



UNIVERSIDADE
E D U A R D O
M O N D L A N E

FACULDADE DE CIÊNCIAS
Departamento de Matemática e Informática

Trabalho de Licenciatura em
Estatística

**Análise dos principais determinantes dos gastos em
saúde pelos agregados familiares em Moçambique em
2022**

Autor: Fernando Micas Maposse

Maputo, Outubro de 2024



UNIVERSIDADE
E D U A R D O
MONDLANE

FACULDADE DE CIÊNCIAS
Departamento de Matemática e Informática

Trabalho de Licenciatura em
Estatística

**Análise dos principais determinantes dos gastos
em saúde pelos agregados familiares em
Moçambique em 2022**

Autor: Fernando Micas Maposse

Supervisor: Lino Marques C.dos Santos, MSc, Kiev

Co-supervisor(es): Jonas Gonçalves Nassabe, MSc, Turim

Maputo, Outubro de 2024

Dedicatória

Dedico este trabalho a Santíssima Trindade, Virgem Maria, Santo António de Pádua, aos meus avôs Manuel Nhabombe, Fernando Maposse, Nora Simone e Virgínia Tembe, aos meus pais Micas Maposse e Alexandrina Nhabombe, a minha tia Marlene Nhabombe.

Declaração de Honra

Declaro por minha honra que este trabalho é resultado da minha própria investigação, que nunca tenha sido submetido a outra instituição de ensino superior para obtenção de um grau académico ou qualquer outra habilitação que não seja o indicado - **Licenciatura em Estatística**, na Faculdade de Ciências da Universidade Eduardo Mondlane.

Maputo, Outubro de 2024

(Fernando Micas Maposse)

Agradecimentos

Agradeço à Deus pelo Dom da vida e todos os santos por interceder por mim em todos os dias da minha vida. Adicionalmente, agradeço a minha família, por sempre me apoiar e me dar forças para crescer em todas as áreas da minha vida, principalmente na área académica.

Agradeço a todos os meus professores do Departamento de Matemática e Informática da Universidade Eduardo Mondlane de forma especial ao Prof. Lino Marques, Prof. João Moreno, Prof. Jonas Nassabe minha fonte de inspiração em todas as áreas da vida, Prof. Rachid Muleia, Prof. Miranda, Prof. Mulenga, Prof. Osvaldo Loquia e a Prof. Rafika Abuldarazak pelos ensinamentos e a acompanhamento durante a minha formação académica

Agradeço a todos meus colegas em especial a Benigna Novela, Cleide Portugal, Agostinho Zeinadine e Arlindo Massingue.

Agradeço ao meu explicador Zito Magaiva por desde cedo buscar me ensinar as bases da ciência e a arte do saber pensar.

Agradeço a minha Namorada Leide Carlos Lameque pelo seu amor e suporte em todos os momentos da minha caminhada académica

À todos meus amigos em especial a Neide Murrure, que contribuíram para a realização deste trabalho, meu muito obrigado.

À todos que directa ou indirectamente contribuíram para o processo de minha formação em todos os níveis de ensino vai o meu agradecimento.

Lista de abreviaturas

AIC	Critério de Informação de Akaike
IC	Intervalo de Confiança
MISAU	Ministério da Saúde
OMS	Organização Mundial da saúde
OR	Odds Ratio
UPA	Unidades Primarias de Amostragem
USAID	Agência dos Estados Unidos para o Desenvolvimento Internacional
BM	Banco Mundial
INE	Instituto Nacional de Estatística

Resumo

Os sistemas de saúde oferecem serviços preventivos e curativos que podem melhorar significativamente a saúde das pessoas. No entanto, o acesso a esses serviços pode exigir que indivíduos gastem uma parte substancial de seu rendimento disponível, o que pode levá-los à pobreza. O objectivo principal deste estudo é analisar os principais determinantes dos gastos em saúde dos agregados familiares em Moçambique. Recorreu-se ao teste de independência do Qui-Quadrado e a regressão logística múltipla a análise da relação entre a variável dependente e as variáveis independentes que influenciaram a ocorrência dos gastos em saúde. Constatou-se que a maioria dos agregados familiares reside em áreas urbanas (58%) em comparação com áreas rurais (42%). A predominância de chefes de agregado do sexo feminino foi observada (55%) e do estado civil solteiro (57%) sendo a maioria solteira (57%) e com baixa escolaridade (66% com ensino primário), 30% dos agregados familiares foi considerada pobre, observou-se que os gastos em saúde se mostrou relacionado com todas as variáveis independentes excepto pelas variáveis sexo, estado civil e o estado de empregabilidade, os principais factores associado aos gastos em saúde foram: condição social, nível académico (chefe do agregado familiar), área de residência, número de membros do agregado familiar, tipo de casa, tipo de consulta, problemas alimentares. De acordo com o teste de Hosmer Lemeshow, o modelo apresenta um bom ajuste. O estudo mostrou ainda que é importante a promoção da equidade social e educacional, o apoio a famílias numerosas, a melhoria das condições habitacionais e a criação de políticas de segurança alimentar. Além disso, é crucial aumentar o foco em políticas de saúde pública preventivas, como vacinação e exames de rotina, especialmente em regiões de baixa renda e com menor infraestrutura, para garantir acesso adequado e prevenir maiores despesas com saúde.

Palavras-chave: Gastos em saúde, agregados familiares, regressão logística

Abstract

Healthcare systems provide preventive and curative services that can significantly improve people's health. However, access to these services may require individuals to spend a substantial portion of their disposable income, which can drive them into poverty. The primary objective of this study is to analyze the main determinants of household health expenditures in Mozambique. The Chi-Square test of independence and multiple logistic regression were used to examine the relationship between the dependent variable and the independent variables that influenced health expenditures. The findings revealed that the majority of households are located in urban areas (58%) compared to rural areas (42%). There was a predominance of female-headed households (55%), and a significant portion of household heads were single (57%) with low levels of education (66% having primary education). Additionally, 30% of households were classified as poor. Health expenditures were found to be related to all independent variables except gender, marital status, and employment status. The main factors associated with health expenditures included social status, the educational level of the household head, place of residence, household size, type of dwelling, type of medical consultation, and food insecurity. According to the Hosmer-Lemeshow test, the model demonstrated a good fit. The study also highlighted the importance of promoting social and educational equity, supporting large families, improving housing conditions, and developing food security policies. Furthermore, it is crucial to increase the focus on preventive public health policies, such as vaccination and routine screenings, especially in low-income regions with less infrastructure, to ensure adequate access and prevent higher healthcare expenses.

Keywords: Health expenditures, households, logistic regression

Índice

1. INTRODUÇÃO	11
1.1. Contextualização	11
1.2. Definição do problema	12
1.3. Objectivos.....	12
1.3.1. Objectivo geral.....	12
1.3.2. Objectivos específicos	12
1.4. Justificação	13
1.5. Estrutura do trabalho	13
2. REVISÃO DA LITERATURA	15
2.1. Moçambique: análise abrangente das questões sociais e necessidades de saúde.....	15
2.2. Sistema de saúde	16
2.2.1. Sistema de saúde de moçambique: antecedentes históricos	16
2.2.2. Sistema de saúde moçambicano.....	18
2.2.3. Sistema de financiamento a saúde em Moçambique	19
2.3. Gastos.....	19
2.3.1. Gastos com alimentação dos agregados familiares.....	20
2.3.2. Linha de pobreza (pl) e os gastos de subsistência do agregado familiar	20
2.3.3. Capacidade de pagamento do agregado familiar	20
2.3.4. Gastos em saúde.....	21
2.4. Tipo de gastos em saúde	21
2.4.1. Gastos de saúde não catastróficas	21
2.4.2. Gastos de saúde catastróficas.....	21
2.5. Relação entre gastos de saúde e factores determinantes.....	22
2.6. Análise da associação das variáveis.....	29
2.7. Estatística multivariada.....	30

2.7.1. Análise de regressão	30
2.7.2. Regressão logística.....	30
2.7.3. Regressão logística simples	31
2.7.4. Regressão logística múltipla	34
2.7.5. Estimaco dos parâmetros do modelo regresso logística múltipla.....	35
2.7.6. Transformaco logit	35
2.7.7. Métodos de selecco das variáveis	36
2.7.8. Critérios para a selecco do modelo	36
2.7.9. Interpretaco dos coeficientes estimados.....	48
2.7.10. Razes para o uso da regresso logística	49
2.7.11. Vantagens para o uso da regresso logística:.....	49
3. MATERIAL E MÉTODOS	51
3.1 Classificaco de pesquisa	51
3.1.1 Quanto à natureza da pesquisa.....	51
3.1.2 Quanto aos procedimentos técnicos.....	51
3.1.3 Quanto à abordagem ou método	51
3.2 Material	51
3.3. Métodos.....	54
3.3.1. Análise exploratória	54
3.3.2. Determinaco do tipo de gastos em sade	54
3.4.1. Regresso logística binária	56
4. RESULTADOS E DISCUSSO.....	59
4.1. Análise descritiva.....	59
4.2. Modelo de Regresso Logística Binária	64
4.2.1. Análise da variável dependente	64
4.2.2. Teste de multicolinearidade	65

4.2.3. Análise de Associação entre as Variáveis independentes e à variável dependente	66
4.2.4. Análise de regressão logística	66
4.2.5. Análise do ajuste do modelo	67
4.3. Discussão dos resultados	72
5. CONCLUSÃO E RECOMENDAÇÕES	76
5.1. Conclusões	76
5.2. Recomendações	77
Limitações	78
Referências	79
Apêndices	85

Lista de tabelas

Tabela 2.1: Tabela de classificação	48
Tabela 3.1: Descrição e classificação das variáveis em estudo	54
Tabela 4.1: Distribuição das frequências absolutas e relativas em relação entre as despesas e o total em média.....	86
Tabela 4.2: Distribuição das frequências relativas de acordo com a ocorrências de gastos catastróficos em saúde.....	88
Tabela 4.3: Distribuição das frequências absolutas e relativas de acordo com a ocorrências de gastos em saúde.....	89
Tabela 4.4: Distribuição da média das despesas de subsistência dos agregados familiares.....	91
Tabela 4.5: Estatística de colinearidade.....	66
Tabela 4.6: Teste de associação entre as variáveis independentes e dependentes	67

Lista de gráficos

Gráfico 4.1: Distribuição dos agregados por área.....	59
Gráfico 4.2: Distribuição dos agregados por província	59
Gráfico 4.3: Empregabilidade do agregado	60
Gráfico 4.4: Condição social do agregado	60
Gráfico 4.5: Género do chefe do agregado	60
Gráfico 4.6: Estado civil do chefe do agregado	60
Gráfico 4.7: Nível de académico do chefe do agregado familiar	61
Gráfico 4.8: Curva Roc	71

Capítulo 1

1. INTRODUÇÃO

1.1. Contextualização

A busca por sistemas de saúde cada vez mais sofisticados, inclusivos e acessíveis para todos é uma prioridade nas agendas de saúde a nível global. No entanto, os países em desenvolvimento, como Moçambique, deparam-se com esta problemática. Grande parte da população moçambicana vive em situação de pobreza, e o aumento dos preços dos bens essenciais dificulta, muitas vezes, o acesso aos serviços de saúde. Além disso, o sistema de saúde moçambicano enfrenta desafios como a fragmentação do financiamento, a dependência de doadores externos. Perante este cenário, torna-se crucial analisar e discutir estratégias que promovam um financiamento mais equitativo e sustentável para o sector da saúde em Moçambique (Ministério da Saúde, 2020).

Segundo o Ministério da Saúde (MISAU, 2015), os gastos totais em serviços de saúde em 2012 eram de cerca de 24.3 mil milhões de meticais, em que os doadores participavam com 58%, o governo com 24%, as fontes não especificadas com 12% e os agregados familiares com 6%. Ainda de acordo com a mesma fonte, os gastos totais em serviços de saúde em 2015 eram de cerca de 42.8 mil milhões de meticais, em que os doadores participavam com 55%, o governo com 27% e os agregados familiares com 11%.

De acordo com o Ministério da Saúde (MISAU, 2021), os gastos totais em serviços de saúde em 2021 atingiu 49,377 milhões de meticais, dos quais 24,517 milhões de Meticais foram financiados com recursos internos 49.6% e 24,860 milhões, que representa 50,4% com recursos externos, verifica-se que com o passar do tempo os gastos com serviços de saúde vem aumentando.

De acordo com a Agência dos Estados Unidos para o Desenvolvimento Internacional (USAID, 2022), cerca de 63% dos moçambicanos vivem abaixo da linha de pobreza, onde os custos directos e indirectos com serviços de saúde são consideravelmente altos para esta população o que pode conduzir a gastos catastróficos com serviços de saúde.

Segundo Xu et al. (2007), estes gastos são definidos como gastos em saúde que excedem uma proporção significativa do rendimento do agregado familiar, e têm potencial para deixar pessoas em situações financeiras precárias, prejudicando assim os esforços para a obtenção de cuidados de saúde equitativos e acessíveis bem como o acesso universal a saúde.

1.2 Definição do problema

Os sistemas de saúde oferecem serviços preventivos e curativos que podem melhorar significativamente a saúde das pessoas. No entanto, o acesso a esses serviços pode exigir que indivíduos gastem uma parte substancial de seu rendimento disponível, o que pode levá-los à pobreza. Em muitos casos, devido a esses custos elevados, alguns agregados familiares acabam renunciando ao atendimento médico e sofrendo problemas de saúde como resultado.

A disparidade na distribuição de riqueza afeta o crescimento socioeconómico da população. Em Moçambique, mais de 60% dos habitantes vivem com cerca de um dólar por dia, e a taxa de desemprego é de aproximadamente 18,4%, o que afecta directamente a capacidade financeira das famílias e as torna vulneráveis à pobreza. Gastos de saúde inesperados podem agravar a situação económica dos agregados familiares já empobrecidos, expondo-os a maiores privações. A alta taxa de pobreza intensifica os riscos de novas privações e limita o acesso a cuidados de saúde acessíveis.

Com o avanço na melhoria da cobertura de saúde para alcançar o acesso universal, surge a preocupação com o peso dos gastos em saúde nos orçamentos familiares. O custo do atendimento médico muitas vezes se torna um fardo financeiro insustentável e desafia a capacidade do governo de responder adequadamente, dado que a saúde é uma das suas responsabilidades. Portanto, é crucial investigar **quais são os principais determinantes dos gastos em saúde dos agregados familiares em Moçambique?**

1.3 Objectivos

1.3.1 Objectivo geral

Analisar os principais determinantes dos gastos em saúde dos agregados familiares em Moçambique.

1.3.2 Objectivos específicos

- ❖ Descrever o perfil dos agregados familiares em Moçambique;

- ❖ Identificar o tipo de gastos dos agregados familiares em Moçambique em despesas de saúde;
- ❖ Verificar a associação entre as variáveis independentes e os gastos em saúde;
- ❖ Aplicar um modelo de regressão logística para identificar os principais determinantes dos gastos dos agregados familiares em Moçambique em despesas de saúde.

1.4 Justificação

O desenvolvimento deste estudo sobre os principais determinantes dos gastos em saúde dos agregados familiares em Moçambique, é importante para subsidiar aos gestores e pesquisadores na formulação de políticas e programas que tenham como objectivo diminuir a ocorrência dos factores determinantes nos gastos em saúde.

Além disso a compreensão dos principais determinantes dos gastos em saúde dos agregados familiares em Moçambique tem implicações profundas tanto para a saúde pública quanto para a justiça social, mas também podem ter efeitos de longo alcance.

Ao analisar os principais determinantes dos gastos em saúde dos agregados familiares, podem ser extraídas lições e recomendações que poderão ser aplicadas em contextos similares, onde os sistemas de saúde enfrentam problemas semelhantes de acesso, equidade e sustentabilidade financeira.

Ao identificar os factores determinantes dos gastos em saúde vai se ter evidências para desenho de políticas com vista a reduzir a ocorrência dos gastos em saúde nos agregados familiares com risco de pobreza bem como identificar acções de melhorias nos serviços de saúde com vista ao acesso universal a saúde.

1.5 Estrutura do trabalho

O presente capítulo, que constitui a introdução, enquadra os aspetos gerais do trabalho, a importância do tema de aplicação, o problema em estudo, assim como a restante estrutura do trabalho;

Capítulo 2: Segue a revisão de literatura, descrevendo os factores determinantes a gastos em saúde, assim como aborda conceitos relacionados ao modelo de regressão logística binária.

Capítulo 3: Apresenta o material e os métodos usados para alcançar os objectivos do trabalho.

Capítulo 4: Neste capítulo é apresentada análise de dados, e, além disso, é apresentado também a discussão dos resultados.

Capítulo 5: São descritas as conclusões, recomendações e limitações que se teve no âmbito da realização do presente trabalho. Em apêndices são apresentadas algumas tabelas, como informação complementar, em detalhe.

Capítulo 2

2. REVISÃO DA LITERATURA

Este capítulo tem como propósito apresentar as bases teóricas sobre o tema deste trabalho que são resultado das investigações desenvolvidas por diversos autores nas diferentes áreas de actividades relacionadas com os gastos em saúde pelos agregados familiares e também apresentar os conceitos teóricos das principais técnicas estatísticas aplicadas nos estudos do gênero.

2.1 Moçambique: análise abrangente das questões sociais e necessidades de saúde

Moçambique é um país localizado no Sudeste da África, caracterizado por uma forma aproximadamente oblonga e banhado pelo Oceano Índico a Leste. O País experimenta um clima tropical, embora as condições climáticas específicas variem com base na altitude e latitude.

Do ponto de vista socioeconômico, Moçambique compartilha semelhanças com outros países da África Subsaariana. Apesar de melhorias recentes, ele continua entre os países mais pobres e menos desenvolvidos do mundo, com um PIB per capita entre os 10 mais baixos. A expectativa de vida ao nascer é de 56,3 anos, e a taxa de fertilidade total é de 4,8 crianças por mulher (estimada para 2022). A posição de Moçambique no Índice de Desenvolvimento Humano de 2022 reflete seus desafios de desenvolvimento (Programa das Nações Unidas para o Desenvolvimento, 2022).

Aproximadamente 70% da população de Moçambique reside em áreas rurais, com uma grande maioria envolvida em atividades agrícolas, a principal ocupação para uma grande parte da população. O comércio informal também é prevalente tanto em áreas rurais quanto urbanas.

Administrativamente, Moçambique compreende 10 províncias e uma capital com status provincial (Maputo). As províncias são subdivididas em distritos, que se dividem em postos administrativos e localidades. Maputo, juntamente com a cidade vizinha de Matola, constitui a maior área urbana do país, conforme o censo recente de 2017 realizado pelo Instituto Nacional de Estatística de Moçambique (INE, 2017).

A diversidade étnica é notável em Moçambique, com variações significativas em normas comportamentais e sociais entre os grupos étnicos do país em relação a estruturas matrimoniais, padrões de descendência e idade ao casar (Ndege, 2007).

As doenças transmissíveis continuam sendo as principais causas de mortalidade em Moçambique, com malária, HIV/AIDS, condições perinatais, doenças diarreicas e infecções respiratórias inferiores respondendo por uma parte significativa das mortes (OMS, 2022).

Moçambique ainda enfrenta altas taxas de mortalidade materna e infantil, com recém-nascidos e bebês com menos de um ano representando uma proporção considerável de todas as mortes (INE, 2022).

2.2 Sistema de saúde

Um sistema de saúde é uma infraestrutura organizada e complexa composta por políticas, práticas e recursos que têm como objetivo primordial promover e preservar a saúde e o bem-estar de uma população específica (OMS, 2000). Este sistema engloba uma variedade de instituições, desde hospitais e clínicas até serviços de saúde comunitários, além de programas de prevenção de doenças, políticas de saúde pública e acesso a medicamentos e tratamentos (OMS, 2008).

É uma estrutura dinâmica que não só responde às necessidades de saúde imediatas da população, mas também busca abordar as causas subjacentes das doenças e promover estilos de vida saudáveis (Evans et al., 2013). Além disso, um sistema de saúde eficaz deve ser acessível a todos os membros da sociedade, independentemente de sua renda, local de residência, gênero ou etnia (Organização Mundial da Saúde, 2010). Isso implica em garantir que os serviços de saúde sejam equitativamente distribuídos e financeiramente acessíveis para que ninguém seja deixado para trás (Starfield et al., 2005).

2.2.1 Sistema de saúde de Moçambique: antecedentes históricos

Durante os anos 70, em meio aos distintos tipos de socialismos africanos e às limitações econômicas ligadas à geopolítica e questões sociais internas (Mondlane, 1969), países como Moçambique enfrentaram desafios na construção de um sistema de saúde pública completo. Esse sistema era fundamentado em agentes comunitários de saúde, postos e centros de saúde, além de hospitais rurais e provinciais maiores.

