

Rosana Gayer Carvalho

**Análise do grau de dificuldade na locomoção  
dos idosos no Brasil utilizando o modelo de  
chances proporcionais parciais**

Niterói – RJ, Brasil

29 de julho de 2016



**Universidade Federal Fluminense**

Rosana Gayer Carvalho

**Análise do grau de dificuldade na locomoção  
dos idosos no Brasil utilizando o modelo de  
chances proporcionais parciais**

Trabalho de Conclusão de Curso

Projeto Final II apresentado para obtenção do grau de Bacharel em  
Estatística pela Universidade Federal Fluminense.

Orientador: Prof. José Rodrigo de Moraes

Niterói - RJ, Brasil

29 de julho de 2016



Universidade Federal Fluminense

Rosana Gayer Carvalho

# **Análise do grau de dificuldade na locomoção dos idosos no Brasil utilizando o modelo de chances proporcionais parciais**

Projeto Final II de graduação sob o título “*Análise do grau de dificuldade na locomoção dos idosos no Brasil utilizando o modelo de chances proporcionais parciais*”, defendido por Rosana Gayer Carvalho e aprovada em 29 de julho de 2016, na cidade de Niterói, no Estado do Rio de Janeiro, pela banca examinadora constituída pelos professores:

---

**Prof. Dr. José Rodrigo de Moraes**

Orientador

Departamento de Estatística – UFF

---

**Profa. Dra. Luciane Ferreira Alcoforado**

Departamento de Estatística – UFF

---

**Dra. Patrícia Viana Guimarães Flores**

Instituto de Estudos em Saúde Coletiva – UFRJ

Niterói, 29 de julho de 2016

C331 Carvalho, Rosana Gayer

Análise do grau de dificuldade na locomoção dos idosos do Brasil utilizando o modelo de chances proporcionais parciais / Rosana Gayer Carvalho. – Niterói,. RJ : [s.n.], 2016.

60 f.

Orientador: José Rodrigo de Moraes.

Trabalho de Conclusão de Curso (Bacharelado em Estatística) - Universidade Federal Fluminense, 2016.

4. Capacidade motora em idosos – Regressão logística. 2. Pesquisa Nacional de Saúde. 3. Estatística. I. Título.

CDD 519.5

## Resumo

A manutenção da saúde e da capacidade funcional na fase idosa é uma preocupação central no campo da saúde pública, tendo em vista os problemas gerados pela incapacidade funcional que afetam tanto o idoso quanto os seus familiares. A incapacidade funcional é definida pela dificuldade ou pela necessidade de ajuda na realização das tarefas básicas do cotidiano ou de tarefas mais complexas. Utilizando os dados de uma pesquisa com plano amostral complexo, denominada Pesquisa Nacional de Saúde 2013, este trabalho tem como objetivo analisar o grau de dificuldade na locomoção dos idosos no Brasil e estabelecer a sua associação com um conjunto de variáveis sociodemográficas e de saúde dos idosos de 60 anos ou mais de idade. Foi utilizado o modelo de regressão logística ordinal, também denominado modelo de chances proporcionais parciais, onde a variável resposta do modelo contém três categorias mutuamente exclusivas que representam diferentes graus de dificuldade na locomoção do idoso: “grave”, “no máximo moderado” e “nenhum”. Para o ajuste do modelo, foram considerados os pesos amostrais e as informações estruturais do plano amostral da pesquisa. Usando o teste de Wald, verificou-se que as seguintes variáveis apresentaram associação significativa com o grau de dificuldade na locomoção dos idosos, a um nível de significância de 5%: faixa-etária, situação conjugal, região de residência, nível de escolaridade, autoavaliação geral de saúde e diagnóstico de alguma doença crônica, física ou mental. Observou-se maior grau de dificuldade entre idosos mais velhos, não casados, sem instrução, residentes na região Sul e que reportaram pior estado de saúde. Além disso, observou-se que os idosos com diagnóstico de alguma doença crônica, física ou mental têm chance 64,5% maior de apresentar algum grau de dificuldade na locomoção (no máximo moderado ou grave), comparativamente aos idosos sem diagnóstico dessas doenças (OR=1,645; p-valor<0,001). Desse modo, é preciso garantir o atendimento das demandas específicas da população idosa e a realização de políticas públicas eficientes, a fim de contribuir para um envelhecimento populacional saudável, evitando assim uma saúde precária e a incapacidade funcional.

**Palavras-chaves:** Regressão logística ordinal. Pesquisa Nacional de Saúde. Idoso. Grau de dificuldade na locomoção. Qualidade de vida.

## **Dedicatória**

Dedico este trabalho a minha falecida e amada mãe, Elizabeth Maria Gayer Carvalho, por ter me incentivado e apoiado enquanto pôde estar ao meu lado, ao meu pai Ricardo de Mello Carvalho e aos meus irmãos, Renato Gayer Carvalho e Ricardo de Mello Carvalho Júnior, por me apoiarem na realização desse grande sonho.

## **Agradecimentos**

Agradeço primeiramente à minha falecida e amada mãe, Elizabeth Maria Gayer Carvalho, por todo incentivo e apoio. Ela não teve a oportunidade de presenciar a concretização deste nosso sonho, mas com toda certeza está feliz e torcendo por mim de onde estiver.

Ao meu pai Ricardo de Mello Carvalho por me guiar nessa caminhada e por me ter fornecido toda a assistência necessária durante todo o período da graduação. Eu posso dizer que a minha formação não seria possível sem você.

Aos meus irmãos Renato Gayer Carvalho e Ricardo de Mello Carvalho Júnior, pelo apoio e principalmente por entenderem a minha ausência em muitos momentos durante este período. E ao meu namorado Walter Camões Júnior, pelo apoio e companhia durante todo o período de estresses e estudos.

Agradeço a todos os meus amigos da vida, por tornar essa caminhada mais fácil e alegre. Aos amigos que fiz durante a graduação, especialmente aos que mais me ajudaram e incentivaram: Isabella Perrut, Lucas Meireles, Fernando Alencar, Roberto Costa, Yasmin Cavaliere, Natan Borges, Thainá Castro, Juliana Chaves, Vanessa Eufrazino, Rebecca Souza, Tiago Sales e William Alecsander, amigos da graduação para a vida.

Agradeço também a todos os professores que me acompanharam durante a graduação, especialmente, ao professor Guillermo Coca Velarde, ao professor Adrian Pizzinga e à professora Ludmilla Jacobson, pelos ensinamentos, conselhos e incentivos. À professora Luciane Alcoforado e a Dra. Patrícia Guimarães por terem aceitado a participar da banca deste trabalho.

Por fim, agradeço imensamente e especialmente, ao meu professor e orientador José Rodrigo de Moraes, pela ajuda, paciência e dedicação na orientação e conclusão deste trabalho. Agradeço por ter sido tão importante na minha vida acadêmica, onde também tive o prazer de ter a sua orientação na monitoria da disciplina obrigatória “Modelos Lineares I”, além de ter sido sua aluna em 3 disciplinas da graduação, “Estatística Geral II”, “Modelos Lineares I” e “Modelos Lineares II”.

## Sumário

<b>Lista de Figuras .....</b>	<b>10</b>
<b>Lista de Quadros.....</b>	<b>11</b>
<b>Lista de Tabelas .....</b>	<b>12</b>
<b>Lista de Siglas .....</b>	<b>13</b>
<b>1. Introdução .....</b>	<b>14</b>
1.1 Mensuração dos níveis de vida .....	14
1.2 Envelhecimento populacional no Brasil .....	14
1.3 Incapacidade funcional de idosos no Brasil.....	15
1.1 Modelo conceitual para incapacidade funcional de idosos .....	16
<b>2. Objetivos.....</b>	<b>19</b>
2.1 Objetivo geral.....	19
2.2 Objetivos específicos .....	19
<b>3. Materiais e métodos .....</b>	<b>20</b>
3.1 Pesquisa Nacional de Saúde (PNS) 2013.....	20
3.2 Questionário da PNS.....	21
3.2.1 Parte domiciliar .....	21
3.2.2 Parte geral.....	21
3.2.3 Parte individual .....	21
3.3 Plano amostral da PNS.....	22
3.3.1 Sistema Integrado de Pesquisas Domiciliares .....	22
3.3.2 Amostra mestra .....	23
3.4 População de estudo.....	24
3.5 Variáveis analisadas.....	24
3.5.1 Grau de dificuldade na locomoção .....	24
3.5.2 Variáveis sociodemográficas e de saúde .....	24
3.6 Modelo de regressão logística ordinal .....	26
3.6.1 Modelo de chances proporcionais .....	26
3.6.2 Modelo de chances proporcionais parciais .....	32
3.6.3 Método de Máxima Pseudo-Verossimilhança (MPV). .....	34
3.6.4 Teste de homogeneidade geral dos coeficientes .....	39
3.6.5 Teste de Wald de significância individual dos parâmetros do modelo .....	40



3.6.6 Intervalo de confiança para os parâmetros e razões de chance .....	42
3.6.7 Medidas de qualidade do ajuste .....	43
<b>4. Resultados .....</b>	<b>46</b>
<b>5. Discussões e conclusões .....</b>	<b>55</b>
<b>Referências bibliográficas .....</b>	<b>58</b>

## Lista de Figuras

<b>Figura 1:</b> Dimensões da incapacidade funcional.....	18
<b>Figura 2:</b> Relação entre uma variável latente e uma variável observada com três categorias.....	30

## Lista de Quadros

<b>Quadro 1:</b> Características sociodemográficas e de saúde dos idosos de 60 anos ou mais de idade.....	24
<b>Quadro 2:</b> Comparações efetuadas no MCP .....	31
<b>Quadro 3:</b> Distribuição absoluta das unidades segundo as categorias observadas e preditas da variável resposta do modelo.....	44

## Lista de Tabelas

<b>Tabela 1:</b> Distribuição percentual (%) dos idosos de 60 anos ou mais por “grau de dificuldade na locomoção” .....	46
<b>Tabela 2:</b> Distribuição percentual dos idosos por “grau de dificuldade na locomoção”, segundo as suas características sociodemográficas e de saúde .....	47
<b>Tabela 3:</b> Resultados do ajuste do modelo de chances proporcionais parciais (MCPP) explicativo do “Grau de dificuldade na locomoção”, considerando todas as variáveis explicativas .....	50
<b>Tabela 4:</b> Resultados do ajuste do modelo de chances proporcionais parciais (MCPP) explicativo do “Grau de dificuldade na locomoção”, contendo as variáveis explicativas selecionadas .....	52

## Lista de Siglas

ONU	Organização das Nações Unidas
OMS	Organização Mundial de Saúde
IBGE	Instituto Brasileiro de Geografia e Estatística
PNS	Pesquisa Nacional de Saúde
MCP	Modelo de Chances Proporcionais
MCPP	Modelo de Chances Proporcionais Parciais
SIPD	Sistema Integrado de Pesquisas Domiciliares
UPA	Unidade Primária de Amostragem
PNADC	Pesquisa Nacional por Amostra de Domicílios Contínua
POF	Pesquisa de Orçamentos Familiares
RM	Região Metropolitana
UF	Unidade da Federação
OR	Odds Ratio
MV	Máxima Verossimilhança
MPV	Máxima Pseudo-Verossimilhança
IC	Intervalo de Confiança

## **1. Introdução**

### **1.1 Mensuração dos níveis de vida**

A Declaração Universal dos Direitos Humanos proclamada pela Assembleia Geral das Nações Unidas, em 1948, reconheceu pela primeira vez a proteção universal dos direitos humanos, estabelecendo no seu 25º artigo que todo indivíduo tem direito a um nível de vida suficiente a fim de assegurar a ele e à sua família saúde e bem-estar. Desse modo, a mensuração do nível de vida de uma população, num determinado tempo e espaço, tem sido uma preocupação de órgãos governamentais nacionais e internacionais (VERMELHO et al., 2009). Em 1952, um comitê convocado pela ONU propôs a utilização de um conjunto de medidas para a mensuração de níveis de vida de uma população, entre as quais estão aquelas relacionadas ao nível de saúde. Até o início dos anos 1960 os indicadores de saúde eram constituídos fundamentalmente por medidas de mortalidade (COSTA, 2006).

Com o conceito ampliado de saúde estabelecido pela OMS em 1946, como estado de completo bem-estar físico, mental e social, e não apenas ausência de doença, diferentes indicadores para mensuração dos níveis de saúde foram paulatinamente desenvolvidos, com foco no estado de saúde e não mais na mortalidade. Entre os indicadores com foco no estado de saúde, pode-se destacar a morbidade, a incapacidade funcional, a autoavaliação do estado geral de saúde e a percepção da qualidade de vida.

De acordo com Parahyba & Veras (2008), essa mudança de foco da mortalidade para o estado de saúde levou a realização de estudos voltados para a compreensão da distribuição dos determinantes e das consequências da incapacidade funcional, especialmente de populações idosas.

### **1.2 Envelhecimento populacional no Brasil**

O envelhecimento da população é um fenômeno mundial que ocorre tanto em países desenvolvidos quanto em países em desenvolvimento (LAURENTI, 2003). No Brasil, devido a redução dos níveis da fecundidade e da mortalidade, ocorreram mudanças na estrutura etária da população brasileira, especialmente a partir de meados

da década de 80, caracterizando uma sociedade em acelerado processo de envelhecimento populacional (IBGE, 2008). Em 2008 a participação relativa da população com 60 anos ou mais de idade correspondia a 9,5%. Estima-se que em 2020, 13,7% será composta por idosos de 60 anos ou mais, passando para 29,8% em 2050. (IBGE, 2008).

No Brasil, o crescimento da população idosa pode constituir um grave problema para o governo e para a sociedade, tendo em vista a mudança de perfil de saúde da população com predomínio de doenças crônico-degenerativas e de suas complicações (SIQUEIRA et al., 2002; ALVES et al., 2008). A presença de doenças crônicas representa a principal causa de incapacidade, além de explicar o aumento da demanda por serviços de saúde e a maior parcela dos gastos no setor saúde (ALMEIDA et al., 2002). De acordo com Veras (2009) a procura de idosos por internações hospitalares são mais frequentes e a duração de ocupação do leito é maior que a dos outros grupos etários.

