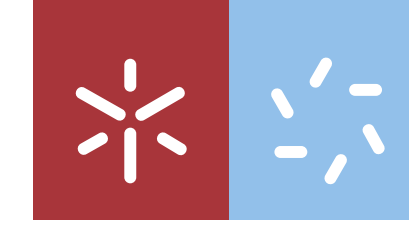




Yoann Rodrigues Folha

Tendências e determinantes acerca do conhecimento do HIV/SIDA em Moçambique

Universidade do Minho
Escola de Ciências





Universidade do Minho
Escola de Ciências

Yoann Rodrigues Folha

**Tendências e determinantes acerca do
conhecimento do HIV/SIDA em
Moçambique**

Dissertação de Mestrado

Mestrado de Estatística

Trabalho efetuado sob a orientação do

Professor Doutor Luís Filipe Meira Machado

e da

Doutora Carla Maria Gonçalves de Macedo Moreira

Fevereiro 2021

DECLARAÇÃO

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.



<https://creativecommons.org/licenses/by/4.0/>

AGRADECIMENTOS

Um primeiro agradecimento aos meus orientadores, o professor Luís Machado e a investigadora Carla Moreira pela disponibilidade que tiveram nas sucessivas reuniões e pelas observações que nelas surgiram.

Agradeço ao doutor Francisco Mbofana pela disponibilidade das bases de dados através das quais se fez o trabalho.

Queria agradecer aos meus pais, simplesmente por tudo.

E um último agradecimento aos professores que tive nesta universidade pela transmissão de conhecimento.

DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducentes à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

RESUMO

Moçambique é um dos países mais tocados pelos vírus do HIV/SIDA com cerca de 2,1 milhões de moçambicanos a viverem com ele (Governo de Moçambique, 2008), apresentando uma prevalência significativa entre a população jovem, razões que motivaram para que se fizesse um estudo sobre o HIV/SIDA nesta parte da população. Mais especificamente, um estudo que se concentrasse nas pessoas com idades compreendidas entre os 15 e os 24 anos e que incidisse no conhecimento que elas possuíam acerca da doença. O estudo teve como base inquéritos realizados nos anos de 2009, 2011 e 2015, sendo que inicialmente estava prevista ter resultados dos inquéritos do ano de 2003. Das bases de dados retiveram-se informações acerca da pessoa que estava a ser inquirida, dos conhecimentos que tem sobre HIV/SIDA e dos comportamentos que a pessoa em questão tem como o uso de preservativo, o teste do HIV/SIDA, entre outros.

O objetivo deste trabalho é identificar os determinantes que mais influenciam um maior ou menor conhecimento acerca do HIV/SIDA e também tentar observar tendências do mesmo ao longo dos anos em que se realizou o estudo. Foi contruído um modelo de regressão logístico de modo a verificar variáveis com uma influência negativa ou positiva no conhecimento do HIV/SIDA, sendo a variável com esta informação binária fazia mais sentido utilizar este tipo de modelo.

Palavras-chave: HIV/SIDA, conhecimento acerca do HIV/SIDA, determinantes, tendências, regressão logística

ABSTRACT

Mozambique is one of the countries most affected by the HIV/AIDS virus with about 2.1 million Mozambicans living with it (Government of Mozambique, 2008), presenting a significant prevalence among the young population, reasons that motivated a study on HIV/AIDS part of the population. More specifically, a study that focuses on people between the age of 15 and 24 and which focused on their age of having about the disease. The study was based on surveys conducted in the years 2009, 2011 and 2015, and initially it was planned to have results of the surveys of the year 2003, the databases withheld information about the person being surveyed, the knowledge he has about HIV/AIDS and the behaviors that the person in which he/she has such as condom use, the HIV/AIDS test, among others.

The objective of this work is to identify the determinants that most influence a greater or lesser knowledge about HIV/AIDS and also to try trends of the same over the years in which the study was carried out. It was a logistic regression model in order to verify variables with a negative or positive influence on the knowledge of HIV/AIDS, and the variable with this binary information is more sense to use this type of model.