A infraestrutura do sistema de saúde pública moçambicano foi significativamente expandida em apenas uma década, passando de 326 unidades de saúde em 1975 para 1.195 em 1985 (Magnus et al., 2004). Essa evolução o transformou em um modelo de Cuidados de Saúde Primários, priorizando a equidade e suprimindo o modelo de serviço médico colonial, que enfatizava o cuidado curativo e urbano.

Nos anos 80 e 90, Moçambique enfrentou profundas transformações sociais, econômicas e políticas. Em 1989, após doze anos de guerra civil e duas greves de doadores em 1983 e 1986, quando a ajuda alimentar foi interrompida (Hanlon, 2004), a Frente de Libertação de Moçambique (FRELIMO) formalmente abandonou o marxismo. Sob pressão do Fundo Monetário Internacional (FMI) e do Banco Mundial (BM), Moçambique assinou um programa de ajuste estrutural (PAE) em 1987, seguido pela adoção de uma nova constituição em 1990, que viabilizou eleições multipartidárias e resultou na privatização de serviços, redução de gastos governamentais e transição para uma economia de mercado.

O contínuo aumento da ajuda externa após o PAE tornou Moçambique um dos maiores beneficiários dos cuidados de saúde na África atualmente (IHME, 2016a). No entanto, a ampla diversidade de atores no setor de saúde - organizações multilaterais, doadores bilaterais, ONGs, fundações ou universidades - levou à fragmentação dentro do setor, devido à falta de coordenação entre os fluxos de ajuda externa e aos interesses concorrentes dos doadores.

Na metade da década de 90, a governança em Moçambique estava enfraquecida, o orçamento do Estado para a saúde era escasso e várias instalações de saúde foram destruídas durante a guerra civil, especialmente nas áreas rurais. Além disso, o clima político e econômico neoliberal predominante introduziu mudanças significativamente adversas para o funcionamento eficaz do sistema público de saúde, como a redução dos salários dos funcionários públicos para um terço do que eram em 1991.

Nos anos seguintes, a grande maioria dos moçambicanos passou a depender de um sistema de saúde pública enfraquecido (Organização Mundial de Saúde, 2016), caracterizado por recursos extremamente limitados. A retórica predominante em relação ao bem-estar público, em um contexto de escassez de recursos públicos e alta dependência internacional, muitas vezes condicionada à ajuda, impôs obstáculos adicionais ao desenvolvimento de uma abordagem abrangente e mínima do sistema público de saúde (Mackintosh, 2000).

A atual Constituição de Moçambique garante o direito à saúde para a população. Entretanto, embora esse direito seja assegurado, o acesso aos serviços de saúde permanece restrito devido aos custos diretos e indiretos, incluindo acessibilidade física, fatores socioculturais, percepção de benefícios e necessidades (dos Anjos & Cabral, 2016; Wagenaar et al., 2016).

2.2.2 Sistema de saúde moçambicano

Em Moçambique, o sector da saúde é composto por várias entidades que desempenham papéis essenciais na prestação de serviços de saúde e na formulação de políticas. O Ministério da Saúde de Moçambique (MISAU) é a principal autoridade responsável pela coordenação e supervisão do sistema de saúde do país (Ministério da Saúde de Moçambique, 2020). Além disso, existem 11 Direções Provinciais de Saúde (DPS) e 146 Serviços Distritais de Saúde de Mulher e Ação Social (SDSMAS), que desempenham um papel fundamental na implementação das políticas de saúde em nível regional e local.

Outras instituições de saúde que recebem financiamento direto do Orçamento do Estado incluem o Centro de Medicamentos e Artigos Médicos (CMAM) e o Conselho Nacional de Combate ao HIV/SIDA (CNCS) (Ministério da Saúde de Moçambique, 2020). Além disso, o país conta com uma rede de hospitais que abrange três Hospitais Centrais, quatro Hospitais Gerais, oito Hospitais Provinciais, um Hospital Distrital e um Hospital Psiquiátrico.

A DPS e os SDSMAS operam sob a supervisão conjunta do Ministério da Saúde (MISAU) e do Ministério da Economia e Finanças (MEF), garantindo uma abordagem integrada na gestão dos recursos e na implementação de políticas de saúde (Ministério da Saúde de Moçambique, 2020).

Um elemento-chave na governança do sistema de saúde é o desenvolvimento de políticas de saúde e a formulação de planos estratégicos pelo Ministério da Saúde. Essas políticas e planos estratégicos orientam as intervenções a serem implementadas visando alcançar resultados desejáveis de saúde para a população moçambicana (Ministério da Saúde de Moçambique, 2020).

Em termos de financiamento do sistema de saúde, Moçambique enfrenta desafios significativos. De acordo com dados do Ministério da Saúde, o investimento total em saúde em Moçambique foi de aproximadamente 4,5% do PIB em 2020 (Ministério da Saúde de Moçambique, 2020). No entanto, uma grande parte desse financiamento é proveniente de

doadores internacionais e organizações não governamentais, o que pode levar a inconsistências na prestação de serviços de saúde e na implementação de políticas de saúde a longo prazo.

Além disso, Moçambique continua a enfrentar desafios de acesso e equidade no sistema de saúde. Segundo o Instituto Nacional de Estatística de Moçambique, em 2021, apenas cerca de 50% da população tinha acesso a serviços de saúde básicos (INE, 2021). Isso destaca a necessidade de investimentos contínuos na expansão e melhoria da infraestrutura de saúde, bem como na capacitação de profissionais de saúde para garantir um acesso equitativo aos serviços de saúde em todo o país.

2.2.3 Sistema de financiamento a saúde em Moçambique

Os gastos com saúde do Governo de Moçambique como percentagem dos Gastos Gerais do Governo (GGE) é de 11,2% (UNICEF, 2023), ainda abaixo da meta de 15% da Declaração de Abuja. Além disso, a despesa total com saúde per capita é de 48 milhões de dólares americanos, abaixo da recomendação da OMS de 65 milhões de dólares americanos, e ainda depende muito da assistência externa (Banco Mundial, 2021).

A maior parte dos gastos com saúde é financiada externamente, com o orçamento do Estado contribuindo com 31%, o Pró-saúde com 8% e os fundos verticais com 61% (Banco Mundial, 2023). É importante salientar que cerca de 57% do gasto total com saúde em Moçambique, principalmente os fundos verticais, não é gerido pelo Governo, nem registado no MEF (Ministério da Economia e Finanças), nem auditado pelo Tribunal Administrativo (fora do orçamento e da CUT) como refere o Banco Mundial.

Quanto à contribuição do desembolso directo, corresponde a 7,2% da despesa total com saúde em 2022 (Banco Mundial, 2023).

De acordo com o Plano Estratégico do Sector da Saúde (PESS), a estratégia de financiamento do sector de saúde continua a ser desenvolvida com foco em quatro dimensões: angariação de fundos, pesquisa de fundos, mecanismos de compra e alocação de recursos (MISAU, 2022).

2.3 Gastos

De acordo Brigham e Ehrhardt (2013), gastos representam os dispêndios ou saídas de recursos financeiros necessários para a aquisição de bens, serviços ou o cumprimento de obrigações. Elas englobam uma variedade de categorias, desde gastos diárias essenciais, como alimentação e saúde, até investimentos estratégicos e obrigações financeiras de longo prazo.

2.3.1 Gasto com alimentação dos agregados familiares

De acordo com a OMS (2005), os gastos dos agregados familiares com alimentação, isto é, todo o montante gasto em todos os géneros alimentícios pelos agregados familiares mais o valor da própria produção de alimentos do agregado familiar. No entanto, exclui os gastos com bebidas alcoólicas, tabaco e consumo de alimentos fora de casa (por exemplo, hotel e restaurantes).

2.3.2 Linha de pobreza (pl) e os gastos de subsistência do agregado familiar

De acordo com a (Dardot & Laval, 2016), gastos de subsistência, são os custos essenciais necessários para garantir a sobrevivência e o bem-estar básico de um agregado familiar. Esses custos geralmente incluem gastos com alimentação, moradia, saúde, educação e outros itens essenciais para uma vida digna.

A linha de pobreza é uma medida quantitativa utilizada para determinar um limiar financeiro abaixo do qual uma pessoa ou agregado familiar é considerada em situação de pobreza (Sen, 1999). Essa medida é frequentemente baseada no rendimento ou no consumo, indicando o nível mínimo de recursos necessários para atender às necessidades básicas, como alimentação, habitação, saúde e educação.

Há muitas maneiras de definir a pobreza. Nenhum deles é perfeito, considerando a solidez na teoria e na prática de viabilidade. Neste trabalho utilizou-se uma linha de pobreza baseada na partilha de alimentos para estimar a subsistência dos agregados familiares. Esta linha de pobreza é definida como o gasto em alimentação das famílias cuja participação nos gastos está no percentil 50 no país.

2.3.3 Capacidade de pagamento do agregado familiar

De acordo com a OMS (2005), a capacidade de pagamento do agregado familiar é definida como um gasto de não subsistência do agregado familiar. No entanto, alguns agregados familiares podem reportar gastos de alimentação que é menor do que os gastos de subsistência. Isso indica que o gasto alimentar do agregado familiar é inferior ao padrão de pobreza estimada para o país. Tal situação também pode ser devido ao facto de que os gastos de alimentos reportados na pesquisa não consideram os subsídios de alimentos, coupons, auto-produção e outros meios de consumo de alimentos. Neste caso específico, o gasto não-alimentar é usado como gasto não-subsistência.

2.3.4 Gastos em saúde

Segundo a OMS (2019), gastos em saúde referem-se aos recursos financeiros totais alocados para serviços de saúde, programas e produtos dentro de um período específico, geralmente em nível nacional ou institucional. Isso engloba gastos em uma ampla variedade de atividades relacionadas à saúde, incluindo, mas não se limitando a hospitais, clínicas, serviços médicos, produtos farmacêuticos, equipamentos médicos, iniciativas de saúde pública e custos administrativos. Os gastos em saúde são uma métrica crítica para avaliar o comprometimento financeiro e o investimento no setor de saúde, refletindo os recursos econômicos dedicados à manutenção e melhoria da saúde de uma população.

Em suma, segundo Folland et al. (2016), gastos em saúde, também conhecidas como gastos com assistência médica ou custos médicos, são os desembolsos financeiros associados ao acesso e recebimento de cuidados médicos e serviços médicos. Esses gastos cobrem uma ampla gama de custos relacionados à saúde.

2.4 Tipo de gastos em saúde

2.4.1 Gastos de saúde não catastróficas

Segundo Cutler et al. (2013), os gastos com saúde não catastróficos referem-se aos gastos regulares e rotineiras com cuidados de saúde que os indivíduos têm para cuidados preventivos, diagnósticos e gerenciamento de condições crônicas, em oposição a eventos de saúde catastróficos que exigem recursos financeiros significativos. Esta categoria de gastos com saúde engloba uma ampla gama de serviços destinados a manter o bem-estar geral e prevenir o surgimento de doenças graves.

2.4.2 Gastos de saúde catastróficas

De acordo com Wagstaff et al. (2018), os gastos de saúde catastróficos referem-se a gastos financeiros impostos a indivíduos ou famílias ao enfrentar custos de saúde elevados e inesperados, que ultrapassam geralmente 5% a 15% da renda familiar ou 40% dos gastos de subsistência (linha de pobreza). Esses pagamentos são considerados catastróficos quando impõem uma pressão financeira significativa a indivíduos ou famílias, levando a consequências adversas, estes gastos não incluem valores reembolsados pelo seguro de saúde.

Segundo Xu et al. (2007), gastos elevados fora do bolso podem desencorajar os indivíduos a buscar os serviços de saúde necessários, resultando em tratamento médico adiado ou não

realizado. Isso pode resultar em piores resultados de saúde e taxas de mortalidade aumentadas, especialmente entre as populações vulneráveis.

De acordo com Wagstaff et al. (2018), O peso financeiro dos pagamentos de saúde catastróficos é um grande impulsionador da pobreza, pois as famílias podem ser forçadas a vender ativos, fazer empréstimos ou reduzir os gastos com necessidades essenciais, como alimentos e educação, para cobrir os custos de saúde. Esse encargo financeiro afeta desproporcionalmente as populações de baixa renda e marginalizadas, exacerbando as disparidades existentes em saúde e socioeconômicas.

De acordo com Wagstaff et al. (2018), a questão dos gastos de saúde catastróficos é uma preocupação global, com variações em sua prevalência e impacto em diferentes regiões e países.

2.5 Relação entre gastos de saúde e factores determinantes

A relação entre os gastos com saúde e os factores determinantes é influenciada por uma variedade de elementos. Diversos factores podem afetar a quantidade de recursos financeiros alocados para a saúde, bem como a eficiência e eficácia desses gastos

Relação entre gastos em saúde e pandemias e epidemias

A relação entre os gastos médicos directos e epidemias/pandemias é uma dinâmica multifacetada que influencia significativamente as finanças da saúde, tanto individualmente quanto em comunidades. Durante essas crises de saúde pública, os indivíduos frequentemente enfrentam aumento de gastos médicos directos devido à intensificação da utilização de serviços de saúde, testes, tratamentos e medidas preventivas.

Epidemias e pandemias desencadeiam um aumento na demanda por cuidados de saúde, levando a custos médicos directos elevados para aqueles que buscam assistência médica (Pauly et al., 2020). Procedimentos diagnósticos, incluindo testes e diagnósticos, contribuem para os gastos médicos directos, pois os indivíduos incorrem em custos com taxas laboratoriais e exames de imagem.

O tratamento de doenças infecciosas durante epidemias envolve medicamentos, internações hospitalares e intervenções médicas, impondo um considerável fardo financeiro aos indivíduos afetados (Pauly et al., 2020). Medidas de saúde pública, como campanhas de vacinação, visam mitigar a propagação de doenças infecciosas. Embora algumas medidas preventivas possam

ser subsidiadas, os indivíduos ainda podem arcar com gastos médicos directos para vacinas e medicamentos preventivos (Yamey et al., 2017).

Populações vulneráveis, sem seguro saúde ou recursos financeiros, enfrentam desafios aumentados para cobrir gastos médicos directos durante epidemias, agravando disparidades de saúde existentes (Xu et al., 2003). A acessibilidade aos serviços de saúde torna-se uma questão crítica durante essas crises, potencialmente desencorajando indivíduos de procurar assistência médica atempada e contribuindo para desafios prolongados e graves de saúde pública.

Relações entre gastos de saúde e tipos de doenças

De acordo com Goldman et al. (2007), Doenças crônicas, que exigem gerenciamento contínuo, frequentemente resultam em custos acumulativos mais altos para a saúde. Condições como diabetes, doenças cardiovasculares e certos tipos de câncer podem resultar em gastos directos significativas do bolso para medicamentos, consultas regulares e tratamentos especializados.

Doenças graves e potencialmente fatais, como certos tipos de câncer ou cirurgias complexas, também podem contribuir para gastos directos significativas em saúde.

Além disso, o tipo de doença pode influenciar o equilíbrio entre custos de cuidados preventivos e tratamento. Medidas preventivas para doenças infecciosas podem ser menos dispendiosas em comparação com os custos prolongados de tratamento associados a condições crônicas (Goldman et al., 2007).

A relação entre o tipo de doença e os gastos de bolso também é influenciada pelo acesso aos serviços de saúde. Indivíduos que enfrentam dificuldades ao acesso à assistência médica podem adiar a busca por tratamento para determinadas condições, o que pode resultar em problemas de saúde mais significativos e custos mais altos a longo prazo (Van Doorslaer et al., 2007).

Além disso, a estrutura dos sistemas de saúde em diferentes países desempenha um papel crucial na determinação dos gastos directos em saúde do próprio bolso. Em países com saúde universal, a carga pode ser menor, enquanto em sistemas com alta dependência de seguros privados, os indivíduos podem suportar custos mais significativos (Kwon, 2012).

Relação entre gastos em saúde e o sexo de chefe do agregado familiar

Segundo Smith e Jones (2020), as famílias chefiadas por mulheres podem enfrentar desafios distintos na gestão dos custos com saúde, pois frequentemente enfrentam taxas mais altas de

pobreza e podem ter recursos financeiros limitados. Isso pode resultar em uma maior proporção da renda sendo destinada a gastos de saúde fora do bolso, causando tensão financeira.

Por outro lado, de acordo com Johnson et al. (2018), as que famílias chefiadas por homens também podem enfrentar dinâmicas específicas relacionadas aos gastos com saúde, influenciadas por fatores como status de emprego, diferenças salariais e tamanho da família.

Relação entre o nível de educação do chefe do agregado familiar e gastos em saúde

A relação entre os gastos em saúde e o nível de educação tem sido objeto de crescente interesse na pesquisa em saúde pública. Segundo Cutler e Lleras-Muney (2006), Níveis mais altos de educação geralmente estão relacionados ao aumento do acesso a recursos de saúde, maior alfabetização em saúde e um processo de tomada de decisão mais informado. Isso pode levar a um uso mais eficiente dos serviços de saúde e a uma melhor capacidade de navegar no sistema de saúde, possivelmente atenuando o ônus financeiro associado aos gastos directos em saúde

Segundo Smith, J. P. (1999), níveis mais elevados de educação geralmente estão associados a estilos de vida mais saudáveis, aumento da conscientização sobre medidas preventivas e melhor aderência a conselhos médicos. Indivíduos com níveis mais altos de educação tendem a ser mais proativos na busca de serviços de saúde e na adoção de hábitos mais saudáveis, reduzindo potencialmente a necessidade de gastos significativas fora do bolso.

Segundo Smith, J. P. (1999), embora educação e renda sejam conceitos distintos, frequentemente estão interconectados. Indivíduos com níveis mais altos de educação tendem a ter maiores rendimentos, o que pode atuar como um fator de proteção contra o ônus financeiro dos custos com saúde. No entanto, disparidades nos gastos de saúde fora do bolso ainda podem persistir entre vários grupos de renda e educação.

Relação entre gastos em saúde e localização geográfica

A localização geográfica desempenha um papel crucial na moldagem dos padrões de gastos directos de saúde. Disparidades no acesso e disponibilidade de serviços de saúde podem levar a variações na utilização e nos custos de saúde em diferentes regiões. Regiões com infraestrutura de saúde bem desenvolvida podem ter uma maior densidade de instalações e profissionais de saúde, levando a um maior acesso aos serviços. No entanto, isso também pode resultar em custos mais elevados (Cutler & Zeckhauser, 2000). Por outro lado, áreas rurais podem enfrentar desafios, como acesso limitado a instalações de saúde, o que pode contribuir para gastos mais elevadas devido a custos de viagem e falta de serviços próximos.

Segundo Chatterjee e Joynt (2018), O status econômico de uma região é um determinante significativo dos gastos directos em saúde. Áreas mais ricas podem ter residentes com melhores condições financeiras para arcar com os custos de saúde, enquanto regiões economicamente desfavorecidas podem enfrentar barreiras financeiras mais altas para acessar serviços de saúde.

As políticas de saúde nacionais e regionais, bem como a cobertura de seguros, podem influenciar significativamente os gastos diretos em saúde. Regiões com cobertura abrangente de seguros de saúde e políticas que priorizam a acessibilidade podem apresentar níveis mais baixos de gastos directos (Goldman et al., 2007).

De acordo com Finkelstein et al. (2012), Crenças culturais e práticas relacionadas à saúde também podem afetar o comportamento de busca por cuidados de saúde. Em algumas regiões, as pessoas podem ser mais propensas a buscar serviços de saúde preventivos ou precoces, enquanto em outras, fatores culturais podem levar a uma utilização atrasada ou infrequente de serviços de saúde.

Relação entre renda mensal e gastos em saúde

De acordo Anderson e Reinhardt (2001), a relação entre gastos directos em saúde e renda mensal representa uma dimensão crítica na compreensão das implicações econômicas dos cuidados com a saúde para indivíduos e famílias. À medida que os custos com saúde continuam a aumentar globalmente, o encargo financeiro sobre os indivíduos torna-se uma preocupação crescente (OMS, 2000).

