



FACULDADE DE CIÊNCIAS
Departamento de Matemática e Informática

Trabalho de Licenciatura em
Estatística

**Análise Dos Factores Determinantes
à Ocorrência da Gravidez
Em Adolescentes Dos 10-19 anos no Norte de
Moçambique**

Autora: Alzira João Timane

Maputo, Março de 2025



FACULDADE DE CIÊNCIAS
Departamento de Matemática e Informática

Trabalho de Licenciatura em
Estatística

**Análise Dos Factores Determinantes
à Ocorrência da Gravidez
Em Adolescentes Dos 10-19 anos no Norte de
Moçambique**

Autor: Alzira João Timane

Supervisor: Zacarias Mutombene, Msc, UP

Maputo, Março de 2025

Dedicatória

Dedico este trabalho a minha filha por ser alguém muito especial para mim, mesmo que não tenha conhecimento do trabalho, mas me inspirei nela para a sua realização.

Declaração de Honra

Declaro por minha honra que sou a autora deste trabalho de licenciatura e que o processo foi concebido para ser submetido apenas para a obtenção do grau de **Licenciado em Estatística**, na Faculdade de Ciências da Universidade Eduardo Mondlane.

Maputo, Março de 2025

(Alzira João Timane)

Agradecimentos

Em primeiro lugar, quero expressar minha mais profunda gratidão a Deus, cuja luz guiou cada passo deste caminho desafiador. Sem sua orientação e graça, esta conquista não teria sido possível.

À minha família, por estar comigo em todos os momentos em oração, ainda que distantes. Suas palavras e gestos de amor e carinho foram, como em todas as outras, meus alicerces nessa etapa tão especial da minha vida, minha gratidão pelo incentivo, por compartilharem comigo os momentos de tristeza e também de alegria, que com a graça de Deus, já foram vencidos.

Agradeço especialmente ao meu supervisor **Mestre Zacarias Mutombene**, pela paciência, experiência e pelos ensinamentos transmitidos durante a formação assim como na realização deste trabalho. Aos meus queridos amigos, companheiros da escola e irmãos na amizade, que estiveram ao meu lado durante todo este percurso, compartilhando tantos momentos de alegria quanto os de tristeza. Suas palavras de encorajamento foram indispensáveis para a minha motivação e perseverança.

Agradeço também a todos docentes, orientadores e demais profissionais que contribuíram para o meu desenvolvimento acadêmico, dedicando seu tempo, conhecimento e experiência para me orientar e inspirar.

A todos que não pude citar, meus agradecimentos.

Lista de Abreviaturas

GA	Gravidez na adolescência
IMASIDA	Inquérito de Imunização, Malária e HIV/SIDA
MISAU	Ministério da Saúde
ONU	Organização das Nações Unidas
OR	<i>Odds Ratio</i>
IDS	Inquérito Demográfico e de Saúde
MC	Métodos Contraceptivos
SPSS	<i>Statistical Package For the Social Sciences</i>
UNICEF	Fundo das Nações Unidas para a Infância
VIF	Factor de Inflação da Variância

Resumo

A gravidez na adolescência é considerada uma situação de risco biológico tanto para as adolescentes como para os recém-nascidos. Trata-se de um dos problemas que atentam a saúde pública da sociedade em geral e carece de atenção especial das entidades responsáveis para a provisão de soluções eficazes para a sua redução. O presente estudo consiste na Análise dos factores determinantes à ocorrência da gravidez em adolescentes dos 10-19 anos no norte de Moçambique, importa referir que este é de extrema importância para a sociedade, pois fornece dados relevantes sobre os factores determinantes para a prevalência da gravidez na adolescência no Norte de Moçambique. O objectivo principal deste trabalho foi analisar os factores determinantes à ocorrência da gravidez em adolescentes dos 10-19 anos no norte de Moçambique. Para o efeito, foi utilizada uma base de dados referente ao IDS 2022-2023 obtida no site da *Demographic Health Survey* (DHS), <https://www.dhsprogram.com>, mediante um pedido de autorização efectuado pela autora. Trata-se de um estudo quantitativo e de natureza aplicada cuja amostra é composta pelas respostas dadas por 3873 mulheres dos 15-49 anos, para a análise de dados foram utilizados o teste de independência do Qui-quadrado de Pearson para verificar a associação das variáveis e a técnica de regressão logística múltipla para verificar a influência das variáveis independentes à ocorrência de Gravidez na adolescência. Concluiu-se que cerca 35.1% das mulheres residem no meio urbano e 64.9% residem no meio rural, após aplicar a técnica de regressão logística múltipla, concluiu-se que os factores determinantes à ocorrência da gravidez em adolescentes dos 10-19 anos no norte de Moçambique são: área de residência, província, nível de escolaridade, estado civil, uso de Métodos contraceptivos, idade na primeira relação sexual, idade na primeira menstruação e idade na primeira coabitação. Quanto menores forem as idades da primeira relação sexual e coabitação, maiores são as chances de engravidar na adolescência. Os dados ajustaram-se bem ao modelo, pois de acordo com a tabela de classificação, cerca de 77.7% das observações foram correctamente classificadas para um nível de probabilidade de 50%.

Plavras-chave: Gravidez precoce; Adolescência; União prematura; Regressão logística.

Abstract

Teenage pregnancy is considered a biological risk situation for both adolescents and newborns. This is one of the problems that threaten the public health of society in general and requires special attention from the entities responsible for providing effective solutions to reduce it. This study with the theme "Analysis of Determinant Factors the Occurrence of Adolescent Pregnancy in Northern Mozambique" is of extreme importance for society, as it provides relevant data on the determining factors for the prevalence of teenage pregnancy in Northern Mozambique. The main objective of this work was to analyze the factors that determine the occurrence of Pregnancy in adolescents aged 10-19 in northern Mozambique. For this purpose, a database referring to the IDS 2022-2023 was used, obtained from the *Demographic Health Survey* (DHS) website, <https://www.dhsprogram.com>, following an authorization request made by the author. This is a quantitative and applied study whose sample is composed of the responses given by 3873 women aged 15-49. For data analysis, Pearson's Chi-square test of independence was used to verify the association of variables and the multiple logistic regression technique to verify the influence of independent variables on the occurrence of GA. It was concluded that approximately 35.1% of women reside in urban areas and 64.9% reside in rural areas. After applying the multiple logistic regression technique, it was concluded that the determining factors for the occurrence of pregnancy in adolescents aged 10 – 19 in northern Mozambique are: area of residence, province, level of education, marital status, use of MC, age at first sexual intercourse, age at first menstruation and age at first cohabitation. The younger the age at first sexual intercourse and cohabitation, the greater the chances of becoming pregnant in adolescence. The data fit the model well, as according to the classification table, about 77.7% of the observations were correctly classified at a probability level of 50%.

Keywords: Early pregnancy; Adolescence; Premature union; Logistic Regression.

Índice

1	Introdução	1
1.1	Contextualização	1
1.2	Definição do Problema	2
1.3	Objectivos	2
1.3.1	Objectivo Geral	2
1.3.2	Objectivos específicos	2
1.4	Justificação do Estudo	3
1.5	Estrutura do Trabalho	3
2	Revisão da Literatura	5
2.1	Conceitos Básicos	5
2.2	Estágio actual dos casos União e gravidez na adolescência em Moçambique	6
2.3	Factores associados à gravidez na adolescência	6
2.3.1	Factores culturais	7
2.3.2	Factores Sociodemográficos	7
2.3.3	Factores Económicos	8
2.4	Principais consequências e desafios relacionados a gravidez na adolescência	9
2.5	Técnicas Estatísticas	9
2.5.1	Teste de associação do Qui-quadrado de Pearson	9
2.5.2	Regressão Logística	10
2.5.3	Regressão Logística Simples	13
2.5.4	Regressão Logística Múltipla	16
3	Material e métodos	18
3.1	Classificação do estudo	18
3.1.1	Quanto a natureza	18
3.1.2	Quanto a Abordagem	18
3.1.3	Quanto aos Objectivos	18
3.1.4	Quanto aos procedimentos	18
3.2	Material	19
3.3	Métodos	20
3.3.1	Análise Exploratória	20

3.3.2	Premissas da Regressão Logística	20
3.3.3	Teste de independência do Qui-quadrado	21
3.3.4	Razão de Chances	22
3.3.5	Testes de significância dos coeficientes	24
3.3.6	Avaliação da qualidade do ajuste do modelo estimado	25
3.3.7	Tabelas de Classificação	28
3.3.8	Interpretação do Modelo Logístico	28
3.4	Procedimentos de Análise	29
4	Resultados e Discussão	30
4.1	Análise Descritiva	30
4.2	Associação entre as variáveis independentes e a ocorrência de gravidez na adolescência	35
4.3	Análise dos Pressupostos do modelo de regressão logística	35
4.4	Modelo de Regressão Logística	36
4.5	Discussão dos Resultados	41
5	Conclusões e Recomendações	42
5.1	Conclusões	42
5.2	Recomendações	43
5.3	Limitações	43
	Referências	44
	Apêndices	48

Lista de Figuras

2.1	Função Logística	13
4.1	Distribuição das mulheres por área de residência (a) e por nível de educação (b).	31
4.2	Distribuição percentual das mulheres por índice de riqueza dos agregados (c) e casos de gravidez na adolescência (d).	31

Lista de Tabelas

3.1	Descrição das Variáveis	20
3.2	Matriz de Classificação	28
4.1	Distribuição de frequências absolutas e percentuais das mulheres com idade compreendida entre os 15 à 49 anos residentes no norte de Moçambique, de acordo com as características socioeconómicas, demográficas e culturais.	34
4.2	Teste de independência entre as variáveis explicativas e a Ocorrência de gravidez na adolescência	35
4.3	Teste de Multicolinearidade	36
4.4	Teste de Durbin-Watson	36
4.5	Estimativa de Parâmetros do Modelo de Regressão Logística Múltipla . . .	37
4.6	Teste Global do Modelo Logístico	39
4.7	Medidas de Avaliação da Qualidade de ajuste do Modelo	39
4.8	Teste Hosmer e Lemeshow	39
4.9	Quadro de Classificação do Modelo Ajustado	40
4.10	Modelo logístico final	40
5.1	Tabela de Contingência para o teste de Hosmer e Lemeshow	48

Capítulo 1

Introdução

1.1 Contextualização

A gravidez na adolescência (GA) é encarada como um obstáculo, um acidente triste, impedindo as raparigas de realizarem os seus sonhos. Quando fica grávida a vida da rapariga muda, pois esta passa a ter as responsabilidades de um adulto, nomeadamente cuidar da criança e possivelmente do marido. Em algumas regiões, estas adolescentes são discriminadas pelas suas famílias e pela comunidade.

Segundo Silva & Surita (2012), a GA é considerada de risco, perigosa, inapropriada e inadequada para os interesses dos jovens, particularmente por afectar maioritariamente meninas que vivem na pobreza, em países pouco desenvolvidos. Nestes países, estima-se que cerca de 21 milhões de meninas entre 15 e 19 anos engravidam, sendo a principal causa de morte nessa faixa etária, (Pinheiro, 2019).

De acordo com a Agência de saúde sexual e reprodutiva das Nações Unidas, (UNFPA) em 2013, a nível global, as tendências apontam uma pequena queda na gravidez na adolescência, de cerca de 23% para 20%, entre 1990 e 2011. A prevalência da gravidez na adolescência, nos países em desenvolvimento, é cerca de 19%, mas os países da África subsaariana continuam a apresentar uma elevada prevalência, com taxas de cerca de 30% ou mais.

Segundo os resultados apresentados no Inquérito de Indicadores de Imunização, Malária e HIV/SIDA (IMASIDA, 2015), a gravidez na adolescência tem vindo a apresentar uma tendência crescente desde 1997 e os resultados do mesmo inquérito sugerem que a percentagem destes fenómenos é maior nas províncias da zona norte do País.

A GA é uma realidade a que quase metade das adolescentes Moçambicanas estão expostas, especialmente aquelas que residem nas áreas rurais, as menos escolarizadas e de famílias

de baixo nível económico. De acordo com UNFPA (2013), a gravidez na adolescência resulta do baixo investimento no capital humano das raparigas, ao qual se adiciona as pressões sociais e desigualdades estruturais que limitam o poder de decisão sobre a sua saúde, comportamento sexual, relações, casamento e maternidade.

1.2 Definição do Problema

A idade na qual a mulher tem sua primeira relação sexual exerce um efeito importante sobre a sequência e o tempo dos subsequentes eventos no processo reprodutivo da mulher. De acordo com o IDS (2022-23), os níveis mais elevados de gravidez na adolescência verificam-se nas províncias de Cabo Delgado (55%) e Niassa (52%). A Cidade de Maputo (12%) e a província de Maputo (18%) são as que apresentam os níveis mais baixos.

A GA é considerada um problema de saúde pública pelo facto de estarem associados à morbilidade e mortalidade, tanto da mulher como da criança. A ocorrência deste fenómeno pode acarretar sérias consequências físicas, psíquicas e sociais de uma adolescente. Quando uma rapariga fica grávida, o seu presente e o futuro mudam radicalmente, pois a probabilidade de abandono escolar aumenta, as oportunidades de emprego diminuem, a sua saúde fica em risco e agrava-se a sua vulnerabilidade à pobreza, exclusão e dependência.

Para além das consequências a nível social e económico das adolescentes, este fenómeno pode influenciar nos padrões de morbilidade, mortalidade, natalidade e consequentemente nas características de uma determinada população residente num determinado País. Devido a pertinência e magnitude deste fenómeno, tornando necessário responder a seguinte questão: Quais são os factores determinantes à ocorrência da gravidez na adolescência no norte de Moçambique?