A maior morbidade e incapacidade na população idosa podem ser devido a adoção de medidas inadequadas de prevenção de doenças e promoção de saúde (GRITTI et al., 2015). Desse modo, é preciso garantir o atendimento das demandas específicas da população idosa e a realização de políticas públicas eficientes, a fim de contribuir para um envelhecimento populacional saudável, evitando assim uma saúde precária e a incapacidade funcional (ALVES & RODRIGUES, 2005; PARAHYBA & VERAS, 2008).

### **1.3 Incapacidade funcional de idosos no Brasil**

A manutenção da saúde e da capacidade funcional na fase idosa é uma preocupação central no campo da saúde pública, tendo em vista os problemas gerados pela incapacidade funcional que afetam tanto o idoso quanto os seus familiares (RABELO & CARDOSO, 2007). A dificuldade de realizar as atividades da vida diária determina a necessidade de um cuidador que se disponibilize a auxiliá-lo. Assim, a dependência de um familiar idoso gera impacto na dinâmica, na economia e na saúde da família que se ocupa dos cuidados (JEDE & SPULDARO, 2009).

A incapacidade funcional é definida pela dificuldade ou pela necessidade de ajuda na realização das tarefas básicas do cotidiano ou de tarefas mais complexas, necessárias para o indivíduo viver de maneira independente na comunidade (ALVES et

al., 2008). A redução da capacidade funcional é o que explica a maior dependência do indivíduo a um nível de assistência mais complexo (VERAS, 2013). A incapacidade afeta o bem-estar e a qualidade de vida do indivíduo e requer maiores cuidados de saúde (RABELO & CARDOSO, 2007).

Sendo assim, quando o idoso se torna dependente, as alterações são inevitáveis e envolvem afeto, finanças, relações de poder e outras, desenvolvendo um processo de reorganização familiar. Pode-se inferir, portanto, que o envelhecimento em dependência e a figura do cuidador estão a exigir novas formas de assistência e novos enfoques por parte das políticas públicas de saúde. Além do que, o próprio envelhecimento da população brasileira está se dando num momento de profunda desordem econômica, deixando, com certeza, a população de baixa renda mais desamparada e carente, demandando posturas de apoios compensatórios a essa realidade. Está se configurando um severo e crítico quadro de exclusão social do idoso, que se torna ainda mais grave quando o idoso perde a sua capacidade funcional (KARSH, 2003).

A incapacidade funcional pode ser definida por meios de escalas de dificuldade e dependência, podendo ser mensurada por meio de indicadores como o grau de dificuldade na realização de determinadas atividades, ou ainda como o grau de assistência ou de dependência na realização das atividades (ALVES, 2008).

A medida de dependência avalia se um indivíduo necessita ou não de ajuda de outra pessoa na realização de atividades, sendo o indivíduo considerado como dependente no caso de necessitar de ajuda ou no caso de não conseguir realizar uma atividade. Na escala de dificuldade, avalia-se o grau de esforço despendido pelos indivíduos para realizar as atividades funcionais, indicando o grau de dificuldade para realiza-las (ALVES, 2008; MORAES, 2012).

No presente estudo, será utilizado o grau de dificuldade na locomoção como uma *proxy* da incapacidade funcional para pessoas idosas, como descrito na seção 3.5.1.

#### **1.4 Modelo conceitual para incapacidade funcional de idosos**

Tendo em vista as consequências da incapacidade funcional na vida dos idosos e de seus familiares, conhecer os fatores que influenciam na incapacidade é fundamental para identificar quais desses fatores precisam ser priorizados para atender as necessidades dos idosos. Através da identificação dos possíveis fatores de riscos associados à incapacidade funcional, é possível adotar medidas de prevenção em busca



de uma maior longevidade, com maior independência, autonomia e qualidade de vida (DUCA et al., 2009).

O envelhecimento humano é um processo de caráter individual que sofre influência de fatores intrínsecos, que se referem às alterações biológicas ocorridas durante o tempo de vida, e de fatores extrínsecos, como sedentarismo, alcoolismo, tabagismo, alterações climáticas e ambientes insalubres. Com o processo de envelhecimento, os indivíduos se tornam mais frágeis fisicamente e mais suscetíveis a doenças, dependendo da exposição dos indivíduos a esses fatores (CABRAL et al., 2010).

A incapacidade funcional tem uma multiplicidade de fatores desencadeantes, como a evolução de doenças de base, a presença de comorbidades, e a assistência inadequada, seja de âmbito familiar, social ou institucional (VERAS, 2013).

As doenças crônicas são um dos principais determinantes de incapacidade funcional (RABELO & CARDOSO, 2007). É de extrema importância que os serviços de saúde realizem ações de monitoramento e acompanhamento, objetivando evitar as complicações oriundas das doenças crônicas.

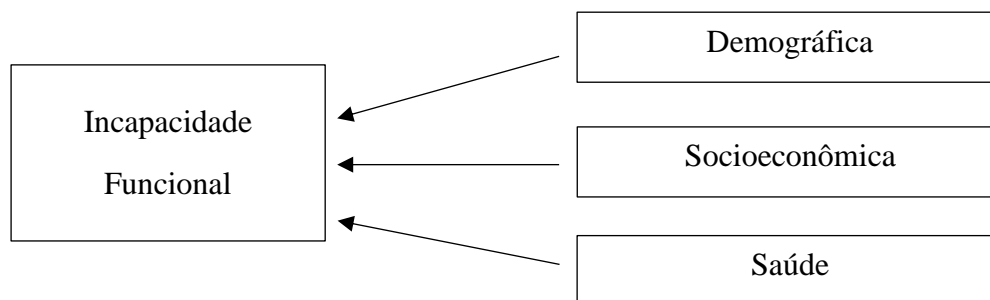
No Brasil, ainda são poucos os trabalhos que abordam a temática sobre incapacidade funcional dos idosos. Entre os estudos que visam identificar os fatores associados à incapacidade funcional, pode-se destacar os de Alves et. al (2010), Duca et al. (2009) e Costa (2006).

De acordo com Alves et al. (2010), usando análise multinível, o fator mais fortemente relacionado com a incapacidade funcional é a autopercepção de saúde, visto que os idosos que autoavaliaram a saúde como ruim eram mais propensos a não realizarem atividades básicas do cotidiano, como subir ladeiras ou escadas. Além disso, em nível individual, outros fatores como doenças crônicas, sexo, ocupação e escolaridade também se mostraram associados com a incapacidade funcional. Ou seja, idosos do sexo feminino, portadores de doenças crônicas e idosos sem ocupação apresentaram maior probabilidade de não conseguirem realizar as atividades. Com relação a escolaridade e a renda, Alves et al. (2010) observaram que idosos mais escolarizados e idosos com maior renda, tiveram menor dificuldade na realização das atividades. Quanto ao nível contextual, no mesmo estudo, a desigualdade de renda também foi considerada como um fator altamente associado a ocorrência de incapacidade funcional.

Baseada em uma análise bruta, Duca et al. (2009) verificaram que os fatores que estão associados a uma maior incapacidade funcional para as atividades básicas são o sexo feminino, cor parda/preta/outra, e a situação conjugal definida como solteiro(a)/separado(a)/viúvo(a). Além disso, observou-se uma relação positiva da incapacidade com a idade e negativa com a escolaridade. Já na análise ajustada, apenas os fatores cor da pele e aumento de idade obtiveram associação com o desfecho de incapacidade funcional.

No estudo de Costa (2006) também foi verificado que o sexo e o aumento de idade são fatores associados à incapacidade funcional, isto é, as mulheres e indivíduos com 60 anos ou mais de idade (idosos), apresentaram maior grau de limitação na realização das atividades básicas do cotidiano.

A Figura 1 ilustra as principais dimensões associadas à incapacidade funcional em idosos.



**Figura 1:** Dimensões da incapacidade funcional

## **2. Objetivos**

### **2.1 Objetivo Geral:**

Analisar o grau de dificuldade na locomoção dos idosos no Brasil e estabelecer a sua associação com um conjunto de variáveis sociodemográficas e de saúde dos idosos de 60 anos ou mais de idade.

### **2.2 Objetivos Específicos:**

- Estimar a prevalência dos níveis de dificuldade na locomoção dos idosos no Brasil;
- Identificar as variáveis sociodemográficas e de saúde associadas ao grau de dificuldade na locomoção dos idosos;
- Avaliar o sentido e a magnitude da associação entre as características dos idosos e o grau de dificuldade na locomoção.

### **3. Materiais e Métodos**

Nesta seção são apresentadas informações sobre a Pesquisa Nacional de Saúde (PNS) 2013, que incluem a descrição da estrutura do questionário e do plano amostral dessa pesquisa. Além disso, são apresentadas a definição da população de estudo, a listagem das variáveis analisadas e a especificação de dois tipos de modelo de regressão logística ordinal: modelo de chances proporcionais (MCP) e modelo de chances proporcionais parciais (MCP). Para efeitos de simplificação, optou-se por fornecer o procedimento de estimação dos parâmetros pelo método de Máxima Pseudo-Verossimilhança (MPV), em particular para o caso do MCP. Por fim, é apresentado o teste de Wald de significância individual dos parâmetros, o teste de homogeneidade dos coeficientes (ou proporcionalidade das chances), o intervalo de confiança para os parâmetros do modelo e para as medidas de razão de chance; e as medidas de qualidade do ajuste.

#### **3.1 Pesquisa Nacional de Saúde (PNS) 2013**

A Pesquisa Nacional de Saúde é uma pesquisa por amostragem domiciliar de abrangência nacional, de periodicidade de 5 anos, que emprega um plano amostral complexo e que faz parte do Sistema Integrado de Pesquisas Domiciliares (SIPD) do Instituto Brasileiro de Geografia e Estatística (IBGE). Do ponto de vista epidemiológico, a PNS trata-se de um estudo transversal observacional.

A PNS foi realizada em 2013 pelo IBGE em convênio com o Ministério da Saúde, e levantou informações sobre as condições de saúde dos moradores de domicílios particulares do Brasil, com enfoque em aspectos relacionados à percepção de saúde, estilos de vida e doenças crônicas. Não foram considerados nesta pesquisa os domicílios localizados em setores especiais, como aldeias indígenas, quartéis, bases militares, alojamentos, acampamentos, embarcações, penitenciárias, colônias penais, presídios, cadeias, asilos, orfanatos, conventos e hospitais.

## **3.2 Questionários da PNS**

A PNS utilizou um questionário padronizado com perguntas fechadas, que pode ser dividido em três partes: domiciliar, geral e individual.

### **3.2.1 Parte domiciliar**

A parte domiciliar do questionário da PNS englobou informações sobre as características domiciliares, como: o tipo de domicílio; material predominante na construção, na cobertura (telhado) e no piso do domicílio; forma e frequência do abastecimento de água do domicílio; coleta de lixo; origem da energia elétrica; esgotamento sanitário; posse de bens no domicílio, entre outras. Essas informações são respondidas por qualquer morador do domicílio.

### **3.2.2 Parte geral**

A parte geral do questionário da PNS englobou informações sobre as características demográficas (como, por exemplo, sexo, idade, cor/raça e estado civil), socioeconômicas (como educação, trabalho e renda) e de saúde (percepção do estado de saúde, doenças crônicas, estilos de vida, saúde bucal, atendimento médico, etc.) de todos os moradores residentes no domicílio. Essa parte do questionário é respondida por qualquer morador do domicílio que tenha condição de informar sobre essas características para todos os moradores.

### **3.2.3 Parte individual**

A parte individual do questionário da PNS englobou informações sobre as características de um morador adulto de 18 anos ou mais de idade selecionado com equiprobabilidade dentre os moradores adultos residentes no domicílio. Além das características demográficas, socioeconômicas e de saúde, foram obtidas para o morador adulto selecionado para a entrevista individual, medidas antropométricas, como peso, altura, circunferência da cintura e pressão arterial. Também foram realizados exames laboratoriais (exame de sangue e urina) com o adulto selecionado.

### 3.3 Plano amostral da PNS

A PNS utiliza um plano amostral conglomerado com três estágios de seleção, onde os *setores ou um conjunto de setores* formam as unidades primárias de amostragem (UPAs)<sup>1</sup>, os *domicílios* formam as unidades secundárias e os *moradores* com 18 anos ou mais de idade formam as unidades terciárias (IBGE, 2014). Cabe ressaltar que para obter as informações contidas na parte domiciliar e geral do questionário da PNS, bastou o emprego de um plano amostral com apenas 2 estágios de seleção. Além da conglomeração, houve estratificação das unidades primárias de amostragem na PNS.

#### 3.3.1 Sistema Integrado de Pesquisas Domiciliares

A implantação do Sistema Integrado de Pesquisas Domiciliares por amostragem (SIPD) foi iniciada pelo IBGE em 2011 com a realização da primeira coleta de dados da Pesquisa Nacional por Amostra de Domicílios Contínua. Em relação aos aspectos da implantação do SIPD, destaca-se a necessidade de utilização de uma mesma infraestrutura amostral, de um mesmo cadastro de seleção e de uma amostra comum, denominada “Amostra Mestra”, sendo esta, um dos elementos fundamentais para a implementação deste sistema (FREITAS & ANTONACI, 2014).

Dentre as principais pesquisas que integram o SIPD, estão a Pesquisa Nacional de Saúde (PNS), a Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) e a Pesquisa de Orçamentos Familiares (POF).

A PNS, por integrar o SIPD, tem como vantagem um maior espalhamento geográfico e uma maior eficiência amostral. Assim, a PNS foi planejada para a obtenção de indicadores com precisão controlada e para possibilitar um monitoramento da maior parte dos indicadores que compunham o suplemento de saúde da PNAD.