Segundo Anderson e Reinhardt (2001), a renda mensal é uma determinante chave do bem-estar financeiro de um indivíduo, desempenha um papel crucial na capacidade de absorver gastos relacionadas à saúde. Ainda de acordo com Selden, Thomas M et.al. (2004), disparidades nos níveis de renda contribuem para o acesso diferenciado aos serviços de saúde, perpetuando desigualdades na saúde. Indivíduos de baixa renda frequentemente se veem presos em um ciclo de vulnerabilidade financeira, onde a necessidade de cuidados de saúde entra em conflito com recursos econômicos limitados

Relação entre gastos em saúde e o número de membros no agregado familiar

Segundo estudos de Smith e Jones (2005), a relação entre os custos de saúde e o número de membros do agregado familiar é uma consideração crucial na compreensão do impacto econômico dos cuidados de saúde para as famílias. À medida que os custos de saúde continuam

a aumentar, especialmente em países em desenvolvimento, a pressão financeira sobre as famílias com um maior número de membros se torna cada vez mais significativa.

De acordo com Smith e Jones (2005), o tamanho do agregado familiar desempenha um papel importante na determinação da capacidade de uma família de lidar com os custos de saúde. Famílias maiores podem enfrentar desafios adicionais devido à necessidade de atender às necessidades de saúde de múltiplos membros, o que pode resultar em um aumento da pressão financeira sobre o orçamento familiar.

Além disso, de acordo com um estudo recente de Garcia et al. (2020), famílias com um maior número de membros podem enfrentar dificuldades adicionais devido à falta de acesso a planos de seguro de saúde acessíveis para todos os membros. Isso pode resultar em uma maior exposição a gastos médicos inesperados e um maior risco de endividamento devido a custos médicos elevados.

Relação entre gastos em saúde e o tipo de residência

Segundo estudos de Silva e Santos (2010), a relação entre o tipo de casa e os custos de saúde é uma variável importante a ser considerada na análise do impacto econômico dos cuidados de saúde para os indivíduos e famílias. O tipo de habitação pode influenciar os gastos com saúde de diversas maneiras, afetando tanto a disponibilidade de recursos financeiros quanto o acesso a condições de moradia adequadas.

De acordo com Silva e Santos (2010), indivíduos que residem em casas arrendadas podem enfrentar desafios específicos em relação aos custos de saúde, devido à possível instabilidade residencial e à falta de controle sobre o ambiente habitacional. Por outro lado, indivíduos que possuem suas próprias casas podem ter uma maior estabilidade residencial e, conseqüentemente, enfrentar menos preocupações relacionadas à moradia, o que pode influenciar positivamente sua saúde geral.

Além disso, estudos de Oliveira et al. (2015) sugerem que o tipo de habitação também pode afetar o acesso a recursos de saúde, como serviços de atenção primária e instalações médicas próximas. Indivíduos que residem em áreas com alta concentração de habitações arrendadas podem enfrentar maior dificuldade em acessar serviços de saúde de qualidade, o que pode impactar sua saúde e bem-estar.

Em resumo, o tipo de casa em que os indivíduos residem pode desempenhar um papel significativo em seus custos de saúde e acesso a serviços de saúde. Compreender essa relação

é essencial para desenvolver políticas e programas eficazes que promovam a saúde e o bem-estar das comunidades.

Relação entre gastos em saúde e o local da consulta

Segundo estudos de Santos e Silva (2010), o local de consulta é uma variável crucial na determinação do acesso aos cuidados de saúde e nos custos associados à busca por tratamento. A escolha entre diferentes locais de consulta, como hospitais públicos, privados e curandeiros (médicos tradicionais), pode ter implicações significativas para os indivíduos e suas famílias em termos de qualidade dos cuidados, acessibilidade e custos.

De acordo com Santos e Silva (2010), os hospitais públicos são frequentemente a principal opção para indivíduos de baixa renda ou sem seguro saúde, pois oferecem serviços subsidiados ou gratuitos. No entanto, esses hospitais podem enfrentar desafios, como longas filas de espera, falta de recursos e infraestrutura inadequada, o que pode afetar a qualidade e a eficiência do atendimento.

Por outro lado, os hospitais privados oferecem uma gama mais ampla de serviços e uma experiência mais personalizada, mas a um custo mais elevado. Indivíduos que optam por hospitais privados podem ter acesso a tratamentos mais avançados e uma melhor relação médico-paciente, mas isso geralmente vem com um ônus financeiro significativo.

Além disso, em algumas comunidades, os curandeiros ou médicos tradicionais desempenham um papel importante na prestação de cuidados de saúde. Esses profissionais podem oferecer tratamentos baseados em práticas tradicionais e conhecimentos locais, que são valorizados por algumas pessoas. No entanto, a eficácia desses tratamentos pode variar e, em muitos casos, os curandeiros podem não ter formação médica formal, o que levanta preocupações sobre segurança e qualidade dos cuidados.

Em resumo, a escolha do local de consulta é uma decisão complexa que envolve considerações sobre acessibilidade, qualidade e custo dos cuidados de saúde. Compreender as opções disponíveis e suas vantagens e desvantagens é essencial para garantir que todos tenham acesso a cuidados de saúde adequados e de qualidade.

Relação entre gastos em saúde e estado civil do chefe do agregado familiar

Segundo estudos de Almeida e Pereira (2020), o estado civil do chefe do agregado familiar, particularmente se ele é solteiro ou casado, tem um impacto significativo nos gastos com saúde

da família. Esta variável reflete diferenças nos padrões de consumo, acesso a cuidados e suporte social, sendo crucial para a formulação de políticas de saúde mais eficazes e inclusivas.

De acordo com Almeida e Pereira (2020), agregados familiares liderados por chefes casados tendem a ter melhores resultados de saúde e menores gastos com saúde em comparação com aqueles liderados por chefes solteiros. Esse fenômeno pode ser atribuído ao suporte emocional e financeiro proporcionado por um parceiro, além de uma maior probabilidade de compartilhamento de responsabilidades financeiras e de cuidados dentro do lar.

Estudos realizados por Silva et al. (2021) indicam que chefes de família casados frequentemente têm melhor acesso a seguros de saúde, o que pode reduzir significativamente os gastos diretos com cuidados médicos. A presença de um cônjuge também pode incentivar comportamentos de saúde mais saudáveis entre os membros da família, como a adesão a regimes de tratamento e a realização de exames preventivos, contribuindo para uma melhor saúde geral e menores gastos médicos.

Por outro lado, agregados familiares liderados por chefes solteiros enfrentam desafios únicos que podem resultar em maiores gastos com saúde. Chefes de família solteiros, especialmente aqueles que são responsáveis por filhos, podem ter menos recursos financeiros disponíveis para cobrir gastos médicos e menos suporte social para ajudar a gerenciar as necessidades de saúde da família. Estudos sugerem que esses chefes de família têm menor probabilidade de buscar cuidados preventivos e podem atrasar a procura por tratamento até que as condições de saúde se tornem mais graves e custosas.

Relação entre gastos em saúde e problemas de alimentação por falta de dinheiro ou outros recursos

Segundo estudos de Oliveira e Silva (2018), os problemas de alimentação devido à falta de dinheiro ou outros recursos representam uma preocupação significativa em muitas comunidades ao redor do mundo. A capacidade de uma pessoa ou família de garantir acesso regular a alimentos nutritivos e adequados está intrinsecamente ligada a fatores socioeconômicos, acesso a serviços de saúde e políticas públicas relacionadas à segurança alimentar.

De acordo com Oliveira e Silva (2018), a falta de recursos financeiros é frequentemente citada como uma das principais causas de insegurança alimentar. Indivíduos ou famílias que enfrentam dificuldades financeiras podem ter dificuldade em adquirir alimentos suficientes ou

de qualidade adequada para atender às suas necessidades nutricionais, o que pode levar a deficiências nutricionais e problemas de saúde relacionados à dieta inadequada.

Além disso, estudos de Santos et al. (2020) destacam que a insegurança alimentar também pode ser exacerbada por outros fatores, como falta de acesso a mercados de alimentos, transporte limitado, problemas de armazenamento de alimentos e falta de conhecimento sobre nutrição e preparação de refeições saudáveis. Esses desafios podem afetar especialmente comunidades rurais ou urbanas de baixa renda, onde o acesso a alimentos frescos e saudáveis pode ser limitado.

Em muitos casos, problemas de alimentação por falta de dinheiro ou outros recursos podem resultar em consequências graves para a saúde, incluindo desnutrição, obesidade, doenças crônicas e comprometimento do desenvolvimento infantil. Esses problemas podem ter um impacto duradouro na qualidade de vida e no bem-estar das pessoas afetadas, além de colocar uma carga adicional sobre os sistemas de saúde e assistência social.

2.6 Análise da associação das variáveis

Teste de independência do qui-quadrado

A estatística do qui-quadrado (χ^2) serve para determinar se duas variáveis são independentes. De acordo com Silva (1992), utiliza-se o teste de independência quando os dados da pesquisa se apresentam sob forma de frequências em categorias discretas. Neste trabalho, aplicou-se o teste do qui-quadrado (χ^2) para testar a associação entre as variáveis independentes e a variável dependente

As hipóteses a serem testadas são:

- ❖ H_0 : As variáveis são independentes, isto é, não há associação entre as variáveis
- ❖ H_1 : As variáveis são dependentes, isto é, há associação entre as variáveis

O teste do qui-quadrado é representado pela seguinte expressão:

$$\chi^2 = \sum_i^L \sum_j^C \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi_{(L-1)(C-1)}^2 \quad (2.1)$$

Onde:

O_{ij} - Representa frequências observadas na linha i da coluna j ;

E_{ij} - Representa frequências esperadas na linha i da coluna j .

L - É o número de linhas.

C - É o número de colunas.

2.7 Estatística multivariada

Segundo Mingoti (2005), a estatística multivariada consiste em um conjunto de métodos estatísticos utilizados em situações nas quais as variáveis são medidas simultaneamente, em cada elemento amostral. Em geral, as variáveis são correlacionadas entre si e quanto maior for o número de variáveis, mais complexa torna-se a análise por métodos comuns de estatística univariada.

De acordo com Bissacot (2015), a estatística multivariada divide-se em dois grupos: um primeiro, constituídos por técnicas exploratórias de sintetização (ou simplificação) da estrutura de variabilidade dos dados, e um segundo, consistindo em técnicas de inferência estatística.

Em geral, os métodos de estatística multivariada são utilizados com o propósito de simplificar ou facilitar a interpretação do fenómeno que esta sendo estudado através da construção de variáveis alternativas que sintetizem a informação original dos dados; Construir grupos de elementos amostrais que apresentem similaridade entre si, possibilitando a segmentação do conjunto de dados original; Investigar as relações de dependência entre as variáveis respostas associadas ao fenómeno e outros factores (variáveis explicativas), muitas vezes, com objectivos de predição; Comparar populações ou validar suposições através de testes de hipóteses.

2.7.1 Análise de regressão

Segundo Maroco (2007), o termo “regressão” foi proposto pela primeira vez, num estudo onde demonstrou que a altura dos filhos não tende a refletir a altura dos pais, mas tende a regredir para a média da população. A análise de regressão é uma técnica básica para modelar e investigar relações entre variáveis.

2.7.2 Regressão logística

De acordo com Montgomery e Runger (2011), a coleção de técnicas estatísticas que são usadas para modelar e explorar relações entre variáveis que estão relacionadas de maneira não determinística é chamada de análise de regressão. Para os autores, a regressão linear funciona

muito bem quando a variável de resposta é quantitativa, enquanto para a situação em que a variável resposta é qualitativa, o modelo de regressão utilizado é a regressão logística, cujos parâmetros são geralmente estimados pelo método de máxima verossimilhança.

Quando a variável dependente é do tipo nominal dicotômico, a regressão logística é a técnica de regressão mais adequada para modelar a ocorrência, em termos probabilísticos de uma das realizações das classes da variável.

De acordo com Amaral (2008), os modelos logísticos, geralmente fornecem bom ajuste com um número menor de parâmetros, comparativamente aos modelos lineares. A transformação de um modelo não linear em linear nos seus parâmetros, facilita o processo de ajuste do modelo, mas em alguns casos pode fazer-se suposições não realistas sobre o termo dos erros (distribuição normal e variância constante), e conseqüentemente perde-se a informação sobre os erros padrão originais. De salientar que existem modelos que não podem ser linearizados por transformação, porque são intrinsecamente não lineares.

2.7.3 Regressão logística simples

Segundo Hosmer e Lemeshow (2000), variável resposta é dicotômica, atribuindo-se o valor 1 ao acontecimento de interesse (sucesso) e 0 ao acontecimento complementar (insucesso). Considerando o modelo abaixo:

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (2.2)$$

Assumindo que $E[\varepsilon_i] = 0$, obtém-se:

$$E[Y_i] = \beta_0 + \beta_1 x_i \quad (2.3)$$

A variável resposta Y tem a distribuição de Bernoulli $(1, \pi)$, com a probabilidade de sucesso $P(Y_i = 1) = \pi_i$ e de fracasso $P(Y_i = 0) = 1 - \pi_i$, desta forma

$$E[Y_i] = \pi_i \quad (2.4)$$

Igualando (2) e (3), temos

$$E[Y_i] = \pi_i = \beta_0 + \beta_1 x_i \quad (2.5)$$

O modelo de regressão logística simples é dado pela seguinte notação:

$$\pi(x_i) = \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \quad (2.6)$$

Transformação Logit

De acordo com Maroco (2007), para o ajuste do modelo recorreu-se a linearização dos parâmetros com a transformação logit, que é designada por função de ligação nos modelos lineares generalizados. Esta transformação define-se como:

$$g(x) = \text{Logit}(\pi(x_i)) = \ln \frac{\pi(x_i)}{1-\pi(x_i)} \quad (2.7)$$

$$g(x) = \ln \frac{\frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}}}{1 - \frac{1}{1 + e^{\beta_0 + \beta_1 x_i}}} \quad (2.8)$$

$$g(x) = \ln e^{\beta_0 + \beta_1 x_i} = \beta_0 + \beta_1 x_i \quad (2.9)$$

O rácio $\frac{\pi(x_i)}{1-\pi(x_i)}$ designa-se por rácio de verosimilhança, chances ou odds. Este rácio é um conceito chave na regressão logística e traduz a razão entre a probabilidade do sucesso e a probabilidade do insucesso.

Características da transformação logit

A função logit possui algumas propriedades do modelo regressão linear tais como:

1. A função logit pode ser contínua;
2. É linear nos parâmetros;
3. Os seus valores podem variar em R.

Estimação dos parâmetros

Segundo Hosmer e Lemeshow (2000), na regressão linear, a variável resposta pode ser expressa como $Y_i = E[Y | X = x_i] + \varepsilon_i$, onde a quantidade ε_i é o erro expressa o desvio de uma observação em relação à média. Parte-se do pressuposto que o erro segue uma distribuição normal, com média zero e variância constante. Mas no caso que se depara com uma variável dicotómica isso não acontece. O erro ε_i , pode assumir dois valores:

- ❖ $Y = 1 \rightarrow 1 - \pi(x_i)$, com a probabilidade $\pi(x_i)$, onde $\pi(x_i) = P(Y = 1|X = x_i) = \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}}$
- ❖ $Y = 0 \rightarrow 1 - \pi(x_i)$, com a probabilidade $\pi(x_i)$, onde $\pi(x_i) = P(Y = 1|X = x_i) = \frac{1}{1 + e^{\beta_0 + \beta_1 x_i}}$

Então, ϵ_i tem uma distribuição com média zero e variância $\pi(x_i)[1 - \pi(x_i)]$. Assim, a distribuição condicional da resposta de Y , é uma distribuição de Bernoulli com parâmetro π_i .

Ajuste do modelo

Para o ajuste do modelo de regressão logístico é preciso estimar os parâmetros desconhecidos, β_0 e β_1 . O método usado para a estimação dos parâmetros é o de máxima verossimilhança. Segundo Hair et al. (2006), a estimação dos parâmetros na regressão logística é feita usando o procedimento de máxima verossimilhança para encontrar as estimativas mais prováveis. Supondo que (x_i, Y_i) , seja uma amostra independente com n pares de observações, Y_i representa o valor da variável resposta dicotômica e x_i é o valor da i -ésima observação em que $n=1, 2, 3, \dots, n$. A função de distribuição de probabilidade de Y_x , é dada por:

$$\phi_{Y_x} = \pi(x_i)^{y_x} (1 - \pi(x_i))^{1-y_x}, \text{ com } y_x \in [0, 1] \quad (2.10)$$

Assumindo-se a independência das observações, a função de verossimilhança é obtida da seguinte maneira:

$$L(\beta) \prod_{i=1}^n (\phi_{x_i}) = \prod_{i=1}^n \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i}, y_i \in [0, 1] \quad (2.11)$$

A expressão log-verossimilhança é definida por:

$$L(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i}, \beta \in R^2 \quad (2.12)$$

O princípio da máxima verossimilhança é estimar o valor de β que maximiza $L(\beta)$. Portanto, aplicando o logaritmo, a expressão é definida como:

$$\begin{aligned} L(\beta) &= \ln[L(\beta)] = \ln \left[\prod_{i=1}^n \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \right] \\ &= \sum_{i=1}^n [y_i \ln(\pi(x_i)) + \ln(1 - \pi(x_i)) - y_i \ln(1 - \pi(x_i))] \end{aligned}$$

$$= \sum_{i=1}^n \left[y_i \ln \left(\frac{\pi(x_i)}{1-\pi(x_i)} \right) + \ln(1 - \pi(x_i)) \right] \quad (2.13)$$

Substituindo pelas equações (4) e (5), obtém-se:

$$L(\beta) = \sum_{i=1}^n \left[y_i (\beta_0 + \beta_1 x_i) + \ln \left(\frac{1}{1 + e^{\beta_0 + \beta_1 x_i}} \right) \right] \quad (2.14)$$

O valor β que maximiza $\ln [L(\beta)]$ é obtido após derivar $l(\beta)$ em relação aos parâmetros (β_0 e β_1) Derivando em ordem aos parâmetros obtém-se:

$$\frac{\partial \ln[L(\beta)]}{\partial \beta_0} = \sum_{i=1}^n \left[y_i - \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \right] = \sum_{i=1}^n [y_i - \pi(x_i)] \quad (2.15)$$

$$\frac{\partial \ln[L(\beta)]}{\partial \beta_1} = \sum_{i=1}^n \left[y_i x_i - x_i \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \right] = \sum_{i=1}^n x_i [y_i - \pi(x_i)] \quad (2.16)$$

Essas expressões São equações não lineares nos parâmetros, conseqüentemente, para a sua resolução necessária a aplicação de métodos iterativos. E o método mais usado, na maioria dos softwares estatísticos é o de Newton- Ralph.

2.7.4 Regressão logística múltipla

Quando se tem mais de uma variável independente, ou seja, múltiplas variáveis independentes, Hosmer e Lemeshow (2000), generalizam o modelo de regressão logística da seguinte forma: Seja um conjunto de p variáveis independentes, denotadas por $x_i^T = (x_{i0}, x_{i1}, \dots, x_{ip})$, o vector da i -ésima linha da matriz (x) das variáveis explicativas, em cada elemento da matriz corresponde ao i -ésimo componente (x_{ij}), onde $i= 1,2,3, \dots, n$ e $j= 0,1,2, \dots, p$, com $x_{i0}=1$. Denota-se por $\beta = (\beta_0, \beta_1, \beta_2, \dots, \beta_p)^T$ o vector de parâmetros desconhecidos e β_j é o i -ésimo parâmetro associado a variável explicativa x_j . A probabilidade de sucesso no modelo de regressão múltipla é dada por:

$$\pi(x_i) = p(Y = 1|X = x_i) = \frac{e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}}}{1 + e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}}} \quad (2.17)$$

E a probabilidade de fracasso é dada por:

$$1 - \pi(x_i) = p(Y = 0|X = x_i) = 1 - \frac{e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}}}{1 + e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}}} \quad (2.18)$$

O logit da regressão logística múltipla é dado pela equação:

$$g(x) = \ln \left(\frac{\pi(x_i)}{1-\pi(x_i)} \right) = x_i^T \beta = \beta_0 + \sum_{j=1}^p \beta_j x_{ij} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (2.19)$$

2.7.5 Estimação dos parâmetros do modelo regressão logística múltipla

Para a estimação dos parâmetros, a regressão logística múltipla é similar a regressão logística simples porque ambas usam o método de máxima verosimilhança. Para encontrar o valor β que maximiza o $l(\beta)$, é utilizado o processo interactivo de Newton-Raphson, para tal, deve-se derivar $l(\beta)$ em função de cada parâmetro.

$$\frac{\partial l(\beta)}{\partial \beta_j} = \sum_{i=1}^n \left[y_i x_{ij} - \frac{e^{x_i^T \beta}}{1 + e^{x_i^T \beta}} \right] \quad (2.20)$$

Deste modo, o vector $U(\beta)$, é dado por:

$$U(\beta) = x_y^T - x_\pi^T = x_{(y-\pi)}^T \quad (2.21)$$

E a matriz de informação de Fisher é dada por:

$$I(\beta) = E \left[-\frac{\partial^2 l(\beta)}{\partial \beta \partial \beta^T} \right] = X^T Q X \quad (2.22)$$

Onde $Q = \text{diag} [\pi_i (1 - \pi_i)]$ e X a matriz de dados, e a sua inversa $[I(\beta)]^{-1}$, a matriz de variâncias e covariâncias das estimativas de máxima verosimilhança dos parâmetros (Maroco, 2007).