1.3 Objectivos

1.3.1 Objectivo Geral

Analisar os factores determinantes à ocorrência da gravidez em adolescentes dos 10-19 anos no norte de Moçambique.

1.3.2 Objectivos específicos

- Descrever o perfil socio-demográfico das adolescentes dos 10-19 anos residentes no Norte de Moçambique entrevistadas durante a realização do IDS 2022-23;
- Identificar os factores associados à ocorrência da Gravidez em adolescentes dos 10-19 anos no norte de Moçambique;

- Estimar a chance de uma adolescente dos 10-19 anos engravidar ou ser mãe no Norte de Moçambique.

1.4 Justificação do Estudo

A gravidez na adolescência é um fenómeno prejudicial no desenvolvimento não só das raparigas, bem como das comunidades existentes no país, pois esta tem consequências a nível social, educacional, económico e na saúde dos indivíduos. Independentemente de ser ou não desejada, a gravidez precoce pode elevar o risco de morte da mãe e do bebé, acarretando ainda riscos de prematuridade, anemia, aborto espontâneo, eclâmpsia, depressão pós-parto, entre outros.

Portanto, este estudo é importante para a sociedade, pois fornecerá dados relevantes sobre os factores determinantes à prevalência da gravidez na adolescência no Norte de Moçambique, mas também irá identificar o perfil socio-demográfico das adolescentes dos 10-19 anos com maiores chances de engravidar durante esta faixa etária. Este facto possibilitará as autoridades competentes a delinear estratégias de combate e prevenção da gravidez em adolescentes e consequentemente reduzir a morbilidade, mortalidade e outros riscos para a saúde destas mulheres assim como do bebé.

1.5 Estrutura do Trabalho

- **Capítulo 1: Introdução**

Composto pelos seguintes aspectos: contextualização do tema, a definição do problema de investigação, os objectivos e a relevância do estudo.

- **Capítulo 2: Revisão de literatura**

Apresentam-se aspectos teóricos sobre a prevalência da GA e uma descrição dos conceitos das regressões logística simples e múltipla assim como os testes de independência do Qui-Quadrado.

- **Capítulo 3: Material e métodos**

No terceiro capítulo, é apresentada a fonte de obtenção de dados, a descrição das variáveis do estudo e por fim foi feita uma descrição da métodos utilizados visando alcançar os objectivos descritos na secção 1.2.

- **Capítulo 4: Resultados e discussão**

Neste capítulo, apresentar-se-á os resultados e o tratamento estatístico dos dados, recorrendo a estatísticas descritivas, o teste qui-quadrado e o modelo de regressão logística múltipla.

- **Capítulo 5: Conclusões e recomendações**

Neste capítulo são apresentadas as principais conclusões do trabalho, recomendações e limitações com base nos resultados alcançados.

Estão incluídas na parte final deste trabalho, as referências bibliográficas consultadas durante a elaboração do mesmo assim como os apêndices que contém materiais complementares e suplementares.

Capítulo 2

Revisão da Literatura

Este capítulo apresenta os conceitos relacionados ao tema do trabalho que foram desenvolvidos por diversos autores e também conceitos teóricos das principais técnicas estatísticas utilizadas.

2.1 Conceitos Básicos

Sexualidade

De acordo com Damiani (2003), a sexualidade é um conjunto de expressões, valores, atitudes, comportamentos, uma necessidade individual, independentemente da idade ou sexo, que o ser humano desenvolve e manifesta através da sua personalidade. Portanto, a sexualidade, sendo um fator biológico, controla o desenvolvimento sexual em todas as etapas do crescimento, o funcionamento e satisfação sexual para além de ser responsável pelas diferenças de comportamentos, estimulação sexual entre os sexos, produzindo efeitos biológicos diversos como por exemplo pulso acelerado, rubor e outros.

Educação sexual

A educação sexual é o nome dado ao processo que visa educar, ou seja, esclarecer jovens e adolescentes a respeito da responsabilidade particular de cada um em relação a temas como sexo, a gravidez, o aborto, métodos contraceptivos e doenças sexualmente transmissíveis. Segundo Mubate (2024), a educação sexual é um conceito amplo que envolve não só aquisição de conteúdos através do ensino, mas também adoção de atitude e mudança de comportamento referentes a sexualidade e a saúde sexual e reprodutiva através da aprendizagem e isto pode acontecer nas várias instituições sociais durante a formação do sujeito.

Adolescência

A adolescência é o período de crescimento e desenvolvimento humano que ocorre após a infância e antes da idade adulta. A OMS define a adolescência como o período de tempo entre os 10 e os 19 anos, mas a idade é apenas uma das características que delimita este período de desenvolvimento.

2.2 Estágio actual dos casos União e gravidez na adolescência em Moçambique

Em Moçambique a taxa de mortalidade materna é de 500 por 100.000 partos, mas em adolescentes entre os 12 e os 14 anos de idade a taxa aumenta para 1.816 óbitos. A média etária das mulheres que morrem devido a causas maternas é de 27 anos e cerca de 20% dos óbitos ocorrem antes dos 20 anos, 14% devidos a aborto inseguro (Pires & Josaphat, 2016).

De acordo com o inquérito demográfico e de saúde de 2022-23, o casamento é o principal indicador da exposição regular das mulheres ao risco de gravidez, por conseguinte, é importante para a compreensão da fecundidade. As populações em que a idade do primeiro casamento é baixa tendem a ter uma maternidade precoce e uma fecundidade elevada. A idade mediana no primeiro casamento é de 18,5 anos entre as mulheres de 25–49 anos e 22,2 entre os homens do mesmo grupo etário. Conforme mencionado anteriormente, os níveis mais elevados de gravidez na adolescência foram verificados nas províncias de Cabo Delgado (55%) e Niassa (52%).

O mesmo inquérito revelou ainda que a percentagem de mulheres de 20–49 anos que tiveram a primeira relação sexual antes dos 15 anos é duas vezes maior (31%) do que a dos homens do mesmo grupo etário (15%). O início da primeira relação sexual nas mulheres de 20–49 anos de idade ocorre na idade mediana mais jovem nas províncias de Cabo Delgado (14,5 anos) e Niassa (14,7 anos) e na idade mediana mais avançada na Cidade de Maputo (17,3 anos) e Província de Maputo (16,9 anos). As províncias do Norte do país apresentam os níveis mais elevados de fecundidade, onde se destaca Niassa, com 6,8 filhos por mulher. Em contrapartida, as províncias do Sul apresentam os níveis mais baixos, destacando-se a Cidade de Maputo, com 2,1, filhos por mulher.

2.3 Factores associados à gravidez na adolescência

A GA, dentro de suas particularidades, é um fenómeno mundial que afecta, em muitos casos, países subdesenvolvidos e com elevados graus de iniquidade social. De acordo com Pinheiro (2019), o controle da gravidez precoce se apresenta como um desafio para os serviços de saúde, uma vez que envolve problemáticas diversas no que diz respeito a com-

portamentos, valores culturalmente estabelecidos, organização social e estruturação das políticas educacionais e de saúde, abaixo são descritos alguns dos factores comumente relacionados.

2.3.1 Factores culturais

A questão cultural exerce alguma influência na ocorrência de uniões prematuras, através de herança de hábitos e costumes comunitárias de união em idade prematura. Os ritos de iniciação são uma prática que vem sendo levada a cabo pelas diversas comunidades moçambicanas um pouco por todo o país, mas com maior incidência na zona norte e centro, onde são envolvidas crianças e adolescentes em diversas práticas tradicionais com cariz educativo.

Mungomane & Barbosa (2022) afirmam que problemática da desistência da rapariga na rede escolar relacionada com a prática dos ritos de iniciação constitui um grande desafio e, por conseguinte, suscita novas perspectivas na gestão do processo educativo por parte do Ministério de Educação e Desenvolvimento Humano, no geral, e particularmente nos Serviços Distritais de Educação Juventude e Tecnologia, espalhados um pouco por todos os distritos de Moçambique.

2.3.2 Factores Sociodemográficos

Segundo Damiani (2003), a gestação em adolescentes encontra-se associada à baixa adesão ao pré-natal, o que pode ocasionar maior prevalência de recém-nascido de baixo peso, parto pré-termo e aumentar a necessidade de suporte psicossocial ocasionado pelo estresse da gravidez nessa fase da vida. Um outro factor que influencia nas taxas de gestação precoce é o preconceito social em relação à concepção, pois seu uso seria uma confissão da vida sexual ativa, e uma adolescente prevenida é discriminada por estar assumindo a responsabilidade quanto à prática sexual. Além disso, a instabilidade emocional própria da fase adolescente também tende a levar ao uso e abandono de métodos anticoncepcionais, devido à ausência de vida sexual regular e à troca de parceiros frequente.

De acordo com Souza et al. (2017), a menor escolaridade foi um factor relatado bem como o fracasso escolar, caracterizado como reprovar ou abandonar, não frequentar a escola e estudar em escola pública e ter mãe com ensino fundamental completo ou incompleto constituem factores associados à GA. O consumo de grandes quantidades de álcool, como ocorre na intoxicação, devido ao seu efeito sobre a consciência, aumenta o risco de relações sexuais desprotegidas. O uso de tabaco e gravidez na adolescência podem estar relacionados com o contexto familiar.

Damiani (2003), o aumento da incidência de casos de GA vem se tornando cada vez mais um problema de saúde pública no Brasil, e um dos factores a liberalidade sexual, que leva ao início precoce da actividade sexual precoce e muitas vezes desprotegida. Outros factores são a dificuldade ao acesso e utilização dos métodos contraceptivos, a informação tardia e mal veiculada principalmente em distritos mais recônditos (Ex. a pílula do dia seguinte), o amadurecimento sexual mais precoce, com a idade média da menarca se antecipando a cada geração o que propicia a GA, a desarticulação e da insegurança da família em propiciar uma educação sexual sadia.

2.3.3 Factores Económicos

Segundo Pires & Josaphat (2016), a pobreza é um factor determinante para a ocorrência de casamentos prematuros, que muitas vezes resulta em gravidezes precoces em Moçambique. De acordo com o relatório do Fundo das Nações Unidas para a Infância de 2010 (UNICEF, 2010), a menina de famílias de baixa renda tem maior probabilidade de casar prematuramente. Em muitas comunidades, com maior incidência na região norte do país, os adultos forçam as meninas como maneira de reduzir as despesas dos agregados familiares.

Para Silva & Clapis (2023), o desemprego, dos pais da menina ou outro membro familiar que atue como líder financeiro da família, é um fator deflagrador para a ocorrência deste fenómeno, pois estarão mais susceptíveis a viabilizar o processo de união prematura da adolescente. na visão dos profissionais de saúde a vivência de uma condição socioeconômica insuficiente para as necessidades da família é fator de risco para a gravidez prematura.

O acesso aos serviços de saúde em muitas comunidades da zona norte do Moçambique é muito limitado, em resposta a essa limitação o MISAU recorre aos Agentes Polivalentes elementares (APE) para prestar alguns cuidados básicos de saúde nas comunidades rurais. Os APE estão formados para distribuir alguns tipos de contraceptivos, alguns são mais activos que outros fazendo com que haja comunidades que desconhecem a disponibilidade de contraceptivos, (Pires & Josaphat,2016).

Embora existam algumas medidas tomadas pelo MISAU, estas muitas vezes não tem sido suficiente devido aos elevados custos associados. É necessário aumentar a cobertura dos cuidados de saúde através de construção de unidades sanitárias locais e formação de profissionais qualificados.

2.4 Principais consequências e desafios relacionados a gravidez na adolescência

Segundo César et al. (2000), a gravidez na adolescência é considerada de alto risco para a mortalidade infantil, em função de aspectos fisiológicos, como peso, estatura, estado nutricional e desenvolvimento do aparelho reprodutivo da mãe. Os riscos biológicos são agravados com as condições socioeconômicas, culturais e psicológicas, uma vez que a maioria das mães adolescentes, em geral, é pobre, de baixa escolaridade e maior desnutrição, com menor acesso aos serviços de saúde. O início precoce da maternidade reduz as oportunidades educativas e laborais das mulheres e consequentemente na qualidade de vida das adolescentes.

Para Pires & Josaphat (2016), os problemas de fertilidade são muitas vezes devidos a complicações durante a primeira gravidez em raparigas que ficam grávidas muito jovens, que tem problemas durante o parto e muitas vezes devem ser submetidas a cesariana, culminando em dificuldades para ter uma segunda gravidez ou até mesmo na morte do bebé.

A gravidez precoce pode culminar, em alguns casos, em mortalidade materna. Por sua vez, Alberto et al. (2011), consideram que a mortalidade infantil também é influenciada por vários fatores. Quanto aos aspectos demográficos, no que tange a idade da mãe, as evidências geralmente apontam que os maiores riscos de morte são registrados para filhos de mulheres nos extremos da idade reprodutiva.

Para da Costa (2021), o impacto de uma gravidez precoce na vida de uma adolescente aumenta consideravelmente quando o pai da criança decide se eximir de qualquer responsabilidade. A corresponsabilidade na maternidade é um dos pontos principais para a reinserção da menina no âmbito educacional e laboral, uma vez que se o pai e/ou companheiro compartilhar as tarefas de cuidado com a criança e do lar, a mulher tem a oportunidade de se dedicar a objectivos pessoais.

2.5 Técnicas Estatísticas

2.5.1 Teste de associação do Qui-quadrado de Pearson

O teste do qui-quadrado de independência está associado a duas variáveis qualitativas, ou seja, uma análise bidimensional e serve para verificar a relação de dependência entre as duas variáveis analisadas. Utiliza-se o teste de independência quando os dados estiverem sob forma de frequências em categorias, para se aplicar este teste de hipóteses, recorre-se a estatística de Qui-quadrado, (Tavares, 2007).