As definições metodológicas e de abrangência das principais pesquisas domiciliares do IBGE, como PNAD e POF, sempre ocorreram isoladamente. Com isso, o SIPD foi implementado como proposta de repensar todo esse conjunto de investigação conjuntamente, onde cada tema investigado passa a ser parte de um sistema integrado de indicadores socioeconômicos e demográficos. As etapas de planejamento, execução,

---

<sup>1</sup> Unidades de área com um determinado tamanho mínimo populacional contabilizado em termos de domicílios.

disseminação e análise dos resultados das diversas pesquisas podem ser conduzidas de forma associada, otimizando recursos e facilitando o atendimento de novas demandas.

### 3.3.2 Amostra Mestra

De acordo com Freitas & Antonaci (2014), a amostra mestra é um conjunto de unidades de áreas que são selecionadas para atender a diversas pesquisas que compõem o SIPD. Essas unidades de áreas, consideradas como unidades primárias de amostragem (UPAs), são selecionadas do cadastro mestre que continha 316.574 setores censitários, que dividem o território nacional, provenientes da Base Operacional Geográfica de 2010. Com a exclusão dos setores especiais, por não fazerem parte da abrangência da PNS, o cadastro passou a ser composto por 310.329 setores censitários.

Devido a existência de setores censitários pequenos, em termos de quantidade de unidades domiciliares, foi necessário efetuar uma agregação dos setores para a composição das UPAs, respeitando os critérios de contiguidade, tamanho mínimo de 60 domicílios e algumas características dos setores, tais como o tipo, situação e divisão administrativa (subdistrito). Desse modo, os 310.329 setores foram agrupados em 292.067 UPAs.

Do cadastro composto por 292.067 UPAs (cadastro de seleção), selecionou-se a amostra mestra, utilizando a estratificação e amostragem proporcional ao tamanho, utilizando o número de domicílios particulares permanentes como medida de tamanho das UPAs. A seleção da amostra mestre obedeceu quatro critérios: 1) *O critério administrativo*, contempla a divisão da UF em capital, resto da Região Metropolitana (RM) e resto da UF; 2) *O critério geográfico*, subdivide as capitais e outros municípios de grande porte em mais estratos, buscando garantir um maior espalhamento territorial, além de facilitar a operação de coleta, controle e redução dos custos de deslocamento entre as unidades da amostra; 3) *O critério de situação censitária* que envolve a divisão em área rural e urbana; 4) *O critério estatístico*, por sua vez, envolve a estratificação estatística das UPAs considerando as informações sociodemográficas do Censo Demográfico 2010, como rendimento total e total de domicílios particulares permanentes, com objetivo de agrupar as UPAs de forma melhorar a precisão das estimativas (FREITAS e ANTONACI, 2014).

A amostra da PNS, por sua vez, é uma subamostra das UPAs da amostra mestra, que foi selecionada respeitando os mesmos critérios de estratificação.

### 3.4 População de estudo

A população de estudo é composta por idosos com 60 anos ou mais de idade residentes em domicílios particulares permanentes<sup>2</sup> no Brasil, exceto os domicílios localizados em setores especiais descritos na seção 3.1. Idosos sem informação sobre o grau de dificuldade na locomoção ou sem informação sobre alguma das variáveis de estudo foram desconsiderados da análise.

### 3.5 Variáveis analisadas

#### 3.5.1 Grau de dificuldade na locomoção

“*Em geral que grau de dificuldade o(a) sr(a) tem para se locomover?*”. É uma pergunta que está presente no módulo N sobre “percepção do estado de saúde” do questionário da PNS 2013, com cinco alternativas de resposta: nenhum, leve, médio, intenso e não consegue.

Para efeito de análise, serão consideradas apenas três categorias mutuamente exclusivas que representam diferentes níveis de dificuldade: grave, no máximo moderado e nenhum. Na categoria “grave” estão sendo considerados os idosos que responderam um grau “intenso” de dificuldade na locomoção ou que “não consegue” se locomover. Na categoria “no máximo moderado” estão sendo considerados os idosos que responderam o grau “leve” ou “médio” de dificuldade na locomoção.

#### 3.5.2 Variáveis sociodemográficas e de saúde

**Quadro 1:** Características sociodemográficas e de saúde dos idosos de 60 anos ou mais de idade.

<i>Variáveis de estudo</i>	<i>Categorias</i>
Sexo	Masculino Feminino

(*Continua*)

<sup>2</sup> Domicílios que servem exclusivamente para habitação e têm como finalidade servir de moradia para uma ou mais pessoas.



**Quadro 1** (*Continuação*): Características sociodemográficas e de saúde dos idosos de 60 anos ou mais de idade.

<i>Variáveis de estudo</i>	<i>Categorias</i>
Faixa etária	De 60 a 69 anos De 70 a 79 anos 80 anos ou mais
Cor/Raça	Branca Não branca
Situação conjugal	Casado Não casado
Grandes regiões brasileiras	Norte Nordeste Sudeste Sul Centro-Oeste
Área de localização do domicílio	Urbana Rural
Nível educacional	Sem instrução Ensino fundamental Ensino médio Ensino superior
Qualidade da construção da moradia (material usado no piso, nas paredes e teto)	Adequado Inadequado
Plano de saúde	Sim Não
Procura o mesmo lugar, mesmo médico ou mesmo serviço de saúde quando precisa de atendimento de saúde	Sim Não
Autoavaliação geral de saúde	Bom Regular Ruim
Diagnóstico de alguma doença crônica, física ou mental	Sim Não

### 3.6 Modelo de regressão logística ordinal

O modelo de regressão logística ordinal é um modelo linear generalizado composto por uma variável resposta ordinal ( $Y$ ) com distribuição de probabilidade multinomial, por uma função de ligação logística e por variáveis explicativas numéricas ou categóricas. Nas seções seguintes são apresentadas as equações de dois tipos de modelos de regressão logística ordinal: modelo de chances proporcionais (MCP) e modelo de chances proporcionais parciais (MCPP).

#### 3.6.1 Modelo de chances proporcionais (MCP)

O MCP é um modelo de regressão ordinal utilizado para explicar variáveis respostas com escala ordinal, isto é, em que as categorias são ordenáveis (LIU & KOIRALA, 2013; MORAES et al., 2012). No MCP o termo  $\alpha_k$  desse modelo varia para cada uma das categorias e cada parâmetro  $\beta$  independe da categoria. Logo, este modelo possui uma suposição de chances proporcionais, também conhecida como suposição de homogeneidade dos coeficientes (efeitos das variáveis explicativas). A suposição de homogeneidade dos coeficientes deve ser satisfeita para a utilização do MCP. Caso contrário, o MCPP, definido na seção 3.6.2 pode ser adotado.

Para uma variável resposta em escala ordinal definida por  $K$  categorias, o MCP é definido na equação (3.1):

$$\ln \left[ \frac{P(Y \leq k | x_1, x_2, \dots, x_p)}{P(Y > k | x_1, x_2, \dots, x_p)} \right] = \alpha_k - (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) \quad (3.1)$$

$$\forall k = 1, 2, \dots, K - 1$$

onde:

$\text{Logit}[P(Y \leq k | x_1, x_2, \dots, x_p)] = \ln \left[ \frac{P(Y \leq k | x_1, x_2, \dots, x_p)}{P(Y > k | x_1, x_2, \dots, x_p)} \right]$  é o logaritmo da razão entre a probabilidade acumulada  $P(Y \leq k | x_1, x_2, \dots, x_p)$  e o seu complementar  $P(Y > k | x_1, x_2, \dots, x_p)$ .

$\alpha_k$  é o parâmetro (ponte de corte) referente a  $k$ -ésima categoria da variável resposta;  $k=1,2,\dots,K-1$ ;

$x_j$  é a  $j$ -ésima variável explicativa (covariável ou variável *dummy* para os níveis dos fatores) do modelo;  $j = 1, \dots, p$ ;

$\beta_j$  é o efeito principal da  $j$ -ésima variável explicativa do modelo;  $j = 1, \dots, p$ .

O MCP representado na equação (3.1) pode ser escrito matematicamente de uma forma alternativa, como mostrado na equação (3.2):

$$\ln \left[ \frac{1}{\frac{P(Y > k | x_1, x_2, \dots, x_p)}{P(Y \leq k | x_1, x_2, \dots, x_p)}} \right] = \alpha_k - (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)$$

$$\ln[1] - \ln \left[ \frac{P(Y > k | x_1, x_2, \dots, x_p)}{P(Y \leq k | x_1, x_2, \dots, x_p)} \right] = \alpha_k - (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)$$

$$- \ln \left[ \frac{P(Y > k | x_1, x_2, \dots, x_p)}{P(Y \leq k | x_1, x_2, \dots, x_p)} \right] = \alpha_k - (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)$$

$$\ln \left[ \frac{P(Y > k | x_1, x_2, \dots, x_p)}{P(Y \leq k | x_1, x_2, \dots, x_p)} \right] = -\alpha_k + (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)$$

$$\ln \left[ \frac{P(Y > k | x_1, x_2, \dots, x_p)}{P(Y \leq k | x_1, x_2, \dots, x_p)} \right] = \alpha_k^* + (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) \quad (3.2)$$

$$\forall k = 1, 2, \dots, K - 1$$

onde:

$\text{Logit}[P(Y > k | x_1, x_2, \dots, x_p)] = \ln \left[ \frac{P(Y > k | x_1, x_2, \dots, x_p)}{P(Y \leq k | x_1, x_2, \dots, x_p)} \right]$  é o logaritmo da razão entre a probabilidade acumulada  $P(Y > k | x_1, x_2, \dots, x_p)$  e seu complementar  $P(Y \leq$

$k | x_1, x_2, \dots, x_p$ ), isto é, representa a chance (em escala logarítmica) da variável resposta  $Y$  assumir categorias maiores do que  $k$ , relativamente as categorias menores ou iguais a  $k$ .

$\alpha_k^*$  é o parâmetro (ponte de corte) referente a  $k$ -ésima categoria da variável resposta;  $k=1,2,\dots,K-1$ ;

$x_j$  é a  $j$ -ésima variável explicativa (covariável ou variável *dummy* para os níveis dos fatores) do modelo;  $j = 1, \dots, p$ ;

$\beta_j$  é o efeito da  $j$ -ésima variável explicativa do modelo;  $j = 1, \dots, p$ .

Apesar da relação de ordem entre as categorias da variável resposta obrigar que a modelagem seja feita em termos de probabilidade acumuladas, no MCP a variável resposta ordinal pode ser entendida como a operacionalização de uma variável contínua não observada, chamada de variável latente ( $Y^*$ ), cujo intervalo de variação é de  $]-\infty, +\infty[$  e que gera os valores (categorias) da variável resposta observada  $Y$ .

Os valores da variável latente  $Y^*$  são divididos em  $K$  grupos mutuamente exclusivos, sendo, portanto necessário definir  $K-1$  pontos de corte, isto é, a quantidade de pontos de corte depende do número de categorias da variável ordinal observada.

De modo geral, a variável ordinal observada ( $Y$ ) está relacionada com uma variável latente ( $Y^*$ ), da seguinte forma:

$$Y = k \quad \text{se} \quad \alpha_{k-1} \leq Y^* \leq \alpha_k \quad (3.3)$$

$$\forall k = 1, 2, \dots, K - 1$$

Devido a essa relação (3.3), o MCP pode ser especificado através do seguinte modelo estrutural, de modo a facilitar a interpretação das estimativas dos parâmetros do modelo:

$$Y^* = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon \quad (3.4)$$

onde:

$x_j$  é a  $j$ -ésima variável explicativa (covariável ou variável *dummy* para os níveis dos fatores) do modelo;  $j = 1, \dots, p$ ;

$\beta_j$  é o efeito principal da  $j$ -ésima variável explicativa do modelo;  $j = 1, \dots, p$ ;

$\varepsilon$  é o erro do modelo com distribuição logística de média 0 e variância  $\frac{\pi^2}{3}$ .

Com relação a interpretação dos parâmetros do MCP, representado em (3.2) se  $\beta_j > 0$ , quando  $x_j$  aumenta, aumenta a chance, em escala logarítmica, da variável resposta assumir valores de ordem superior a  $k$ , ou seja, quando  $x_j$  aumenta,  $Y$  aumenta. Caso contrário ( $\beta_j < 0$ ), quando  $x_j$  aumenta,  $Y$  diminui.

No MCP é possível obter uma medida de razão de chances (*odds ratio*), sendo esta medida de associação igual para todas as categorias da variável resposta  $Y$ . Sendo assim, a razão de chances ( $OR_{jk}^*$ ) do modelo da equação (3.2) é dada pela comparação da chance da variável resposta assumir valores maiores à categoria  $k$  ( $Y > k$ ) para  $x_j = a$  (sendo  $a = 1$  no caso de ter a característica de interesse) com aquela para  $x_j = b$  (sendo  $b = 0$  no caso de não ter a característica de interesse), representada por:

$$OR_{jk}^* = \frac{P(Y > k | x_1, x_2, \dots, x_j = a, \dots, x_p) / P(Y \leq k | x_1, x_2, \dots, x_j = a, \dots, x_p)}{P(Y > k | x_1, x_2, \dots, x_j = b, \dots, x_p) / P(Y \leq k | x_1, x_2, \dots, x_j = b, \dots, x_p)}$$

ou seja,

$$OR_{jk}^* = \frac{e^{\alpha_k^* + (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j a + \dots + \beta_p x_p)}}{e^{\alpha_k^* + (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_j b + \dots + \beta_p x_p)}} = e^{(a-b)\beta_j} = e^{(1-0)\beta_j} = e^{\beta_j} \quad (3.5)$$

$$\forall k = 1, 2, \dots, K - 1$$

Se o coeficiente for positivo  $\beta_j$  ( $\beta_j > 0$ ) então o exponencial do coeficiente é maior que a unidade ( $OR_{jk}^* > 1$ ), o que significa que no grupo 1 ( $x_j = a = 1$ ) a chance de se pertencer a categorias maiores que  $k$  ( $Y > k$ ), relativamente a categorias menores ou iguais a  $k$ , é  $(OR_{jk}^* - 1) \cdot 100\%$  maior que no grupo de referência ( $x_j = b = 0$ ).