2.7.6 Transformação logit

Para corrigir o problema dos valores de probabilidade serem números que não pertencem ao intervalo $[0,1]$, pode-se linearizar a função logística pela transformação designada logit. A transformação logit para estudos de regressão logística central é definida como.

$$g(x) = \ln \left(\frac{P(Y=1|X=x_i)}{P(Y=0|X=x_i)} \right) \quad (2.23)$$

Desenvolvendo a equação acima tem-se:

$$g(x) = \ln (e^{(\beta_0 + \sum_{i=1}^n \beta_i x_i)}) = \beta_0 + \sum_{i=1}^n \beta_i x_i \quad (2.24)$$

A função $g(x)$ apresenta as seguintes propriedades: é linear em seus parâmetros, continua, varia no intervalo de $]-\infty, +\infty[$ com correspondência no intervalo de $[0,1]$, para π_i . No contexto de modelos lineares generalizados, uma função monótona e derivável que relaciona a média ao preditor é denominada de função ligação.

2.7.7 Métodos de seleção das variáveis

De acordo com Moura (2018), a seleção das variáveis do modelo é baseada em algum algoritmo que verificam a importância de dada variável e a sua inclusão ou não no modelo. Assim tem-se de forma bastante difundida estes três métodos aqui apresentados:

- ❖ **Método enter** todas as variáveis pré-selecionadas são forçadas a ficar no modelo, não tem exclusão de variável insignificante.
- ❖ **Método forward** cada variável é adicionada individualmente, sendo a primeira a que adiciona maior poder de explicação ao modelo e assim susceptivelmente até que nenhuma das variáveis restantes aumente o poder de explicação do modelo.
- ❖ **Método backward** contrário do forward ele começa com todas as variáveis e retira individualmente a variável que adiciona o menor poder de explicação ao modelo até que restem somente as variáveis que expliquem significativamente uma parcela da nossa variável dependente. Este é o método de seleção de variáveis usado neste trabalho.
- ❖ **Método stepwise** incorpora os modelos forward e backward, inicia com o forward porém a cada variável adicionada as variáveis anteriores são revisadas e verifica-se se seu poder de explicação do modelo permanece significativo.

2.7.8 Critérios para a seleção do modelo

Escolher o melhor modelo é controverso, mas um bom modelo deve conseguir equilibrar a qualidade do ajuste e a complexidade, sendo esta, em geral, medida pelo número de parâmetros presentes no modelo; quanto mais parâmetros, mais complexo o modelo, sendo pois mais difícil interpretar o modelo. A seleção do “melhor” modelo torna-se então necessário.

Critério de informação de Akaike

De acordo com Vaz (2020), Akaike (1974), propôs utilizar a informação de Kullback-Leibler para a seleção de modelos. Ele estabeleceu uma relação entre a máxima verossimilhança e a informação de Kullback-Leibler desenvolvendo então um critério para estimar a informação de Kullback-Leibler, o posteriormente chamado, Critério de Informação de Akaike (AIC). O Critério de Informação de Akaike realiza um processo de minimização que não envolve testes estatísticos e pode ser expresso em função do desvio do modelo, e é baseado na função de verossimilhança. Fundamenta-se no conceito de entropia, oferecendo uma medida relativa das informações perdidas, quando um determinado modelo é usado para descrever a realidade. Com isso, Akaike (1974) definiu seu critério de informação como

$$AIC = -2(\text{Função maximizada}) + 2(\text{número de parâmetros}) \quad (2.25)$$

onde K é o número de parâmetros no modelo estatístico, e L é o valor maximizado da função de verossimilhança para o modelo estimado. O AIC não é uma prova sobre o modelo, no sentido de testar hipóteses, mas uma ferramenta para a seleção de modelos; não é um teste de hipóteses, não há significância e nem valor-p. Dado um conjunto de dados e vários modelos concorrentes, pode-se classificá-los de acordo com o seu AIC, quanto menor o valor de AIC, possivelmente melhor será o modelo. Não se deve atribuir um valor cima do qual um determinado modelo é “rejeitado”.

Por si só, o valor do AIC para um determinado conjunto de dados não tem qualquer significado. O AIC torna-se útil quando são comparados diversos modelos.

Testes de Significância do Modelo

Após a estimação dos coeficientes do modelo, avalia-se a qualidade da estimação, ou seja, testa-se se as variáveis explicativas pertencentes ao modelo são significativas para explicar o comportamento da variável resposta. De acordo com Paula (2013), citada por Vaz (2020),

os testes de hipóteses para os modelos lineares generalizados baseiam-se em três estatísticas: teste de razão de verossimilhança, teste de Wald e o teste de Escore.

De acordo com Vaz (2020), no livro de Buse (1982) podemos encontrar uma forma bastante didática para a interpretação geométrica dos testes da razão de verossimilhanças, escore e Wald. E segundo o mesmo livro, a estatística de razão de verossimilhança é o critério mais poderoso.

Estatísticas associadas a regressão logística

Segundo Guimaraes (2006), depois de calcularmos coeficientes, o passo seguinte consiste em determinar se as variáveis independentes são associadas significativamente à variável de resultado. Os métodos de inferência nos modelos lineares generalizados baseiam-se, fundamentalmente, na teoria de máxima verosimilhança. Os testes mencionados são assintoticamente equivalentes, sendo a razão de verosimilhança o critério que define um teste uniforme mais poderoso.

Teste de wald

Segundo Moura (2018), o teste Wald é um teste estatístico paramétrico que testa se cada coeficiente é significativamente diferente de zero. Desta forma, este teste verifica se cada uma das variáveis independentes apresenta uma relação estatisticamente significativa com a variável dependente. A estatística do teste é:

$$Wald = \frac{\widehat{\beta}_i}{SE(\widehat{\beta}_i)} \quad (2.26)$$

Onde:

- ❖ $\widehat{\beta}_i$ é o i-ésimo coeficiente da regressão
- ❖ $SE(\widehat{\beta}_i)$ é o erro padrão do i-ésimo coeficiente da regressão

$$\begin{cases} H_0: \beta_0 = 0 \\ H_1: \beta_1 \neq 0 \end{cases}$$

Regra de decisão: Rejeita-se a hipótese nula se o nível de significância associado ao teste (p-valor) for menor ou igual a 5%.

Razão de Chances (Odds Ratio)

A razão de chances é dada pelo quociente entre a odds do acontecimento de interesse ocorrer ($Y = 1$) nos indivíduos com $x = 1$ e a odds desse acontecimento ocorrer nos indivíduos com $x = 0$. O odds do acontecimento de interesse ocorrer nos indivíduos com $x = 1$ é definida por

$\frac{\pi_1}{1-\pi_1}$ Analogamente, a odds do acontecimento de interesse ocorrer nos indivíduos com $x = 0$ é definida por $\frac{\pi_0}{1-\pi_0}$. Assim, o odds ratio é uma forma de comparar se a probabilidade de o acontecimento de interesse ocorrer é a mesma para os indivíduos com $x = 1$ ou $x = 0$.

$$OR = \frac{\text{Probabilidade da ocorrência do evento}}{\text{Probabilidade da não ocorrência do evento}} \quad (2.27)$$

A medida de associação Odds Ratio (OR), denominado razão de chances, é utilizada usualmente na regressão logística invariável para complementar o teste à significância da covariável (x).

A codificação de x permite a interpretação trivial dos parâmetros. Na prática é usado a tabela de contingência para o cálculo de odds ratio. As probabilidades de o evento de interesse ocorrer para as duas categorias de x , são dadas pelas expressões

$$\pi_{(1)} = \frac{\exp(\beta_0 + \beta_1)}{1 + \exp(\beta_0 + \beta_1)} \quad (2.28)$$

$$\pi_{(0)} = \frac{\exp(\beta_0)}{1 + \exp(\beta_0)} \quad (2.29)$$

Consequentemente, o valor da razão de chances é dado pela expressão

$$OR = \frac{\pi_{(1)}[1-\pi_{(0)}]}{\pi_{(0)}[1-\pi_{(1)}]} \quad (2.30)$$

Onde:

- ❖ $OR > 1$ indica que a probabilidade de o evento ocorrer é maior do que a probabilidade do evento não ocorrer.
- ❖ $OR < 1$ indica uma diminuição na ocorrência do evento.
- ❖ $OR = 1$ indica que a probabilidade de o evento ocorrer não afeta a probabilidade do evento não ocorrer.

Teste de Razão de Verossimilhança

Na regressão linear, para avaliação da qualidade do modelo estimado aplica-se a soma dos quadrados dos erros, enquanto, na regressão logística aplica-se o teste de razão de verossimilhança. Segundo Cabral (2013), com o teste de razão de verossimilhança, pretende-se testar simultaneamente se os coeficientes de regressão associados a β são todos nulos com

exceção de β_0 . A comparação dos valores observados e dos valores esperados usando a função de verosimilhança é baseada na seguinte expressão:

$$D = -2 \ln \sum_{i=1}^n \left[y_i \ln \left(\frac{\hat{\pi}}{y_i} \right) + (1 - y_i) \ln \left(\frac{1 - \hat{\pi}}{1 - y_i} \right) \right] \quad (2.31)$$

Onde, Modelo saturado corresponde ao modelo com todas as variâncias e interações e o Modelo corrente, corresponde ao modelo com apenas as variáveis desejadas para o estudo. Assim sendo, a hipótese a testar é:

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0$$

$$H_1 : \exists i = 1, 2, 3, \dots, k: \beta_k \neq 0$$

Onde a estatística do teste é definida por:

$$G = D (\text{modelo} \cdot \text{sem} \cdot p \cdot \text{variáveis}) - D (\text{modelo} \cdot \text{com} \cdot p \cdot \text{variáveis}) \quad (2.32)$$

Ao rejeita-se a hipótese nula, pode-se concluir que pelo menos um dos coeficientes é estatisticamente diferente de zero. Salientar que antes de chegar-se a conclusão final, deve-se testar se cada um dos coeficientes é significativamente diferente de zero, sendo assim, o teste apropriado é o Teste de Wald.

Diagnóstico do Ajuste

A análise de diagnóstico é uma etapa fundamental no ajuste de modelos de regressão. Quando um modelo é ajustado a um conjunto de dados tem de se ter em atenção a análise de medidas das diferenças entre os seus valores observados da variável resposta, y , e os resíduos. Vale destacar que os resíduos têm papel fundamental na verificação do ajuste de um modelo.

Qualidade do ajuste do modelo

Na regressão linear, a dimensão do efeito das variáveis independentes sobre a variável dependente é obtida partindo do coeficiente de determinação. No caso da regressão logística, não é possível aplicar o mesmo procedimento, pois, a variância da variável dependente, depende da probabilidade em que ocorra os seus valores. Segundo Maroco (2007), a qualidade do ajustamento no modelo de regressão logística pode ser avaliada recorrendo a várias estatísticas.

Onde:

$LL(M)$ é -2 vezes o logaritmo da função de verosimilhança do modelo estimado e $LL(M_0)$ é -2 vezes o logaritmo do modelo nulo. Um coeficiente de determinação que considera a dimensão da amostra é o coeficiente proposto por Cox e Snell dado por:

$$R_{CS}^2 = 1 - \exp\left(\frac{LL(M) - LL(M_0)}{n}\right) \quad (2.33)$$

O coeficiente de ajuste proposto por Nagelkerke tem vantagem de assumir o valor 1 como valor máximo.

$$R_N^2 = \frac{R_{CS}^2}{1 - \exp\left(\frac{LL(M) - LL(M_0)}{n}\right)} \quad (2.34)$$

Embora estes coeficientes não possam ser interpretados de forma independente ou comparados em diferentes conjuntos de dados, eles são teis para comparar e avaliar diferentes modelos de previsão com a mesma variável resposta no mesmo conjunto de dados. Em geral, estes coeficientes de determinação assumem valores baixos e valores que variam entre 0.2 e 0.4 indicam um bom ajustamento.

Teste de Hosmer e Lemeshow

Segundo Hair et al. (2010), este teste mede o grau de ajuste do modelo logístico, este indicador corresponde a um teste do qui-quadrado que consiste em dividir o número de observações em cerca de dez classes e, em seguida, comparar as frequências preditas com as observadas. A finalidade desse teste é verificar se existem diferenças significativas entre as classificações realizadas pelo modelo e a realidade observada. Busca-se não rejeitar a hipótese de que não existem diferenças entre os valores preditos e observados.

$$HL = \sum_{i=1}^g \frac{(y_i - n_i \bar{\pi})^2}{n_i (\bar{\pi})(1 - \bar{\pi})} \sim \chi_{g-2}^2 \quad (2.35)$$

Ajuste do modelo

Supondo que se tem uma amostra com observações independentes do $p+1$ vector (x_i, y_i) , $i = 1, 2, 3, \dots, n$ onde y_i é o valor da variável dicotômica e x_i o i -ésimo valor de variáveis independentes. Para ajustar o modelo é preciso estimar $\beta T = (\beta_0, \beta_1, \dots, \beta_p)$. De acordo com

Cabral (2013), o método utilizado em caso multivariado é o de máxima verosimilhança. Assumindo a independência das observações a função de verosimilhança é expressa por:

$$L(\beta) = \ln[L(\beta)] = \sum_{i=1}^n [y_i \ln(\pi(x_i)) + (1 - y_i) \ln(1 - \pi(x_i))] \quad (2.36)$$

= ... =

$$\sum_{i=1}^n [y_i \beta_0 + y_i \beta_1 x_{i1} + \dots + y_i \beta_p x_{ip} - \ln(1 + e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}})] \quad (2.37)$$

Matriz de informação de Fisher

A matriz de covariância dos coeficientes estimados é obtida a partir das derivadas parciais de segunda ordem do logaritmo da função de verosimilhança:

$$\frac{\partial^2 \ln[L(\beta)]}{\partial \beta_j^2} = - \sum_{i=1}^n [x_{ij}^2 \pi_i (1 - \pi_i)] \quad (2.38)$$

$$\frac{\partial^2 \ln[L(\beta)]}{\partial \beta_j \partial \beta_k} = - \sum_{i=1}^n [x_{ij} x_{ik} \pi_i (1 - \pi_i)] \quad (2.39)$$

Onde, $j, k = 0, 1, 2, \dots, p$ e π_i representa $\pi(x_i)$.

Se for formada uma matriz quadrada de dimensão $(p + 1)$, constituída pelo simétrico dos valores médios em termos referidos nas duas equações anteriores, obtém-se $I(\beta)$, a chamada

Matriz de Informação.

De acordo com Cabral (2013), as variâncias dos coeficientes e as covariâncias entre os coeficientes estimados são obtidos por inversão desta matriz. Designar-se-á por $\sigma^2(\beta_j)$ o j -ésimo elemento da diagonal principal da matriz, $I^{-1}(\beta) \equiv \Sigma(\beta)$ a variância de $\hat{\beta}_j$ e por $\sigma(\beta_j, \beta_u)$ a covariância entre β_j, β_u .

Os estimadores da variância e da covariância, são obtidos de $\Sigma(\beta)$ quando se substitui β pelo seu estimador $\hat{\beta}$. Serão utilizados $\hat{\sigma}^2(\hat{\beta}_j)$ e $\hat{\sigma}(\hat{\beta}_j, \hat{\beta}_u)$, com $j, u = 0, 1, \dots, p$, para designar os valores da respectiva matriz.

Usando a notação matricial pode-se escrever $\hat{I}(\hat{\beta}) = X'V X$, onde X é a matriz $n \times (p + 1)$ contendo, além do vector 1 os valores observados para as variáveis independentes, e V é uma matriz diagonal $n \times p$, de elemento genérico $\hat{\pi}_i(1 - \hat{\pi}_i)$:

Suposições da Regressão logística

De acordo com Kleinbaum (1988), na regressão logística recomenda-se que sejam satisfeitos os pressupostos como a ausência de multicolinearidade entre as variáveis independentes, a ausência de outliers, e amostra mínima recomendada é de 50 observações.

Pressupostos de inexistência de relação linear exata

Segundo Fonseca (2003), em modelos de regressão com duas ou mais variáveis explicativas é usual que tais variáveis apresentem algum tipo de interdependência e essa relação de interdependência chamamos multicolinearidade. A presença de multicolinearidade independe da existência de relação de dependência entre a variável dependente e termos independentes (Ferrari e Glauber, 1969).

Segundo Gujarati e Porter (2001), originalmente, a expressão multicolinearidade significava a existência de uma relação linear “perfeita” ou exata entre alguma ou todas as variáveis explanatórias do modelo de regressão. Hoje, no entanto, o termo multicolinearidade é usado em um sentido mais amplo, para incluir o caso de multicolinearidade perfeita, bem como o caso em que as variáveis X do modelo de regressão estão fortemente intercorrelacionadas, mas não perfeitamente. O aparecimento da multicolinearidade pode ocorrer por diversos motivos, entre eles o método como se deu a coleta dos dados empregados no modelo; as restrições ao modelo ou à população ao qual pertence essa amostra; a especificação do modelo ou a utilização de um modelo sobre determinado.

Consequência da multicolinearidade

Segundo Hoffmann (2016 p. 184), as principais consequências desse fato são:

- ❖ As variâncias e covariâncias das estimativas dos parâmetros serão muito elevadas, isto é, as estimativas obtidas podem ter erros muito grandes e esses erros podem estar altamente correlacionados entre si. A baixa precisão das estimativas torna difícil distinguir as influências das diversas variáveis explanatórias;

- ❖ Pôde-se eliminar variáveis da análise porque os coeficientes não se mostram estatisticamente diferentes de zero, essas variáveis podem, na realidade, ser importante e a amostra disponível é que não permite detectar sua influência.
- ❖ As estimativas dos coeficientes variam muito de amostra para amostra. A adição de algumas observações à amostra pode alterar no valor da estimativa obtida.

Segundo Farrar e Glauber (1967), se o conjunto de variáveis independentes for totalmente independente os coeficientes de regressão não poderão ser estimados uma vez que a matriz resultante da multiplicação da matriz transporta de variáveis independentes da matriz das variáveis independentes será singular e não será possível a inversão dessa matriz necessária para cálculo dos coeficientes de regressão ou seja não existe $(X^T X)^{-1}$. Nesse caso a multicolinearidade é, obviamente, severa e o modelo deve ser revisto.

Detectando a Multicolinearidade

Há alguns indicadores de presença de multicolinearidade que não são calculados pela maioria dos softwares que fazem regressão linear múltipla (Newbold, 1994) entretanto é importante que se fique atento a elementos que indicam a sua presença.

Segundo Montgomery et al. (2008), alguns indicativos de presença de multicolinearidade são mostrados abaixo:

1. Valores altos do coeficiente de correlação;
2. grandes alterações nas estimativas dos coeficientes de regressão quando uma variável independente for adicionada ou retirada do modelo, ou quando uma observação foi alterada ou eliminada;
3. A rejeição da hipótese $H_0 : \beta_1 = \beta_2 = \dots = \beta_m = 0$, por meio da realização do teste F, mas nenhuma rejeição das hipóteses $H_0 : \beta_i = 0, i = 1, 2, 3, \dots, m$, por meio da realização dos testes **T**, sobre os coeficientes individuais da regressão;
4. obtenção de estimativas para os coeficientes de regressão com sinais algébricos contrários àqueles que seriam esperados a partir de conhecimentos teóricos disponíveis de experiências anteriores sobre o fenômeno estudado;

5. Obtenção de intervalos de confiança com elevadas amplitudes para os coeficientes de regressão associados a variáveis independentes importantes.

Segundo Miloca e Conejo (2014), a presença de multicolinearidade pode ser detectada de várias maneiras. Entre as formas conhecidas, dois testes formais mais comumente utilizados são o valor de tolerância (TOL) ou seu inverso, chamado de inflação de variância (VIF) definido pelas equações mostradas abaixo:

$$\text{VIF}(\beta_i) = \frac{1}{1-r_i^2} \quad (2.40)$$

$$\text{TOL} = \frac{1}{\text{VIF}(\beta_i)} \quad (2.41)$$

Análise dos Resíduos

No estudo dos modelos de regressão, a análise dos resíduos é fundamental para verificar a validade dos pressupostos subjacentes ao modelo final ajustado.