De acordo com Guimarães (2019) este teste de hipóteses destina-se a encontrar um valor da dispersão para duas variáveis categóricas nominais e avaliar a associação existente entre variáveis qualitativas. O teste de Qui-Quadrado é considerado um teste não paramétrico, pois não depende de parâmetros populacionais (média e variância, por exemplo). O princípio básico deste teste é comparar proporções, ou seja, possíveis divergências entre as frequências observadas e esperadas para um certo evento.

O teste Qui-Quadrado pode ser utilizado para:

- Verificar se a frequência com que um determinado acontecimento observado em uma amostra se desvia significativamente ou não da frequência com que ele é esperado;
- Comparar a distribuição de diversos acontecimentos em diferentes amostras, com a finalidade de avaliar se as proporções observadas destes eventos mostram ou não diferenças significativas ou se as amostras diferem significativamente quanto às proporções desses acontecimentos.

2.5.2 Regressão Logística

Segundo Fávero e Belfiore (2017), a técnica de regressão logística foi desenvolvida por volta da década de 1960 em resposta ao desafio de realizar previsões ou explicar a ocorrência de fenômenos quando a variável dependente fosse de natureza binária. Um dos estudos que contribuiu para o maior destaque da técnica foi o *Framingham Heart Study* cujo objetivo era identificar os factores que propiciam doenças cardiovasculares em 5209 indivíduos dos 30 aos 60 anos residentes em Massachusetts.

Assim, os resultados desse estudo permitiram identificar, graças a regressão logística, a hipertensão arterial, elevadas taxas de colesterol, tabagismo, obesidade, diabetes e sedentarismo como factores que propiciam as doenças cardiovasculares.

A regressão logística tem sido utilizada em diversas áreas. Sales *et al.* (2005), aplicaram a regressão logística para estudar a prevalência e o valor prognóstico da anemia em uma população hospitalizada por insuficiência cardíaca descompensada.

Gonçalves *et al.* (2013) aplicaram a técnica de regressão logística no desenvolvimento de um modelo de previsão de *credit scoring* com dados de uma instituição financeira. Tomás *et al.* (2021) aplicaram a regressão logística na área educacional para analisar os factores que influenciam a participação dos alunos em um processo de avaliação docente, identificando os perfis mais e menos propensos a participar, incluindo características individuais

e dos cursos aos quais estão vinculados.

Conceito

De acordo com Gonzalez (2018), a regressão logística é uma técnica estatística cujo objetivo é produzir, a partir de um conjunto de observações, um modelo que permita a previsão de valores tomados por uma variável categórica, frequentemente binária, em função de uma ou mais variáveis independentes que podem ser contínuas e/ou binárias.

Para Fávero e Belfiore (2017), de um modo geral, o modelo de regressão logística destina-se a aferir a probabilidade de ocorrência de um evento dada uma observação aleatória e a identificar características dos elementos pertencentes a cada grupo determinado pela variável categórica.

Segundo Hair *et al.* (2005), existem várias razões pelas quais a regressão logística é uma alternativa atraente à análise discriminante. Primeiro porque é menos afectada do que a análise discriminante pelas desigualdades da variância/covariância ao longo dos grupos, uma suposição básica da análise discriminante.

Segundo porque a regressão logística pode lidar facilmente com variáveis independentes categóricas, enquanto que na análise discriminante, o uso de variáveis dicotómicas cria problemas com igualdades de variância/covariância. Finalmente, os resultados da regressão logística acompanham paralelamente os da regressão múltipla em termos de sua interpretação e das medidas diagnosticadas caso a caso disponíveis na análise de resíduos.

Vantagens do modelo logístico

Segundo Guimarães (2019) e Kleinbaum *et al.* (1998), a regressão logística possui as seguintes vantagens:

- Facilidade para lidar com variáveis independentes categóricas;
- Fornece resultados em termos de probabilidade;
- Facilidade de classificação de indivíduos em categorias;
- Robustez e facilidade de interpretação;
- Requer pequeno número de suposições;

- Alto grau de confiabilidade.

Conforme foi mencionado anteriormente, a regressão logística permite prever ou calcular a probabilidade de um evento ocorrer, dada uma observação aleatória. Segundo Belfiore (2015), neste modelo, a variável dependente Y é frequentemente binária, logo, nestes casos ela segue uma distribuição de Bernoulli tendo uma probabilidade desconhecida p , os valores de Y são dados por:

$$Y = \begin{cases} 1, & \text{se ocorrer sucesso} \\ 0, & \text{se ocorrer fracasso} \end{cases}$$

A probabilidade de ocorrência do sucesso é $0 \leq p \leq 1$ e a probabilidade de ocorrência do fracasso é $q = 1 - p$.

Função Logit

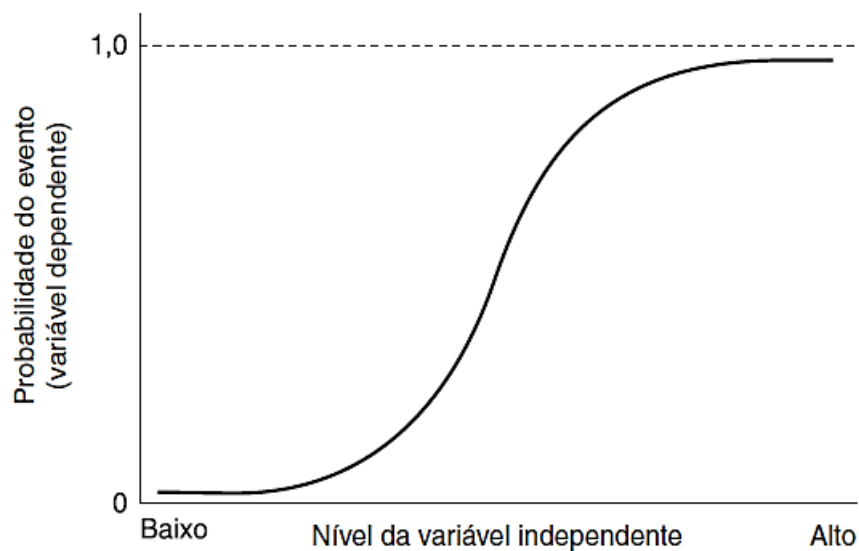
Na regressão logística, é feita a estimação da probabilidade desconhecida p , dada uma combinação linear de variáveis independentes. Para o efeito, é preciso conectar essas variáveis à distribuição de Bernoulli, presente na variável dependente e essa ligação é chamada logit.

De acordo com Gonzalez (2018), para ligar a combinação linear de variáveis à distribuição de Bernoulli, é necessária uma função que as una ou faça o mapeamento da combinação linear de modo a retornar qualquer valor na distribuição de probabilidade de Bernoulli cujo o domínio será de 0 a 1. A razão de probabilidade é chamada de chance ou *odds* em inglês, e é representada por:

$$\ln(odds) = \ln\left(\frac{p}{1-p}\right) \quad (2.1)$$

A função logística apresenta um formato da letra “S” e é obtida pelo inverso da função logit (anti-logit) e que permite transformá-la em probabilidade.

Figura 2.1: Função Logística

Fonte: Hair *et al.* (2005)

2.5.3 Regressão Logística Simples

Segundo Hosmer e Lemeshow (1989), a regressão logística simples é aplicada quando se tem apenas uma variável independente e a variável dependente é dicotômica, isto é, atribui-se o valor 1 para o evento de interesse (sucesso) e o valor 0 para o acontecimento complementar (fracasso).

Seja a função linear $g(x) = \beta_0 + \beta_1 x_i$, com uma variável independente (x) e coeficientes β_0 e β_1 , igualando a equação 2.1 à função $g(x)$, segue que:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_i \quad (2.2)$$

como o objectivo do modelo logístico é estimar p , então, isolando o p na igualdade 2.2, segue que:

$$\begin{aligned} \frac{p}{1-p} &= e^{\beta_0 + \beta_1 x_i} \\ \hat{p} &= \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \\ \hat{p} &= \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_i)}} \end{aligned} \quad (2.3)$$

A equação 2.3 é chamada de equação de regressão estimada e é essencialmente a função que representa o objectivo da regressão logística, pois $\hat{p}$ é a probabilidade estimada para quaisquer valores dos coeficientes e variáveis que venham a ser colocadas na equação.

As probabilidades do sucesso e fracasso no modelo de regressão logística simples são dadas respectivamente por:

$$\pi(x_i) = p(Y_i = 1|X_i = x_i) = \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}}$$

$$1 - \pi(x_i) = p(Y_i = 0|X_i = x_i) = \frac{1}{1 + e^{\beta_0 + \beta_1 x_i}}$$

Onde: $\beta = (\beta_0, \beta_1)^T$ é o vector de parâmetros desconhecidos.

Nos modelos de regressão modela-se o valor médio da variável dependente dados os valores das variáveis independentes. Esta quantidade é designada média condicional denotada por $E(Y_i|X = x_i)$. Devido a natureza da variável resposta, a amplitude da média condicional varia no intervalo $[0;1]$, ou seja, $0 \leq E(Y_i|X = x_i) \leq 1$.

A variável dependente Y_i dado x_i é modelada por $Y_i = \pi_i + \varepsilon_i$, mas ε_i só pode assumir $\varepsilon_i = 1 - \pi_i$ para $Y_i = 1$ ou $\varepsilon_i = -\pi_i$ para $Y_i = 0$.

Onde: ε_i é o termo de erro, que segue uma distribuição binomial com média zero e $Var[\varepsilon_i] = \pi_i(1 - \pi_i)$.

Estimação dos parâmetros no modelo de regressão logística simples

Nos modelos de regressão linear a estimação dos parâmetros é feita através do método de mínimos quadrados que consiste em minimizar a soma dos quadrados da diferença entre os valores observados e estimados. No entanto, devido a ausência de relação linear na regressão logística, o método de mínimos quadrados não é apropriado para estimar os coeficientes do modelo, sendo utilizado o método da máxima verosimilhança.

Segundo Gonzalez (2018), o método da máxima verosimilhança consiste em procurar, a partir de um conjunto de observações, estimadores para o modelo de regressão logística simples, que são os valores de $\hat{\beta}_0$ e $\hat{\beta}_1$ que maximizam o logaritmo da função de máxima verosimilhança. Este método de estimação permite encontrar estimadores dos parâmetros do modelo que tem maior probabilidade de replicar o padrão de observações nos dados da amostra.

Para o modelo de regressão logística simples com $Y_i \sim Ber(\pi_i)$, a distribuição de probabilidade é dada por:

$$f(y_i, \pi_i) = \pi_i^{y_i} (1 - \pi_i)^{1-y_i}$$

De acordo com Sousa (2018), a definição de estimativa de máxima verosimilhança de β , baseada em uma amostra aleatória $(x_1, x_2, \dots, x_n)$, é o valor de β que maximiza $L(x_1, x_2, \dots, x_n; \beta)$, a função de verosimilhança é dada por:

$$L(\beta) = \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i}, \beta \in R^2 \quad (2.4)$$

onde β é um escalar ou vector, representando o parâmetro populacional desconhecido, aplicando o logaritmo natural na expressão 4, segue que:

$$\begin{aligned} l(\beta) &= \ln [L(\beta)] = \ln \left[\prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \right] \\ l(\beta) &= \sum_{i=1}^n [y_i \ln(\pi_i) + (1 - y_i) \ln(1 - \pi_i)] \\ l(\beta) &= \sum_{i=1}^n [y_i \ln(\pi_i) + \ln(1 - \pi_i) - y_i \ln(1 - \pi_i)] \\ l(\beta) &= \sum_{i=1}^n \left[y_i \ln \left(\frac{\pi_i}{1 - \pi_i} \right) + \ln(1 - \pi_i) \right] \\ l(\beta) &= \sum_{i=1}^n \left[y_i \ln (\beta_0 + \beta_1 x_i) + \ln \left(\frac{1}{1 + e^{\beta_0 + \beta_1 x_i}} \right) \right] \end{aligned}$$

após a aplicação do logaritmo, obteve-se a seguinte expressão

$$l(\beta) = \sum_{i=1}^n [y_i \ln (\beta_0 + \beta_1 x_i) - \ln (1 + e^{\beta_0 + \beta_1 x_i})] \quad (2.5)$$

Para encontrar o valor de β que maximiza $L(\beta)$ deriva-se $l(\beta)$ em relação a cada parâmetro (β_1, β_0) , obtendo-se duas equações:

$$\frac{\partial l(\beta)}{\partial (\beta_0)} = \sum_{i=1}^n \left[y_i - \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \right] \quad (2.6)$$

$$\frac{\partial l(\beta)}{\partial (\beta_1)} = \sum_{i=1}^n \left[y_i * x_i - \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \right] \quad (2.7)$$

que igualadas a zero, geram o seguinte sistema de equações:

$$\begin{cases} \sum_{i=1}^n (y_i \pi_i) = 0 \\ \sum_{i=1}^n (y_i \pi_i) x_i = 0 \end{cases} \quad (2.8)$$

em que $i = 1, 2, \dots, n$ e $\pi_i = \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}}$

Como as equações 2.6 e 2.7 não são lineares em β_0 e β_1 , são necessários procedimentos iterativos para a resolução, estes encontram-se disponíveis em vários programas computacionais, assim, o método mais utilizado, na maioria dos softwares estatísticos é o método de Newton-Raphson.

2.5.4 Regressão Logística Múltipla

A regressão logística múltipla é uma extensão do modelo logístico simples, ou seja, o objectivo é estudar a relação estatística entre uma variável dependente dicotómica, e mais de uma variável independente.