Keywords: HIV/AIDS, knowledge about HIV/AIDS, determinants, trends, logistic regression

Índice

Página

1 Introdução

1.1 Apresentação do tema e dos objetivos do trabalho.....	1
1.2 Dados dos inquéritos realizados.....	2
1.3 Metodologia estatística utilizada.....	5

2 Análise dos dados dos inquéritos sobre o HIV/SIDA

2.1 Análise estatística na forma descritiva	12
2.2 Análise estatística na forma de modelo de regressão logística	26
2.3 Conclusões	30

3 Referências	31
---------------------	----

Lista de tabelas

	Página
Tabela 1 Lista de variáveis relativas ao respondente	3
Tabela 2 Lista de viáveis relativas ao conhecimento do HIV/SIDA	4
Tabela 3 Lista de variáveis relativas ao comportamento do respondente	4
Tabela 4 Informação acerca das características dos respondentes presentes na base de dados.	12
Tabela 5 Informação acerca dos comportamentos e conhecimento acerca do HIV.	13
Tabela 6 Tendências dos comportamentos dos jovens moçambicanos entre 2009 e 2015.	14
Tabela 7 Tendências no conhecimento dos jovens moçambicanos acerca do HIV/SIDA ao longo dos anos do estudo e divididos pelo sexo dos respondentes e por área de residência e categoria etária.	15
Tabela 8 Tendências do conhecimento dos jovens moçambicanos sobre o HIV/SIDA ao longo dos anos do estudo.	15
Tabela 9 Tabela cruzada das proporções de jovens moçambicanos a utilizarem preservativos com o mais recente parceiro sexual.	16
Tabela 10 Tabela cruzada com as proporções de jovens moçambicanos a terem feito o teste de HIV.	17
Tabela 11 Tabela cruzada com as proporções de jovens moçambicanos que já ouviram falar do HIV/SIDA.	18
Tabela 12 Resultados do modelo de regressão logístico.	27
Tabela 13 Comparação do valor AIC do modelo obtido pelo método stepwise com o modelo completo.	28
Tabela 14 Teste de verosimilhança entre os modelos completos e o modelo obtido pelo método stepwise.	29

Lista de figuras

	Página
Figura 1 Distribuição etária dos respondentes que têm conhecimentos acerca do HIV e os que não têm ao longo dos anos.	20
Figura 2 Evolução da frequência absoluta do conhecimento do HIV ao longo dos anos pelos níveis de escolaridade.	21
Figura 3 Evolução da frequência absoluta do conhecimento acerca do HIV ao longo dos anos pelos diferentes estatutos socioeconômicos.	22
Figura 4 Evolução da frequência absoluta, ao longo dos anos, dos inquiridos que têm e os que não têm conhecimentos acerca do HIV nas diferentes áreas residência.	22
Figura 5 Distribuição etária dos respondentes que sabem e os que não sabem que o preservativo o risco de ter HIV nos diferentes anos do estudo.	23
Figura 6 Evolução, ao longo dos anos, da frequência absoluta dos inquiridos que sabem e os que não sabem que o preservativo reduz o risco de ter HIV nos diferentes níveis de escolaridade.	24
Figura 7 Relação da frequência absoluta dos níveis de educação e de estatuto socioeconômico com a área de residência (1 = Urbano, 2 = Rural).	25

Lista de siglas

HIV - Human Immunodeficiency Virus (Vírus da imunodeficiência humana)

SIDA – Síndrome da Imunodeficiência Adquirida

DST – Doença sexualmente transmissível

AIC – Akaike information criterion (Critério de informação de Akaike)

1 Introdução

1.1 Apresentação do tema e dos objetivos do trabalho

Moçambique é um país com aproximadamente 29 milhões de pessoas, onde o primeiro caso de AIDS foi diagnosticado em 1986 (Machel, 2010). Em termos epidemiológicos, vivem uma epidemia generalizada de HIV sendo o país africano de expressão portuguesa e um dos países do mundo onde o número de infetados com o vírus HIV/SIDA é maior. Segundo dados do governo de Moçambique estima-se que 2,1 milhões de moçambicanos vivem com o vírus, a estimativa mais recente da prevalência do HIV nas pessoas com idades entre os 15 e os 49 é de 13,5% (Governo de Moçambique, 2018), cerca de 15% das grávidas no mesmo intervalo etário tem o vírus. A prevalência varia significativamente por sexo, idade e geografia, sendo estimada em 15,4% nas mulheres e meninas, e 10,1% nos homens e meninos (Audet et al., 2010). A incidência é mais alta nas meninas adolescentes e em mulheres jovens de 15 a 24 anos assim como nos homens entre os 20 e 29 anos. A prevalência entre meninas adolescentes de 15 a 19 anos é estimada em 6,5% e 13,3% entre as jovens de 20 e 24 anos, em comparação, ela é de 1,5% e 5,3 nas respetivas classes etárias no sexo masculino (Ministério Moçambicano da Saúde, 2008).