Segundo Cox e Snell (1989), os resíduos representam os afastamentos das estimativas médias em relação à média geral preditas onde toda a interferência estatística no modelo de regressão (testes de hipóteses) se baseia nestes pressupostos e se houver violação dos mesmos, a utilização do modelo deve ser posta em causa.

Existem dois tipos de resíduos possíveis que poderão ser utilizados para avaliar a qualidade ajustamento: os resíduos de Pearson e Deviance

❖ Resíduos de Pearson

Os resíduos de Pearson para o i -ésimo indivíduo são definidos por:

$$r(y_i, \hat{\pi}_i) = \frac{y_i - \hat{\pi}_i}{\sqrt{\hat{\pi}_i(1-\hat{\pi}_i)}}, i = 1, 2, 3, \dots, n \quad (2.42)$$

A estatística do teste global baseada nos resíduos de Pearson designa-se por estatística de Qui-quadrado de Pearson é calculado da seguinte forma:

$$X^2 = \sum_{i=1}^n r(y_i, \hat{\pi}_i)^2 \sim X^2(n-p-1) \quad (2.43)$$

Uma estatística alternativa é obtida à custa dos resíduos de Deviance, ainda sob a mesma hipótese nula onde H_0 significa “O modelo encontrado explica bem os dados”.

❖ Resíduos de Deviance

Os resíduos de Deviance para o i -ésimo indivíduo são definidos da seguinte forma por:

$$d(y_i, \hat{\pi}_i) = d_i = \pm \{2[y_i \ln(\frac{y_i}{\hat{\pi}_i}) + (1 - y_i) \ln(\frac{1-y_i}{1-\hat{\pi}_i})]\}^{\frac{1}{2}} \quad (2.44)$$

A estatística do teste a utilizar é:

$$D = \sum_{i=1}^n d(y_i, \hat{\pi}_i)^2 \sim X^2 (n-p-1) \quad (2.45)$$

❖ Resíduos de Pearson padronizados

Um procedimento mais adequado consiste em dividir os resíduos pelo valor estimado do seu desvio padrão. Este valor é aproximado por $\sqrt{\hat{\pi}_i(1-\hat{\pi}_i)(1-h_i)}$, onde h_i é i -ésimo elemento da diagonal do elemento da matriz \hat{H} dada por:

$$\hat{H} = \hat{V}^{\frac{1}{2}} (X' \hat{V} X)^{-1} X' \hat{V}^{\frac{1}{2}} \quad (2.46)$$

Onde \hat{V} é a matriz diagonal com n elementos $\hat{\pi}_i(1-\hat{\pi}_i)$, X é uma matriz ($n \times p$), X' é a transposta de X e $\hat{V}^{\frac{1}{2}}$ é a matriz diagonal em que os seus elementos são iguais a raiz quadrada dos elementos da matriz \hat{V}

Posto isto, os resíduos padronizados são dados por:

$$r_{sj} = \frac{r_j}{\sqrt{1-h_j}} \quad (2.47)$$

Uma estatística de diagnóstico extremamente útil é aquela que examina o efeito provocado pela eliminação de uma observação, nos coeficientes estimados, para os coeficientes do modelo e pode ser representada da seguinte forma:

$$\Delta \hat{\beta}_i = \frac{r_{sj}^2 h_j}{\sqrt{1-h_j}} \quad (2.48)$$

Tabela de classificação

Um modelo de regressão pode ser estatisticamente significativo e não representar a realidade em estudo. Uma das formas de avaliar a eficiência classificativa do modelo é através de tabela classificativa.

Para a construção destas tabelas precisamos de calcular as probabilidades estimadas para a ocorrência do endpoint e, de seguida, determinamos o cut-off, c , para estas probabilidades. A partir do cut-off vamos assumir que indivíduos com probabilidades estimadas superiores a experimentam o endpoint e indivíduos com probabilidades abaixo do cut-off não o experimentam. O valor de cut-off usualmente utilizado é 0.5, contudo nem sempre é o mais adequado. A sensibilidade dá-nos a proporção de verdadeiros positivos. A especificidade fornece a proporção de falsos negativos, valor este que é determinado pela probabilidade de prevermos a não ocorrência do endpoint entre os indivíduos em que este não foi observado (Gonçalves, 2013)

Tabela 2.1: Tabela de classificação

		Estimados	
		Endpoint=1	Endpoint=0
Observados	Endpoint=1	A	B
	Endpoint=0	C	D
Desempenho		Sensibilidade: $A / (A + B)$	Especificidade: $D / (C + D)$

Determinado o cut-off, podemos construir a tabela cujas linhas apresentam os valores observados e as colunas os valores estimados para a variável resposta. Num modelo perfeito todos os casos estariam sobre a diagonal principal. Contudo, na prática é muito difícil obter um modelo perfeito e, como tal, teremos de classificar a sua capacidade predictor. Esta é considerada boa se a sensibilidade e a especificidade forem superiores a 80%, razoável se estes dois valores estiverem entre 50% e 80% e medíocre se ambos forem inferiores a 50%.

Curva Roc

A análise ROC (Receiver Operating Characteristic) é uma ferramenta que permite avaliar o desempenho de um modelo de regressão binária, ela consiste em determinar uma curva que

corresponde à razão entre a sensibilidade e um menos a especificidade (proporção de falsos positivos), isto é, proporção de incorretamente classificada como tendo a característica a modelar, em relação ao total de sujeitos que não tem característica a modelar) a curva ROC é desenvolvida num plano unitário de abcissa $(1 - e)$ e ordenadas, sobre o qual são traçados os diferentes pontos de corte das probabilidades relativas a sensibilidade e a especificidade.

Deve-se escolher como ponto de corte aquele que mais se aproxima do canto superior esquerdo do plano unitário, ou seja, aquele que maximiza a sensibilidade e a especificidade. Quanto mais próximo do canto superior esquerdo o ponto for maior será o poder de discriminação do modelo. Sobre o plano unitário é traçado também uma bissetriz ($y = x$), quanto mais próxima a curva estiver da bissetriz, menor é o poder de discriminação do modelo

A área de baixo da curva ROC, varia, obviamente entre 0 e 1 e é utilizado para medir a eficiência de como o modelo consegue efectuar a classificação ou discriminação dos sujeitos. Se a área for igual a 0,5 o modelo não consegue discriminar os sujeitos com versos sem característica melhor do que a escolha por um mero acesso (isto é, aceita com a probabilidade 0,5). Quanto mais próximo a área estiver de 1 (isto é, da razão entre a sensibilidade igual a 1 e especificidade igual a 0), maior é a capacidade do modelo para discriminar os sujeitos que apresentam a característica modelada (sucesso) dos indivíduos que não apresentam essa característica (insucesso). (Hosmer & Lemeshow, 1989).

2.7.9 Interpretação dos coeficientes estimados

De acordo com Hair et al. (2006), uma das vantagens da regressão logística é que se precisa apenas saber se um evento ocorreu para usar um valor dicotómico como a variável dependente. A partir desse valor dicotómico, o procedimento prevê sua estimativa da probabilidade de que o evento ocorra ou não. Se a probabilidade prevista for maior que 50%, então a previsão será sim, caso contrário, será não.

O procedimento usado para calcular o coeficiente da regressão logística compara a probabilidade de o evento não ocorrer. E a expressão pode ser denotada da seguinte maneira:

Os coeficientes estimados são medidos das variações das probabilidades, chamada de razão de desigualdade. Além disso, eles são apresentados em logaritmos, o que faz com que seja transformado de volta, para que o seu efeito relativo sobre as probabilidades seja avaliado de forma mais simplificada. Um coeficiente com sinal positivo, significa que aumenta a probabilidade de o evento ocorrer, enquanto, um coeficiente com sinal negativo, diminui a probabilidade do evento ocorrer.

De acordo com Hosmer e Lemeshow (2000), para interpretar os valores associados aos coeficientes do modelo, é conveniente proceder a análise de acordo com a natureza das variáveis independentes. Quanto à natureza das variáveis independentes, pode considerar-se três tipos nomeadamente: variáveis dicotómicas, polipotómicas e contínuas.

2.7.10 Razões para o uso da regressão logística

Segundo Maroco (2007), aplica-se a regressão logística quando se pretende prever a presença ou ausência de uma característica ou resultado em valores de um conjunto de variáveis predictoras. É adequado para modelos em que variável dependente é dicotómica, os coeficientes da regressão logística podem ser usados para estimar razões de chances para cada uma das variáveis independentes no modelo.

Segundo Menard, S. (2002), a regressão logística é amplamente utilizada para tarefas de previsão e classificação, sendo aplicada em áreas como marketing, medicina, finanças e ciências sociais, é comumente usada em estudos médicos e sociais para analisar e modelar dados com variáveis binárias ou dicotómicas.

Segundo Agresti, A., & Finlay, B. (2008) A regressão logística é robusta mesmo quando algumas suposições não são totalmente atendidas, tornando-a uma escolha sólida em muitas situações práticas e, é menos sensível à multicolinearidade (correlação alta entre variáveis independentes) em comparação com outras técnicas de regressão, como a regressão linear.

Com este trabalho quer-se determinar a presença de factores determinantes a gastos em saúde em Moçambique, onde tem-se a variável dependente dicotómica, e apresenta dois acontecimentos possíveis, sucesso e fracasso, e a(s) variável(is) independentes(s) que influencia(m) a ocorrência de determinado acontecimento.

As variáveis independentes podem ser categóricas ou contínuas. Na variável dependente, é comum atribuir-se o código “1” ao resultado mais importante ou aquele que se pretende relacionar ao acontecimento de interesse (“Sucesso”) e o código “0” ao resultado complementar (“Fracasso”). O modelo de regressão logística é utilizado quando a variável dependente é binária, categórica ordenada ou mesmo categórica desordenada.

2.7.11 Vantagens para o uso da regressão logística:

De acordo com Maier (2012), o modelo de regressão logística apresenta as seguintes vantagens:

- ❖ Requer um menor número de suposições iniciais;
- ❖ Admite variáveis métricas e não métricas, simultaneamente;

- ❖ Facilita a construção de modelos destinados à previsão de riscos em diversas áreas do conhecimento;
- ❖ Comparada a outras técnicas de dependência, a regressão logística acolhe com mais facilidade variáveis categóricas;
- ❖ Mostra-se mais adequada à solução de problemas que envolvem estimação de probabilidade, pois trabalha com uma escala de resultados que vai de 0 a 1;
- ❖ O modelo é mais flexível quanto a suposições iniciais, tende a ser mais útil e a apresentar resultados mais confiáveis;
- ❖ Os resultados podem ser interpretados com relativa facilidade;
- ❖ Apresenta facilidade computacional, tendo sido incluída em vários pacotes estatísticos amplamente difundidos no mundo todo.

Capítulo 3

3. MATERIAL E MÉTODOS

Este capítulo tem como propósito detalhar os materiais e métodos empregados para atingir os objetivos do trabalho. Dentre as abordagens utilizadas, destacam-se duas principais pesquisas: a pesquisa descritiva, destinada a fornecer uma descrição abrangente da população em estudo, e a pesquisa explicativa, voltada para identificar os factores determinantes relacionados aos gastos catastróficos dos agregados familiares moçambicanos.

3.1 Classificação de pesquisa

3.1.1 Quanto à natureza da pesquisa

Esta é uma pesquisa explicativa, porque o estudo é focado em identificar os factores que determinam ou que contribuem para a ocorrência de um fenómeno. Segundo Gil (2008), pesquisa explicativa é focada em identificar os factores que determinam ou que contribuem para a ocorrência dos fenómenos. Esse tipo de pesquisa visa explicar a razão, o porquê das coisas, e envolve a realização de experimentos e testes de hipóteses.

3.1.2 Quanto aos procedimentos técnicos

Sob essa perspectiva, a pesquisa é um estudo bibliográfico, pois utiliza materiais já publicados, como livros e revistas, para identificação dos conceitos teóricos sobre o tema em estudo. Segundo Gil (1991), citado por Nascimento (2016), a pesquisa bibliográfica é um trabalho de natureza exploratória, que proporciona bases teóricas ao pesquisador, auxiliando no exercício reflexivo e crítico sobre o tema em estudo.

3.1.3 Quanto à abordagem ou método

Segundo Nascimento (2016), quanto ao método ou abordagem metodológica, a pesquisa pode seguir um ou dois métodos, a saber: abordagem ou método quantitativo e abordagem ou método qualitativo. No que concerne à abordagem, esta pesquisa é quantitativa, pois envolve a quantificação na colecta e tratamento dos dados, utilizando técnicas estatísticas para a análise dos mesmos.

3.2 Material

O presente estudo se fundamentou em uma base de dados secundária, especificamente proveniente do Inquérito sobre o Orçamento Familiar (IOF) do ano de 2022. Essa base abrange

informações: demográficas, gastos dos agregados familiares referentes ao consumo alimentício e não alimentício, Educação, saúde, seguros, pobreza, calamidades, posse de terra, refeições, Segurança alimentar, turismo, rendas e de assistência social, pelo que foram selecionadas apenas as variáveis da tabela para poder-se fazer a análise e alcançar os objectivos do estudo.

Tabela 3.1: Descrição e classificação das variáveis em estudo

Nome da variável	Descrição	Codificação	Tipo de variável
Província	Nome da província do agregado familiar	1.Niassa, ...,11. Maputo cidade	Qualitativa
Região	Nome da zona regional do agregado	1.Sul, 2.Centro, 3. Norte	Qualitativa
Bens_Alimentares	Valor mensal de bens alimentares		Quantitativa
Bens_não_alimentares	Valor mensal de bens não alimentares		Quantitativa
Gastos_em_saúde	Valor mensal dos gastos em saúde		Quantitativo
Consumo_exp	Valor mensal total dos gastos do agregado familiar		Quantitativo
Número_de_membros	Número de membros do agregado		Quantitativo
Condição_social	Condição social do agregado social	1.Agregado pobre, 2.Agregado não pobre	Qualitativa
Área_de_residência	Área de residência do agregado familiar	1.Urbano, 2. Rural	Qualitativa
Género	Género do/a chefe agregado familiar	1.Homem, 2. Mulher	Qualitativa
Estado_civil	Estado civil do chefe do agregado familiar	1. Casado, 2. Recusou, 3. Solteiro	Qualitativa
Empregabilidade	Esta variável indica o status de emprego atual do chefe do agregado familiar	1.Desempregado, 2. Subempregado, 3. Empregado	Qualitativa
Idade	Idade do chefe do agregado familiar		Quantitativa

Nível_acadêmico	Nível acadêmico do chefe do agregado familiar	1. Ensino Primário, 2. Ensino Secundário, 3. Ensino Técnico, 4. Não sabe, 5. Nunca frequentou escola, 6. Pré-escolar, 7. Superior	Qualitativa
Saúde	Agregados com ocorrência de gastos em saúde	1. Com gastos diretos em saúde, 0. Sem gastos diretos em saúde	Qualitativa
Tipo_de_casa	Tipo de casa em que o agregado reside	1. Arrendada, 2. Cedida, 3. Própria	Qualitativa
Número_de_visitas_ao_hospital	Número de visitas que os membros do agregado fazem por mês		Quantitativa
Tipo_de_consulta	Tipo de consulta hospitalar	1. Consultas em clínicas especiais e consultórios privados excepto dentários, 3. Consultas em serviços públicos para doentes sem internamento, excepto dentários, 4. Consultas em serviços públicos para doentes sem internamento, excepto dentários	Qualitativa
Problemas_com_Alimentação	Agregados que tiveram problemas de alimentação por falta de valores monetários	1. Teve problemas de alimentação, 2. Não teve problemas de alimentação	Qualitativa
AF30	Agregado familiar que ficou sem comida em sua casa por falta de dinheiro ou outros meios?	1. Não, 2. Não sabe, 3. Sim	Qualitativa
AF31	Agregado familiar que, teve fome, mas não comeu, porque não tinha dinheiro ou outros meios para obter comida?	1. Não, 2. Não sabe, 3. Sim	Qualitativa

Softwares

Para realizar a análise dos dados, foi empregue o software SPSS 25 e Microsoft Excel. Para a produção do relatório final, utilizou-se o software Microsoft Word, proporcionando um documento estruturado e informativo.

3.3 Métodos

A seguir estão delineados os métodos empregados na análise dos dados, notadamente:

Regra de decisão

Neste trabalho, para a tomada de decisão será considerado o nível de significância de 5%, portanto, rejeita-se a hipótese nula se $p\text{-valor} < 0.05$.

3.3.1 Análise exploratória

- ❖ Para a descrição do perfil socioeconómico e demográfico da amostra, as variáveis foram resumidas por meio de gráficos de barras, gráficos circulares e tabela de frequências.
- ❖ Para examinar a relação entre as variáveis independentes e a ocorrência de gastos em saúde, foi empregue o teste de independência do Qui-Quadrado, conforme destacado por Lopes (2003), esse teste é aplicado quando se deseja investigar a associação entre duas ou mais variáveis qualitativas.

3.3.2 Determinação do tipo de gastos em saúde

Para a determinação do tipo de gastos em saúde que a país tem no presente estudo referente aos factores determinantes aos gastos em saúde, recorreu-se a metodologia concebida pela Organização Mundial de Saúde (OMS, 2005), amplamente utilizada para avaliar o impacto das políticas de financiamento à saúde, monitoria do progresso das reformas entre regiões assim como entre países e para analisar os padrões de utilização de serviços de saúde, os procedimentos foram:

1. Conversão de todas as variáveis relacionadas com gastos para variáveis mensais, caso os dados estivessem em outras unidades de medição como por exemplo semana, dia ou ano;
2. Cálculo da linha de linha de pobreza (PI) e os gastos de subsistência do agregado familiar

A fim de minimizar os erros de medição, usou-se os gastos médios de alimentos das famílias cuja contribuição dos gastos em alimentação no total dos gastos das famílias está entre percentil

45 e 55 do total da amostra. Considerando a escala de economia do consumo das famílias, a escala de equivalência familiar foi utilizada ao invés de tamanho real dos agregados familiares.

A percentagem dos gastos em alimentação nos gastos totais foram calculados determinando a contribuição dos gastos com alimentação ($Foodexp_h$) no total dos gastos para cada agregado familiar, dividindo gastos com alimentação do agregado familiar pelos seus gastos totais (exp_h).

Os gastos com alimentação equivalente ($eqfood_h$) foram calculados a partir da divisão dos gastos alimentares das famílias pelo tamanho do agregado familiar equivalente.

Foram identificadas as contribuições dos gastos alimentares nos gastos totais dos agregados familiares que estão no 45 ° ao 55 ° percentil de toda a amostra, foi calculada a média dos gastos com alimentação entre 45 ° ao 55 ° percentil. Isto deu os gastos de subsistência per capita (equivalente), que também é linha de pobreza (PI).

Após o cálculo da linha de pobreza foi calculado o gasto de subsistência para cada agregado familiar (se_h).

Em seguida que foi determinada a condição social do agregado familiar, em que um agregado familiar é considerado pobre ($Poor_h$), quando seus gastos totais são menores do que seus gastos de subsistência.

Em seguida foi feito o cálculo da capacidade de pagamento do agregado familiar (ctp_h), que consiste na subtração dos gastos totais do agregado familiar pelos gastos de subsistência, caso os gastos alimentares forem maiores que os gastos de subsistência, e a subtração dos gastos totais do agregado familiar pelos gastos alimentares, caso os gastos de subsistência forem maiores que os gastos de alimentares.

Em seguida foi feito o cálculo da razão entre os gastos directos em saúde e a capacidade de pagamento do agregado familiar ($oopctp_h$).

Em seguida foi determinada o tipo de gasto em saúde do agregado familiar, se o quociente dos gastos directos em saúde sobre a capacidade de pagamento do agregado familiar for igual ou superior a 5% o agregado familiar tem gastos em saúde do tipo catastrófico e caso contrário o agregado familiar tem gastos em saúde do tipo não catastrófico ($cata_h$).

3.4.1 Regressão logística binária

Além da análise feita anteriormente, fez-se análise descritiva com ajuda de tabelas de frequência e aplicou-se um modelo de regressão logística. Após a análise descritiva passou-se a aplicação da técnica estatística denominada regressão logística, porém esta técnica pressupõe ausência de multicolinearidade entre as variáveis predictoras e para verificar a multicolinearidade recorreu-se aos valores da tolerância e inflação da variância de um factor (VIF).