Seja $X = (x_1, x_2, \dots, x_t)$ um conjunto de variáveis independentes cuja combinação linear é definida como:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_t x_{it} \quad (2.9)$$

Para obter o modelo de regressão de logística múltipla, a probabilidade estimada ($\hat{p}$), basta isolar o p na igualdade 2.9, utilizando o antilogaritmo resultando assim na expressão da probabilidade estimada:

$$\hat{p} = \frac{e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_t x_{it}}}{1 + e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_t x_{it}}} \quad (2.10)$$

De acordo com Hosmer e Lemeshow (1989), no modelo de regressão logística, as probabilidades do sucesso e fracasso são dadas, respectivamente, pelas seguintes expressões:

Estimação de parâmetros na regressão logística múltipla

A estimação de parâmetros é feita pelo método de máxima verosimilhança, cujo objectivo é encontrar o valor de β que maximiza $l(\beta)$ através da derivação parcial em relação a cada parâmetro. Tratando-se de uma regressão logística múltipla, a logaritmização da expressão, resulta em:

$$l(\beta) = \ln[L(\beta)] = \sum_{i=1}^t \sum_{j=1}^n y_i [\ln(\beta_0 + \beta_i x_{ij}) - \ln(1 + e^{\beta_0 + \beta_i x_{ij}})] \quad (2.11)$$

portanto, a estimativa de β_0 é a derivada de parcial de $l(\beta)$ em relação a β_0 será:

$$\frac{\partial l(\beta)}{\partial \beta_0} = \sum \left[y_i x_{ij} - \frac{e^{(x_i^T \beta)}}{1 + e^{(x_i^T \beta)}} x_{ij} \right] \quad (2.12)$$

A matriz de covariância dos coeficientes estimados é obtida a partir das derivadas

parciais de segunda ordem do logaritmo da função de verosimilhança $[l(\beta)]$:

$$\frac{\partial^2 l(\beta)}{\partial \beta_j^2} = - \sum_{i=1}^n [x_{ij}^2 \pi_i (1 - \pi_i)] \quad (2.13)$$

$$\frac{\partial^2 l(\beta)}{\partial \beta_j \partial \beta_k} = - \sum_{i=1}^n [x_{ij} x_{ik} \pi_i (1 - \pi_i)] \quad (2.14)$$

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

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

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ 1 & x_{31} & x_{32} & \cdots & x_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

$$V = \begin{bmatrix} \hat{\pi}_1 (1 - \hat{\pi}_1) & 0 & \cdots & 0 \\ 0 & \hat{\pi}_2 (1 - \hat{\pi}_2) & \cdots & 0 \\ 0 & 0 & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{\pi}_n (1 - \hat{\pi}_n) \end{bmatrix}$$

Capítulo 3

Material e métodos

3.1 Classificação do estudo

3.1.1 Quanto a natureza

É um estudo aplicado, pois foi conduzido com o propósito de resolver um problema existencial. “O investigador é movido pela necessidade de contribuir para fins práticos mais ou menos imediatos, buscando soluções para problemas concretos”, (Cervo & Bervian, 2002).

3.1.2 Quanto a Abordagem

Trata-se de um estudo quantitativo, pois todas análises serão feitas recorrendo a ferramentas estatísticas e matemáticas.

3.1.3 Quanto aos Objectivos

Quanto aos objectivos, trata-se de um estudo explicativo, pois este visa analisar factores associados à ocorrência de GA no norte de Moçambique. Segundo Gil (1991), essas têm aplicação mais restrita, empregam o método experimental de pesquisa, e dotadas de complexidade, servindo para identificar atributos ou factores que determinam a ocorrência de fenómenos.

3.1.4 Quanto aos procedimentos

Trata-se de uma pesquisa realizada após a observação ou ocorrência de fenómeno ou do experimento. Ocorre quando o pesquisador não tem controlo sobre as variáveis, mas trabalha como se estivesse realizando um experimento. Segundo Yin (2005), o propósito básico desta pesquisa é o mesmo da pesquisa experimental: verificar a existência de relações entre variáveis.

3.2 Material

Para a realização deste trabalho, foram utilizados dados fornecidos pelo Ministério da Saúde (MISAU), referente ao Inquérito Demográfico e de Saúde (IDS 2022-23). Os dados foram obtidos no site <https://www.dhsprogram.com>, mediante um pedido de autorização efectuado no âmbito da realização do presente estudo.

No âmbito de realização do inquérito, todas as mulheres de 15–49 anos que eram residentes habituais ou visitantes na casa na noite anterior às entrevistas no agregado familiar foram incluídas no IDS 2022–23 e foram elegíveis para serem entrevistadas.

O IDS 2022-23 compreendeu duas etapas, a primeira envolveu a selecção da amostra de conglomerados (clusters), consistindo em áreas de enumeração (AE) delineadas para a população com base no IV Recenseamento Geral de População e Habitação (IV RGPH) de 2017 seleccionando 619 áreas de enumeração que foram seleccionadas com probabilidade proporcional à dimensão, sendo a medida de tamanho, o número de agregados familiares em cada estrato explícito. Das 619 AEs, 232 eram de áreas urbanas e 387 de áreas rurais.

No segundo estágio, 26 agregados familiares foram seleccionados sistematicamente com probabilidades iguais de cada área de enumeração. Com base nesse procedimento, foram seleccionados 16.045 agregados familiares (AFs) para o IDS 2022–23.

Assim, o tamanho da amostra geral do IDS foi de 16.045 agregados familiares, estimada como suficiente para se analisar os principais indicadores de saúde da mulher (por exemplo, utilização de métodos contraceptivos), com representatividade urbana e rural e a níveis de regiões e de províncias de Moçambique este tamanho de amostra foi determinado previamente pelo Instituto Nacional de Estatística de Moçambique com 95% de confiança para as estimativas requeridas e erro amostral de 5%.

Na tabela 3.1, são apresentadas as variáveis seleccionadas para o presente estudo, sua descrição bem como sua codificação. A base de dados possui um tamanho amostral composto por 3873 mulheres com idades entre os 15 à 49 anos (inclusive) residentes na zona norte de Moçambique que foi seleccionada a partir da base geral do IDS. Importa referir que a variável dependente "Gravidez na adolescência" foi categorizada com base nas idades da primeira gravidez das mulheres que compõem a amostra sendo **Sim** para quem teve a sua primeira gravidez entre os 10 e 19 anos e **Não** para o caso contrário.

Tabela 3.1: Descrição das Variáveis

Código	Descrição	Codificação	Classificação
Y	Gravidez na adolescência	0: Não; 1: Sim	Qualitativa Nominal
X ₁	Área de Residência	0: Urbana ; 1: Rural	Qualitativa Nominal
X ₂	Província	1: Nampula 2: Niassa; 3: Cabo Delgado	Qualitativa Nominal
X ₃	Nível de Escolaridade	0: Nenhum; 1: Primário 2: Secundário; 3: Superior	Qualitativa Ordinal
X ₄	Estado Civil	0: Solteiro; 1: Casada/União Marital; 2: Divorciada; 3: Viúva	Qualitativa Nominal
X ₅	Religião	0: Não; 1: Sim	Qualitativa Nominal
X ₆	Quintil de Riqueza	1: Mais Baixo; 2: Segundo; 3: Médio; 4: Quarto; 5: Mais Elevado;	Qualitativa Ordinal
X ₇	Exposição aos meios de Comunicação	0: Nenhuma; 1: Baixa; 2: Média; 3: Alta	Qualitativa Ordinal
X ₈	Uso de Métodos Contraceptivos	0: Não; 1: Sim	Qualitativa Nominal
X ₉	Uso de Internet	0: Não; 1: Sim	Qualitativa Nominal
X ₁₀	Ouviu falar sobre o PF	0: Não; 1: Sim	Qualitativa Nominal
X ₁₁	Consumo de tabaco	0: Não; 1: Sim	Qualitativa Nominal
X ₁₂	Faixa etária na primeira relação sexual	8 à 14 anos; 15 à 19 anos 20 ou mais anos	Qualitativa Ordinal
X ₁₃	Faixa etária na primeira menstruação	7 à 14 anos; 15 à 19 anos 20 ou mais anos	Qualitativa Ordinal
X ₁₄	Faixa etária na primeira coabitação	10 à 14 anos; 15 à 19 anos 20 ou mais anos	Qualitativa Ordinal

Os dados foram processados e analisados com o auxílio dos programas SPSS versão 26 e Microsoft Office Excel 2019, o $\text{\LaTeX}$ foi utilizado para a digitalização do relatório final e importa salientar que todas as hipóteses foram testadas a um nível de significância de 5% e para efeitos de avaliação da regra de decisão foi usado o p-valor associado a estatística do teste.

3.3 Métodos

3.3.1 Análise Exploratória

A análise exploratória foi utilizada para a caracterização das variáveis que estão na base de dados que será feita através de medidas descritivas tais como: medidas de posição, medidas de dispersão e medidas de forma de distribuição.

3.3.2 Premissas da Regressão Logística

De acordo com Fávero e Belfiore (2017), os pressupostos da regressão Logística são:

- Valor esperado dos resíduos é igual a zero;
- Ausência de Heterocedasticidade; e
- Ausência de Multicolinearidade.

Ausência de Heterocedasticidade

Segundo Maroco (2007), citado por Chichava (2020), a verificação da heteroscedasticidade é realizada a partir do gráfico de resíduos da variável dependente ou através do teste formal de Durbin-Watson, sua presença pode ser observada quando a distribuição dos pontos em torno da recta representativa da média apresenta um comportamento bem definido, indicando uma variação gradativa da dispersão.

Este teste foi utilizado para verificar se a variância do termo de erro é constante. As hipóteses testadas são as seguintes:

$$\begin{cases} H_0 : \rho = 0 \\ H_1 : \rho > 0 \end{cases}$$

A estatística do teste é dada por:

$$D = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=2}^n e_i^2}$$

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

Multicolinearidade

Segundo Gujarati e Dawn (2011), o termo multicolinearidade significa a existência de mais de uma relação linear entre algumas variáveis explicativas, mas também pode ser entendida como a existência de altas correlações entre as variáveis independentes de um modelo de regressão. Este pressuposto foi verificado através do Factor de Inflação da variância (VIF), quanto mais próximo de 0, menor será a multicolinearidade sendo que o valor habitualmente considerado como o limite acima do qual existe multicolinearidade é 10.

3.3.3 Teste de independência do Qui-quadrado

Neste estudo, a estatística do χ^2 foi utilizada para verificar dependências ou não existente entre as variáveis independentes e categóricas do estudo e a variável dependente. Importa ressaltar que foram utilizados os testes Qui-quadrado de Pearson para permitir a comparação dos resultados fornecidos por cada um dos testes.

As hipóteses 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.

A estatística de teste do qui-quadrado é representada pela seguinte expressão:

$$\chi^2 = \sum_{i=1}^l \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2_{(l-1)(c-1)} \quad (3.1)$$

Onde:

O_{ij} - representa a frequência observada na linha i da coluna j .

E_{ij} - representa a frequência esperada na linha i da coluna j .

l - é o número de linhas.

c - é o número de colunas.

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

Neste estudo, a estatística do χ^2 irá auxiliar a verificar possíveis dependências ou não existente entre as variáveis independentes e categóricas do estudo e a variável dependente.

3.3.4 Razão de Chances

Segundo Kleinbaum *et al.* (1998), os coeficientes de regressão no modelo logístico desempenham um papel importante ao fornecer informações sobre as relações da previsão no modelo com a variável dependente. Para obter essas relações é necessário envolver um parâmetro chamado *odds ratio*.

Silveira (2021) considera que a **Razão de Chances** (Odds Ratio, OR) é definida como a razão entre a chance de ocorrência de um evento em um grupo e a ocorrência deste evento em outro grupo. A razão de chances é amplamente utilizada para medir o efeito em estudos epidemiológicos.

Assim, considerando a existência de dois grupos e as respectivas probabilidades de um evento ocorrer em cada um deles dadas por p e q a razão de chances é obtida por:

$$OR = \frac{\frac{p}{1-p}}{\frac{q}{1-q}} = \frac{p(1-q)}{q(1-p)} \quad (3.2)$$

Assumindo uma variável binária, o *odds ratio* é dado 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 $x = 0$. A *odds* do acontecimento de interesse ocorrer nos indivíduos com $x = 1$ e $x = 0$ são dadas respectivamente pelas expressões $\frac{\pi(1)}{1 - \pi(1)}$ e $\frac{\pi(0)}{1 - \pi(0)}$.

Na regressão logística, as probabilidades do acontecimento de interesse ocorrer para as duas categorias de x , são dadas por:

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

e

$$\pi(0) = \frac{1}{1 + e^{\beta_0 + \beta_1 x_i}}$$

Consequentemente, o valor do odds ratio é dado pela expressão:

$$OR = \frac{\pi(1)[1 - \pi(0)]}{\pi(0)[1 - \pi(1)]} = e^{\beta_1} \quad (3.3)$$

Tornando evidente a relação entre o odds ratio e o coeficiente do modelo. O valor do OR depende da codificação adoptada para a variável binária x , que pode ser definida por quaisquer dois valores. Considerando a codificação usando os valores genéricos a e b , o valor do odds ratio é dado por:

$$OR = \frac{\pi(a)[1 - \pi(b)]}{\pi(b)[1 - \pi(a)]} = e^{\beta_1(a-b)} \quad (3.4)$$

A interpretação do *odds ratio* não pode ser feita sem antes conhecer a codificação de x . Habitualmente a codificação adoptada é definida em termos de 0 e 1, por permitir uma interpretação trivial dos parâmetros. Na prática o cálculo do *odds ratio* é feito a partir de dados reais geralmente organizados em tabelas de contingência. Destas tabelas pode-se obter uma estimativa do *odds ratio*, ($\hat{OR}$), à qual aplica-se o logaritmo obtendo uma estimativa razão de chance com a seguinte expressão:

$$\ln(\hat{OR}) = \ln(e^{\hat{\beta}_1}) = \hat{\beta}_1 \quad (3.5)$$

Todos valores de *odds ratio* maiores que 1 indicam aumento das chances, e *odds ratio* menores que 1 indicam redução das chances. Um valor do *odds ratio* igual a 1 indica que a variável não influencia a variável dependente, sendo que a variação percentual da razão das chances é dada pela expressão $(OR - 1) \times 100\%$ (Maroco, 2007).