Quando o coeficiente  $\beta_j$  for negativo ( $\beta_j < 0$ ) então o exponencial do coeficiente é menor que a unidade ( $OR_{jk}^* < 1$ ). Isto significa que no grupo 1 ( $x_j = a =$

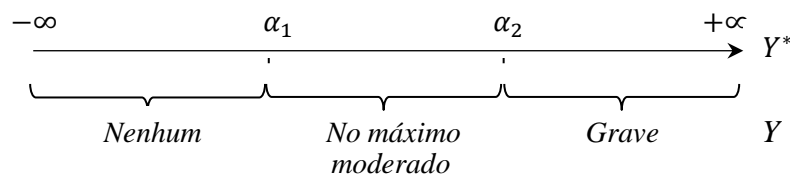
1) a chance de se pertencer a categorias maiores que  $k$  ( $Y > k$ ), relativamente a categorias menores ou iguais a  $k$ , é  $(OR_{jk}^* - 1) \cdot 100\%$  menor que no grupo de referência ( $x_j = b = 0$ ).

Por fim, quando o coeficiente for igual a zero ( $\beta_j = 0$ ), o exponencial do coeficiente é a unidade ( $OR_{jk}^* = 1$ ), indicando que não existe diferença nas chances de se pertencer a categorias maiores que  $k$  ( $Y > k$ ), relativamente a categorias menores ou iguais a  $k$ , entre o grupo 1 e o grupo de referência.

Para fins explicativos, a teoria apresentada anteriormente será contextualizada considerando a variável resposta ordinal utilizada no presente estudo. A variável resposta ordinal “grau de dificuldade na locomoção”, que contém três categorias (grave, no máximo moderado, nenhum) está relacionada com uma variável latente ( $Y^*$ ), da seguinte forma:

$$Y = \begin{cases} 1 \rightarrow \text{Nenhum} & \text{se } -\infty \leq Y^* \leq \alpha_1 \\ 2 \rightarrow \text{No máximo moderado} & \text{se } \alpha_1 \leq Y^* \leq \alpha_2 \\ 3 \rightarrow \text{Grave} & \text{se } \alpha_2 \leq Y^* \leq +\infty \end{cases}$$

A Figura 2 ilustra a divisão dos valores da variável latente em três categorias. Neste caso, existem dois pontos de corte  $\alpha_1$  e  $\alpha_2$  crescentes, tal que  $\alpha_1 < \alpha_2$ .



**Figura 2:** Relação entre uma variável latente e uma variável observada com três categorias.

O MCP (3.2) leva em conta a natureza ordinal da variável resposta “grau de dificuldade na locomoção” como ilustrado no Quadro 2.

**Quadro 2:** Comparações efetuadas no MCP

<b>Grau de dificuldade na locomoção</b>	
<b>Comparações com três categorias (<math>K=3</math>)</b>	
$k = 1$ : Nenhum	2,3 vs 1
$k = 2$ : No máximo moderado	3 vs 1,2
$k = 3$ : Grave	-

Para esta variável resposta com 3 categorias, o MCP (3.2) pode ser representado por duas equações lineares, dadas nas equações (3.6) e (3.7):

Para  $k=1$ :

$$\begin{aligned}
 \ln \left[ \frac{P(Y > 1)}{P(Y \leq 1)} \right] &= \ln \left[ \frac{P(Y = 2) + P(Y = 3)}{P(Y = 1)} \right] = \\
 &= \ln \left[ \frac{P(\text{No máximo moderado}) + P(\text{Grave})}{P(\text{Nenhum})} \right] = \\
 &= \alpha_1^* + (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) \tag{3.6}
 \end{aligned}$$

Para  $k=2$ :

$$\begin{aligned}
 \ln \left[ \frac{P(Y > 2)}{P(Y \leq 2)} \right] &= \ln \left[ \frac{P(Y = 3)}{P(Y = 1) + P(Y = 2)} \right] = \\
 &= \ln \left[ \frac{P(\text{Grave})}{P(\text{Nenhum}) + P(\text{No máximo moderado})} \right] = \\
 &= \alpha_2^* + (\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p) \tag{3.7}
 \end{aligned}$$

### 3.6.2 Modelo de chances proporcionais parciais (MCP)

Geralmente a suposição de chances proporcionais para todas as variáveis explicativas consideradas na modelagem é difícil de ser atingida na prática, então neste caso uma alternativa adequada é utilizar o MCP (WILLIAMS, 2006). Como o MCP não impõe que todas as variáveis atendam a suposição de chances proporcionais, esse modelo permite que algumas ou todas variáveis possam ser modeladas sem a suposição de chances proporcionais. No caso dessa suposição não ser satisfeita, os efeitos das variáveis explicativas diferem através das equações; caso contrário, os seus efeitos são homogêneos para todas as equações.

Sendo assim, o MCP tem a parametrização equivalente à equação do MCP (3.2), mas no MCP não é necessário que todos os efeitos das variáveis explicativas sejam iguais para todas as equações, isto é, para todas as categorias  $k=1,2,\dots,K-1$  da variável resposta, sendo portanto denotados genericamente por  $\beta_{1k}, \beta_{2k}, \dots, \beta_{pk}$ . Então, o MCP pode ser representado através da seguinte equação (3.8):

$$\ln \left[ \frac{P(Y > k | x_1, x_2, \dots, x_p)}{P(Y \leq k | x_1, x_2, \dots, x_p)} \right] = \gamma_k^* + (\beta_{1k}x_1 + \beta_{2k}x_2 + \dots + \beta_{pk}x_p) \quad (3.8)$$

$$\forall k = 1, 2, \dots, K - 1$$

onde:

$Logit[P(Y > k | x_1, x_2, \dots, x_p)] = \ln \left[ \frac{P(Y > k | x_1, x_2, \dots, x_p)}{P(Y \leq k | x_1, x_2, \dots, x_p)} \right]$  é o logaritmo da razão entre a probabilidade acumulada  $P(Y > k | x_1, x_2, \dots, x_p)$  e seu complementar  $P(Y \leq k | x_1, x_2, \dots, x_p)$ , isto é, representa a chance (em escala logarítmica) da variável resposta  $Y$  assumir categorias maiores do que  $k$ , relativamente às categorias menores ou iguais a  $k$ .

$\gamma_k^*$  é o parâmetro (ponte de corte) referente a  $k$ -ésima categoria da variável resposta;  $k=1,2,\dots,K-1$ ;

$x_j$  é a  $j$ -ésima variável explicativa (covariável ou variável *dummy* para os níveis dos fatores) do modelo;  $j = 1, \dots, p$ ;

$\beta_{jk}$  é o efeito da  $j$ -ésima variável explicativa do modelo referente à  $k$ -ésima categoria da variável resposta;  $j = 1, \dots, p$ ;  $k = 1, 2, \dots, K - 1$ .



Analogamente ao MCP, no MCPP é possível obter uma medida de razão de chances (*odds ratio*), representada por:

$$OR_{jk}^* = \frac{e^{\gamma_k^* + (\beta_{1k}x_1 + \beta_{2k}x_2 + \dots + \beta_{jk}a + \dots + \beta_{pk}x_p)}}{e^{\gamma_k^* + (\beta_{1k}x_1 + \beta_{2k}x_2 + \dots + \beta_{jk}b + \dots + \beta_{pk}x_p)}} = e^{(a-b)\beta_{jk}} = e^{(1-0)\beta_{jk}} = e^{\beta_{jk}} \quad (3.9)$$

$$\forall k = 1, 2, \dots, K - 1$$

No MCPP, a interpretação da medida de razão de chances é feita da mesma forma que no MCP quando a suposição de chances proporcionais é verdadeira para alguma variável explicativa. Porém quando a suposição de chances proporcionais não é satisfeita, a interpretação se modifica, devendo levar em conta a categoria da variável resposta, da seguinte forma:

Se o coeficiente for positivo  $\beta_{jk}$  ( $\beta_{jk} > 0$ ) e conseqüentemente  $OR_{jk}^* > 1$ , no grupo 1 ( $x_j = a = 1$ ) a chance de se pertencer a categorias maiores que  $k$  ( $Y > k$ ), relativamente a categorias menores ou iguais a  $k$ , é  $(OR_{jk}^* - 1) \cdot 100\%$  maior que ao grupo de referência ( $x_j = b = 0$ ).

Se o coeficiente  $\beta_{jk}$  for negativo ( $\beta_{jk} < 0$ ) e conseqüentemente  $OR_{jk}^* < 1$ , no grupo 1 ( $x_j = a = 1$ ) a chance de se pertencer a categorias maiores que  $k$  ( $Y > k$ ), relativamente a categorias menores ou iguais a  $k$ , é  $(OR_{jk}^* - 1) \cdot 100\%$  menor que ao grupo de referência ( $x_j = b = 0$ ).

Quando o coeficiente for igual a zero ( $\beta_{jk} = 0$ ) e  $OR_{jk}^* = 1$ , não existe diferença nas chances de se pertencer a categorias maiores que  $k$  ( $Y > k$ ), relativamente a categorias menores ou iguais a  $k$ , entre ambos os grupos.

No caso de utilização do MCPP para explicar a variável resposta ordinal “*grau de dificuldade*”, as duas equações são dadas por:

Para  $k=1$ :

$$\ln \left[ \frac{P(Y > 1)}{P(Y \leq 1)} \right] = \ln \left[ \frac{P(Y = 2) + P(Y = 3)}{P(Y = 1)} \right] =$$

$$\begin{aligned}
&= \ln \left[ \frac{P(\text{No máximo moderado}) + P(\text{Grave})}{P(\text{Nenhum})} \right] = \\
&= \gamma_1^* + (\beta_{11}x_1 + \beta_{21}x_2 + \dots + \beta_{p1}x_p)
\end{aligned} \tag{3.10}$$

Para  $k=2$ :

$$\begin{aligned}
&\ln \left[ \frac{P(Y > 2)}{P(Y \leq 2)} \right] = \ln \left[ \frac{P(Y = 3)}{P(Y = 1) + P(Y = 2)} \right] = \\
&= \ln \left[ \frac{P(\text{Grave})}{P(\text{Nenhum}) + P(\text{No máximo moderado})} \right] = \\
&= \gamma_2^* + (\beta_{12}x_1 + \beta_{22}x_2 + \dots + \beta_{p2}x_p)
\end{aligned} \tag{3.11}$$

A seguir é descrito o método de estimação dos parâmetros do MCP pelo método de Máxima Pseudo-Verossimilhança (MPV). Entretanto cabe ressaltar que o procedimento demonstrado na seção 3.6.3 é similar no caso do MCPP, sendo a única diferença o fato dos parâmetros do MCPP poderem variar de acordo com as categorias da variável resposta quando a suposição de proporcionalidade das chances não for satisfeita.

Neste trabalho foi utilizado o software STATA 11.0 para todas as análises estatísticas desenvolvidas com base nos dados da PNS 2013. Para ajustar o MCPP foi utilizado o comando *gologit29*, incluindo a opção “*autofit*”. Esta opção permite avaliar se o efeito de cada variável difere através das equações, ou seja, se a suposição de homogeneidade individual dos parâmetros é satisfeita ou não (WILLIAMS, 2006).

### 3.6.3 Método de Máxima Pseudo-Verossimilhança (MPV)

Um modelo estatístico pode ser ajustado por métodos padrões como o método da máxima verossimilhança (MV), quando as observações amostrais são consideradas independentes e identicamente distribuídas (IID). Porém a estimação dos parâmetros do

modelo feita através do método de MV ignora o desenho amostral e os pesos amostrais das unidades da amostra e, além disso, na prática das pesquisas amostrais, raramente é adequada a hipótese IID para observações amostrais.

Então quando se trata de uma pesquisa que utiliza um plano amostral complexo, a qual não se coleta informações para toda a população, uma forma alternativa para o processo de ajuste de modelo é utilizar o método de Máxima Pseudo-Verossimilhança (MPV), que permite incorporar os pesos amostrais, os conglomerados e os estratos na modelagem estatística.

Neste trabalho, como os dados são provenientes da Pesquisa Nacional de Saúde 2013 que tem um plano amostral complexo, será utilizado o método de MPV para o ajuste do modelo.