Criação da variável dependente

Na estimação do modelo de regressão logística binária, a variável resposta resultante da categorização da variável referentes a valores gastos em saúde foi agrupada de modo a formar uma variável resposta dicotômica definida do seguinte modo:

- ❖ 1, se o agregado familiar tem gastos directos em saúde
- ❖ 0, O agregado familiar não tem gastos directos em saúde

Crítérios de seleção do modelo

Para a seleção do modelo, foi usado o método Forward que começa com o modelo nulo e a cada etapa vai incorporando no modelo as variáveis que apresentem um p-valor menor que o nível de significância de remoção, ou seja, o $\alpha = 0.05$ usado no estudo.

Estimação do modelo logístico

Para a estimação do modelo de regressão logística foi usado o método de máxima verossimilhança. Para a estimação do modelo final primeira estimou-se um modelo inicial sem variável para fornecer um padrão para comparações.

Após a estimação do modelo inicial, passa-se então, a considerar todas variáveis selecionadas, a fim de identificar quais farão parte do modelo final. A estatística utilizada para a inclusão definitiva das variáveis no modelo final é a estatística de Wald com os níveis de significância de 0.05, a qual fornece o indicativo de quais variáveis são estatisticamente significativas dentro do modelo. A estatística de Wald tem com hipóteses:

- ❖ H_0 : O coeficiente da regressão logística é igual a zero.
- ❖ H_1 : O coeficiente da regressão logística é diferente a zero.

Avaliação do modelo

Segundo Hair et al. (2010), para a avaliação do ajuste geral do modelo primeiro recorre-se 2LL. O procedimento consiste em comparar o valor do modelo final desta medida com o valor dos demais modelos estimados no processo forward. Uma maior redução deste valor indica um bom ajuste do modelo. O poder explicativo do modelo é dado pelo valor de R^2 de Nagelkerke quanto maior este valor será o ajuste.

- ❖ H_0 : Não há diferença significativas entre os resultados previstos e os observados.
- ❖ H_1 : Há diferença significativas entre os resultados previstos e os observados.

Qualidade de ajuste

Para testar a qualidade de ajuste do modelo aplicou-se o teste proposto por Hosmer e Lemeshow (2000). O teste de Hosmer- Lemeshow testa se o modelo obtido explica de forma adequada os dados observados avaliando o modelo ajustado através das distâncias entre as probabilidades ajustadas e as probabilidades observadas. Estes dados são separados em grupos de acordo com as probabilidades previstas. Neste caso, a hipótese a testar é que o modelo se ajusta bem aos dados.

Diagnostico do modelo logístico

Segundo Hosmer et al. (2013), para avaliar o desempenho do modelo logístico usou-se a curva ROC. A área abaixo da curva de ROC, fornece uma medida de discriminação, que indica a possibilidade de um factor ser determinante a ocorrência de gastos catastróficos em saúde. Seja **A**, o valor que corresponde a área abaixo da curva de ROC, como regra geral tem-se as seguintes linhas de orientação:

3. Se $A \leq 0.5$ não há discriminação
4. Se $0.6 \leq A < 0.7$ discriminação é fraca
5. Se $0.7 \leq A < 0.8$ discriminação é aceitável
6. Se $0.8 \leq A < 0.9$ discriminação é excelente
7. $A \geq 0.9$ discriminação excepcional

Interpretação do modelo

Para interpretar os coeficientes estimados usou-se a razão de chances que é uma medida de associação muito importante e muito utilizada na interpretação do modelo de regressão

logística, pois, ela permite conhecer quais as chances que um evento tem para acontecer se sob as mesmas condições ele não acontece.

Para um valor de EXP (β) superior a 1 ($\beta > 1$) aumenta a probabilidade de ocorrência do evento, enquanto um valor de EXP (β) inferior a 1 ($\beta < 1$) diminui a probabilidade e se EXP (β) for igual a um ($\beta = 1$), a probabilidade de chances dessa ocorrência não altera, isto é, não afecta as chances.

Capítulo 4

4. RESULTADOS E DISCUSSÃO

Neste capítulo, são descritas a caracterização da amostra, os resultados alcançados com o uso da análise descritiva e da técnica de regressão logística binária, além da discussão correspondente.

4.1 Análise descritiva

Caracterização da Amostra em Estudo

A amostra analisada é composta por 15346 agregados familiares em que cada agregado familiar se encontra representado pelo seu chefe do agregado. De acordo com a gráfico 4.1 verifica-se que maior parte dos agregados familiares reside na zona urbana em detrimento da rural sendo que a zona urbana tem uma percentagem de 58% e a zona rural 42%. O gráfico 4.2 mostra que a Cidade de Maputo com 12% e as províncias da Zambézia com 11% e Nampula também com 11% são as que apresentam as frequências mais elevadas. As províncias de Cabo Delgado e Niassa são as que apresentam as menoríssimas percentagens da amostra com cerca de 7% cada.

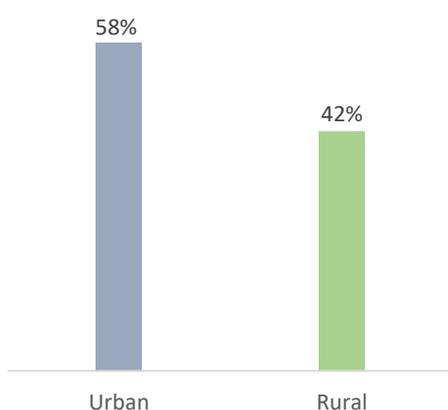


Gráfico 4.1: Distribuição dos agregados por área

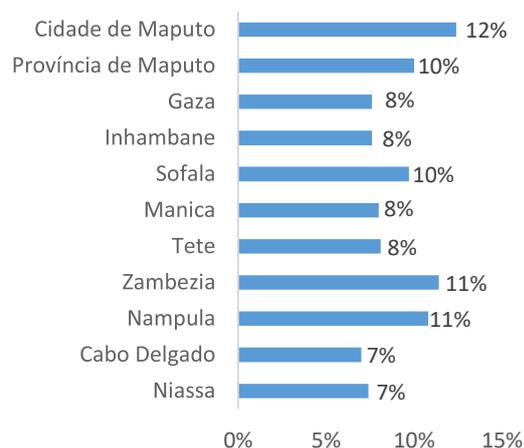


Gráfico 4.2: Distribuição dos agregados por província

A situação de empregabilidade é fundamental para a melhoria da situação socioeconómica do agregado familiar. Os agregados familiares cujo seu chefe ou algum membro encontra-se empregado ou com um trabalho remunerado tem melhores condições para responder a necessidade de subsidiar os gastos em saúde. Ademais, o local de trabalho é um dos factores a

considerar na análise deste fenómeno. De acordo com o gráfico 4.3, cerca de 18% dos chefes dos agregados está desempregado, 10% está subempregado e 72% está empregado.

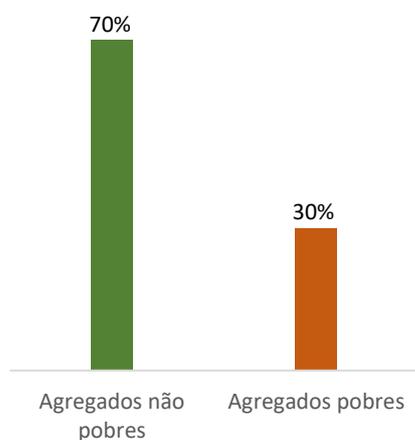


Gráfico 4.4: Condição social do agregado

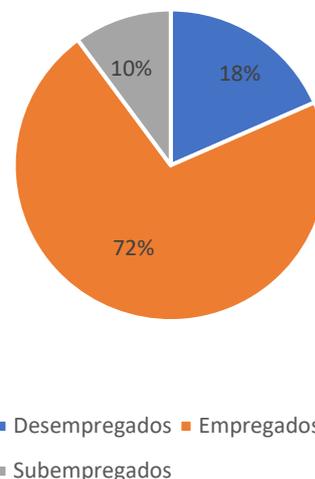


Gráfico 4.3: Empregabilidade do agregado

A situação socioeconómica do agregado familiar é um indicador da capacidade de pagamento do agregado familiar bem como para o financiamento dos seus gastos em saúde. Os agregados familiares classificados como sendo pobres tendem a ter mais dificuldades para pagar os gastos em saúde e tendem a transferir o valor que seria alocado para alimentação e outros gastos de subsistência para cobrir os gastos em saúde. O gráfico 4.4 mostra que 30% dos AF foram considerados pobres e 70% não pobres.

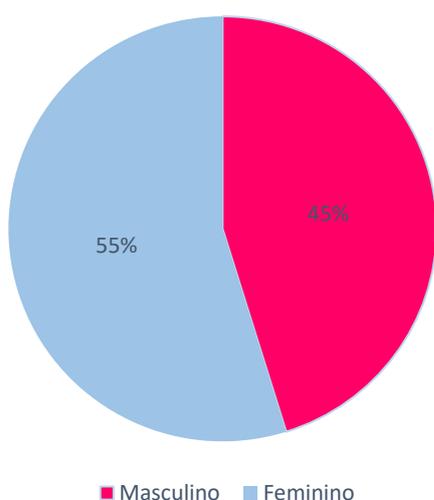


Gráfico 4.5: Género do chefe do agregado

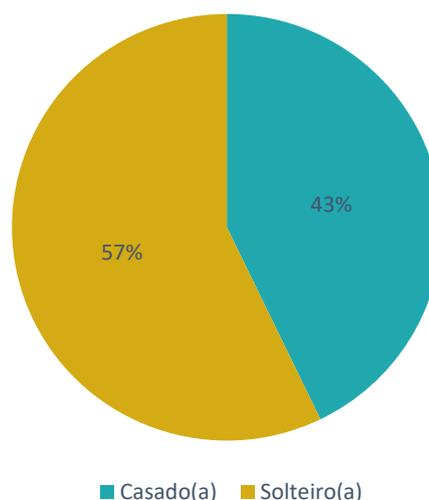


Gráfico 4.6: Estado civil do chefe do agregado

De acordo com a gráfico 4.5 e 4.6 verifica-se a predominância de chefes dos agregados familiares do sexo feminino, representando 55% do total, em comparação com 45% do sexo masculino, além disso, 57% dos chefes dos agregados familiares são solteiros, enquanto 43% são casados.

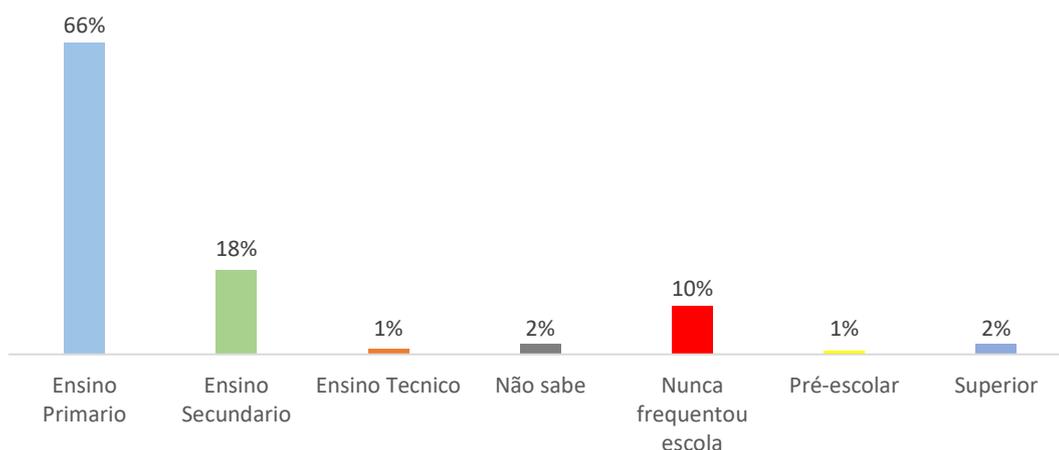


Gráfico 4.7: Nível de académico do chefe do agregado familiar

De acordo com o gráfico 4.7 verifica-se que a maioria dos chefes de agregados familiares tem apenas o ensino primário com 66%, seguido pelo ensino secundário com 18%, apenas 1% possui ensino técnico, 2% possui o ensino superior e 10% nunca frequentaram a escola.

De acordo com a tabela 4.1, observa-se que os gastos alimentares representam a maior parcela dos gastos totais em todas as categorias, destacando-se a significativa diferença entre áreas urbanas e rurais. Nas áreas rurais, os gastos alimentares representam uma percentagem 75.9% em comparação com as áreas urbanas com uma percentagem de 62.14%.

A análise por gênero do chefe do agregado familiar mostra uma distribuição bastante equilibrada entre agregados familiares chefiados por homens e mulheres, com gastos alimentares ligeiramente maiores para mulheres representado uma percentagem de 68.08% em comparação aos homens com 67.76%. Os gastos não alimentares são semelhantes entre os gêneros dos chefes dos agregados familiares com 32% para homens e 32% para mulheres, e os gastos com saúde também não mostram variação significativa com 0.24% para ambos os gêneros.

A condição social e a presença de gastos catastróficos em saúde revelam as maiores disparidades nos gastos. Agregados não pobres gastam menos em gastos alimentares com 66.65%, 33.19% em gastos não alimentares e 0.16% em gastos em saúde, em comparação aos agregados pobres com 71.08% de gastos alimentares 28.49% com gastos não alimentares e 0.44% com gastos em saúde. Além disso, os agregados com gastos catastróficos em saúde têm uma elevada proporção de gastos alimentares com uma percentagem de 79.01% e um significativo aumento nos gastos com saúde com uma percentagem de 5.1%. Em contraste, os agregados sem esses gastos catastróficos gastam apenas 0.07% da sua renda em saúde.

De acordo com a tabela 4.2, em termos de distribuição por província, os dados indicam que a província de Cabo Delgado e Inhambane apresentam as maiores proporções de agregados familiares com gastos catastróficos em saúde com 5% e 3% respectivamente. Em Niassa 3% dos agregados enfrentam tais gastos, enquanto em Nampula essa proporção é de 2%. Nas províncias de Tete, Manica, Sofala, Gaza, Província de Maputo e Cidade de Maputo, apenas 1% dos agregados familiares relatam gastos catastróficos em saúde, e a província da Zambézia destaca-se por não ter registado agregados com esses gastos.

Em relação a área de residência, verificou-se que os agregados em áreas rurais enfrentam uma maior percentagem de gastos catastróficos em saúde com 2% comparados aos das áreas urbanas com 1%.

A distribuição por gênero do chefe do agregado familiar mostra que agregados chefiados por homens apresentam uma proporção ligeiramente maior de gastos catastróficos com 2% em comparação aos chefiados por mulheres com 1%.

Considerando o estado civil do chefe do agregado familiar, os gastos catastróficos são mais comuns entre agregados chefiados por solteiros com 2%, enquanto aqueles chefiados por casados e que recusaram informar o estado civil têm uma incidência de 1%.

Em relação ao nível acadêmico, agregados familiares onde o chefe do agregado familiar nunca frequentou a escola apresentam a maior proporção de gastos catastróficos com 7%. Aqueles com nível pré-escolar e que não sabem o nível acadêmico têm incidências de 4% e 3%, respectivamente. Agregados com ensino técnico não apresentam gastos catastróficos com 0%.

No que concerne à condição social, 5% dos agregados pobres enfrentam gastos catastróficos em saúde, enquanto nenhum dos agregados não pobres enfrentam tais gastos.

Regionalmente, a região norte tem a maior incidência de gastos catastróficos com 3%, em comparação com 1% nas regiões centro e sul.

Quanto ao tipo de casa, agregados em casas próprias enfrentam uma maior incidência de gastos catastróficos com 2% comparados aos que vivem em casas arrendadas ou cedidas com 1%.

Analisando o tipo de alimentação, agregados com alimentação menos adequada e mais que adequada têm uma incidência de 2% de gastos catastróficos, enquanto aqueles com alimentação adequada têm 1%.

No que se refere a problemas alimentares, 2% dos agregados que não tiveram problemas de alimentação enfrentam gastos catastróficos, comparado a 1% daqueles que tiveram problemas.

O local de consulta também influencia a incidência de gastos catastróficos. Agregados que utilizam clínicas especiais, consultórios privados e serviços de curandeiros, apresentam 7% de incidência de gastos catastróficos enquanto aqueles que utilizam serviços públicos apresentam uma incidência de 2%.

Sobre a experiência de insegurança alimentar nos últimos 12 meses, agregados que ficaram sem comida por falta de dinheiro ou outros meios têm uma incidência de 2% de gastos catastróficos. Aqueles que não sabem se passaram por essa situação têm 3%, enquanto aqueles que não passaram por essa situação têm 1%.

Em relação ao número de visitas ao hospital, agregados com 3 a 5 visitas ao hospital têm uma incidência de 3% de gastos catastróficos em saúde, enquanto aqueles com 0 a 2 visitas ao hospital têm uma incidência 2% de gastos catastróficos em saúde e agregados com mais de 6 visitas não registam gastos catastróficos.

Considerando o número de pessoas no agregado familiar, aqueles com 1 a 5 membros têm uma incidência de 2% de gastos catastróficos em saúde. Agregados com 6 a 10 membros, 11 a 15 membros e mais de 16 membros apresentam uma incidência de 1% de gastos catastróficos em saúde. A média percentual de agregados com gastos catastróficos em saúde em Moçambique é de 2%.

4.2 Modelo de regressão logística binária

4.2.1 Análise da variável dependente

De acordo com a tabela 4.3, observa-se que famílias que vivem na zona sul tem menor percentagem de gastos em saúde com cerca de 30% em relação a famílias que vivem na zona centro e norte com a zona centro apresentando 34%.

Observa-se que a população rural gasta proporcionalmente menos em saúde, com 29%, em comparação com a população urbana, com 34%.

Observa-se que os agregados que vivem em casas próprias tem mais gastos com uma em saúde com uma percentagem de 28% em comparação os agregados que vivem em casas cedidas ou de aluguer com uma percentagem de 29% e 28% respectivamente.

Observa-se que agregados familiares cujo chefes do agregado tem nível superior apresentam maiores gastos em saúde com cerca de 39% comparativamente aos demais níveis académicos, e os chefes dos agregados familiares sem nível académico apresentam menores gastos em saúde comparativamente a todos os níveis académicos com uma percentagem de 29%

Os dados indicam uma pequena diferença entre os gastos em saúde no que concerne ao género dos chefes dos agregados familiares, os agregados com chefes homens apresentam 33% de gastos, enquanto as mulheres têm 31%.

Quanto a condição social agregados familiares pobres gastam proporcionalmente menos em saúde, com 28%, do que agregados familiares não pobres, com 34%.

A análise dos gastos em saúde por tipo de residência revela que agregados familiares que vivem em casas próprias têm um gasto uma menor percentagem em gastos de saúde, com 34%, em comparação com aquelas que vivem em casas arrendadas, com 14%.

Em termos de alimentação, agregados familiares com alimentação menos adequada gastam menos em saúde, com uma percentagem de 28%, do que agregados familiares com alimentação mais que adequada, que tem uma percentagem de 36%.

A maioria dos gastos em saúde concentra-se em agregados que fazem de 3 a 5 visitas ao hospital com uma percentagem de 49%, enquanto aquelas com menos visitas de 0 a 2 mostram uma proporção significativamente menor de gastos 32%. Todavia famílias que realizaram mais de 6 visitas ao hospital têm 42% de gastos.

Além disso, agregados familiares com menor número de membros 1 a 5 também mostram uma proporção menor de gastos em saúde (30%), em detrimento dos agregados com maior número de membros que chegam a ter gastos que variam de 37 a 50%

A análise dos gastos em saúde também varia significativamente com o local de consulta dos agregados familiares. Em clínicas especiais e consultórios privados os agregados apresentam 100% de gastos, serviços públicos apresentam 83% de gastos e em serviços de curandeiros apresentam 72% de gastos.

4.2.2 Teste de multicolinearidade

Tabela 4.5: Estatística de colinearidade

Modelo	Tolerance	VIF
Condição social	0.944	1.059
Género	0.982	1.018
Estado civil	0.956	1.046
Empregabilidade	0.966	1.035
Nível académico	0.954	1.049
Província	0.102	9.824
Área de residência	0.882	1.134
Região	0.104	9.631
Número de membros	0.952	1.051
Tipo de casa	0.996	1.004
Número de visitas ao hospital(mensal)	0.996	1.004
Tipo de consulta	0.982	1.018
Tipo alimentação	0.954	1.048
Problema alimentação	0.944	1.059
Af30.	0.426	2.345
Af31.	0.428	2.337
Variável dependente: dep		

Para a realização da regressão logística há pressupostos por verificar, onde da tabela 4.5 acima observa-se que não há multicolinearidade entre as variáveis predictoras dado que os valores de VIF são inferiores que 10 e os valores da tolerância para todas as variáveis predictoras são inferiores que unidade. Deste modo procede-se com análise.