3.3.5 Testes de significância dos coeficientes

Com o objectivo de garantir a qualidade de um modelo de regressão estimado, é imperativo incluir apenas as variáveis independentes que sejam significativas para explicar a variável dependente. No entanto, essa verificação deve ser feita através de testes de hipóteses apropriados, neste estudo, foram utilizados os testes de Wald e o da Razão de Verossimilhança.

Teste de Wald

Assim como a regressão múltipla, a regressão logística também possui um teste de significância dos coeficientes, ou seja, testar a hipótese de que o coeficiente é diferente de zero, onde a razão de desigualdade não muda e a probabilidade não é afectada.

Neste contexto, o teste de Wald avalia se cada coeficiente é significativamente diferente de zero, ou seja, este teste avalia se a relação de uma determinada variável independente com a variável dependente é estatisticamente significativa.

O teste de Wald é obtido por comparação entre a estimativa de máxima verossimilhança do parâmetro β_i e a estimativa do seu erro-padrão. A estatística de teste é dada por:

$$W = \left(\frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \right)^2 \quad (3.6)$$

Onde:

W: Estatística de Wald

$\hat{\beta}_j$: Coeficiente estimado da variável independente j

$SE(\hat{\beta}_j)$: Erro padrão do coeficiente estimado.

As hipóteses a serem testadas são:

H_0 : O coeficiente da regressão logística é igual a zero.

H_1 : O coeficiente da regressão logística não é igual a zero.

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

Mesquita (2014), alerta sobre a existência de alguns casos em que o teste de Wald costuma

não rejeitar a hipótese nula quando esta deveria ser rejeitada. Sendo assim, recomenda-se que o teste da razão de verosimilhança seja utilizado nos casos em que houver dúvidas acerca da eficiência do teste de Wald.

Teste da Razão de verosimilhança

Conforme Cabral (2013), por meio desta medida, testa-se se os coeficientes de regressão associados a β são todos nulos, com exceção de β_0 . Sendo assim, compara-se os valores observados e os valores esperados usando a função de verosimilhança, conforme a seguinte expressão:

$$D = -2 \ln \left[\frac{\text{Função de máxima verosimilhança do modelo corrente}}{\text{Função de máxima verosimilhança do modelo saturado}} \right] \quad (3.7)$$

$$D = -2 \sum_{i=1}^n \left[y_i \ln \left(\frac{\hat{\pi}_i}{y_i} \right) + (1 - y_i) \ln \left(\frac{1 - \hat{\pi}_i}{1 - y_i} \right) \right] \quad (3.8)$$

O modelo é dito saturado se contém todas as variáveis, enquanto que o ajustado corresponde ao modelo apenas com as variáveis desejadas para o estudo. Esta função D, também chamada de *deviance* (desvio), sempre é positiva quanto menor, melhor é qualidade de ajuste do modelo.

As hipóteses a serem testadas são:

$$\begin{cases} H_0 : \beta_1 = \dots = \beta_t = 0 & (\text{o modelo não é adequado}) \\ H_1 : \exists_{j=1, \dots, p} \beta_j \neq 0 & (\text{o modelo é adequado}) \end{cases}$$

Assim, na hipótese nula (H_0) a ser testada, os parâmetros do modelo serão igualados a 0. O modelo saturado que mantém o valor dos seus respectivos coeficientes, representará a hipótese H_1 .

Para estimar a significância de uma variável independente, comparam-se o valor de D com e sem a variável independente na equação. A alteração no valor de D esperada pela inclusão da variável independente no modelo e obtida através de:

$$G = D \left(\frac{\text{modelo sem a variável}}{\text{modelo com a variável}} \right) \quad (3.9)$$

Ao registrar a hipótese nula, tem-se que a variável independente testada, é significativa para o modelo.

3.3.6 Avaliação da qualidade do ajuste do modelo estimado

Segundo Silva (2016), sempre que se estima um modelo de regressão é fundamental, antes de passar a extrair conclusões, verificar se o modelo se ajusta efectivamente aos dados

usados para estimá-lo.

De acordo com Hair *et al.* (2005), a qualidade de ajuste para um modelo de regressão logística pode ser avaliada de duas maneiras, uma é a avaliação de ajuste usando valores pseudo R^2 , semelhantes à aqueles encontrados na regressão múltipla. A segunda abordagem é examinar a precisão preditiva (com a matriz de classificação).

As duas técnicas examinam ajuste de modelo sob diferentes perspectivas, mas devem conduzir a conclusões semelhantes. Portanto, para avaliar a qualidade de ajuste do modelo logístico serão utilizados os testes de *Hosmer e Lemeshow*, as medidas Pseudo R^2 de Cox e Snell.

Medidas Pseudo R^2 de Cox e Snell

Segundo Gonzalez (2018), a estatística R^2 é uma medida que avalia em termos percentuais, a qualidade de ajuste de um modelo de regressão linear aos dados observados. Na regressão logística, não existe uma estatística que seja equivalente ao R^2 .

O mesmo autor afirma ainda que a denominação de R^2 deve-se ao facto de parecer com um R^2 do modelo de regressão linear, mas apesar dessa similaridade não podem ser interpretados da mesma forma.

De acordo com Silveira *et al.* (2021), esta medida é denominada pseudo R^2 , pelo fato de apresentar semelhanças com o R^2 dos modelos de regressão linear. Entretanto, deve-se ressaltar que, apesar da similaridade, a interpretação de ambos é diferente. Existem muitas maneiras de calcular o pseudo R^2 , sendo o pseudo R^2 de Cox e Snell um dos mais frequentemente usados pelos softwares estatísticos. A medida é definida por:

$$R^2 = 1 - \left(\frac{L(\beta)_0}{L(\beta)_M} \right)^{\frac{2}{n}} \quad (3.10)$$

onde n é o tamanho da amostra, $L(\beta)_0$ o valor da função verosimilhança para um modelo sem variáveis explicativas e $L(\beta)_M$ a verosimilhança do modelo sendo estimado.

A medida resulta em um valor que varia de 0 até aproximadamente a 1. Esses valores são utilizados para comparar os modelos nos quais as variáveis independentes melhor explicam as variações na variável dependente, ao comparar dois modelos busca-se por aquele que apresente um pseudo R^2 mais elevado.

Teste Nagelkerke R^2

Conforme abordado, o indicador Cox-Snell R^2 baseia-se no Valor de Verossimilhança, porém é situado em uma escala que se inicia em zero, mas não chega ao limite superior 1. O teste Nagelkerke R^2 é um ajuste desenvolvido no indicador anterior para que o mesmo chegue ao limite máximo de 1.

Para Corrar *et al.* (2011), o teste Nagelkerke R^2 situa-se numa escala que vai de 0 a 1 e tem a mesma finalidade do indicador Cox-Snell R^2 . Quando aplicado de modo prático, a principal diferença é que o teste Nagelkerke R^2 torna-se mais compreensível que o Cox-Snell R^2 .

Teste de Hosmer e Lemeshow

De acordo com Hair *et al.* (2005), fornece uma medida ampla de precisão preditiva que é baseada não no valor de verossimilhança, mas na real previsão da variável dependente.

Este teste é largamente utilizado na regressão logística com o objetivo de testar a qualidade do ajuste, isto é, o teste avalia se o modelo obtido pode explicar adequadamente os dados observados baseando-se na divisão dos dados de acordo com as probabilidades estimadas. As observações são separadas em g grupos de acordo com as probabilidades previstas, a estatística de teste é dada por:

$$\hat{C} = \sum_{k=1}^g \frac{(o_k - e_k)^2}{e_k \left(1 - \frac{e_k}{n_k}\right)} \quad (3.11)$$

onde: $o_k = \sum_{j=1}^{n_k} y_{kj}$ representa o número de casos registrados no k -ésimo decil; e

$e_k = \sum_{j=1}^{n_k} \hat{\pi}_{kj}$ representa o número esperado de casos no k -ésimo decil.

Salotti e Yamoto (2005), citado por Nhambire (2012), definem que as hipóteses desse teste são:

$$\begin{cases} H_0 : \text{Não há diferenças significativas entre os resultados previstos e os observados.} \\ H_1 : \quad \text{Há diferenças significativas entre os resultados previstos e os observados.} \end{cases}$$

Sob H_0 , o modelo estimado ajusta-se aos dados, isto é, não existem diferenças significativas entre os valores previstos e os observados, em que a regra de decisão é:

- Se $\hat{C} \leq \chi^2_{(1-\alpha;g-2)}$ ou $P\left(\hat{C} > \chi^2_{(1-\alpha;g-2)}\right) \geq \alpha$, não rejeitar H_0 .
- Se $\hat{C} < \chi^2_{(1-\alpha;g-2)}$ ou $P\left(\hat{C} < \chi^2_{(1-\alpha;g-2)}\right) \geq \alpha$, rejeitar H_0 .

A finalidade do teste é avaliar se o nível de significância (Sig) referente a esse indicador é superior ao nível de significância da investigação, ou seja, a não rejeição da hipótese nula.

3.3.7 Tabelas de Classificação

De acordo com Hosmer e Lemeshow (2000), uma maneira prática de qualificar o ajuste do modelo de regressão logística é pela projeção do modelo na tabela de classificação (ou Matriz de Confusão). Para o efeito, precisa-se criar uma tabela com o resultado da classificação cruzada da variável resposta, de acordo com uma variável dicotômica em que os valores se derivam das probabilidades logísticas estimadas na regressão.

No entanto, é preciso definir uma regra de previsão, que dirá se houve acerto ou não da probabilidade estimada com os valores reais, pois as probabilidades variam de 0 a 1 enquanto os valores reais binários possuem valores fixos de 0 ou 1.

Geralmente, supor que se as probabilidades aproximam-se de 1 o indivíduo estimado pode ser classificado como $\hat{Y} = 1$ bem como de forma contrária, se o modelo estimar probabilidades perto de 0, classificado como $\hat{Y} = 0$. Neste caso é necessário definir um ponto de corte, usualmente na literatura utiliza-se 0.5 mas dependendo do estudo proposto pode não ser limitado a este nível (Hosmer e Lemeshow 2000).

Após determinar o ponto de corte, é importante avaliar o poder de discriminação do modelo, pelo seu desempenho portanto em classificar os eventos dos não eventos. Cria-se a Matriz de Confusão com as observações de Verdadeiro Positivo (VP), Falso Positivo (FP), Falso Negativo (FN) e Verdadeiro Negativo (VN) e em seguida determinam-se alguns parâmetros como a Precisão, Sensibilidade e Especificidade.

Tabela 3.2: Matriz de Classificação

		Valor Observado	
		$Y = 0$	$Y = 1$
Valor Estimado	$\hat{Y} = 0$	VP	FP
	$\hat{Y} = 1$	FN	VN

Fonte: Fawcett (2006).

3.3.8 Interpretação do Modelo Logístico

Interpreta-se o modelo fazendo a análise de sinais dos parâmetros verificando se estão de acordo com o esperado relativamente ao relacionamento lógico com a variável dependente, em seguida calcula-se os seus anti-logaritmos para encontrar a razão das chances ou seja,

encontrar a medida de associação que dá o valor da chance de um evento acontecer se sob as mesmas condições ele não acontece. A estimação dos parâmetros foi realizada através do método da máxima verosimilhança e importa referir que a partir desta técnica, foi estimada a probabilidade de ocorrência de gravidez em adolescentes dos 10-19 anos com base nas variáveis independentes.

3.4 Procedimentos de Análise

Os procedimentos utilizados para analisar os dados deste estudo foram os seguintes:

- Inicialmente foi realizada uma análise exploratória dos dados de modo a descrever de forma mais detalhada as características sócio-demográficas e culturais da amostra estudada recorrendo a tabelas de distribuição de frequências e gráficos;
- Em segundo lugar foi realizado o teste de hipóteses de qui-quadrado de Pearson e Fisher para verificar possíveis associações entre a ocorrência de gravidez na adolescência e as variáveis independentes;
- Em seguida foram verificados os pressupostos da regressão logística (multicolinearidade e independência) e posteriormente foram estimados os parâmetros do modelo de regressão logística e com base nessas estimativas realizou-se os testes de significância de Wald e de verosimilhança, seguindo-se para a avaliação de qualidade de ajuste do Modelo estimado. Para o efeito foram realizados do teste de Hosmer e Lemeshow, Pseudo R^2 e através da tabela de classificação.
- Por fim apresentou-se o modelo logístico final, sendo que neste é composto apenas pelas variáveis consideradas significativas para estimar a probabilidade de uma mulher engravidar na adolescência na região norte de Moçambique.

Capítulo 4

Resultados e Discussão

Neste capítulo, são apresentados todos os resultados obtidos de acordo com a metodologia apresentada no capítulo 3 e onde foi realizada a análise, interpretação, bem como a discussão dos resultados, de forma a responder os objectivos estabelecidos.

4.1 Análise Descritiva

Caracterização da Amostra

As figuras 4.1 e 4.2, ilustram respectivamente a distribuição percentual das mulheres por área de residência, nível de educação, índice de riqueza dos agregados familiares e os casos de gravidez na adolescência.