A variável resposta  $Y$  do modelo de chances proporcionais (MCP) possui  $K$  categorias, assim, pode-se definir  $K$  variáveis binárias do seguinte modo:

$$D_{ik} = \begin{cases} 1 & \text{se } Y_i = k \\ 0 & \text{se } Y_i \neq k \end{cases}$$

No caso de  $y_1, y_2, \dots, y_N$  serem valores observados das respectivas variáveis  $Y_1, Y_2, \dots, Y_N$ , independentes e cada uma com distribuição multinomial, a função de probabilidade da variável aleatória  $Y_i$  é dada por:

$$P(Y_i = y_i | \mathbf{x}_i) = p_{i1}^{D_{i1}} p_{i2}^{D_{i2}} \dots p_{iK}^{D_{iK}} \quad (3.12)$$

$$i = 1, 2, \dots, N$$

onde:

$\mathbf{x}_i = (x_{1i}, x_{2i}, \dots, x_{pi})$  é o vetor das variáveis explicativas (variáveis do tipo *dummy* para os níveis dos fatores) da  $i$ -ésima unidade da população, de dimensão  $p \times 1$ ;

$p_{ik}$  é a probabilidade do  $i$ -ésimo elemento da amostra pertencer à  $k$ -ésima categoria da variável resposta  $Y$ , sendo calculada do seguinte modo:

$$\begin{aligned}
p_{ik} &= p_{ik}(\boldsymbol{\beta}) = P(Y = k|\mathbf{x}_i) = P(Y \leq k|\mathbf{x}_i) - P(Y \leq k-1|\mathbf{x}_i) = \\
&= \frac{e^{\alpha_k - \mathbf{x}_i \boldsymbol{\beta}}}{1 + e^{\alpha_k - \mathbf{x}_i \boldsymbol{\beta}}} - \frac{e^{\alpha_{k-1} - \mathbf{x}_i \boldsymbol{\beta}}}{1 + e^{\alpha_{k-1} - \mathbf{x}_i \boldsymbol{\beta}}} \quad (3.13)
\end{aligned}$$

O método de MPV consiste em maximizar a função de verossimilhança populacional ou o logaritmo desta função. Então se todos os elementos da população fossem conhecidos, a função de verossimilhança populacional seria dada por:

$$\begin{aligned}
L(\boldsymbol{\beta}|\mathbf{y}) &= \prod_{i=1}^N P(Y_i = y_i|\mathbf{x}_i) = \prod_{i=1}^N p_{i1}^{D_{i1}} p_{i2}^{D_{i2}} \dots p_{iK}^{D_{iK}} \\
L(\boldsymbol{\beta}|\mathbf{y}) &= \prod_{i=1}^N \prod_{k=1}^K p_{ik}^{D_{ik}} \quad (3.14)
\end{aligned}$$

e o logaritmo da função de verossimilhança populacional seria dado por:

$$\begin{aligned}
\ln L(\boldsymbol{\beta}|\mathbf{y}) &= \ln \prod_{i=1}^N \prod_{k=1}^K p_{ik}^{D_{ik}} = \sum_{i=1}^N \sum_{k=1}^K \ln p_{ik}^{D_{ik}} \\
\ln L(\boldsymbol{\beta}|\mathbf{y}) &= \sum_{i=1}^N \sum_{k=1}^K D_{ik} \ln p_{ik} \quad (3.15)
\end{aligned}$$

Primeiramente calcula-se a primeira derivada parcial da equação (3.15), do seguinte modo:

$$\frac{\partial}{\partial \boldsymbol{\beta}} \ln L(\boldsymbol{\beta}|\mathbf{y}) = \frac{\partial}{\partial \boldsymbol{\beta}} \sum_{i=1}^N \sum_{k=1}^K D_{ik} \ln p_{ik} = \sum_{i=1}^N \frac{\partial}{\partial \boldsymbol{\beta}} \sum_{k=1}^K D_{ik} \ln p_{ik} = \sum_{i=1}^N \mathbf{u}_i(\boldsymbol{\beta}) \quad (3.16)$$

Igualando (3.16) a  $\mathbf{0}$ , obtêm-se o sistema de equações de verossimilhança populacionais:

$$\sum_{i=1}^N \mathbf{u}_i(\boldsymbol{\beta}) = \mathbf{0} \quad (3.17)$$

onde,  $\mathbf{u}_i(\boldsymbol{\beta}) = \frac{\partial}{\partial \boldsymbol{\beta}} \sum_{k=1}^K D_{ik} \ln p_{ik}$  é o vetor de escores da  $i$ -ésima unidade populacional, de dimensão  $P \times 1$ .

O lado esquerdo do sistema de equações de verossimilhança populacionais (3.17) é a soma dos vetores de escores da população, como mostrado a seguir:

$$T(\boldsymbol{\beta}) = \sum_{i=1}^N u_i(\boldsymbol{\beta}) \quad (3.18)$$

A partir de (3.17) é possível obter o estimador de MV do vetor  $\boldsymbol{\beta}$  de parâmetros de interesse, que na abordagem de MPV representa o vetor de pseudos-parâmetros denotado por  $\boldsymbol{\beta}_{PP}$ . Os pseudos-parâmetros somente seriam calculáveis se os valores das variáveis explicativas e da variável resposta fossem conhecidos para todos os  $N$  elementos da população, mas na realidade encontra-se disponível apenas uma amostra selecionada da população usando o plano amostral descrito na seção 3.3, sendo, portanto necessário determinar estimadores consistentes para os pseudos-parâmetros. Segundo Pessoa & Silva (1998), um estimador adequado para o vetor  $\boldsymbol{\beta}_{PP}$  também é adequado para o vetor  $\boldsymbol{\beta}$  de parâmetros de interesse.

Assim, para obter estimadores para os pseudos-parâmetros estima-se o vetor de totais populacionais  $T(\boldsymbol{\beta}) = \sum_{i=1}^N u_i(\boldsymbol{\beta})$  usando o vetor de estimadores lineares ponderados de totais, denotado por  $\hat{T}(\boldsymbol{\beta}) = \sum_{i=1}^n w_i u_i(\boldsymbol{\beta})$ , onde  $w_i$  é o peso amostral referente a  $i$ -ésima unidade de amostra. O estimador de MPV do vetor  $\boldsymbol{\beta}_{PP}$ , conseqüentemente do vetor  $\boldsymbol{\beta}$ , pode ser obtido através da solução do seguinte sistema de equações de pseudo-verossimilhanças:

$$\hat{T}(\hat{\boldsymbol{\beta}}_{MPV}) = \sum_{i=1}^n w_i u_i(\hat{\boldsymbol{\beta}}_{MPV}) = \mathbf{0} \quad (3.19)$$

Ao obter o estimador de MPV do vetor  $\boldsymbol{\beta}$ , denotado por  $\hat{\boldsymbol{\beta}}_{MPV}$ , pode-se estimar a probabilidade do  $i$ -ésimo elemento da amostra pertencer à categoria  $k$ , dado o vetor  $\mathbf{x}_i$  de valores observados das variáveis explicativas, da seguinte forma:

$$\begin{aligned} \hat{p}_{ik} &= \hat{p}_{ik}(\hat{\boldsymbol{\beta}}_{MPV}) = P(\hat{Y} = k | \mathbf{x}_i) = P(\hat{Y} \leq k | \mathbf{x}_i) - P(\hat{Y} \leq k - 1 | \mathbf{x}_i) = \\ &= \frac{e^{\hat{\alpha}_k - \mathbf{x}_i \hat{\boldsymbol{\beta}}_{MPV}}}{1 + e^{\hat{\alpha}_k - \mathbf{x}_i \hat{\boldsymbol{\beta}}_{MPV}}} - \frac{e^{\hat{\alpha}_{k-1} - \mathbf{x}_i \hat{\boldsymbol{\beta}}_{MPV}}}{1 + e^{\hat{\alpha}_{k-1} - \mathbf{x}_i \hat{\boldsymbol{\beta}}_{MPV}}} \end{aligned} \quad (3.20)$$

Para estimar a variância assintótica dos estimadores de MPV dos parâmetros de um modelo de regressão ordinal, Liu & Koirala (2013) destacam o uso da expressão geral desenvolvida por Binder (1983) para modelos lineares generalizados ajustados com dados amostrais complexos. Com base em (3.21), observa-se que a matriz de variância-covariância dos estimadores de MPV é obtida através da multiplicação da matriz de variância-covariância do vetor dos totais amostrais dos escores ponderados, pelo inverso da matriz de derivadas segundas, com relação aos parâmetros estimados pelo método de MPV. Assim, a estimação dessa variância é dada por:

$$\widehat{Var}(\widehat{\boldsymbol{\beta}}_{MPV}) \cong [\widehat{J}(\widehat{\boldsymbol{\beta}}_{MPV})]^{-1} \widehat{Var}[\widehat{T}(\widehat{\boldsymbol{\beta}}_{MPV})] [\widehat{J}(\widehat{\boldsymbol{\beta}}_{MPV})]^{-1} \quad (3.21)$$

onde:

$\widehat{Var}[\widehat{T}(\widehat{\boldsymbol{\beta}}_{MPV})] = \widehat{Var}[\sum_{i=1}^n w_i u_i(\widehat{\boldsymbol{\beta}}_{MPV})]$  é o estimador consistente para a matriz de variância-covariância do estimador do total de escores populacionais, ou do vetor de totais amostrais dos escores ponderados.

$[\widehat{J}(\widehat{\boldsymbol{\beta}}_{MPV})]^{-1}$  é a inversa da matriz simétrica,

$$\widehat{J}(\widehat{\boldsymbol{\beta}}_{MPV}) = \left. \frac{\partial \widehat{T}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}_{MPV}} = \sum_{i=1}^n w_i \left. \frac{\partial u_i(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} \right|_{\boldsymbol{\beta}=\widehat{\boldsymbol{\beta}}_{MPV}}, \text{ de dimensão } P \times P.$$

No caso de amostras suficientemente grandes, o vetor de estimadores de MPV dos parâmetros do modelo de regressão logística ordinal segue uma distribuição assintoticamente normal multivariada:

$\widehat{\boldsymbol{\beta}}_{MPV} \sim N[\boldsymbol{\beta}, \widehat{VAR}(\widehat{\boldsymbol{\beta}}_{MPV})]$ , e, portanto pode-se afirmar que:

$$[\widehat{VAR}(\widehat{\boldsymbol{\beta}}_{MPV})]^{-\frac{1}{2}} (\widehat{\boldsymbol{\beta}}_{MPV} - \boldsymbol{\beta}) \sim N(\mathbf{0}, \mathbf{I}) \quad (3.22)$$

### 3.6.4 Teste de homogeneidade geral dos coeficientes

O teste de homogeneidade geral dos parâmetros, também chamado de teste de proporcionalidade das chances ou teste de paralelismo, permite avaliar se a suposição de homogeneidade dos coeficientes é satisfeita ou não para um determinado nível de significância. Em outras palavras, permite avaliar se os parâmetros do modelo são iguais para as diferentes equações. A seguir são fornecidas as hipóteses a serem testadas, a estatística de teste, a região crítica e o critério de decisão do teste.

1) *Hipóteses a serem testadas:*

$$\begin{cases} H_0: \boldsymbol{\beta}^{(1)} = \boldsymbol{\beta}^{(2)} = \dots = \boldsymbol{\beta}^{(K-1)} \\ H_1: \boldsymbol{\beta}^{(k)} \neq \boldsymbol{\beta}^{(l)} \text{ para algum } k, l \text{ (} k \neq l; \text{ tal que: } k, l = 1, 2, \dots, K-1) \end{cases}$$

onde:

$\boldsymbol{\beta}^{(k)} = (\beta_{1k}, \beta_{2k}, \dots, \beta_{pk})'$  é o vetor, de dimensão  $1 \times p$ , de parâmetros desconhecidos referentes a  $k$ -ésima categoria da variável resposta do modelo;  $k=1, 2, \dots, K-1$ .

As hipóteses acima podem ser expressas usando notação matricial do seguinte modo (IBM, 2012):

$$\begin{cases} H_0: \mathbf{C}\boldsymbol{\beta} = \mathbf{0} \\ H_1: \mathbf{C}\boldsymbol{\beta} \neq \mathbf{0} \end{cases}$$

onde  $\mathbf{C}$  é uma matriz de contrastes de dimensão  $R \times P$ ;  $\boldsymbol{\beta}$  é um vetor de parâmetros desconhecidos de dimensão  $P \times 1$ ; e  $\mathbf{0}$  é um vetor de dimensão  $R \times 1$ .

Por meio de palavras, as hipóteses podem ser definidas do seguinte modo:

$$\begin{cases} H_0: \text{A hipótese de homogeneidade geral é satisfeita.} \\ H_1: \text{A hipótese de homogeneidade geral não é satisfeita.} \end{cases}$$

2) *Estatística de teste:*

Para testar as hipóteses, utiliza-se a estatística de Wald, construída considerando a restrição de que  $\mathbf{C}\boldsymbol{\beta} = \mathbf{0}$ , dada por:

$$W = (\mathbf{C}\widehat{\boldsymbol{\beta}}_{MPV})' \cdot [\mathbf{C} \widehat{VAR}(\widehat{\boldsymbol{\beta}}_{MPV}) \mathbf{C}']^{-1} \cdot (\mathbf{C}\widehat{\boldsymbol{\beta}}_{MPV}) \quad (3.23)$$

A estatística  $W$  tem distribuição assintótica Qui-quadrada com  $r$  graus de liberdade, sendo  $r = \text{posto}[\mathbf{C} \widehat{VAR}(\widehat{\boldsymbol{\beta}}_{MPV}) \mathbf{C}']$ .

3) *Região Crítica:*

A região crítica do teste de Wald conduzido para  $H_0: \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$  é dada por  $RC = \{w \in \mathbb{R} | w \geq \chi_{\alpha, r}^2\}$ , e o p-valor =  $P(W \geq w_{obs})$ .

4) *Critério de decisão:*

Com relação ao critério de decisão, rejeita-se  $H_0: \mathbf{C}\boldsymbol{\beta} = \mathbf{0}$  ao nível de significância de  $\alpha=5\%$ , se o valor observado ( $w_{obs}$ ) da estatística de teste pertencer a região crítica. Equivalentemente, rejeita-se  $H_0$  se o p-valor =  $P(W \geq w_{obs}) \leq \alpha = 5\%$ . No caso de se rejeitar  $H_0$ , conclui-se que a hipótese de homogeneidade geral dos parâmetros (proporcionalidade das chances) não é satisfeita; e no caso de não rejeição de  $H_0$ , conclui-se que hipótese de homogeneidade geral é satisfeita.

### 3.6.5 Teste de Wald de significância individual dos parâmetros do modelo

Para avaliar a significância individual dos parâmetros de um modelo de regressão logística ordinal utiliza-se o teste de Wald, considerando o nível de significância de 5%. A seguir são fornecidas as hipóteses a serem testadas, a estatística de teste, a região crítica e o critério de decisão.



1) *Hipóteses a serem testadas:*

$$\begin{cases} H_0: \beta_{jk} = 0 \\ H_1: \beta_{jk} \neq 0 \end{cases}$$

onde,  $\beta_{jk}$  é o efeito principal do  $j$ -ésima variável explicativa referente a  $k$ -ésima categoria da variável resposta do modelo.

2) *Estatística de teste:*

Sob  $H_0: \beta_{jk} = 0$ , a estatística de teste é dada por:

$$Z = \frac{\hat{\beta}_{jk}}{\sqrt{\text{VAR}(\hat{\beta}_{jk})}} \sim N(0,1) \quad (3.24)$$

Ou alternativamente por:

$$W = Z^2 \sim \chi_1^2 \quad (3.25)$$

onde,  $\hat{\beta}_{jk}$  é o estimador de MPV do parâmetro  $\beta_{jk}$ .