Dados publicados pelo ministério da saúde de Moçambique indicava que a prevalência tinha aumentado de 11,5% para 13,2% em 2015, tendo o Diretor Nacional de Saúde Pública, Francisco Mbofana, que chefiou a equipa de investigação afirmado que os intervalos de confiança se sobrepunham e que por isso esse aumento não poderia ser estatisticamente significativo, apontando a expansão dos tratamentos como uma causa para esse aumento do número de infetados pois eles vivem mais tempo. Por seu turno, a Secretária Executiva Adjunta do Conselho Nacional de Combate ao HIV/SIDA, Idalina Libombo sublinhava uma baixa no nível de conhecimento do HIV principalmente na camada mais jovem sublinhando a necessidade de ter esta parte da população como público alvo (Ministério moçambicano da saúde, 2017).

Um inquérito realizado em Moçambique indicava que 11,1% das mulheres e 3,5% dos homens entre os 15 e os 24 anos tinham HIV. A prevalência era superior nas mulheres, aumentando com a idade e sendo maior nas áreas urbanas (10,1%) do que nas rurais (6,4%). O inquérito mostrava também que três quartos dos participantes sabiam que se podiam proteger do vírus usando preservativo (INSIDA, 2009).

A literatura evidencia a correlação entre o aumento do conhecimento compreensivo e a redução da incidência e prevalência do HIV/SIDA em regiões caracterizadas pela elevada prevalência de HIV. Tendo em conta as últimas tendências na prevalência do HIV entre adolescentes e jovens em Moçambique, é crucial determinar o nível de conhecimento do HIV entre os jovens de forma a se monitorizar os programas e prevenção, identificando lacunas e afinando as atividades de divulgação.

Este trabalho tem como objetivo descrever tendências recentes de comportamentos e conhecimentos relacionados ao HIV entre adolescentes e jovens (15

e 24 anos) em Moçambique entre 2009 e 2015, com base nos questionários nacionais realizados em 2009, 2011 e 2015.

Existem outros trabalhos que trataram este mesmo tema, como é o exemplo de um estudo que também procurava analisar as tendências e determinantes acerca do conhecimento do HIV nas mulheres do Bangladesh entre os anos 2007 e 2014 (Sheik et Khan, 2017) ou outro em que estudaram o mesmo tema, mas para os jovens entre os 15 e os 24 anos na Nigéria entre 2003 e 2013 (Oginni, Abebajo et Ahonsi, 2017). Para além dos anos do estudo serem próximos daqueles que se verificam no nosso trabalho também utilizou o mesmo método de análise estatística que será apresentado na Seção 2.2 do relatório, ou seja, o modelo de regressão logística.

1.2 Dados dos inquiridos realizados

Os dados utilizados neste trabalho referem-se a questionários feitos nos anos 2009, 2011 e 2015. Tendo, no total, uma base de dados com 42024 observações de pessoas com idades entre os 15 e os 64 anos. Dessas 42024 observações, 11212(26,7%) são do ano 2009, 17780(42,3%) de 2011 e 13032(31%) de 2015. Dois terços são de inquiridos do sexo feminino, temos 22709 observações (54%) que correspondem a pessoas que vivem em zonas rurais. Com respeito ao nível da educação, 21,4% dos inquiridos não frequentaram a escola, 51,9% tiveram o ensino primário, 24,3% o ensino secundário e 2,4 foram até ao ensino superior, por fim, 97,7% dos respondentes disseram ter conhecimento acerca do vírus do SIDA.

Nesta mesma base de dados temos que 15906 observações (37,9%), são de pessoas entre os 15 e os 24 anos, o intervalo de idades do estudo a ser realizado, logo, a primeira coisa que se fez foi, da base de dados que nos foi disponibilizada, analisar apenas as observações de inquiridos com idades compreendidas entre os 15 e os 24 anos.