Em relação a área de residência pode-se observar que 64.9% das mulheres residem em zonas rurais e 35.1% em zonas urbanas. Quanto ao nível de educação, cerca de 35.3% sequer foram a escola, 45.2% possuem o nível primário, 19% o secundário contra apenas 6% que possuem o nível superior. Relativamente ao quintil de riqueza, cerca de 25.74% dos agregados familiares são muito pobres, 19.21% possuem o nível médio de riqueza e apenas 10.46% das pessoas são consideradas ricas, mas também verificou-se que cerca de 59.82% das mulheres entrevistadas ficaram grávidas durante a adolescência contra apenas 40.18% que não engravidaram na adolescência.

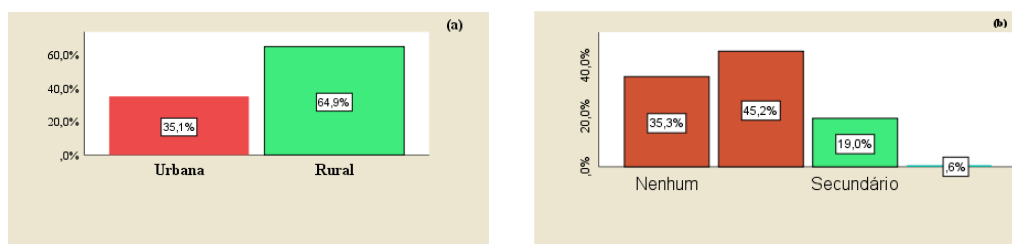


Figura 4.1: Distribuição das mulheres por área de residência (a) e por nível de educação (b).

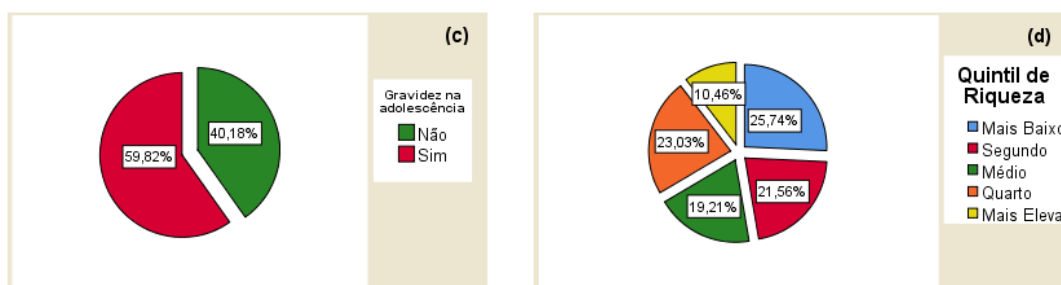


Figura 4.2: Distribuição percentual das mulheres por índice de riqueza dos agregados (c) e casos de gravidez na adolescência (d).

A tabela 4.1 apresenta a distribuição de frequências absolutas e percentuais das mulheres com idade compreendida entre os 15 à 49 anos residentes no norte de Moçambique, de acordo com as características socioeconómicas, demográficas e culturais.

A gravidez na adolescência foi mais frequente tanto nas zonas rurais assim como urbanas com cerca de 61,4% e 57% dos casos respectivamente. De um total de 1314 mulheres residentes em cabo delgado, cerca de 61,6% engravidaram na adolescência contra 38,4% que não engravidaram nesta faixa etária. Das três províncias do norte do país, as províncias de Niassa e Cabo delgado apresentaram menor número de casos de gravidez na adolescência com 429 e 505 casos respectivamente.

Em relação ao nível de escolaridade, observou-se gravidez na adolescência em 57,0% das mulheres sem nenhum nível de escolaridade, 65,3% com nível primário, 52,3% com nível secundário e 50% em mulheres com o nível superior. De um total de 2717 mulheres casadas ou em uma união marital, 31,3% não ficaram grávidas na adolescência enquanto que cerca de 68,7% engravidaram nesta faixa etária, no entanto, num total de 646 mulheres solteiras, apenas 19,8% engravidaram na adolescência.

A religião possui um papel fundamental na vida dos indivíduos, pois esta influencia no comportamento dos mesmos em relação a reprodutividade. De acordo com os resultados verificou-se que maior parte (1298) dos casos de gravidez na adolescência foram observados em mulheres pertencentes a religião islâmica com cerca de 60,7%, seguido da católica

com 59.1% e outras religiões com 56.7%. Quanto ao quintil de riqueza pode-se observar que dentre 997 mulheres com o quintil mais baixo, 622 (62.4%) ficaram grávidas na adolescência enquanto que 375 (37.6%) não engravidaram durante o mesmo período. Dentre 835 mulheres com o segundo quintil de riqueza, cerca de 503 (60.2%) engravidaram na adolescência e 332 (39.8%) não engravidaram na adolescência. Importa destacar uma redução substancial dos casos de gravidez na adolescência no quintil de riqueza mais elevado, pois num total de 405 mulheres, 201 (49.6%) engravidaram na adolescência contra 204 (50.4%) que não ficaram grávidas nesta faixa etária.

Em um universo de 2094 mulheres que sem qualquer exposição aos meios de comunicação social, houve 1247 (59.6%) que engravidaram na adolescência e 847 (40.4%) não engravidaram durante a adolescência. Em mulheres que possuem uma baixa exposição aos meios de comunicação, cerca de 655 (62.1%) ficaram grávidas durante a adolescência e 400 (37.9%) não engravidaram. No norte de Moçambique, apenas 110 mulheres pertencentes a esta amostra possuem uma exposição alta aos meios de comunicação das quais 60 (54.5%) engravidaram durante a adolescência enquanto que 50 (45.5%) não ficaram grávidas no mesmo período.

Relativamente ao uso de Métodos Contraceptivos (MC) pode-se observar que 3237 mulheres não usam qualquer método contraceptivo enquanto que apenas 636 usam pelo menos um métodos de contracepção. Nas mulheres que não usam quaisquer métodos, cerca de 1842 (56.9%) engravidaram na adolescência e 1395 (43.1%) que não ficaram grávidas. A percentagem de mulheres que engravidaram na adolescência aumenta de forma significativa nas que usam pelo menos um método contraceptivo com cerca de 475 (74.7%), representando um aumento de 17.5% em comparação com as que não usam quaisquer métodos.

A tecnologia tem evoluído de forma considerável nos últimos anos, como consequência disso a camada jovem principalmente passa longas horas em redes sociais abdicando-se de acompanhar os meios tradicionais de informação tais como a televisão e a rádio, este facto obriga as autoridades de saúde a inovar e usar a internet como um dos meios de difusão de informação para o combate de certas doenças e assim como na prevenção das mesmas. De um total de 3636 mulheres de que não tem acesso a internet, cerca de 2190 (60.2%) engravidaram na adolescência e 1446 (39.8%) não ficaram grávidas. Quanto ao planeamento familiar, em um total de 2651 mulheres que nunca ouviram falar, 1561 (58.9%) ficaram grávidas durante a adolescência contra 1090 (41.1%) que não engravidaram na adolescência enquanto que de um universo de 1222 mulheres já ouviram falar sobre o PF, 756 (61.9%) engravidaram e 466 (38.1%) não ficaram grávidas na adolescência.

De um total de 3760 mulheres que não consumiam tabaco, 2259 (60.1%) engravidaram

na adolescência e apenas 1501 (39.9%) não ficaram grávidas nesta faixa etária. Nas 113 mulheres que consumiam tabaco, existe um ligeiro equilíbrio entre as percentagens das mulheres que engravidaram e que não engravidaram durante a adolescência com cerca de 58 (51.3%) e 55 (48.7%) casos respectivamente. Observou-se com enorme insatisfação e preocupação que a idade mínima que algumas mulheres iniciaram as suas actividades sexuais foi de 8 anos. Dentre 2007 mulheres com idades entre os 15 aos 19 anos, apenas 975 (49%) engravidaram na adolescência enquanto que 1032 (51.4%), importa destacar o facto de que as mulheres que iniciaram as suas actividades sexuais entre os 8 aos 14 anos de idade, 1336 (75%) engravidaram na adolescência e apenas 446 (25%) não engravidaram na adolescência.

Com base nestes resultados pode-se verificar que em um total de 3170 mulheres que viram a sua primeira menstruação entre os 7 e 14 anos de idade, cerca de 1984 (62.4%) engravidaram na adolescência e 1186 (37.4%) não engravidaram na adolescência. Dentre 17 mulheres que viram sua primeira menstruação com pelo menos 20 anos idade, 10 (58.8%) ficaram grávidas durante a adolescência e 7 não engravidaram nesta faixa etária. Em relação a idade na primeira coabitação, observou-se que nesta amostra, a idade mínima foi de 10 anos, sendo que num total de 703 mulheres que tiveram a sua primeira coabitação na faixa etária dos 10 aos 14 anos de cerca de 643 (91.5%) das mesmas ficaram grávidas durante a adolescência e apenas 60 (8.5%) não engravidaram durante o período adolescente.

Tabela 4.1: Distribuição de frequências absolutas e percentuais das mulheres com idade compreendida entre os 15 à 49 anos residentes no norte de Moçambique, de acordo com as características socioeconómicas, demográficas e culturais.

Covariáveis	Gravidez na Adolescência		Total
	Não ($N_1 = 1556$)	Sim ($N_2 = 2317$)	($N = 3873$)
Área de Residência			
Urbana	585 (43.0%)	774 (57.0%)	1359
Rural	971 (38.6%)	1543 (61.4%)	2514
Província			
Niassa	429 (38.5%)	684 (61.5%)	1113
Cabo Delgado	505 (38.4%)	809 (61.6%)	1314
Nampula	622 (43.0%)	824 (57.0%)	1446
Nível de Escolaridade			
Nenhum	587 (43.0%)	779 (57.0%)	1366
Primário	607 (34.7%)	1142 (65.3%)	1749
Secundário	350 (47.7%)	384 (52.3%)	734
Superior	12 (50.0%)	12 (50.0%)	24
Estado Civil			
Solteiro	518 (80.2%)	128 (19.8%)	646
Casada/União marital	851 (31.3%)	1866 (68.7%)	2717
Viúva	31 (38.3%)	50 (61.7%)	81
Divorciada	156 (36.4%)	273 (63.6%)	429
Religião			
Católica	599 (40.9%)	866 (59.1%)	1465
Islâmica	840 (39.3%)	1298 (60.7%)	2138
Outras	117 (43.3%)	153 (56.7%)	270
Quintil de Riqueza			
Mais Baixo	375 (37.6%)	622 (62.4%)	997
Segundo	332 (39.8%)	503 (60.2%)	835
Médio	284 (38.2%)	460 (61.8%)	744
Quarto	361 (40.5%)	531 (59.5%)	892
Mais Elevado	204 (50.4%)	201 (49.6%)	405
Exposição aos meios de comunicação			
Nenhuma	847 (40.4%)	1247 (59.6%)	2094
Baixa	400 (37.9%)	655 (62.1%)	1055
Média	259 (42.2%)	355 (57.8%)	614
Alta	50 (45.5%)	60 (54.5%)	110
Uso de métodos contraceptivos			
Não	1395 (43.1%)	1842 (56.9%)	3237
Sim	161 (25.3%)	475 (74.7%)	636
Uso de Internet			
Não	1446 (39.8%)	2190 (60.2%)	3636
Sim	110 (46.4%)	127 (53.6%)	237
Ouviu falar sobre o PF			
Não	1090 (41.1%)	1561 (58.9%)	2651
Sim	466 (38.1%)	756 (61.9%)	1222
Consumo de Tabaco			
Não	1501 (39.9%)	2259 (60.1%)	3760
Sim	55 (48.7%)	58 (51.3%)	113
Idade na primeira relação sexual			
15 à 19 anos	1032 (51.4%)	975 (49%)	2007
8 à 14 anos	446 (25.0%)	1336 (75%)	1782
20 anos ou mais	78 (92.9%)	6 (7%)	84
Idade na primeira menstruação			
7 à 14 anos	1186 (37.4%)	1984 (62.6%)	3170
15 à 19 anos	363 (52.9%)	323 (47.1%)	686
20 anos ou mais	7 (41.2%)	10 (58.8%)	17
Idade na primeira coabitação			
15 à 19 anos	948 (40.1%)	1418 (59.9%)	2366
10 à 14 anos	60 (8.5%)	643 (91.5%)	703
20 anos ou mais	548 (68.2%)	256(31.8%)	804

4.2 Associação entre as variáveis independentes e a ocorrência de gravidez na adolescência

Na tabela 4.2 são apresentadas as estatísticas de teste do χ^2 de Pearson assim como o P-valor associado ao teste de Pearson de cada variável explicativa com o objectivo verificar uma possível dependência ou não com a ocorrência de gravidez na adolescência.

De acordo com a tabela 4.2 observa-se que a um nível de significância de 5% existem evidências suficientes para concluir que as variáveis, Religião, Nível de exposição aos meios de comunicação, Ouviu falar sobre o PF e Consumo de Tabaco não estão associadas ocorrência de gravidez na adolescência, pois apresentaram P-valor > 0.05, respectivamente 0.345, 0.202, 0.079 e 0.072 enquanto que as variáveis que não estão destacadas em negrito apresentam um P-valor menor ou igual a 0.05 e consequentemente estão associadas a ocorrência de gravidez na adolescência.

Tabela 4.2: Teste de independência entre as variáveis explicativas e a Ocorrência de gravidez na adolescência

Variável	Qui-quadrado de Pearson	P-valor
Área de Residência	7.179	.007
Província	7.745	.021
Escolaridade	44.399	.000
Estado Civil	521.602	.000
Religião	2.128	.345
Quintil de Riqueza	21.573	.000
Nível de exposição aos meios de comunicação	4.613	.202
Uso de métodos contraceptivos	9.924	.000
Uso de Internet	4.087	.043
Ouviu falar sobre o PF	3.095	.079
Consumo de Tabaco	3.496	.062
Idade na primeira relação sexual	372.695	.000
Idade na primeira menstruação	56.396	.000
Idade na primeira coabitação	554.789	.000

4.3 Análise dos Pressupostos do modelo de regressão logística

A análise de resíduos permite verificar se os pressupostos do modelo logístico foram atendidos, trata-se de uma etapa importante porque influencia na qualidade do modelo que será construído a posterior.