3) *Região crítica:*

$$RC = \{z \in \mathbb{R} \mid |z| > z_{\alpha/2}\}$$

4) *Critério de decisão:*

Se o valor observado ( $z_{obs}$ ) da estatística de teste  $Z$  pertencer à região crítica, rejeita-se  $H_0$  ao nível de significância de 5% e conclui-se que  $\beta_{jk}$  é significativamente diferente de zero. Caso contrário, não há evidências para rejeitar  $H_0$  ao nível de significância de 5%, e neste caso  $\beta_{jk}$  não é significativamente diferente de zero.

Equivalentemente, o critério de decisão pode ser baseado no p-valor do teste bilateral, sendo calculado por:

$$p - \text{valor} = 2 \cdot P(Z > |z_{obs}|) \quad (3.26)$$

Nesta abordagem, rejeita-se  $H_0$  ao nível de significância de 5% se o  $p - \text{valor} \leq 0,05$  e conclui-se que  $\beta_{jk}$  é significativamente diferente de zero. Mas se o  $p - \text{valor} > 0,05$ , não há evidências para rejeitar  $H_0$  ao nível de significância de 5%, e neste caso  $\beta_{jk} = 0$ .

### 3.6.6 Intervalo de confiança para os parâmetros e razões de chances

Conforme a expressão (3.22) os estimadores de MPV dos parâmetros de um modelo de regressão logística ordinal seguem uma distribuição assintoticamente normal para amostras suficientemente grandes. Desse modo, pode-se obter um intervalo de confiança (IC) para o parâmetro  $\beta_{jk}$ , considerando um nível de confiança de  $100(1 - \alpha)\%$ , do seguinte modo:

$$IC_{\beta_{jk}, 100(1-\alpha)\%} \cong [\hat{\beta}_{jk} - z_{\alpha/2} \sqrt{\widehat{VAR}(\hat{\beta}_{jk})}; \hat{\beta}_{jk} + z_{\alpha/2} \sqrt{\widehat{VAR}(\hat{\beta}_{jk})}] \quad (3.27)$$

onde,

$\hat{\beta}_{jk}$  é o estimador de MPV de  $\beta_{jk}$ , sendo que  $\beta_{jk}$  é o efeito principal do  $j$ -ésima variável explicativa referente a  $k$ -ésima categoria da variável resposta do modelo.

$\widehat{VAR}(\hat{\beta}_{jk})$  é o estimador da variância de  $\hat{\beta}_{jk}$ .

Usando os limites do IC definido em (3.27), pode-se obter o IC para a razão de chance  $OR_{jk} = e^{\beta_{jk}}$ , através de (3.28).

$$IC_{OR_{jk}, 100(1-\alpha)\%} \cong [e^{\hat{\beta}_{jk} - z_{\alpha/2} \sqrt{\widehat{VAR}(\hat{\beta}_{jk})}}; e^{\hat{\beta}_{jk} + z_{\alpha/2} \sqrt{\widehat{VAR}(\hat{\beta}_{jk})}}] \quad (3.28)$$

### 3.6.7 Medidas de qualidade do ajuste

Existem basicamente duas formas de avaliar a qualidade do ajuste de um modelo de regressão logística ordinal, sendo uma forma através do cálculo das medidas *pseudo*  $R^2$  e a outra através do cálculo das taxas de classificações corretas.

Através do cálculo das medidas de *pseudo*  $R^2$ , as mais usuais o *pseudo*  $R^2$  de *McFadden*, *Cox & Snell* e *Nagelkerke*. Os *pseudo*- $R^2$  são baseados na comparação do logaritmo da estimativa da função de verossimilhança populacional para o modelo de interesse (modelo selecionado)<sup>3</sup>, denotado por  $\ln \hat{L}(\hat{\beta} | \mathbf{y})$ , com o logaritmo da estimativa da função de verossimilhança populacional para o modelo com apenas o intercepto (modelo nulo)<sup>4</sup>, sendo este denotado por  $\ln \hat{L}(\hat{\beta}^{(0)} | \mathbf{y})$ . Essas medidas indicam o ganho estimado de informação pelo modelo selecionado, comparativamente ao modelo nulo. Quanto maior o valor dessas medidas, melhor o ajuste do modelo.

O *pseudo*- $R^2$  de *McFadden* para um modelo de regressão logística ordinal pode ser calculado da seguinte forma:

$$R_{MF}^2 = 1 - \frac{\ln \hat{L}(\hat{\beta} | \mathbf{y})}{\ln \hat{L}(\hat{\beta}^{(0)} | \mathbf{y})} \quad (3.29)$$

onde,  $0 \leq R_{MF}^2 \leq 1$ .

O *pseudo*- $R^2$  de *Cox & Snell*, por sua vez, é calculado através da equação (3.30). Para desfechos categóricos, essa medida de qualidade do ajuste tem um valor máximo inferior a 1 (unidade), inclusive no caso de um ajuste perfeito.

$$R_{CS}^2 = 1 - e^{\left\{ \frac{2[\ln \hat{L}(\hat{\beta}^{(0)} | \mathbf{y}) - \ln \hat{L}(\hat{\beta} | \mathbf{y})]}{N} \right\}} \quad (3.30)$$

onde,  $0 \leq R_{CS}^2 < 1$ .

<sup>3</sup> Logaritmo da função de pseudo-verossimilhança do modelo selecionado.

<sup>4</sup> Logaritmo da função de pseudo-verossimilhança do modelo nulo.

O *pseudo R<sup>2</sup> de Nagelkerke* é uma versão corrigida do *pseudo-R<sup>2</sup> de Cox & Snell* por ajustar a escala da estatística para cobrir toda a amplitude de 0 a 1, podendo assumir o valor 1 no caso do ajuste ser perfeito.

O *pseudo-R<sup>2</sup> de Nagelkerke* é calculado do seguinte modo:

$$R_N^2 = \frac{R_{CS}^2}{1 - e^{\left\{ \frac{2 \ln \hat{L}(\hat{\beta}^{(0)} | y)}{N} \right\}}} \quad (3.31)$$

onde,  $0 \leq R_N^2 \leq 1$ .

A outra forma de avaliar a qualidade do ajuste do modelo é medindo a capacidade preditiva do modelo, através do cálculo das taxas de classificações corretas (%), em termos gerais e dentro de cada categoria observada da variável resposta (medidas de sensibilidade).

Para medir a capacidade preditiva do modelo logístico ordinal, pode-se construir uma tabela de contingência para as categorias observadas e preditas da variável resposta do modelo. As categorias preditas são obtidas a partir das probabilidades estimadas de cada unidade pertencer a cada uma das  $K$  categorias da variável resposta do modelo. Define-se a categoria predita, que cada unidade pertence, como sendo aquela que apresenta a maior probabilidade.

Uma vez que neste trabalho considerou-se uma variável resposta  $Y$  com  $K=3$  categorias, pode-se obter o quadro 3:

**Quadro 3:** Distribuição absoluta das unidades segundo as categorias observadas e preditas da variável resposta do modelo.

Categorias observadas	Categorias preditas			Total
	$\hat{Y} = 1$	$\hat{Y} = 2$	$\hat{Y} = 3$	
$Y = 1$	$\hat{N}_{11}$	$\hat{N}_{12}$	$\hat{N}_{13}$	$\hat{N}_{1\circ}$
$Y = 2$	$\hat{N}_{21}$	$\hat{N}_{22}$	$\hat{N}_{23}$	$\hat{N}_{2\circ}$
$Y = 3$	$\hat{N}_{31}$	$\hat{N}_{32}$	$\hat{N}_{33}$	$\hat{N}_{3\circ}$
<b>Total</b>	$\hat{N}_{\circ 1}$	$\hat{N}_{\circ 2}$	$\hat{N}_{\circ 3}$	$\hat{N}$

onde,

$\hat{N}_{ij}$  é o número estimado de unidades referentes a  $i$ -ésima categoria observada e a  $j$ -ésima categoria predita da variável resposta. As quantidades nas posições  $i=j$  indicam o número de unidades classificadas corretamente nas respectivas categorias;

$\hat{N}_{i\circ}$  é o total estimado de unidades da  $i$ -ésima categoria observada da variável resposta;

$\hat{N}_{\circ j}$  é o total estimado de unidades da  $j$ -ésima categoria predita da variável resposta;

$\hat{N} = \sum_{i=1}^3 \sum_{j=1}^3 \hat{N}_{ij}$  é o total estimado de unidades na população.

A taxa global de classificações corretas expressa em termos percentuais (%) é dada por:

$$TG = \left( \frac{\hat{N}_{11} + \hat{N}_{22} + \hat{N}_{33}}{\hat{N}} \right) \cdot 100 \quad (3.32)$$

As taxas de classificações corretas dentro de cada categoria observada da variável resposta do modelo representam as sensibilidades das categorias 1, 2 e 3 da variável resposta, sendo representadas em termos percentuais (%) por:

$$S_1 = P(\hat{Y}_1 = 1 | Y_1 = 1) = \left( \frac{\hat{N}_{11}}{\hat{N}_{1\circ}} \right) \cdot 100 \quad (3.33)$$

$$S_2 = P(\hat{Y}_2 = 2 | Y_2 = 2) = \left( \frac{\hat{N}_{22}}{\hat{N}_{2\circ}} \right) \cdot 100 \quad (3.34)$$

$$S_3 = P(\hat{Y}_3 = 3 | Y_3 = 3) = \left( \frac{\hat{N}_{33}}{\hat{N}_{3\circ}} \right) \cdot 100 \quad (3.35)$$

## 4. Resultados

Nesta seção são apresentados os resultados da distribuição percentual dos idosos por “grau de dificuldade na locomoção”, segundo as características sociodemográficas e de saúde dos idosos de 60 anos ou mais de idade. Adicionalmente, são fornecidos os resultados do ajuste do modelo de chances proporcionais parciais (MCP) adotado neste trabalho, tendo em vista que o teste de homogeneidade geral dos coeficientes foi significativo ( $p\text{-valor} < 0,001$ ). Assim, conclui-se que a suposição de homogeneidade dos coeficientes (ou equivalentemente de proporcionalidade das chances) não foi satisfeita para todas as variáveis explicativas consideradas no ajuste do modelo.

**Tabela 1:** Distribuição percentual (%) dos idosos de 60 anos ou mais por “grau de dificuldade na locomoção”.

<b>Grau de dificuldade na locomoção</b>	<b>Percentual dos idosos (<math>\hat{N}=26.404.863</math>)</b>
Nenhum	73,9
No máximo moderado	19,3
Grave	6,8
<b>Total</b>	<b>100</b>

Fonte: Elaborado pela autora com base nos dados da PNS 2013.

Do total de idosos, 73,9% não reportaram apresentar dificuldade de locomoção, enquanto 19,3% dos idosos disseram apresentar no máximo um grau de dificuldade moderado, e os 6,8% restantes disseram ter grave dificuldade na locomoção (Tabela 1).

A tabela 2 fornece a distribuição percentual dos idosos por “grau de dificuldade na locomoção”, segundo as suas características sociodemográficas e de saúde.

**Tabela 2:** Distribuição percentual dos idosos por “grau de dificuldade na locomoção”, segundo as suas características sociodemográficas e de saúde.

Variáveis explicativas	Percentual de idosos (%)	Grau de dificuldade na locomoção		
		Nenhum $\hat{N}_1=19.498.883$	No máximo moderado $\hat{N}_2=5.104.039$	Grave $\hat{N}_3=1.801.941$
<b>Sexo</b>				
Masculino	43,6	76,9	17,0	6,1
Feminino	56,4	71,5	21,1	7,4
<b>Faixa etária</b>				
De 60 a 69 anos	56,4	82,9	13,2	3,9
De 70 a 79 anos	30,0	69,6	22,9	7,5
80 anos ou mais	13,6	45,8	36,8	17,4
<b>Cor</b>				
Branca	53,6	73,8	19,5	6,7
Não branca	46,4	73,9	19,1	7,0
<b>Situação conjugal</b>				
Casado	57,4	77,8	16,9	5,3
Não casado	42,6	68,6	22,6	8,8
<b>Regiões brasileiras</b>				
Norte	5,4	74,2	18,4	7,4
Nordeste	25,2	70,9	20,9	8,2
Sudeste	47,9	76,3	18,1	5,6
Sul	15,1	69,7	22,8	7,5
Centro-Oeste	6,4	76,9	14,8	8,3
<b>Área do domicílio</b>				
Urbana	85,2	74,4	19,0	6,6
Rural	14,8	70,5	21,5	8,0
<b>Nível educacional</b>				
Sem instrução	32,1	65,4	23,7	10,9
Ensino fundamental	45,5	74,5	19,6	5,9
Ensino médio	12,2	83,7	13,4	2,9
Ensino superior	10,2	85,5	11,8	2,7
<b>Qualidade da moradia</b>				
Adequado	78,8	74,8	18,9	6,3
Inadequado	21,2	70,2	20,9	8,9
<b>Plano de saúde</b>				
Sim	32,0	76,3	18,0	5,7
Não	68,0	72,7	20,0	7,3

(Continua)

**Tabela 2** (*Continuação*): Distribuição percentual dos idosos por “grau de dificuldade na locomoção”, segundo as suas características sociodemográficas e de saúde.

Variáveis explicativas	Percentual de idosos (%)	Grau de dificuldade na locomoção		
		Nenhum $\hat{N}_1=19.498.883$	No máximo moderado $\hat{N}_2=5.104.039$	Grave $\hat{N}_3=1.801.941$
<b>Procura o mesmo lugar ou serviço</b>				
Sim	79,5	73,7	19,5	6,8
Não	20,5	74,2	18,8	7,0
<b>Autoavaliação de saúde</b>				
Bom	44,4	86,8	9,9	3,3
Regular	43,5	70,1	23,6	6,3
Ruim	12,1	40,0	38,4	21,6
<b>Diagnóstico de alguma doença</b>				
Sim	40,0	65,3	25,8	8,9
Não	60,0	79,5	15,0	5,5

Fonte: Elaborado pela autora com base nos dados da PNS 2013.