4.2.3 Análise de Associação entre as Variáveis independentes e à variável dependente

Para a análise de existência de relação entre a variável de interesse e as demais variáveis do estudo foi usado o teste Qui-quadrado, cujos resultados encontram-se na tabela 4.6. Consideraram-se como variáveis significativas aquelas que apresentaram P-valor menor que o nível de significância ($\alpha = 5\%$)

Tabela 4.6: Teste de associação entre as variáveis independentes e dependentes

Co-variáveis	χ^2	P-valor
Condição social	42.38	0
Género	3.58	0.58
Estado civil	0.223	0.895
Empregabilidade	0.224	0.894
Nível académico	72.659	0
Província	144.831	0
Área de residência	34.926	0
Região	21.252	0
Número de membros	203.253	0
Tipo de casa	222.369	0
Número de visitas ao hospital(mensal)	40.55	0
Tipo de consulta	26.304	0
Tipo alimentação	102	0
Problema alimentação	1.582	0
Af30.	17.526	0
Af31.	15.574	0

De acordo com o teste de independência do Qui-Quadrado, cujo resultados do mesmo encontram-se ilustrados na tabela 4.6 temos que a um nível de significância de 5%, todas as variáveis independentes estão relacionadas com a variável dependente excepto: género, estado civil e o local de trabalho

4.2.4 Análise de regressão logística

Verificando alguns pressupostos da regressão logística, e os mesmos mostrando que os dados podem ser aplicados na regressão logística, procede-se com a análise.

Tabela 4.7: Modelo estimado de regressão logística sem variáveis independentes

		B	S.E.	Wald	df	Sig.	Exp(B)
Passo 0	Constant	1.573	0.034	2080.688	1	0	4.819

De acordo com a tabela 4.7, verificasse que o modelo logístico sem variáveis independente é estatisticamente significativo apresentando um p-valor menor que o nível de significância de $\alpha = 0.05$. Neste caso o modelo só com a constante pode ser considerado para prever a ocorrência de gastos em saúde num agregado familiar, mas a análise não pode terminar somente com o modelo reduzido, é necessário a inclusão de variáveis independentes para encontrar as demais variáveis que possam possivelmente influenciar a ocorrência dos gastos em saúde pelos agregados familiares.

4.2.5 Análise do ajuste do modelo

Tabela 4.8: Teste de validade do modelo

		Chi-square	df	Sig.
8	Passo	27.391	2	0
	Bloco	635.904	14	0
	Modelo	635.904	14	0

De acordo com a tabela 4.8 esta apresenta o oitavo passo das interações, utilizando diferentes métodos para aplicar o teste qui-quadrado (etapa, bloco, modelo), estas estatísticas servem para indicar a existência de pelo menos uma variável independente significativa no modelo, apesar de não identifica-la.

- ❖ **Hipótese Nula (H_0):** Não há efeito significativo das variáveis predictoras no modelo; ou seja, os coeficientes dos preditores são iguais a zero.
- ❖ **Hipótese Alternativa (H_1):** Existe pelo menos um efeito significativo das variáveis predictoras no modelo; pelo menos um coeficiente dos preditores é diferente de zero.

Com um valor de qui-quadrado de 635.904 e 14 graus de liberdade, e um nível de significância de 0.000, rejeitamos a hipótese nula. Isso significa que há evidências estatísticas suficientes para concluir que pelo menos uma das variáveis predictoras no modelo tem um efeito significativo sobre a variável dependente. Em outras palavras, o modelo como um todo é estatisticamente significativo para explicar a variabilidade nos dados.

Tabela 4.9: Teste de Hosmer e Lemeshow

	Chi-square	df	Sig.
Passo 8	3.086	8	0.929

De acordo com a Tabela 4.9, esta apresenta o teste de Hosmer e Lemeshow com as seguintes hipóteses (Hosmer & Lemeshow, 1989):

- ❖ **Hipótese Nula (H_0):** O modelo se ajusta bem aos dados (bom ajuste).
- ❖ **Hipótese Alternativa (H_1):** O modelo não se ajusta bem aos dados (mau ajuste).

O valor de qui-quadrado obtido é 3.086 com 8 graus de liberdade, e a significância é de 0.929. Isso indica que não há evidências suficientes para rejeitar a hipótese nula. Em outras palavras, com base nos dados e no nível de significância de 5%, o modelo de regressão logística tem um bom ajuste aos dados observados.

4.2.6 Interpretação das Estimativas para a Razão de Chances

Os resultados referentes às razões de chance são apresentados na tabela 4.10 abaixo.

(Ref)-categoria de referência

Em comparação aos agregados pobres os agregados não pobres têm 0.61 vezes menos chances de ter gastos com saúde.

Em comparação com chefes de agregado familiar com ensino superior, chefes de agregado familiar com ensino primário têm 2.92 vezes mais chances de ter gastos com saúde, chefes de agregado familiar com ensino secundário têm 6.24 vezes mais chances de ter gastos com saúde, chefes de agregado familiar com ensino técnico têm 1.61 vezes mais chances de ter gastos com saúde, chefes de agregado familiar que não sabem seu nível de escolaridade têm 0.49 vezes menos chances de ter gastos com saúde, chefes de agregado familiar com apenas educação pré-escolar têm 6.41 vezes mais chances de ter gastos com saúde.

Em comparação com os agregados de áreas rurais, agregados residentes de áreas urbanas têm 0.38 vezes menos chances de ter gastos com saúde.

A cada aumento no número de membros do agregado familiar aumenta as chances do mesmo agregado ter gastos com saúde em cerca de 1.09 vezes.

Em comparação aos residentes em casas própria, residentes em casas arrendadas têm 3.39 vezes mais chances de ter gastos com saúde, residentes em casas cedidas têm 2.27 vezes mais chances

Em comparação com os agregados familiares que utilizamos serviços de curandeiros, agregados familiares que fazem consultas em clínicas privadas têm 0.004 vezes menos chances de ter gastos com saúde e agregados familiares que fazem consultas em serviços públicos têm 0.003 vezes menos chances de ter gastos com saúde.

Em comparação com os agregados que não tiveram problemas de alimentação, agregados familiares que não tiveram problemas de alimentação têm 0.64 vezes menos chances de ter gastos com saúde.

Tabela 4.10: Modelo final

	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I.for EXP(B)	
							Lower	Upper
Condição social								
Agregado não pobre	-0.498	0.079	39.421	1	0	0.608	0.521	0.71
Agregado pobre (<i>Ref</i>)	-	-	-	-	-	-	-	-
Nível académico (Chefe do agregado familiar)								
Ensino primário	1.072	0.152	50.003	1	0	2.921	2.17	3.931
Ensino secundário	1.83	0.723	6.412	1	0.011	6.237	1.512	25.72
Ensino técnico	0.477	0.232	4.224	1	0.04	1.611	1.022	2.539
Não sabe	-0.719	0.095	57.226	1	0	0.487	0.404	0.587
Nunca frequentou escola	0.697	0.549	1.61	1	0.204	2.007	0.684	5.889
Pré-escolar	1.877	0.591	9.887	1	0.002	6.405	2.013	20.38
Superior (<i>Ref</i>)	-	-	-	-	-	-	-	-
Área de residência								
Urbano	-0.969	0.079	152.036	1	0	0.379	0.325	0.443

Rural (<i>Ref</i>)	-	-	-	-	-	-	-	-
Número de membros	0.086684	0.016741	26.81019	1	2.24E-07	1.090552	1.0553491	1.1269289
Tipo de casa								
Arrendada	1.221939	0.210791	33.60423	1	6.75E-09	3.393762	2.2452025	5.1298792
Cedida	0.818214	0.14767	30.70063	1	3.01E-08	2.266448	1.69687	3.027213
própria (<i>Ref</i>)	-	-	-	-	-	-	-	-
Tipo de consulta								
Consultas em clínicas	-5.461484	4.003511	1.86097	1	0.172513	0.004247	1.661E-06	10.861815
Consultas em serviços públicos	-5.837061	4.008164	2.120789	1	0.145312	0.002917	1.13E-06	7.529254
Serviços de curandeiros (<i>Ref</i>)	-	-	-	-	-	-	-	-
Problema alimentação por falta de dinheiro ou outros recursos								
Não teve problemas de alimentação	-0.453549	0.10881	17.37437	1	3.07E-05	0.635369	0.5133422	0.7864039
Teve problemas de alimentação (<i>Ref</i>)	-	-	-	-	-	-	-	-
Intercepto	6.833052	4.008461	2.905856	1	0.088258	928.0187		

Validação do modelo

Para a validação do modelo analisou-se o seu poder discriminatório, a sensibilidade, a especificidade e a taxa de acertos. Para tal usou-se a matriz de confusão, gráfico da curva ROC e o valor da área sob a curva.

Tabela 4.11: Tabela de classificação

Observado		Previsto		Percentagem correta	
		Gastos em saúde			
		Não	Sim		
Passo 8	Gastos em saúde	Não	73	943	7.2
		Sim	66	4830	98.7
Percentagem global				82.9	

De acordo com a tabela 4.11, o modelo classifica de forma correcta 82.9% dos casos. Portanto, é possível perceber que o modelo classificou corretamente 4830 dos 4896 agregados que tiveram gastos em saúde, ou seja, uma especificidade de 98,7%, o modelo de regressão logística possui uma alta precisão ao prever gastos em saúde (classe "Sim"), o modelo classificou de forma correta 73 dos 1016 agregados dos agregados que não tenham gastos em saúde, ou seja, uma sensibilidade de 7.2%.

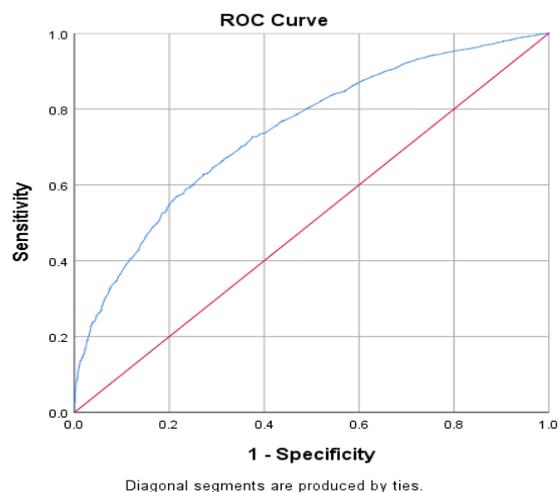


Gráfico 4.8: Curva Roc

Tabela 4.12: Área sob a curva Roc e o respectivo IC

Área	Erro	Assíptota	Intervalo de confiança a 95%	
			Limite inferior	Limite superior
0.741	0.008	0	0.725	0.757

De acordo com a gráfico 4.8, ao examinar-se a curva ROC, observa-se que a curva azul se afasta significativamente da linha diagonal vermelha, que representa um modelo de classificação aleatória. Este comportamento indica que o modelo possui um bom desempenho na distinção entre as classes positivas e negativas. A inclinação da curva em direção ao canto superior esquerdo do gráfico reforça essa observação, mostrando uma alta taxa de verdadeiros positivos em comparação com os falsos positivos.

De acordo com a tabela 4.12, quanto ao desempenho global com a AUC de 0.741 indicando que o modelo tem uma boa capacidade de discriminar entre as classes positivas e negativas. Isso significa que, em 74.1% das vezes, o modelo classificará corretamente uma instância positiva como mais provável do que uma instância negativa.

Quanto a precisão e consistência: O intervalo de confiança de 95% (0.725 a 0.757) é relativamente estreito, o que sugere que a estimativa da AUC é precisa e confiável. Esse intervalo estreito indica que o desempenho do modelo é consistente e que a AUC não varia significativamente com pequenas mudanças nos dados.

4.3 Discussão dos resultados

A amostra estudada, é composta por 15.346 agregados familiares, apresenta uma composição demográfica e socioeconômica diversificada. Predominantemente, os agregados residem em áreas urbanas, refletindo a intensificação do processo de urbanização observado em muitos países africanos contemporâneos. A distribuição econômica mostra que cerca de 70% dos agregados são categorizados como "não pobres", enquanto cerca de 30% são classificados como pobres, sublinhando a existência de disparidades econômicas dentro do grupo estudado. Em relação à educação, a maioria dos chefes de agregados possui apenas educação primária, seguida por uma parcela menor com formação secundária, com níveis ainda mais reduzidos de educação técnica e superior, o que aponta para desafios persistentes no acesso a níveis educacionais mais elevados.

Os resultados indicam que os gastos alimentares dominam o orçamento dos agregados familiares em Moçambique, especialmente nas áreas rurais, onde representam 75.9% dos gastos, em comparação com 62.14% nas áreas urbanas. Esse padrão reflete a menor disponibilidade e acessibilidade a bens não alimentares e serviços de saúde nas zonas rurais. A análise por sexo revela uma distribuição equilibrada dos gastos entre agregados chefiados por homens e mulheres, com gastos alimentares ligeiramente superiores entre as mulheres (68.08% versus 67.76%). Os gastos com saúde são uniformemente baixos para ambos sexos (0.24%).

Agregados familiares com gastos catastróficos em saúde gastam 5.1% do seu rendimento em saúde comparativamente aos agregados sem gastos em saúde que gastam apenas 0.07% de sua renda em gastos de saúde, os agregados com gastos catastróficos apresentaram maior percentagem de gastos com alimentação. Esse cenário vai de encontro com estudos como o de Almeida et al. (2015), que demonstraram que gastos catastróficos em saúde podem levar à pobreza, enfatizando a necessidade de políticas públicas que protejam os mais vulneráveis contra os choques financeiros decorrentes de problemas de saúde.

Aproximadamente 2% dos agregados familiares em Moçambique enfrentam gastos catastróficos em saúde. Esta proporção está alinhada com estudos em países vizinhos da região, como Tanzânia e Malawi, onde também foram observadas incidências significativas de gastos catastróficos, variando entre 1% a 3% (Kamala et al., 2019). Esses números refletem uma tendência regional de vulnerabilidade financeira associada a emergências de saúde, impactando desproporcionalmente famílias de baixa renda. Importa destacar que a esta percentagem de agregados familiares acresce-se outra cuja seu gasto em saúde é subsidiado pelo estado e por outras organizações que financiam os serviços de saúde e outros medicamentos de alto custo.

Em comparação aos agregados familiares pobres, os agregados não pobres têm 0,61 vezes menos chances de ter gastos com saúde. Este resultado sugere que a condição econômica influencia significativamente os gastos com saúde, provavelmente devido ao acesso diferenciado a serviços de saúde e à capacidade financeira para arcar com esses custos. De acordo com o e estudos feito por (Wagstaff & van Doorslaer, 2003) em África a pobreza está associada a um maior risco de gastos com saúde, uma vez que os agregados familiares pobres tendem a gastar uma maior proporção de sua renda em cuidados de saúde.

De acordo com o estudo realizado na Nigéria por Adisa (2015) revelou que a escolaridade dos chefes de família está positivamente associada aos gastos com saúde. Famílias com chefes de

família mais educados têm maior probabilidade de buscar cuidados de saúde formais e, conseqüentemente, gastar mais com saúde. Este achado é consistente com os resultados deste estudo, que mostra que chefes de agregados familiares com ensino secundário têm 6,24 vezes mais chances de ter gastos com saúde comparados aos com ensino superior, e chefes com ensino primário têm 2,92 vezes mais chances. No entanto, no estudo de referência, chefes com ensino técnico têm menos chances de gastar com saúde em comparação com aqueles com ensino superior, o que pode não ser observado em todos os contextos africanos.

De acordo com o estudo realizado na Etiópia e no Quênia por (Sepehri & Chernomas, 2001; Mugo & Orago, 2013), indicam que famílias urbanas tendem a gastar mais com saúde devido à maior acessibilidade a serviços de saúde e à maior conscientização sobre a importância dos cuidados de saúde. Em contraste, com este estudo que mostra que agregados familiares urbanos têm 0,38 vezes menos chances de ter gastos com saúde comparados aos de áreas rurais. Isso pode ser explicado por diferenças nos sistemas de saúde urbana e rural, onde os serviços urbanos podem ser mais acessíveis e baratos, ou onde há maior cobertura de seguros de saúde.

De acordo com o estudo realizado por Burger e Christian (2018) descobriram que na África do Sul, o aumento no tamanho do agregado familiar está associado a maiores gastos de saúde, especialmente em famílias com crianças pequenas e idosos. Este achado está em consonância com este estudo, onde cada aumento no número de membros do agregado familiar moçambicano aumenta as chances do mesmo agregado ter gastos com saúde em cerca de 1,09 vezes.

De acordo com o estudo realizado na Tanzânia por Mtei e Mulligan (2007) mostrou que agregados familiares que residem em habitações arrendadas têm maiores gastos com saúde, possivelmente devido à falta de estabilidade econômica. De maneira similar, neste estudo verificasse que residentes em casas arrendadas têm 3,39 vezes mais chances de ter gastos com saúde comparados aos em casas próprias. Este achado sugere uma possível correlação entre insegurança habitacional e saúde.

Dalinjong e Laar (2012) observaram em Gana que famílias que utilizam serviços de saúde privados tendem a gastar mais com saúde do que aquelas que utilizam serviços públicos ou tradicionais. No entanto, este estudo revela que agregados familiares que fazem consultas em clínicas privadas têm 0,004 vezes menos chances de ter gastos com saúde comparados aos que

utilizam serviços de curandeiros. Este achado pode refletir diferenças nos custos percebidos ou nas políticas de pagamento.

De acordo com o estudo realizado no Uganda por Ndugga e Namiyonga (2017) correlacionaram a insegurança alimentar com maiores gastos de saúde, pois a desnutrição aumenta a vulnerabilidade a doenças. Este achado é consistente com este estudo, onde agregados familiares que não tiveram problemas de alimentação têm 0,64 vezes menos chances de ter gastos com saúde comparados aos que tiveram problemas de alimentação, reforçando a ligação entre nutrição e saúde.

Capítulo 5

5. CONCLUSÃO E RECOMENDAÇÕES

5.1 Conclusões

O objetivo geral deste estudo foi analisar os principais determinantes dos gastos em saúde dos agregados familiares em Moçambique. A amostra, composta por 15.346 agregados familiares, revelou que a maioria destes reside em áreas urbanas, é chefiada por mulheres, apresenta estado civil solteiro e possui um nível de escolaridade reduzido. Foram identificados três grandes tipos de gastos: alimentares, não alimentares e de saúde. No caso dos gastos em saúde, estas foram subdivididas em catastróficas e não catastróficas. Observou-se que apenas uma pequena parcela dos agregados enfrenta gastos considerados catastróficos, predominando entre as famílias mais pobres e com níveis de escolaridade mais baixos.

A análise das associações revelou que os gastos em saúde estão fortemente relacionados com várias características dos agregados familiares, como a condição social, o nível académico, a província, a área de residência, a região, o número de membros do agregado familiar, o tipo de casa, o número de visitas ao hospital, o tipo de consulta médica, tipo de alimentação e problemas alimentares. Por outro lado, não se verificou uma relação significativa com o sexo do chefe do agregado familiar, o estado civil e o estado de empregabilidade.

A aplicação do modelo de regressão logística permitiu identificar os principais determinantes dos gastos em saúde dos agregados familiares em Moçambique. Entre eles destacam-se a condição social, em que famílias de baixo rendimento têm maiores chances de ter gastos com saúde; o nível académico, onde agregados familiares com chefes com menor nível de escolaridade têm maiores chances de ter gastos em saúde; o número de membros do agregado familiar, em que quanto maior o número do agregado familiar maior é a chance do agregado ter gastos em saúde. Também o tipo de residência se revelou determinante, em que agregados que residem em casa arrendadas e cedidas tem maiores chances de ter gastos em saúde. Além disso, os problemas alimentares, em que agregados familiares que tem problemas alimentares tem maiores chances de terem gastos em saúde.

5.2 Recomendações

- ❖ Considerando que a condição social e o nível de escolaridade foram identificados como determinantes dos gastos em saúde, recomenda-se a promoção da equidade social e educacional através de programas governamentais e ONGs que visem melhorar o acesso à educação e aumentar o poder aquisitivo das famílias.
- ❖ Considerando que número de membros do agregado familiar foi identificado como um determinante dos gastos em saúde, recomenda-se a Implementação de programas de apoio às famílias numerosas, considerando que essas tendem a ter maiores gastos em saúde.
- ❖ Considerando que o tipo de residência e problemas alimentares foram identificados como determinantes dos gastos em saúde, recomenda-se a melhoria das condições habitacionais e de financiamento a habitação e a criação de políticas de segurança alimentar que garantam uma dieta adequada e acessível, prevenindo problemas de saúde.