Verificação da Multicolinearidade

A presença de multicolinearidade em um modelo de regressão pode prejudicar a capacidade preditiva do mesmo, na tabela 4.3, são apresentados os valores de tolerância e o factor de inflação da variância (VIF) de cada variável independente. Verifica-se que o pressuposto

de ausência de multicolinearidade não foi violado, pois o VIF de cada uma das variáveis é menor que 10.

Tabela 4.3: Teste de Multicolinearidade

Variáveis	Multicolinearidade	
	Tolerância	VIF
Área de Residência	.584	1.711
Província	.964	1.037
Nível de Escolaridade	.673	1.485
Estado Civil	.928	1.077
Quartil de Riqueza	.532	1.880
Uso de MC	.921	1.086
Uso de Internet	.807	1.239
Idade na primeira relação sexual	.893	1.120
Idade na primeira menstruação	.961	1.041
Idade na primeira coabitação	.902	1.109

A tabela 4.4 ilustra os resultados do teste de Durbin-Watson onde é possível verificar que o pressuposto de Homocedasticidade não foi violado pois o P-valor < 0.05 .

Tabela 4.4: Teste de Durbin-Watson

Estatística de teste	P-valor
1.979	0.000

4.4 Modelo de Regressão Logística

Verifica-se que os pressupostos não foram violados, assim sendo, pode-se estimar o modelo de regressão logístico que inclui todas as variáveis consideradas significativas pelo teste qui-quadrado apresentado na tabela 4.2.

A tabela 4.5 ilustra as estimativas de máxima verossimilhança dos parâmetros, o erro padrão, a estatística de teste de Wald a significância dos parâmetros baseada na estatística de Wald, a razão de chances e os intervalos de confiança à 95%. Na tabela, os números em negrito indicam que determinada categoria da variável é estatisticamente significativa com base no teste de Wald a um nível de significância de 5%.

Tabela 4.5: Estimativa de Parâmetros do Modelo de Regressão Logística Múltipla

Variáveis	β	EP(β)	Wald	P-Valor	OR	IC 95%
Residência-Urbana (ref)						
Rural	-0.236	121	3.789	0.049	0.789	[0.622;1.002]
Província-Niassa (ref)						
Cabo Delgado	-0.163	0.105	2.424	0.120	0.849	[0.691;1.043]
Nampula	-0.044	0.106	0.168	0.682	0.957	[0.777;1.179]
Escolaridade-Nenhum (ref)						
Primário	0.359	0.093	14.946	0.000	1.432	[1.194;1.719]
Secundário	0.218	.146	2.232	0.135	1.244	[0.934;1.657]
Superior	-0.114	0.515	.049	0.825	.893	[0.325;2.448]
Estado Civil-Solteiro (ref)						
Casada/União marital	2.538	0.125	411.781	0.000	12.649	[9.899;16.162]
Viúva	2.431	0.297	67.055	0.000	11.373	[0.355;20.351]
Divorciada	2.450	0.167	215.882	0.000	11.593	[8.361;16.075]
Q. Riqueza-Mais baixo (ref)						
Segundo	-0.055	0.119	0.215	0.643	0.946	[0.750;1.194]
Médio	0.070	0.125	0.312	0.577	1.072	[0.839;1.371]
Quarto	0.216	0.147	2.158	0.142	1.241	[0.930;1.656]
Mais elevado	-0.031	0.205	0.023	0.879	0.969	[0.648;1.449]
Uso de MC - Não (ref)						
Sim	0.817	0.123	44.196	0.000	2.263	[1.779;2.879]
Uso de Internet - Não (ref)						
Sim	-0.033	0.202	0.027	0.870	0.967	[0.652;1.436]
Idade na primeira relação sexual						
15-19 anos (ref)						
8 à 14 anos	0.762	0.093	67.162	0.000	2.143	[1.786;2.572]
20 anos ou mais	-2.056	0.459	20.017	0.000	0.128	[0.052;0.315]
Idade na primeira menstruação						
7 à 14 anos (ref)						
15 à 19 anos	-0.178	0.107	2.771	0.096	0.837	[0.679;1.032]
20 anos ou mais	0.440	0.607	0.524	0.469	1.552	[0.472;5.100]
Idade na primeira coabitação						
15 à 19 anos (ref)						
10 à 14 anos	0.965	0.153	39.499	0.000	2.624	[1.942;3.544]
20 anos ou mais	-1.853	0.101	335.983	0.000	0.157	[0.129;0.191]
Intercepto	-1.786	0.194	84.593	0.000	0.168	...

ref-categoria de referência

Através da tabela 4.5 pode-se observar que as mulheres residentes em zonas rurais possuem menos chances de engravidar na adolescência relativamente às que residem em áreas urbanas ($\beta = -0.236$, $OR = 0.789$) se os efeitos dos restantes factores forem mantidos constantes. Observa-se através do teste de Wald que a área de residência é estatisticamente significativa, pois o P-value é menor ou igual a 0.05.

De acordo com os resultados, o teste de Wald conclui que as províncias de Cabo Delgado ($\beta = -0.163$, $OR = 0.849$) e Nampula ($\beta = 0.044$, $OR = 0.957$) não apresentaram coeficientes estatisticamente significativos para o modelo, com P-valores > 0.05 . Observa-se que as mulheres residentes nas províncias de Cabo Delgado e Nampula têm menos chances de engravidar na adolescência relativamente às residentes em Niassa.

Em relação ao nível de escolaridade, as mulheres com os níveis secundário e superior, não apresentaram respectivamente coeficientes estatisticamente significativos ($P - valor = 0.135$ e $P - valor = 0.825$). No entanto as que possuem o nível primário além de apresentar coeficientes estatisticamente significativos para o modelo, têm mais chances de engravidar na adolescência relativamente às que não possuem qualquer nível de escolaridade. Quanto ao estado civil, as mulheres Casadas ($\beta = 2.538$, $OR = 12.649$), Divorciadas ($\beta = 2.431$, $OR = 11.373$) e Viúvas ($\beta = 2.450$, $OR = 11.593$) apresentam cada uma delas no mínimo onze vezes mais chances de engravidar durante a adolescência.

As mulheres que usam métodos contraceptivos tem cerca de 2.263 mais vezes de engravidar na adolescência em relação às que não usam quais métodos contraceptivos. Em relação a idade na primeira relação sexual, pode-se verificar que as mulheres que iniciaram as suas actividades sexuais entre os 8 e 14 anos, possuem cerca de 2.143 mais chances de engravidar na adolescência enquanto que as que possuem 20 anos ou mais ($\beta = -2.056$, $OR = 0.128$) têm menos chances de engravidar na adolescência comparativamente às que pertencem a faixa etária dos 15 aos 19 anos.

Relativamente a idade na primeira coabitação verificou-se que as mulheres com idades entre 10 e 14 anos ($\beta = 0.965$, $OR = 2.624$) tem cerca de 2.624 mais vezes de engravidar na adolescência relativamente às que possuem entre 15 e 19 anos, no entanto observou-se menos chances para as que tem 20 anos ou mais. Importa destacar ainda que de acordo com estes resultados apresentados, conclui-se através do teste de Wald que os coeficientes dos factores quintil de riqueza, uso de internet e idade na primeira menstruação, não foram estatisticamente significativos para estimar os parâmetros do modelo de regressão logística múltipla pois apresentaram p-valores > 0.05 .

Avaliação da Significância dos Coeficientes e do Ajuste do Modelo

A tabela 4.6 apresenta o teste global do modelo nulo e os modelos em cada um dos passos (Step), Bloco (Block) e o modelo final (Model). A Partir dos valores de ($\chi^2 = 1450.707$ e $p\text{-valor}=0.000$) apresentados na tabela, pode-se rejeitar a hipótese nula, e concluir que pelo menos uma variável explicativa influencia na variável de interesse nomeadamente, a gravidez na adolescência.

Tabela 4.6: Teste Global do Modelo Logístico

		Qui-quadrado	Graus de Liberdade	P-valor
Step 1	Step	1450.707	21	.000
	Block	1450.707	21	.000
	Model	1450.707	21	.000

A tabela 4.7 ilustra as medidas de avaliação da qualidade de ajuste do modelo, onde verifica-se que o valor de -2 Log likelihood do modelo geral foi de 3767.906 e o Pseudo R^2 de Nagelkerke indica que o modelo explica cerca de 42.2% das variações registadas na variável dependente.

Tabela 4.7: Medidas de Avaliação da Qualidade de ajuste do Modelo

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	3767.906	0.312	0.422

As tabelas 4.8 e 5.1 (apresentada em apêndices) ilustram os resultados do teste de Hosmer e Lemeshow cujo objectivo é verificar a existência ou não de diferenças significativas entre as classificações realizadas pelo modelo e a realidade observada, este teste criou 10 grupos de tamanho aproximadamente igual onde em cada grupo é achado a frequência da variável dependente (Y), criando assim uma tabela de contingência. Em seguida foi calculada a frequência esperada em cada grupo a partir da probabilidade estimada do modelo de regressão logística ajustado. Comparando as frequências, a tabela de contingência sugere que a frequência observada é aproximadamente igual à frequência esperada para cada valor da variável dependente. Neste caso, o teste de Hosmer-Lemeshow indica um bom ajuste dos dados ao modelo de regressão logística múltipla, com $p\text{-value} = 0.056$.

Tabela 4.8: Teste Hosmer e Lemeshow

Step	Qui-quadrado	Graus de Liberdade	P-valor
1	15.780	8	.056

A tabela 4.9 ilustra a matriz de confusão ou quadro de classificação, que por sua vez, mostra a percentagem de observações correctamente classificadas segundo um nível de

probabilidade. Portanto, cerca de 77.7% das observações foram correctamente classificadas pelo modelo de regressão ajustado para um nível de probabilidade de 50%.

Tabela 4.9: Quadro de Classificação do Modelo Ajustado

		Predicted		% Correct
	Observado	Gravidez na adolescência		
		Não	Sim	
Step 1	Não	1032	524	66.3
	Sim	341	1976	85.3
% Global				77.7

Após testar a significância de cada variável (tabela 4.5) com o objectivo de identificar as variáveis significativas para o modelo, foram retidas todas as variáveis que apresentaram p-valores inferiores à 5%. A tabela ilustra os coeficientes das variáveis que fazem parte do modelo final e a sua respectiva codificação.

Tabela 4.10: Modelo logístico final

n	Variável	Categoria	Coeficiente
1	Área de Residência	Rural	0.720
2	Nível de Escolaridade	Primário	1.417
3	Estado Civil	Casada/União marital	12.451
4	Estado Civil	Viúva	11.304
5	Estado Civil	Divorciada	11.476
6	Uso de métodos contraceptivos	Sim	2.292
7	Idade na primeira relação sexual	8 à 14 anos	2.142
8	Idade na primeira relação sexual	20 ou mais anos	0.133
9	Idade na primeira coabitação	10 à 14 anos	2.663
10	Idade na primeira coabitação	20 ou mais anos	0.156
11	Intercepto	...	0.171

Assim, a probabilidade de uma mulher residente na zona norte de Moçambique ficar grávida na adolescência será dada por:

$$\pi(x_i) = \frac{e^{(0.171 + \sum_{n=1}^{10} Coeficiente_i * Resposta_i)}}{1 + e^{(0.171 + \sum_{n=1}^{10} Coeficiente_i * Resposta_i)}} \quad (4.1)$$

4.5 Discussão dos Resultados

Após analisar os possíveis factores associados à ocorrência de gravidez na adolescência entre mulheres dos 15 aos 49 anos, constatou-se que a área de residência, província, nível de escolaridade, estado civil, quintil de riqueza, uso de MC uso de internet, idade na primeira relação sexual, idade na primeira menstruação e idade na primeira coabitação possuem uma correlação estatisticamente significativa com a variável dependente gravidez na adolescência.

É consensual em todos os estudos que o início precoce da actividade sexual influencia para a ocorrência de gravidez na adolescência. As mulheres que tiveram sua idade da primeira coabitação entre os 10 e 14 anos possuem cerca de duas vezes mais chances de engravidar na adolescência em relação as que passaram a coabitar com os seus maridos entre os 15 e 19 anos. Estes resultados estão de acordo com os do inquérito demográfico e de saúde de 2022-23 revelando que o casamento é o principal indicador da exposição regular das mulheres ao risco de gravidez, sendo que as populações em que a idade do primeiro casamento é baixa tendem a ter uma maternidade precoce e uma fecundidade elevada. Verificou-se que quanto menor o nível de escolaridade da rapariga, maiores serão as chances de GA tal como foi constatado por Mungomane & Barbosa (2022), onde também os mesmos associam este fenómeno com a prática dos ritos de iniciação nas comunidades.

Relativamente aos factores económicos, os resultados deste estudo revelaram que o quintil de riqueza não é estatisticamente significativo para estimar a chance de ocorrência de GA na zona norte de Moçambique, este facto contradiz aos estudos realizados por UNICEF (2010) e Pires & Josaphat (2016), pois estes afirmaram que a pobreza é um factor determinante para a ocorrência de casamentos prematuros, que muitas vezes resulta em gravidezes precoce em Moçambique e em muitas comunidades, com maior incidência na região norte do país, os adultos forçam as meninas como maneira de reduzir as despesas dos agregados familiares. Entretanto, Silva & Clapis (2023) consideram que o desemprego, dos pais da menina ou outro membro familiar que atue como líder financeiro da família, é um fator deflagrador para a ocorrência de GA.