A base de dados que nos foi entregue estava dividida por ano. Mais, para os anos 2011 e 2015, estava igualmente dividida de acordo com o género dos inquiridos. Por esse motivo, tivemos acesso a um total de 5 bases de dados com milhares de observações e cada uma com centenas de variáveis. Convém também salientar que, em alguns casos as designações para variáveis comuns eram diferentes. Todos estes factos dificultaram o nosso trabalho sendo que em certos casos houve mesmo a necessidade de criar variáveis de modo a complementar a informação contida em algumas variáveis. Foi necessário fazer a junção das 5 bases dados. Tendo em conta uma maior similaridade, optou-se por começar por se juntar as bases de dados dos anos de 2011 e 2015 criando nelas a variável do sexo do inquirido. De seguida fez-se a junção das 3 restantes bases de dados dos respetivos anos criando nelas a variável ano da entrevista.

Para fazer as referidas junções de bases de dados foi necessário ter variáveis comuns entre as diferentes bases de dados, fazer com que estas tenham o mesmo

posicionamento em todas as bases de dados e sobretudo que elas tenham interesse no trabalho que se pretende desenvolver. Para a seleção de variáveis a analisar, de entre as centenas de variáveis existentes nas bases de dados originais, as variáveis escolhidas para a base de dados utilizada no trabalho podem ser divididas em três categorias de acordo com as Tabelas 1, 2 e 3.

Tabela 1 Lista de variáveis relativas ao respondente.

Variável	Categorias
Idade	
Idade categorizada	1 – 15-19 anos 2 – 20-24 anos
Sexo	1 – Masculino 2 – Feminino
Ano da entrevista	2009 2011 2015
Área de residência	1 – Urbano 2 – Rural
Nível educacional	0 – Sem escolaridade 1 – Primário 2 – Secundário 3 – Superior
Estatuto socioeconómico	1 – Muito pobre 2 – Pobre 3 – Médio 4 – Rico 5 – Muito rico
Se o respondente trabalha	0 – Não 1 – Sim
Estado civil	0 – Nunca casou 1 – Casado/a 2 – Vive com alguém 3 – Viúvo/a 4 – Divorciado 5 – Vive sozinho/a
Teve uma DST no último ano	0 – Não 1 – Sim

Tabela 2 Lista de variáveis relativas ao conhecimento do HIV/SIDA.

Variável	Categorias
Conhecimento do HIV/SIDA	0 – Não 1 – Sim
Ter apenas 1 parceiro reduz o risco de ter SIDA	0 – Não 1 – Sim
O uso de preservativo reduz o risco de ter SIDA	0 – Não 1 – Sim

Tabela 3 Lista de variáveis relativas ao comportamento do respondente.

Variável	Categorias
Idade do 1º ato sexual	
Idade do 1º ato sexual categorizada	0 – Ainda não teve relações sexuais 1 – 8-14 anos 2 – 15-19 anos 3 – 20-24 anos
Nº de parceiros/as sexuais, excluindo esposo/a, no último ano	0 – 0 1 – 1 2 – 2 ou mais
Uso de preservativos no último ato sexual	0 – Não 1 – Sim
Teste do HIV/SIDA	0 – Não 1 – Sim
Uso de preservativos no último ato extraconjugal, no último ano	0 – Não 1 – Sim
Tampo desde o último ato sexual (em meses)	

Como referido na introdução, o objetivo do trabalho é descrever tendências do conhecimento e comportamentos dos jovens moçambicanos entre os 15 e os 24 anos relativamente o vírus do HIV/SIDA. Numa primeira parte ela será feita através de estatística descritiva, verificar as proporções de certos comportamentos e do conhecimento em relação ao vírus do HIV/SIDA e analisar as suas diferenças nos anos de estudo, no sexo, na área de residência, entre outras variáveis. Numa segunda parte deste trabalho, as tendências do conhecimento e comportamentos dos jovens moçambicanos será analisada através da construção de modelos estatísticos, nomeadamente pela construção de um modelo de regressão logística e de um modelo de efeitos mistos, onde se procura analisar o efeito das diferentes variáveis explicativas numa variável resposta binária (ter ou não o conhecimento do HIV). Convém salientar que originalmente esta variável resposta continha uma observação cuja resposta a esta pergunta foi “não sabe”, codificada na base de dados aparece com o valor de 9, sendo as correspondentes observações removidas da base de dados.