De acordo com a tabela 2, observa-se que mais da metade dos idosos é do sexo feminino (56,4%), casados (57,4%) e de cor branca (53,6%). A faixa etária com maior percentual é a faixa de 60 a 69 anos (56,4%) e quase metade dos idosos reside na região Sudeste do Brasil (47,9%). A maioria da população idosa reside na área urbana (85,2%) e mora em domicílios considerados adequados quanto à qualidade de construção (78,8%). O nível educacional mais frequente entre os idosos é o ensino fundamental (45,5%) e apenas 10,2% idosos têm nível superior.

Com relação às características de saúde, 68,0% dos idosos não tem plano de saúde, mas a maior parte deles procura o mesmo serviço, o mesmo médico ou mesmo lugar quando precisam de atendimento (79,5%). Quanto à autoavaliação geral de saúde, 55,6% dos idosos reportaram um estado de saúde regular ou ruim, enquanto a minoria avaliou a sua saúde como boa (44,4%). Além disso, 40,0% dos idosos tiveram diagnóstico de alguma doença crônica, física ou mental.

A partir da análise bivariada fornecida na tabela 2, pode-se observar um maior percentual de idosos reportando uma grave dificuldade de locomoção entre as mulheres (7,4%) e entre aqueles com 80 anos ou mais de idade (17,4%). Ainda com relação a idade dos idosos, observa-se que quanto maior a faixa etária, maior a porcentagem de



idosos reportando grave dificuldade na locomoção. Também se observou maior porcentagem de idosos com grave dificuldade na locomoção entre os idosos não casados (8,8%) e os de cor não branca (7,0%).

Quanto as grandes regiões brasileiras, nota-se que nas regiões Norte e Sul aproximadamente 7% dos idosos reportaram grave dificuldade na locomoção, enquanto que nas regiões Nordeste e Centro-Oeste este percentual foi cerca de 8%, e no Sudeste foi de 5,6%. Quanto a área de localização do domicílio e a qualidade da construção da moradia, verifica-se maior percentual de idosos com grave dificuldade de locomoção na área rural (8,0%) e em domicílios considerados inadequados no que se refere à qualidade de construção da moradia (8,9%).

Com relação ao nível educacional foi possível notar que o percentual de idosos que reportaram grave dificuldade de locomoção aumenta à medida que o nível de escolaridade diminui, sendo que a porcentagem de idosos com grave dificuldade na locomoção é de 10,9% entre os sem instrução e de 2,7% entre aqueles que possuem nível superior.

Nos aspectos relacionados à saúde dos idosos, verificou-se na análise bivariada que a porcentagem dos que reportaram grave dificuldade de locomoção é maior nos idosos sem plano de saúde (7,3%) comparativamente aos que tem plano de saúde (5,7%). Já a porcentagem de idosos com grave dificuldade de locomoção é quase a mesma em relação aos que procuram o mesmo lugar, serviço ou médico (6,8%) e aos idosos que não procuram (7,0%).

Ainda com relação aos aspectos de saúde, observa-se que há um aumento considerável na porcentagem de idosos que reportaram uma grave dificuldade na locomoção à medida que piora a sua autoavaliação geral de saúde. A porcentagem de idosos com grave dificuldade na locomoção que autoavaliaram a saúde como boa é de 3,3%, como regular é de 6,3% e como ruim é de 21,6%. Observa-se também que a porcentagem de idosos com grave grau de dificuldade na locomoção é maior entre os idosos que tiveram o diagnóstico de alguma doença crônica, física ou mental (8,9%), comparativamente aos idosos que não tiveram esse diagnóstico (5,5%).

A tabela 3 apresenta os principais resultados do ajuste do MCPD com todas as doze variáveis explicativas consideradas no presente estudo.

**Tabela 3:** Resultados do ajuste do modelo de chances proporcionais parciais (MCP) explicativo do “Grau de dificuldade na locomoção”, considerando todas as variáveis explicativas.

Variáveis explicativas	Comparações					
	Equação 1: (No máximo moderado + Grave) versus Nenhum			Equação 2: Grave versus (Nenhum + No máximo moderado)		
	$OR_{j1}$	$P$ -valor	$IC_{OR_{j1};95\%}$	$OR_{j2}$	$P$ -valor	$IC_{OR_{j2};95\%}$
<b>Sexo</b>						
Masculino	0,885	0,192	(0,736-1,064)	0,885	0,192	(0,736-1,064)
Feminino	1	-	-	1	-	-
<b>Faixa etária</b>						
De 60 a 69 anos	0,185	<0,001	(0,149-0,229)	0,185	<0,001	(0,149-0,229)
De 70 a 79 anos	0,367	<0,001	(0,293-0,459)	0,367	<0,001	(0,293-0,459)
80 anos ou mais	1	-	-	1	-	-
<b>Cor</b>						
Branca	1,160	0,111	(0,967-1,393)	1,160	0,111	(0,967-1,393)
Não branca	1	-	-	1	-	-
<b>Situação conjugal</b>						
Casado	0,788	0,004	(0,671-0,925)	0,788	0,004	(0,671-0,925)
Não casado	1	-	-	1	-	-
<b>Regiões brasileiras</b>						
Norte	1,141	0,431	(0,822-1,585)	1,141	0,431	(0,822-1,585)
Nordeste	1,076	0,576	(0,832-1,393)	1,076	0,576	(0,832-1,393)
Sudeste	1,035	0,803	(0,792-1,351)	1,035	0,803	(0,792-1,351)
Sul	1,391	0,048	(1,003-1,928)	1,391	0,048	(1,003-1,928)
Centro-Oeste	1	-	-	1	-	-
<b>Área do domicílio</b>						
Urbana	0,989	0,925	(0,793-1,235)	0,989	0,925	(0,793-1,235)
Rural	1	-	-	1	-	-
<b>Nível educacional</b>						
Sem instrução	1,542	0,013	(1,094-2,172)	1,542	0,013	(1,094-2,172)
Ensino fundamental	1,274	0,145	(0,920-1,764)	1,274	0,145	(0,920-1,764)
Ensino médio	0,946	0,754	(0,671-1,334)	0,946	0,754	(0,671-1,334)
Ensino superior	1	-	-	1	-	-
<b>Qualidade da moradia</b>						
Adequado	0,922	0,473	(0,738-1,151)	0,922	0,473	(0,738-1,151)
Inadequado	1	-	-	1	-	-

(Continua)

**Tabela 3** (Continuação): Resultados do ajuste do modelo de chances proporcionais parciais (MCP) explicativo do “Grau de dificuldade na locomoção”, considerando todas as variáveis explicativas.

Variáveis explicativas	Comparações					
	Equação 1: (No máximo moderado + Grave) versus Nenhum			Equação 2: Grave versus (Nenhum + No máximo moderado)		
	$OR_{j1}$	<i>P</i> -valor	$IC_{OR_{j1};95\%}$	$OR_{j2}$	<i>P</i> -valor	$IC_{OR_{j2};95\%}$
<b>Plano de saúde</b>						
Sim	1,092	0,353	(0,906-1,312)	1,092	0,353	(0,906-1,312)
Não	1	-	-	1	-	-
<b>Procura o mesmo lugar, médico ou serviço</b>						
Sim	0,954	0,628	(0,788-1,155)	0,954	0,628	(0,788-1,155)
Não	1	-	-	1	-	-
<b>Autoavaliação de saúde</b>						
Bom	0,111	<0,001	(0,087-0,142)	0,154	<0,001	(0,108-0,218)
Regular	0,285	<0,001	(0,233-0,350)	0,285	<0,001	(0,233-0,350)
Ruim	1	-	-	1	-	-
<b>Diagnóstico de alguma doença</b>						
Sim	1,633	<0,001	(1,367-1,951)	1,232	0,138	(0,935-1,623)
Não	1	-	-	1	-	-

Fonte: Elaborado pela autora com base nos dados da PNS 2013.

Na tabela 3 foram apresentados os resultados do ajuste do MCP considerando as 12 variáveis explicativas, mas as seguintes variáveis não apresentaram efeito estatisticamente significativo ao nível de significância de 5% ( $p$ -valor<0,05): “*sexo*”, “*cor*”, “*área de localização do domicílio*”, “*qualidade de construção de moradia*”, “*plano de saúde*” e “*procura o mesmo lugar, mesmo médico ou mesmo serviço de saúde*”.

A tabela 4 apresenta os resultados do ajuste do MCP selecionado, que contém apenas as variáveis que apresentaram pelo menos uma das categorias, com efeito, estatisticamente significativo ao nível de 5%.

**Tabela 4:** Resultados do ajuste do modelo de chances proporcionais parciais (MCP) explicativo do “Grau de dificuldade na locomoção”, contendo as variáveis explicativas selecionadas.

Variáveis explicativas	Comparações					
	Equação 1: (No máximo moderado + Grave) versus Nenhum			Equação 2: Grave versus (Nenhum + No máximo moderado)		
	$OR_{j1}$	<i>P</i> -valor	$IC_{OR_{j1};95\%}$	$OR_{j2}$	<i>P</i> -valor	$IC_{OR_{j2};95\%}$
<b>Faixa etária</b>						
De 60 a 69 anos	0,181	<0,001	(0,147-0,222)	0,181	<0,001	(0,147- 0,222)
De 70 a 79 anos	0,359	<0,001	(0,288-0,448)	0,359	<0,001	(0,288-0,448)
80 anos ou mais	1	-	-	1	-	-
<b>Situação conjugal</b>						
Casado	0,759	<0,001	(0,653-0,883)	0,759	<0,001	(0,653-0,883)
Não casado	1	-	-	1	-	-
<b>Regiões brasileiras</b>						
Norte	1,101	0,563	(0,795-1,525)	1,101	0,563	(0,795-1,525)
Nordeste	1,062	0,644	(0,824-1,368)	1,062	0,644	(0,824-1,368)
Sudeste	1,041	0,765	(0,800-1,354)	1,041	0,765	(0,800-1,354)
Sul	1,448	0,019	(1,064-1,971)	1,448	0,019	(1,064-1,971)
Centro-Oeste	1	-	-	1	-	-
<b>Nível educacional</b>						
Sem instrução	1,433	0,026	(1,043-1,968)	1,433	0,026	(1,043-1,968)
Ensino fundamental	1,204	0,239	(0,884-1,639)	1,204	0,239	(0,884-1,639)
Ensino médio	0,926	0,662	(0,657-1,306)	0,926	0,662	(0,657-1,306)
Ensino superior	1	-	-	1	-	-
<b>Autoavaliação de saúde</b>						
Bom	0,113	<0,001	(0,089-0,145)	0,156	<0,001	(0,110-0,221)
Regular	0,289	<0,001	(0,236-0,355)	0,289	<0,001	(0,236-0,355)
Ruim	1	-	-	1	-	-
<b>Diagnóstico de alguma doença</b>						
Sim	1,645	<0,001	(1,380-1,962)	1,247	0,117	(0,947-1,639)
Não	1	-	-	1	-	-

Fonte: Elaborado pela autora com base nos dados da PNS 2013.

Na tabela 4, é possível observar primeiramente que as variáveis “*faixa etária*”, “*situação conjugal*”, “*grandes regiões brasileiras*” e “*nível educacional*” apresentam efeitos iguais (homogêneos) para as duas equações. Já as variáveis “*autoavaliação geral de saúde*” e “*diagnóstico de alguma doença crônica, física ou mental*” apresentam níveis que diferem entre as duas equações.

Com base no MCPP explicativo do grau de dificuldade na locomoção, conclui-se que quanto maior a faixa etária, maior a chance de o idoso apresentar pior grau de dificuldade na locomoção. Os idosos não casados têm chance 31,8% maior de apresentar pior grau de dificuldade na locomoção comparativamente aos idosos casados (OR=1/0,759=1,318). Os idosos residentes na região Sul têm uma chance 44,8% maior de apresentar pior grau de dificuldade na locomoção do que os idosos da região Centro-Oeste (OR=1,448; p-valor=0,019).

A chance dos idosos sem instrução apresentarem pior grau de dificuldade na locomoção é 44,3% maior que os idosos com nível superior (OR=1,433; p-valor=0,026).

Os idosos com diagnóstico de alguma doença crônica, física ou mental têm chance de 64,5% maior de apresentar algum grau de dificuldade na locomoção (no máximo moderado ou grave), comparativamente aos idosos sem diagnóstico dessas doenças (OR=1,645; p-valor<0,001).

A chance dos idosos com saúde boa apresentar uma grave dificuldade na locomoção é 84,4% menor que os idosos com saúde ruim (OR=0,156; p-valor<0,001). E a chance dos idosos com saúde boa apresentar algum grau de dificuldade na locomoção (no máximo moderado ou grave) é 88,7% menor que os idosos com saúde ruim (OR=0,113; p-valor<0,001). Além disso, idosos com saúde autorreferida regular têm chance 71,1% menor de apresentar pior grau de dificuldade na locomoção comparativamente aos idosos com saúde ruim (OR=0,289; p-valor<0,001).

No que se refere a qualidade do ajuste do MCPP, primeiramente foram obtidos os logaritmos das funções de pseudo-verossimilhanças do modelo selecionado e do modelo nulo disponibilizados pelo programa STATA, e a partir dessas medidas foram calculados os *pseudo-R<sup>2</sup>* usando as expressões apresentadas em (3.29), (3.30) e (3.31).

Assim, verificou-se que o *pseudo-R<sup>2</sup> de MCFadden* é igual a 14,4%; o *pseudo-R<sup>2</sup> de Cox & Snell* é igual a 18,8% e o *pseudo-R<sup>2</sup> de Nagelkerke* igual a 24,6%. Assim, com base no *pseudo-R<sup>2</sup> de Nagelkerke*, pode-se dizer que o ganho de informação obtido com o modelo selecionado foi de 24,6%, comparativamente ao modelo nulo.

Quanto à taxa global de classificações corretas, verificou-se que 74,8% dos idosos foram classificados corretamente pelo modelo MCPP selecionado. Além disso, o modelo conseguiu prever todos os diferentes níveis de dificuldade na locomoção, mas com diferentes percentuais de acertos. Dentre os idosos com nenhum grau de dificuldade na locomoção, 96,1% foram classificados corretamente pelo modelo; dos idosos com no máximo moderado grau dificuldade, 16,4% foram classificados corretamente; e dos idosos com grave grau de dificuldade, o percentual foi de 10,2%.