Observou-se que os factores como província, exposição aos meios de comunicação e religião não foram estatisticamente significativas para estimar a chance de ocorrência de gravidez na adolescência embora Pinheiro (2019) tenha considerado que a GA esteja também associada aos valores culturais e organização social da sociedade. De acordo com os resultados deste estudo, as mulheres residentes em áreas rurais possuem menos chances de ocorrência de GA, um facto inesperado e não está de acordo com os estudos de César et al. (2000) e Alberto et al. (2011), pois afirmam que em áreas urbanas com maior acesso aos serviços de saúde e planeamento familiar reduz o risco de ocorrência deste fenómeno.

Capítulo 5

Conclusões e Recomendações

5.1 Conclusões

O principal objectivo deste estudo foi analisar os factores determinantes à ocorrência da Gravidez em adolescentes dos 10-19 anos no norte de Moçambique. Para o efeito, construiu-se um modelo de regressão logística binária e foi utilizada a base de dados do Inquérito demográfico e de Saúde 2022-2023. Tendo em conta os objectivos estabelecidos e de acordo com os resultados obtidos, obteve-se as seguintes constatações:

- Observou-se que 35.1% das mulheres residem no meio urbano e 64.9% residem no meio rural, cerca de 35.3% sequer foram a escola, 45.2% possuem o nível primário, 19% o secundário e apenas 6% que possuem o nível superior, sendo que 25.74% dos agregados familiares tem índice de riqueza mais baixo e 10.46% apresentam índice mais alto.
- Constituem factores determinantes à ocorrência da gravidez em adolescentes dos 15-19 anos no norte de Moçambique área de residência, província, nível de escolaridade, estado civil, uso de MC, idade na primeira relação sexual, idade na primeira menstruação e idade na primeira coabitação. Quanto menores forem as idades da primeira relação sexual e coabitação, maiores são as chances de engravidar na adolescência.
- Conclui-se ainda que o modelo de regressão logística obtido ajustou-se aos dados e mostrou uma capacidade de classificações correctas de (77.7%), modelo este que foi usado para prever a probabilidade de ocorrência de GA na zona norte de Moçambique.

5.2 Recomendações

- Sensibilizar as entidades institucionais, governamentais e reguladoras, para a importância dos resultados divulgados no âmbito da realização do IDS. Esta questão é crucial para realçar os impactos socioeconómicos da gravidez prematura e seus factores associados, mas também para delimitar soluções de médio e longo prazo no sentido de combater a evolução deste fenómeno.
- Promover a educação sexual, que deve ser debatida nas escolas, famílias, redes sociais e espaços de convivência dos adolescentes. A educação sexual deve abordar a importância do comportamento sexual responsável, o respeito ao outro, a igualdade e equidade de género, e a prevenção de infecções sexualmente transmissíveis. Garantir o acesso a métodos contraceptivos, como camisinhas, diafragma, espermicida, anticoncepcionais injetáveis, pílulas, DIU, entre outros. O uso da caminha é recomendado independentemente do tipo de contraceptivo que for indicado.
- Sensibilizar os pais e encarregados de educação, médicos e outras entidades de saúde para a necessidade de abordar este problema de forma mais ampla através de debates públicos assim como a expansão dos serviços de planeamento familiar, criar uma estratégia a nível nacional de prevenção casos de gravidez na adolescência e minimização das principais consequências.

5.3 Limitações

A principal limitação durante a aplicação da regressão logística foi a impossibilidade de obtenção de informação sobre mortalidade materna e algumas variáveis relacionadas aos factores étnicos, comportamentais, ambientais assim como a impossibilidade de realizar uma análise cruzada com o inquérito anterior devido a falta de recursos como o tempo e recursos humanos.

Referências

- [1] Agência de saúde sexual e reprodutiva das Nações Unidas (UNFPA 2013). Gravidez na Adolescência - Desafios e Respostas de Moçambique.
- [2] Alberto, S. A., Lima, L. C. de., Rodrigues, R. N., & Machado, C. J.. (2011). Fatores associados aos óbitos neonatais e pós-neonatais em Moçambique. *Revista Brasileira De Estudos De População*, 28(1), 203–216. <https://doi.org/10.1590/S0102-30982011000100011>.
- [3] Belfiore, P. (2015). Estatística aplicada: Para cursos de Administração, Contabilidade e Economia com Excel e SPSS. Elsevier Brasil.
- [4] Cabral, C.I.S. (2013). Aplicação do Modelo de Regressão Logística num Estudo uso de métodos contraceptivos (Dissertação de Mestrado). Universidade de Lisboa.
- [5] Chichava, A. (2020). Análise dos Factores Determinantes no Uso de Métodos Contraceptivos na População Feminina em Moçambique (Monografia do curso de Licenciatura em Estatística). Universidade Eduardo Mondlane, Maputo.
- [6] Cervo, Amando L; Bervian, A. (2002). Metodologia Científica. 5. ed. São Paulo: Prentice Hall.
- [7] César, C. C., Miranda-Ribeiro, P., Abreu, D. M. X. Efeito-idade ou efeito-pobreza? Mães adolescentes e mortalidade neonatal em Belo Horizonte. *Revista Brasileira de Estudos de População*, Campinas, v.17, n.1/2, p.177-196, 2000.
- [8] Corrar, L., Paulo, E., Dias Filho, J. M., & Rodrigues, A. (2011). Análise multivariada para os cursos de administração, ciências contábeis e economia.
- [9] Damiani, F. E. (2003). Gravidez na adolescência: a quem cabe prevenir? *Revista Gaúcha de Enfermagem*, 24(2), 161-161.
- [10] da Costa, M. M. M., & de Freitas, M. V. P. (2021). A gravidez na adolescência e a feminização da pobreza a partir de recortes de classe, gênero e raça. *Revista Direitos Culturais Santo ngelo*— v, 16(40), 5-23.

- [11] de Jesus Silva, MM, & José Clapis, M. (2023). *Risco de Depressão na Gravidez na Percepção de Dois Profissionais de Saúde*. Enfermagem Em Foco , 14 , 1–8. <https://doi.org/10.21675/2357-707X.2023.v14.e-202321>
- [12] Fávero, L. P., & Belfiore, P. (2017). *Manual de análise de dados: estatística e modelagem multivariada com Excel, SPSS e Stata*. Elsevier Brasil.
- [13] Gil, António (1991). Métodos e técnicas de pesquisa social. São Paulo: Atlas.
- [14] Gonzalez, L. d. (2018). *Regressão logística suas Aplicações* . acessado julho 30, 2023, em <https://monografias.ufma.br/jspui/bitstream/123456789/3572/1/Leandro-Gonzalez>.
- [15] Gonçalves, H., González-Chica, D. A., Menezes, A. M., Hallal, P. C., Araújo, C. L., & Dumith, S. C. (2013). Conhecimento sobre a transmissão de HIV/AIDS entre adolescentes com 11 anos de idade do Sul do Brasil. pp. 420-431. acesso agosto 11, 2024.
- [16] Guimarães, A. M., (2019). *Estatística: Teste Exato de Fisher e Teste de Qui-Quadrado usando R*. Acessado 13 de julho, 2023, em: <https://medium.com/omixdata/estat%C3%ADstica-teste-exato-de-fisher-e-teste-de-qui-quadrado-usando-r-4ee496da37fc>.
- [17] Gujarati, D. N., & Porter, D. C. (2011). *Econometria básica-5*. Amgh Editora.
- [18] Hair, J.R., Anderson, R.E., Tatham, R.L., e Black, W.C. (2005). *Análise multivariada de dados*. Porto Alegre: Bookman.
- [19] Hosmer, D. W., Jovanovic, B., & Lemeshow, S. (1989). Best subsets logistic regression. *Biometrics*, 1265-1270.
- [20] Hosmer, D. W.; Lemishow, S. (2000). *Applied Logistic Regression*. 2nd ed. New York
- [21] Instituto Nacional de Estatística (INE) e ICF. 2024. *Inquérito Demográfico e de Saúde em Moçambique 2022–23*. Maputo, Moçambique e Rockville, Maryland, EUA: INE e ICF.
- [22] Kleinbaum, D., Kupper, L., Muller, K., & Nizam, A. (1998). *Applied Regression Analysis and other Multivariate Methods* (3 ed.).
- [23] Maroco, J. (2007). *Análise estatística Com Utilização do SPSS*, 3ª Edição. Lisboa: Edição Silabo.
- [24] Ministério da Saúde (MISAU), Instituto Nacional de Estatística (INE), e ICF, 2015. *Inquérito de Indicadores de Imunização, Malária e HIV/SIDA em Moçambique 2015*. Maputo, Moçambique. Rockville, Maryland, EUA: INS, INE, e ICF.

- [25] Pires, P., & Josaphat, J.(2016). Avaliação dos factores determinantes de casamento prematuro e de gravidez na adolescência, Nampula, Moçambique, 2016.
- [26] Mungomane, I. H. L., & Barbosa, A. G. (2022). 21.As implicações da desistência escolar da rapariga a partir dos valores educativos dos ritos de iniciação: caso escola primária, distrito de Gurué: Makachamiho anam'mphachuwela nahano ahiya ohussera n'nthowa na ikano onahusseraiye alakiwaka: Eneneryo ya vascola Primaria Y yawana Gurue. JINGA SEPÉ: revista nternacional e ulturas, Línguas fricanas rasileiras (ISSN: 2764-1244), 2(Especial I), 383-397. Disponível: <https://revistas.unilab.edu.br/index.php/njingaesape/article/view/1019>.
- [27] Mesquita, P.S.B. (2014). Um modelo de regressão logística para avaliação do nível de conhecimento em relação ao uso de métodos de planeamento familiar (Dissertação de Mestrado). Universidade Estadual do Norte Fluminense.
- [28] Mubate, J. L. J. F.. (2024). Educação sexual no programa do Ensino Secundário Geral em Moçambique: análise dos Planos Estratégicos da Educação 1999-2021. Ensaio: Avaliação E Políticas Públicas Em Educação, 32(122), e0244081. <https://doi.org/10.1590/S0104-40362024003204081>.
- [29] Ministério da Saúde (MISAU), Instituto Nacional de Estatística (INE) e ICF International (ICFI). Moçambique Inquérito Demográfico e de Saúde 2011. Calverton, Maryland, USA: MISAU, INE e ICFI.
- [30] Ministério da Saúde (MISAU), Instituto Nacional de Estatística (INE), e ICF, 2015. Inquérito de Indicadores de Imunização, Malária e HIV/SIDA em Moçambique 2015. Maputo, Moçambique. Rockville, Maryland, EUA: INS, INE, e ICF.
- [31] Nhambire, A. (2012).Avaliação do atendimento dos utentes dos serviços públicos da Direcção de Identificação Civil da Cidade de Maputo (Monografia do curso de Licenciatura em Estatística). Universidade Eduardo Mondlane, Maputo.
- [32] OMS (2014). Maternal, newborn, child and adolescent health: Adolescent development.
- [33] Sales, A. L. F., Villacorta, H., Reis, L., & Mesquita, E. T. (2005). Anemia como fator prognóstico em uma população hospitalizada por insuficiência cardíaca descompensada. *Arquivos Brasileiros de Cardiologia*, 84, 237-240.
- [34] Salotti, B. M. e Yamoto, M. (2005). Informação contábil: estudos sobre sua divulgação no mercado de capitais. São Paulo: Atlas,2006.
- [35] Silva, P. (2016). *Modelos de Regressão Linear e Logística utilizando o software R*. Porto, Portugal.

- [36] Silva, J. L. P. e ., & Surita, F. G. C.. (2012). Gravidez na adolescência: situação atual. *Revista Brasileira De Ginecologia E Obstetrícia*, 34(8), 347–350. <https://doi.org/10.1590/S0100-72032012000800001>.
- [37] Silveira, M. B. G., Barbosa, N. F. M., Peixoto, A. P. B., Xavier, É. F. M., & Júnior, S. F. A. X. (2021). Aplicação da regressão logística na análise dos dados dos fatores de risco associados à hipertensão arterial. *Research, Society and Development*, 10(16), e20101622964-e20101622964.
- [38] Souza, M. de L. de, Lynn, F. A., Johnston, L., Tavares, E. C. T., Bruggemann, O. M., & Botelho, L. J. (2017). Fertility rates and perinatal outcomes of adolescent pregnancies: a retrospective population-based study. *Revista Latino-Americana de Enfermagem*. (Online), 25, e2876-e2876. <https://doi.org/10.1590/1518-8345.1820.2876>.
- [39] Sousa, W. B. D. (2018). Estimadores de máxima verosimilhança: casos que não satisfazem as condições de regularidade.
- [40] Tavares, M. (2007). Estatística aplicada à administração. *Secretaria de Educação à Distância do Ministério da Educação Sistema Universidade Aberta do Brasil-UAB. Diretoria do Departamento de Políticas em Educação a Distância-DPEAD*. Brasília: MEC.
- [41] UNICEF. Pobreza infantil e disparidades em Moçambique 2010. Relatório das Nações Unidas. Unicef. Maputo. 2011. Disponível em: <http://www.unicef.org.mz/>.
- [42] Yin, R. K. (2005). Estudo de caso: planejamento e métodos. 3. ed. Porto Alegre: Bookman

Apêndices

Tabela 5.1: Tabela de Contingência para o teste de Hosmer e Lemeshow

	GA=Não		GA=Sim		Total
	Observado	Esperado	Observado	Esperado	
1	348	340.106	38	45.894	386
2	315	308.243	72	78.757	387
3	251	267.056	136	119.944	387
4	188	199.779	205	193.221	393
5	135	134.992	253	253.008	388
6	97	106.175	288	278.825	385
7	99	79.964	288	307.036	387
8	65	59.655	321	326.345	386
9	41	37.791	345	348.209	386
10	17	22.238	371	365.762	388