fundação Calouste Gulbenkian, Portugal, 2003.
- [15] SAKAMOTO, Y.; ISHGURU, M.; KITAMURA, G., **Akaike information criterion statistics**. KTK, Scientific Publisher, Japão, 1986.
- [16] SILVA, O.M. **A epopeia ignorada**. CEDAS, São Paulo-SP, 1987.
- [17] OLIVEIRA, P.T.M.S. **Estimação e testes de hipóteses em calibração comparativa**.
Dissertação de mestrado, IME-USP, São Paulo – SP, 2001.