1.3 Metodologia estatística utilizada

Nesta Secção serão apresentadas de forma sucinta as diferentes metodologias utilizadas neste trabalho. Nomeadamente, o teste de Kruskal-Wallis para comparar os valores medianos para as idades do primeiro ato sexual nos diferentes anos (Tabela 3), escolheu-se este teste por permitir esta comparação para mais de dois grupos independentes sem recorrer ao pressuposto da normalidade dos dados. O teste de comparação de proporções, utilizado para comparar as proporções nos três anos (Tabelas 3 e 4). Fez-se também uma breve introdução aos modelos lineares generalizados de modo a servir de base para apresentação e discussão dos resultados do modelo de regressão logística, utilizado na Secção 2.2 deste relatório. Finalmente, será efetuada uma breve apresentação ao modelo de efeitos mistos.

Teste de Kruskal-Wallis

O teste de Kruskal-Wallis é um teste não paramétrico usado para comparar dois ou mais grupos independentes. Este teste pode ser visto como uma generalização do teste de Mann-Whitney. O equivalente teste paramétrico é o teste F usado na análise de variâncias (ANOVA). Por ser um teste não paramétrico ele não assume a distribuição normal dos resíduos, ao contrário do teste F.

As hipóteses do teste de Kruskal-Wallis são

H_0 : Não existem diferenças dos valores medianos dos grupos.

H_1 : Nem todas as grupos têm medianas idênticas.

Ora, considerando que a amostra contém N observações, g grupos e que os diferentes grupos têm, respetivamente, n_1, n_2, \dots, n_g observações, a estatística de teste é dada pela seguinte equação (Daniel, Wayne W, 1990).

$$H = (N - 1) \frac{\sum_{i=1}^g n_i (\bar{r}_i - \bar{r})^2}{\sum_{i=1}^g \sum_{j=1}^{n_i} (r_{ij} - \bar{r})^2}$$

Em que, \bar{r}_i representa a média do grupo i , \bar{r} a média dos g grupos e r_{ij} o valor da j -ésima observação do i -ésimo grupo.

$$\bar{r}_i = \frac{\sum_{j=1}^{n_i} r_{ij}}{n_i}$$

$$\bar{r} = \frac{1}{N} \sum_{i=1}^g n_i \bar{r}_i$$

O valor H tem distribuição qui-quadrado com $g-1$ graus de liberdade, assumindo um nível de significância igual a α , então $P(H > Q_\alpha) = \alpha$ (sob H_0), isto é, a probabilidade de H ser superior a determinado valor Q_α é igual ao nível de significância, e assim, se o valor obtido de H for efetivamente superior Q_α , então rejeita-se a hipótese nula de que os g grupos são igualmente distribuídos.

O valor de prova (p-valor) pode ser obtido calculando $P(X_{g-1}^2 > H \mid H_0)$, ou seja, a probabilidade de distribuição qui-quadrado com g-1 graus de liberdade ser superior a H, assumindo a hipótese nula como verdadeira.

Teste de comparação de proporções

Neste teste foi-se pela abordagem da estatística de teste de Qui-Quadrado, de seguida serão mostrados alguns exemplos deste teste (Fernandes,1999).

1) Primeiro vai-se considerar uma variável aleatória que segue uma distribuição binomial com parâmetros n e p_1 à qual se denominar X_1 , então a variável aleatória

$$Z = \frac{X_1 - np_1}{\sqrt{np_1(1-p_1)}} \sim N(0,1)$$

e a distribuição limite de $Q = Z^2$ é χ_1^2

$$Q = Z^2 = \frac{(X_1 - np_1)^2}{np_1(1-p_1)}$$

2) Considerando duas amostras aleatórias independentes e identicamente distribuídas, X_1, X_2, \dots, X_{n1} e Y_1, Y_2, \dots, Y_{n2} , onde cada X_i ($i = 1, \dots, n1$) cada Y_j ($j = 1, \dots, n2$) têm uma distribuição de Bernoulli com parâmetros p_1 e p_2 respetivamente.

$$X_1, X_2, \dots, X_{n1} \sim \text{Bernoulli}(p_1) \quad Y_1, Y_2, \dots, Y_{n2} \sim \text{Bernoulli}(p_2)$$

Com isto temos que

$$E[X_i] = p_1$$

$$E[Y_i] = p_2$$

$$\text{Var}(X_i) = p_1(1-p_1)$$

$$\text{Var}(Y_i) = p_2(1-p_2).$$

As variáveis \hat{p}_1 e \hat{p}_2 são estimadores de máxima verosimilhança dos parâmetros p_1 e p_2 , respetivamente, e eles seguem uma distribuição aproximadamente normal

$$\hat{p}_1 \sim N\left(p_1, \frac{p_1(1-p_1)}{n_1}\right) \quad \hat{p}_2 \sim N\left(p_2, \frac{p_2(1-p_2)}{n_2}\right)$$