## 5. Discussões e conclusões

Neste trabalho foi abordada a temática incapacidade funcional de idosos no Brasil, que é um conceito utilizado para avaliar as condições de saúde de idosos. Embora haja diferentes formas de mensuração deste conceito, neste trabalho utilizou-se o grau de dificuldade de locomoção como *proxy* da incapacidade funcional, considerando três categorias (“nenhum”, “no máximo moderado”, “grave”).

Através da análise ajustada, não se observou associação entre o sexo do idoso e o grau de dificuldade na locomoção. Embora alguns estudos tenham mostrado que a incapacidade funcional em mulheres idosas é maior do que em homens (GIACOMIN et al., 2008; ALVES et al., 2010); por outro lado, outros estudos, como os de Duca et al. (2009) e Rosa et al. (2003), também observaram ausência de associação entre sexo e incapacidade. Cabe destacar que Duca et al. (2009) e Rosa et al. (2003) utilizaram estudos transversais realizados com idosos de 60 anos ou mais de idade, por meio de modelagem multivariada; sendo que o primeiro estudo considerou uma amostra probabilística de idosos na cidade de Pelotas-RS e o segundo uma amostra probabilística do município de São Paulo.

No presente estudo, observou-se que quanto maior a faixa etária, maior a chance do idoso apresentar pior grau de dificuldade na locomoção, concordando com os estudos de Alves et al. (2010), Duca et al. (2009) e Costa (2006). Assim, o progresso da idade, aliado ao próprio processo biológico de envelhecimento, está relacionado diretamente com os maiores níveis de incapacidade funcional.

A chance dos idosos casados apresentarem pior grau de dificuldade de locomoção é menor do que a dos idosos não casados. Rosa et al. (2003), também verificaram que idosos casados apresentam menor chance de incapacidade, entretanto esta associação somente foi verificada pelos autores na análise bivariada quando comparada com a categoria “viúvo”.

Neste estudo encontraram-se diferenciais de incapacidade funcional entre idosos da região Sul e da região Centro-Oeste, isto é, idosos residentes na região Sul do Brasil têm chance maior de apresentar pior grau de dificuldade na locomoção do que os idosos residentes da região Centro-Oeste. A maior incapacidade entre os idosos residentes na região Sul poderia ser atribuída a maior longevidade nesta região que, por sua vez,

poderia levar a uma maior probabilidade de sobrevivência com incapacidade funcional, sobretudo entre os idosos com idade mais avançada.

Quanto ao nível de escolaridade, observou-se que idosos menos escolarizados apresentam maior chance de apresentarem incapacidade funcional. Relação similar foi observada por Pereira et al. (2012) e Alves et. al (2010), o que pode ser explicada em parte pela maior dificuldade de acesso dos idosos mais pobres aos serviços de saúde e, conseqüentemente, ao menor acesso a tratamentos e medicamentos. Outra explicação para a associação do nível de escolaridade com o grau de dificuldade na locomoção de idosos é que indivíduos com nível educacional mais elevado tendem a se expor menos aos fatores de risco para doenças e adotam hábitos e comportamentos mais saudáveis.

Houve associação do grau de dificuldade na locomoção com a autoavaliação geral de saúde dos idosos, o que corrobora com os estudos de Alves et. al (2010) e Giacomini et al.(2008) no sentido de que idosos que reportam pior estado de saúde tendem a apresentar maior incapacidade funcional. Esta associação pode ser explicada pelo fato da autoavaliação geral de saúde ser um importante preditor de morbidade e mortalidade (HÖFELMANN & ANJOS, 2012). Além disso, observou-se associação do grau de dificuldade na locomoção com o diagnóstico de alguma doença crônica, física ou mental, mostrando que as doenças crônicas apresentaram forte influência na incapacidade funcional dos idosos, como também verificado por outros estudos (ALVES et. al., 2010; ROSA et al., 2003).

Quanto à limitação deste estudo, aponta-se a não inclusão na modelagem estatística da variável renda, que é um indicador importante de acesso a bens e serviços, mas a sua não consideração no estudo deve-se a não divulgação dessa informação pelos responsáveis da PNS 2013. Outra limitação se refere à dificuldade de comparabilidade entre os estudos sobre incapacidade funcional, devido à diversificação de indicadores utilizados para mensuração da incapacidade funcional e a utilização de diferentes modelos estatísticos para identificar possíveis fatores de risco. Neste trabalho, adotou-se o modelo de chances proporcionais parciais (MCP), enquanto outros utilizaram o modelo de regressão logística binária (PEREIRA et al., 2012; ROSA et al., 2003), modelo de regressão de Poisson (DUCA et al., 2009), o modelo de regressão logística multinomial não ordinal (GIACOMINI et al., 2008) e modelo multinível (ALVES et al., 2010). Outra limitação se refere a pouca literatura sobre o MCP no contexto de amostragem complexa.



Com relação às potencialidades do trabalho, destaca-se o uso do modelo de regressão logística ordinal, mas especificamente do MCP, que leva em consideração a natureza ordinal da variável resposta do modelo, diferentemente do modelo logístico binário, modelo log-linear de Poisson e do modelo logístico multinomial (não ordinal) utilizados em outros estudos sobre incapacidade funcional. Outra potencialidade refere-se à inclusão no ajuste do modelo dos principais aspectos do plano amostral (conglomerado, estratificação e pesos amostrais) da PNS 2013, evitando assim a obtenção de estimativas viciadas para os parâmetros do modelo, medidas de associação (razões de chance) e de precisão (MORAES et al., 2012; PESSOA & SILVA, 1998).

Como sugestão de estudos futuros, considera-se relevante fazer uma pesquisa com a finalidade de levantar os principais comandos do *software R* para o ajuste de diferentes modelos lineares generalizados nos contextos de amostragem complexa. Além disso, com base nos dados desse TCC, pode-se ajustar um modelo para explicar a incapacidade funcional usando o *software R*, considerando e não considerando o plano amostral da pesquisa. Dessa forma, com esta sugestão, espera-se estimular trabalhos que visam disseminar o uso do *R* na análise de dados amostrais complexos, que para dados desta natureza vem sendo frequentemente realizada com o emprego da opção *Complex Sample* do SPSS e do *survey (svy)* do STATA. Também seria possível avaliar a importância do plano amostral e o seu impacto nas estimativas dos parâmetros do modelo e de suas medidas de precisão.

Enfim, com a realização deste trabalho de conclusão de curso, conclui-se que há maior grau de dificuldade na locomoção entre os idosos mais velhos, não casados, sem instrução, residentes na região Sul, com diagnóstico de doenças e que reportaram pior estado de saúde. Deste modo, destaca-se a necessidade de atender as demandas específicas da população idosa e a realização de ações voltadas, sobretudo, para promover mudanças das condições educacionais e de saúde da população em geral, a fim de contribuir para um envelhecimento populacional saudável, evitando assim uma saúde precária e a incapacidade funcional na fase idosa.

## Referências bibliográficas

- ALMEIDA, M.F., BARATA, R.B., MONTERO, C.V., SILVA, Z.P. **Prevalência de doenças crônicas auto-referidas e utilização de serviços de saúde, PNAD/1998, Brasil.** *Ciênc. Saúde coletiva* 2002; 7(4): 743-756.
- ALVES, L.S., RODRIGUES R.N. **Determinantes da autopercepção de saúde entre idosos do Município de São Paulo, Brasil.** *Rev Panam Salud Publica* 2005; 17(5/6): 333-41.
- ALVES, L.C., LEITE, I.C., MACHADO, C.J. **Conceituando e mensurando a incapacidade funcional da população idosa: uma revisão de literatura.** *Cienc Saude Coletiva* 2008; 13(4): 1199-207.
- ALVES, L.C., LEITE, I.C., MACHADO, C.J. **Fatores associados à incapacidade funcional dos idosos no Brasil: análise multinível.** *Rev Saúde Pública* 2010; 44(3): 468-78.
- BINDER D.A. **On the variances of asymptotically normal estimators from complex surveys.** *International Statistical Review* 1983; 51(3): 279-292.
- CABRAL S.O.L, OLIVEIRA C.C.C, VARGAS M.M , NEVES A.C.S. **Condições de ambiente e saúde em idosos residentes na zona rural e urbana em um município da região Nordeste.** *Geriatrics & Gerontologia* 2010; 4(2):76-84
- COSTA, A.J.L. **Metodologias e indicadores para avaliação da capacidade funcional: análise preliminar do Suplemento Saúde da Pesquisa Nacional por Amostra de Domicílios PNAD, Brasil, 2003.** *Ciênc. saúde coletiva* 2006; 11(4): 927-940.
- DUCA, G.F.D; SILVA, M.C.; HALLAL,P.C. **Incapacidade funcional para atividades básicas e instrumentais da vida diária em idosos.** *Rev Saúde Pública* 2009; 43(5): 796-805.
- FREITAS, M.P.S; ANTONACI, G.A. **Sistema integrado de pesquisas domiciliares. Amostra mestra 2010 e amostrada PNAD contínua, 2014.**
- GIACOMIN, K.C., PEIXOTO, S.V., UCHOA, E.; LIMA-COSTA, M.F. **Estudo de base populacional dos fatores associados à incapacidade funcional entre idosos na Região Metropolitana de Belo Horizonte, Minas Gerais, Brasil.** *Cad. Saúde Pública*, 2008; 24(6):1260-1270.

GRITTI, C.C., BENE, A.Z., PINHEIRO, D.M., BIANCHIN, M.A, LAMARI, N.M. **Doenças crônicas não transmissíveis e antecedentes pessoais em reinternados e contribuição da terapia ocupacional.** *Cad. Saúde Colet.* 2015, 23(2): 214-219.

HÖFELMANN, Doroteia Aparecida; ANJOS, Juliana Cristine dos. **Autoavaliação de Saúde e Câncer de Mama em Mulheres de Cidade do Sul do Brasil.** *Revista Brasileira de Cancerologia* 2012; 58(2): 209-222.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Projeção da População do Brasil por Sexo e Idade 1980-2050.** Revisão 2008. Rio de Janeiro; 2008.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Pesquisa Nacional de Saúde 2013: Percepção do estado de saúde, estilos de vida e doenças crônicas – Brasil, Grandes Regiões e Unidade da Federação,** 2014.

JEDE, M.; SPULDARO, M. **Cuidado do idoso dependente no contexto familiar: uma revisão de literatura.** *RBCEH, Passo Fundo* 2009, 6(3): 413-421.

KARSCH, U.M. **Idosos dependentes: famílias e cuidadores.** *Cad. Saúde Pública* 2003, 19(3): 861-866.

**IBM: IBM Knowledge Center, 2012.** Disponível em: [https://www.ibm.com/support/knowledgecenter/SSLVMB\\_21.0.0/com.ibm.spss.statistics.help/alg\\_csordinal\\_testparallel.htm](https://www.ibm.com/support/knowledgecenter/SSLVMB_21.0.0/com.ibm.spss.statistics.help/alg_csordinal_testparallel.htm). Acesso em: 25/06/2016.

LAURENTI, R. **A mensuração da qualidade de vida.** *Rev. Assoc. Med. Bras.* 2003; 49(4): 349-66 361.

LIU X.; KOIRALA H. **Fitting Proportional Odds Models to Educational Data with Complex Sampling Designs in Ordinal Logistic Regression.** *Journal of Modern Applied Statistical Methods* 2013; 12(1): 235-248.

MORAES, E.N. **Atenção à saúde do Idoso: Aspectos Conceituais.** *Brasília: Organização Pan-Americana da Saúde,* 2012.

MORAES, J.R.; MOREIRA, J.P.L.; LUIZ, R.R. **Efeito do plano amostral em modelo logístico ordinal: uma análise do estado de saúde autorreferido de adultos no Brasil usando a Pesquisa Nacional por Amostra de Domicílios de 2008.** *Cad. Saúde Pública* 2012; 28(5):913-924.

PARAHYBA M.I.; VERAS, R. **Diferenciais sociodemográficos no declínio funcional em mobilidade física entre os idosos no Brasil.** *Ciência & Saúde Coletiva* 2008; 13(4): 1257-1264.

- PEREIRA, N.G., BASTOS, G.A.N., DUCA, G.F.D., BÓS, A.J.G. **Indicadores demográficos e socioeconômicos associados à incapacidade funcional em idosos.** *Cad. Saúde Pública, Rio de Janeiro* 2012, 28(11): 2035-2042.
- PESSOA D. G. C.; SILVA P. L. N. **Análise de Dados Amostrais Complexos.** 1998
- RABELO, D.F; CARDOSO, C.M. **Auto-eficácia, doenças crônicas e incapacidade funcional na velhice.** *Psico-USF* 2007; 12(1): 75-81.
- ROSA T.E., BENICIO M.H., LATORRE M.R., RAMOS L.R. **Fatores determinantes da capacidade funcional entre idosos.** *Rev Saude Publica.* 2003; 37(1): 40-8.
- SIQUEIRA, R.L.; BOTELHO, M.I.V.; COELHO, F.M.G. **A velhice: algumas considerações teóricas e conceituais.** *Ciênc. Saude Coletiva* 2002; 7(4): 899-906.
- VERAS, R. **Envelhecimento populacional contemporâneo: demandas, desafios e inovações.** *Rev Saúde Pública* 2009; 43(3): 548-54.
- VERAS, R.P. **Modelos de atenção à saúde do idoso: repensando o sentido da prevenção.** *Revista de Saúde Coletiva* 2013; 23(4): 1189-1213.
- VERMELHO, L.L.; COSTA, A.J.L.; KALE, P.L. **Indicadores de Saúde.** In: MEDRONHO, R.A.; BLOCH, K.V.; LUIZ, R.R.; WERNECK, G.L. (eds.). **Epidemiologia.** São Paulo: Atheneu, 2009.
- WILLIAMS R. **Generalized ordered logit/partial proportional odds models for ordinal dependent variables.** *The Stata Journal* 2006; 6(1): 58–82.