Limitações

Este capítulo tem como objetivo apresentar as principais limitações encontradas ao longo da realização do estudo. A limitação identificada foi:

- ❖ Carência de estudos nacionais que abordam sobre os factores determinantes a ocorrências de gastos em saúde.

Referências

- ❖ Adisa, O. (2015). Investigating determinants of catastrophic health spending among poorly insured elderly households in urban Nigeria. *International Journal for Equity in Health*, 14(1), 79.
- ❖ Agência dos Estados Unidos para o Desenvolvimento Internacional (USAID). (2022). *Health Financing Profile: Mozambique*.
- ❖ Agresti, A., & Finlay, B. (2008). *Statistical Methods for the Social Sciences*. Pearson.
- ❖ Almeida, J., et al. (2015). Distribuição regional dos gastos em saúde. *Revista de Saúde Pública*, 49(3), 211-220.
- ❖ Almeida, R., & Pereira, M. (2020). Impacto do estado civil do chefe do agregado familiar nos gastos com saúde: Uma análise comparativa. *Revista de Economia da Saúde*, 15(2), 145-163.
- ❖ Asian Development Bank. (2020). *Education, health, and development in Asia*.
- ❖ Banco Mundial. (2021). *Relatório sobre o financiamento da saúde em Moçambique*.
- ❖ Banco Mundial. (2021). *Mozambique Economic Outlook*. Washington, DC: World Bank Group.
- ❖ Banco Mundial. (2022). *Atualização econômica de Moçambique*.
- ❖ Bloom, D. E., Cafiero, E. T., Jané-Llopis, E., Abrahams-Gessel, S., Bloom, L. R., Fathima, S., ... & O'Farrell, L. (2011). *The global economic burden of noncommunicable diseases*. Geneva: World Economic Forum.
- ❖ Burger, R., & Christian, C. (2018). Access to health care in post-apartheid South Africa: Availability, affordability, acceptability. *Health Economics, Policy and Law*, 13(1), 1-15.
- ❖ Canton, A. W. P. (1988). *Aplicação de modelos estatísticos na avaliação de produtos*. Tese de Livre-Docência, Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, São Paulo, 1988.

- ❖ Costa, R., et al. (2018). Diferenças de gênero e estado civil nos gastos em saúde. *Journal of Health Economics*, 27(2), 134-145.
- ❖ Dalinjong, P. A., & Laar, A. S. (2012). The National Health Insurance Scheme: Perceptions and experiences of health care providers and clients in two districts of Ghana. *Health Economics Review*, 2(1), 13.
- ❖ Dardot, P., & Laval, C. (2016). *A nova razão do mundo: Ensaio sobre a sociedade neoliberal*. Boitempo Editorial.
- ❖ Desen, A. (1999). *Development as freedom*. Oxford University Press.
- ❖ Evans, D. B., Hsu, J., & Boerma, T. (2013). Cobertura de saúde universal e acesso universal. *Boletim da Organização Mundial da Saúde*, 91(8), 546-546A.
- ❖ FAO. (2020). *Food security and health expenditures: An integrated analysis*.
- ❖ Garcia, C., et al. (2020). Financial burden of health costs on larger household sizes. *Health Services Research*, 15(3), 321-335.
- ❖ Gil, A. C. (2008). *Métodos e técnicas de pesquisa social* (6ª ed.). São Paulo: Atlas.
- ❖ Hosmer, D., & Lemeshow, S. (1989). *Applied logistic regression*. New York: John Wiley & Sons.
- ❖ IMF. (2018). *Mozambique: Poverty reduction strategy paper*. Washington, DC: International Monetary Fund.
- ❖ IMF. (2019). *Mozambique: Poverty reduction strategy paper*. Washington, DC: International Monetary Fund.
- ❖ Instituto Nacional de Estatística de Moçambique. (2018). *Inquérito de Indicadores de Imunização, Malária e HIV/SIDA em Moçambique 2015*.
- ❖ Instituto Nacional de Estatística de Moçambique. (2021). *Inquérito Demográfico e de Saúde de Moçambique*.
- ❖ Instituto Nacional de Estatística de Moçambique. (2022). *Inquérito Sobre Orçamento Familiar*.
- ❖ Instituto Nacional de Estatística de Moçambique. (2022). *Pesquisa Demográfica e de Saúde*.

- ❖ James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer.
- ❖ Kamala, B., et al. (2019). Catastrophic health expenditure among households in Malawi: Determinants and implications for equity in access to healthcare. *BMC Health Services Research*, 19(1), 1-11.
- ❖ Kruk, M. E., Goldmann, E., & Galea, S. (2009). Borrowing and selling to pay for health care in low-and middle-income countries. *Health Affairs*, 28(4), 1056-1066.
- ❖ Kurowski, C., Wyss, K., Abdulla, S., & Mills, A. (2004). Scaling up priority health interventions in Tanzania: The human resources challenge. *Health Policy and Planning*, 19(1), 1-5.
- ❖ Loquia, O., & Abdulrazac, R. (2021). *Estatística multivariada: Análise factorial*. Maputo, Moçambique.
- ❖ Macha, J., et al. (2017). Health financing and catastrophic payments for health care: Evidence from Tanzania. *Health Policy and Planning*, 32(8), 1106-1115.
- ❖ Maier, R. C. (2012). *Análise das relações existentes entre qualidade de vida e qualidade de vida no trabalho através de um modelo de regressão logística*. Ponta Grossa, Brasil.
- ❖ Maroco, J. (2007). *Análise estatística com utilização do SPSS* (3ª ed.). Edições Sílabo.
- ❖ Menard, S. (2002). *Applied logistic regression analysis*. Sage Publications.
- ❖ Mendes, L., et al. (2014). Impacto do nível educacional no uso de serviços de saúde. *Revista Brasileira de Educação Médica*, 38(4), 573-580.
- ❖ Ministério da Saúde de Moçambique. (2015). *Moçambique contas nacionais de saúde (CNS) 2015*.
- ❖ Ministério da Saúde de Moçambique. (2020). *Relatório anual do Ministério da Saúde de Moçambique*.
- ❖ Ministry of Health (Mozambique). (2022). *Health sector strategic plan 2014-2019*. Maputo, Mozambique: Ministry of Health.

- ❖ Mtei, G., & Mulligan, J. A. (2007). Community health funds in Tanzania: A literature review. *Ifakara Health Research and Development Centre*.
- ❖ Mugo, M. G., & Orago, A. (2013). Utilization of health care services among urban slum residents in Kenya. *Public Health Research*, 3(6), 192-198.
- ❖ Moura, G. M. (2018). *Regressão logística aplicada à análise de risco de crédito*. Brasil.
- ❖ Ndege, R. (2007). Normas sociais e práticas entre grupos étnicos em Moçambique: Implicações para programas de desenvolvimento. *Journal of African Studies*, 12(3), 45-61.
- ❖ Neter, J., Kutner, M. H., Nachsheim, C. J., & Wasserman, W. (1996). *Applied linear statistical models*. Chicago: Irwin.
- ❖ Nascimento, F. P. (2016). *Classificação da pesquisa. Natureza, método ou abordagem metodológica, objetivos e procedimentos*. Brasília, Brasil.
- ❖ Oliveira, A., et al. (2015). Acesso a serviços de saúde e tipo de habitação: Uma análise de dados populacionais. *Cadernos de Saúde Pública*, 20(2), 185-200.
- ❖ Oliveira, R., & Silva, M. (2018). Problemas de alimentação por falta de recursos: Uma revisão da literatura. *Revista de Saúde Pública*, 15(2), 245-260.
- ❖ Organização Mundial da Saúde (OMS). (2000). *Relatório mundial de saúde 2000: Sistemas de saúde: Melhorando o desempenho*.
- ❖ Organização Mundial da Saúde. (2005). *Distribuição dos pagamentos de saúde e despesas catastróficas*. Genebra, Suíça.
- ❖ Organização Mundial da Saúde. (2008). *Negócio de todos: Fortalecendo os sistemas de saúde para melhorar os resultados de saúde: Estrutura da OMS para ação*.
- ❖ Organização Mundial da Saúde. (2010). *Financiamento de sistemas de saúde: O caminho para a cobertura universal*. Organização Mundial da Saúde.
- ❖ Organização Mundial da Saúde. (2021). *Mozambique: Health Systems Strengthening*. Geneva: World Health Organization.
- ❖ Ribeiro, A. P., et al. (2014). Morbidade por doenças crônicas e gasto com saúde em Moçambique. *Revista Brasileira de Epidemiologia*, 21(2), 345-357.

- ❖ Sanhudo, P., & Mendes, J. (2017). Educação, estado civil e o impacto nos gastos com saúde: Análise de dados moçambicanos. *Revista de Estudos da Saúde*, 19(1), 57-72.
- ❖ Santos, M. S. (2019). *Modelagem de equações estruturais*. Rio de Janeiro: Elsevier.
- ❖ Sapkota, V. P., & Houser, R. F. (2015). Association between childhood malnutrition and household food insecurity in Nepal. *Maternal & Child Nutrition*, 11(1), 58-69.
- ❖ Shrimme, M. G., et al. (2015). Catastrophic expenditure to pay for surgery worldwide: A modelling study. *The Lancet Global Health*, 3(4), e34.
- ❖ Silva, C. A., & Braga, D. G. (2017). Estudo da relação entre escolaridade e a renda no Brasil: Uma análise com a utilização de modelos de regressão logística. *Revista Brasileira de Economia*, 22(3), 477-498.
- ❖ Srivastava, D., & McGuire, A. (2016). Patient access to health care and medicines across low-income countries. *Social Science & Medicine*, 152, 152-163.
- ❖ Tanser, F., Gijsbertsen, B., & Herbst, K. (2006). Modelling and understanding primary health care accessibility and utilization in rural South Africa: An exploration using a geographical information system. *Social Science & Medicine*, 63(3), 691-705.
- ❖ UNICEF. (2016). *Situação da criança em Moçambique 2016*.
- ❖ UNICEF. (2021). *Relatório de saúde e nutrição de Moçambique*.
- ❖ Van Doorslaer, E., O'Donnell, O., Rannan-Eliya, R. P., Somanathan, A., Adhikari, S. R., Garg, C. C., ... & Zhao, Y. (2006). Effect of payments for health care on poverty estimates in 11 countries in Asia: An analysis of household survey data. *The Lancet*, 368(9544), 1357-1364.
- ❖ Van Minh, H., Phuong, N. T. K., Saksena, P., James, C. D., & Xu, K. (2013). Financial burden of household out-of-pocket health expenditure in Viet Nam: Findings from the National Living Standard Survey 2002-2010. *Social Science & Medicine*, 96, 258-263.
- ❖ Wagstaff, A., Flores, G., Hsu, J., Smits, M., Chepynoga, K., Buisman, L. R., ... & Eozenou, P. (2018). Progress on catastrophic health spending in 133 countries: A retrospective observational study. *The Lancet Global Health*, 6(2), e169-e179.

- ❖ Xu, K., Evans, D. B., Kawabata, K., Zeramdini, R., Klavus, J., & Murray, C. J. L. (2003). Household catastrophic health expenditure: A multicountry analysis. *The Lancet*, 362(9378), 111-117.
- ❖ Xu, K., Evans, D. B., Carrin, G., Aguilar-Rivera, A. M., Musgrove, P., & Evans, T. (2007). Protecting households from catastrophic health spending. *Health Affairs*, 26(4), 972-983.

Apêndices

Tabela 4.1: Distribuição das frequências absolutas e relativas em relação entre as despesas e o total em média

Variáveis	Em relação entre as despesas e o total em média		
	Despesas alimentares	Despesas não alimentares	Despesas com saúde
Província de residência			
Niassa	78.87%	20.94%	0.19%
Cabo delgado	74.75%	24.88%	0.36%
Nampula	76.89%	22.84%	0.27%
Zambézia	72.13%	27.76%	0.12%
Tete	76.07%	23.63%	0.30%
Manica	70.30%	29.52%	0.18%
Sofala	69.66%	30.24%	0.10%
Inhambane	67.98%	31.92%	0.10%
Gaza	64.89%	34.89%	0.22%
Província de Maputo	54.48%	45.35%	0.17%
Cidade de Maputo	50.40%	49.03%	0.56%
Área de residência			
Urbano	62.14%	37.59%	0.27%
Rural	75.90%	23.90%	0.20%
Género			
Masculino	67.76%	32%	0.24%
Feminino	68.08%	32%	0.24%
Estado civil chefe do agregado familiar			
Casado(a)	67.95%	31.81%	0.24%
Recusou	68.05%	31.70%	0.25%
Solteiro(a)	67.78%	32.00%	0.22%
Condição social			
Agregado não pobre	66.65%	33.19%	0.16%
Agregado pobre	71.08%	28.49%	0.44%
Tipo de gastos em saúde			
Agregado sem gastos catastróficos em saúde	67.52%	32.41%	0.07%
Agregados com gastos catastróficos em saúde	79.01%	16.37%	5.1%

Tabela 4.2: Distribuição das frequências absolutas e relativas de acordo com a ocorrência de gastos catastróficos em saúde

Variável	Gastos catastróficos	
	Agregado sem gastos catastróficos em saúde	Agregado com gastos catastróficos em saúde
Província		
Niassa	97%	3%
Cabo delgado	95%	5%
Nampula	98%	2%
Zambézia	100%	0%
Tete	99%	1%
Manica	99%	1%
Sofala	99%	1%
Inhambane	97%	3%
Gaza	99%	1%
Província de Maputo	99%	1%
Cidade de Maputo	99%	1%
Área de residência		
Urbano	99%	1%
Rural	98%	2%
Género		
Masculino	98%	2%
Feminino	99%	1%
Estado civil chefe do agregado familiar		
Casado(a)	99%	1%
Recusou	99%	1%
Solteiro(a)	98%	2%
Nível académico		
Ensino primário	99%	1%
Ensino secundário	99%	1%
Ensino técnico	100%	0%
Não sabe	97%	3%
Nunca frequentou escola	93%	7%
Pré-escolar	96%	4%
Superior	99%	1%
Condição social		
Agregado não pobre	100%	0%
Agregado pobre	95%	5%

Região		
Região sul	99%	1%
Região centro	99%	1%
Região norte	97%	3%
Tipo de casa		
Arrendada	99%	1%
Cedida	99%	1%
Própria	98%	2%
Tipo alimentação		
Adequada	99%	1%
Mais que adequada	98%	2%
Menos adequada	98%	2%
Problema alimentar		
Não teve problemas de alimentação	98%	2%
Teve problemas de alimentação	99%	1%
Local de consulta		
Consultas em clínicas especiais e consultórios privados excepto dentários	93%	7%
Consultas em serviços públicos para doentes sem internamento	98%	2%
Serviços de curandeiros	93%	7%
Af30. Nos últimos 12 meses, alguma vez o sr.(a) já ficou sem comida em sua casa por falta de dinheiro ou outros meios?		
Não	99%	1%
Não sabe	100%	0%
Sim	98%	2%
Af31. Nos últimos 12 meses, alguma vez o sr(a), ou outro membro deste agregado familiar, sentiu fome, mas não comeu, porque não tinha dinheiro ou outros meios para obter comida?		
Não	99%	1%
Não sabe	97%	3%
Sim	98%	2%
Número de visitas hospital		
0 a 2 visitas	98%	2%
3 a 5 visitas	97%	3%

Mais de 6 visitas	100%	0%
Número de pessoas no agregado		
1 a 5 membros	98%	2%
6 a 10 membros	99%	1%
11 a 15 membros	100%	0%
Mais de 16 membros	100%	0%
Media percentual	98%	2%

Tabela 4.3: Distribuição das frequências absolutas e relativas de acordo com a ocorrências de gastos em saúde

Província	Sem gastos	%sem gastos	Com gastos	%com gastos
Niassa	662	59%	469	42%
Cabo delgado	813	75%	266	25%
Nampula	1184	71%	480	29%
Zambézia	1061	61%	693	40%
Tete	859	69%	390	31%
Manica	845	69%	376	31%
Sofala	999	67%	486	33%
Inhambane	832	72%	329	28%
Gaza	782	67%	379	33%
Província de maputo	1098	71%	439	29%
Cidade de maputo	1315	69%	589	31%
Área de residência				
Urbano	5875	66%	3000	34%
Rural	4575	71%	1896	29%
Género				
Masculino	4664	67%	2265	33%
Feminino	5786	69%	2631	31%
Estado civil chefe do agregado familiar				
Casado(a)	2476	68%	1171	32%
Recusou	4647	68%	2184	32%
Solteiro(a)	3327	68%	1541	32%
Nível académico				
Ensino primário	6967	69%	3116	31%
Ensino secundário	1797	66%	921	34%

Ensino técnico	117	61%	74	39%
Não sabe	183	52%	171	48%
Nunca frequentou escola	1107	71%	445	29%
Pré-escolar	74	66%	39	35%
Superior	205	61%	130	39%

Condição social

Agregado não pobre	7236	67%	3641	34%
Agregado pobre	3214	72%	1255	28%

Região

Região sul	4027	70%	1736	30%
Região centro	3764	66%	1945	34%
Região norte	2659	69%	1215	31%

Tipo de casa

Arrendada	1163	86%	189	14%
Cedida	802	69%	366	31%
Própria	8485	66%	4341	34%

Tipo de alimentação

Adequada	5849	72%	2320	28%
Mais que adequada	125	71%	52	29%
Menos adequada	4476	64%	2524	36%

Problema alimentar

Não teve problemas de alimentação	1995	69%	893	31%
Teve problemas de alimentação	8455	68%	4003	32%

Local de consulta

Consultas em clínicas especiais e consultórios privados excepto dentários	0	0%	72	100%
Consultas em serviços públicos para doentes sem internamento, excepto dentários	976	17%	4720	83%
Serviços de curandeiros	40	28%	104	72%

Af30. Nos últimos 12 meses, alguma vez o sr(a) já ficou sem comida em sua casa por falta de dinheiro ou outros meios?

Não	5197	70%	2266	30%
Não sabe	29	76%	9	24%
Sim	5224	67%	2621	33%

Af31. Nos últimos 12 meses, alguma vez o sr(a), ou outro membro deste agregado familiar, sentiu fome, mas não comeu, porque não tinha dinheiro ou outros meios para obter comida?

Não	5197	70%	2266	30%
-----	------	-----	------	-----

Não sabe	29	59%	9	41%
Sim	5224	67%	2621	33%
Número de visitas hospital				
0 a 2 visitas	10362	68%	4814	32%
3 a 5 visitas	74	51%	72	49%
Mais de 6 visitas	14	58%	10	42%
Número de pessoas no agregado				
1 a 5 membros	7436	70%	3137	30%
6 a 10 membros	2849	63%	1648	37%
11 a 15 membros	158	60%	104	40%
Mais de 16 membros	7	50%	7	50%

Tabela 4.4: Distribuição da média das despesas de subsistência dos agregados familiares

Variáveis	Gastos medios de subsistência dos agregados (Mts)
Província de residência	
Niassa	30463
Cabo delgado	30217
Nampula	29474
Zambézia	29600
Tete	29278
Manica	30522
Sofala	29736
Inhambane	27644
Gaza	29012
Província de Maputo	28341
Cidade de Maputo	27585
Área de residência	
Urbano	29313
Rural	29009
Género	
Masculino	29325
Feminino	29070
Estado civil chefe do agregado familiar	
Casado(a)	29023
Recusou	29335

Solteiro(a)	29096
Condição social	
Agregado não pobre	28884
Agregado pobre	29918
Tipo de gastos em saúde	
Agregado sem gastos catastróficos em saúde	29269
Agregados com gastos catastróficos em saúde	27000
Media total	29199

Tabela 4.13: Tamanho da amostra de UPAs planificada, agregados familiares por área de residência e província.

Província	Amostra em UPAs			Nº Esperado de Agregados familiares		
	Urbano	Rural	Total	Urbano	Rural	Total
Total	756	740	1 496	9 072	6 660	15 732
Niassa	48	72	120	576	648	1 224
Cabo Delgado	44	72	116	528	648	1 176
Nampula	72	100	172	864	900	1 764
Zambézia	64	112	176	768	1 008	1 776
Tete	52	72	124	624	648	1 272
Manica	52	68	120	624	612	1 236
Sofala	76	64	140	912	576	1 488
Inhambane	52	60	112	624	540	1 164
Gaza	52	60	112	624	540	1 164
Maputo	84	60	144	1 008	540	1 548
Cidade de Maputo	160	-	160	1 920	-	1 920

Fonte: INE