Subtraindo os dois estimadores temos a seguinte distribuição

$$\hat{p}_1 - \hat{p}_2 \sim N\left(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right)$$

Ou seja,

$$\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}} \sim N(0,1)$$

Para realizarmos o teste para duas proporções com aproximação normal vamos considerar as seguintes hipóteses.

$$H_0: p_1 = p_2 \quad \text{vs} \quad H_1: p_1 \neq p_2$$

Considerando a hipótese nula ($p_1 = p_2 = p$), $\hat{p}_1 - \hat{p}_2$ tem uma distribuição normal com média igual a zero e desvio padrão

$$\sigma = \sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}$$

Não conhecendo o valor de p , é estimado como uma média ponderada de \hat{p}_1 e \hat{p}_2 .

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$$

A estatística de teste é calculada da seguinte forma.

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}} \sim N(0,1)$$

Fixando um nível de significância α , se $P\left(Z > Z_{\frac{\alpha}{2}}\right) < \frac{\alpha}{2}$ ou $P\left(Z < -Z_{\frac{\alpha}{2}}\right) < \frac{\alpha}{2}$, então rejeita-se hipótese nula e aceita-se que $p_1 \neq p_2$.

Modelos lineares generalizados

Os modelos lineares podem ser expressos da seguinte forma

$$y = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k + \varepsilon$$

Em que y é a variável resposta, os X 's são as variáveis explicativas, os β 's são os parâmetros do modelo e ε o erro, aleatório e não observável, ao que assumimos $E[\varepsilon]=0$ e $\text{Var}[\varepsilon]=\sigma^2$.

Uma representação alternativa dos modelos lineares clássicos pode ser feita da seguinte maneira.

Considerando μ_1, \dots, μ_n em que $E[y_i | x_i] = \mu_i$ para $i = 1, \dots, n$, então $y_i | x_i \sim N(\mu_i, \sigma^2)$, e assim

$$E[y_i | x_i] = \mu_i = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k$$

No caso dos modelos lineares generalizados, eles são uma extensão dos modelos lineares clássicos em que

$$Y = X\beta + \varepsilon$$

Y é o vetor da variável resposta de dimensão n , X é uma matriz de dimensão $n \times (n+1)$ associada a um vetor $\beta = (\beta_0, \beta_1, \dots, \beta_n)^T$ de parâmetros e ε o vetor dos erros aleatórios com distribuição que se supõe ser $N(0, \sigma^2 I)$, onde I é uma matriz identidade.

Tem-se que

$$E[Y|X] = \mu = X\beta$$

Ou seja, o valor esperado da variável resposta é uma função linear das covariáveis.

Um modelo linear generalizado está assente em três componentes fundamentais:

- Componente aleatória: a variável resposta Y que se quer modelar

- Componente sistemática: preditor linear do modelo, em que são inseridas as covariáveis por meio de combinação linear de parâmetros. O valor esperado μ_i está relacionado com o preditor linear η_i

$$\eta_i = x_i^T \beta$$

Através da relação $\mu_i = h(\eta_i)$, onde h é uma função monótona e diferencial e $g = h^{-1}$, onde g é a função de ligação que relaciona a média de y_i ao preditor linear, ou seja,

$$g(\mu_i) = \eta_i$$

- Função de ligação: ligação entre os componentes aleatório e sistemático

$$g(\mu_i) = \eta_i = x_i^T \beta$$

Modelo de regressão logística

Os modelos de regressão logística são modelos que, em geral, são utilizados quando a variável resposta é binária, ou seja, só tem dois resultados possíveis (sucesso ou insucesso).

Considerando Y_i uma variável resposta binária (0 = ausência de sucesso; 1 = presença de sucesso) cuja distribuição de probabilidades é a de Bernoulli (p_i)

$$f(y_i|p_i) = p_i^{y_i}(1 - p_i)^{(1-y_i)}$$

Onde $p_i = P(Y_i = 1) = E(Y_i)$

Modelando $E[y_i|x_i] = p_i$ e utilizando a função de ligação $g(\cdot)$ temos

$$g(p_i) = \eta_i = x_i^T \beta$$

A função de ligação a ser escolhida deverá ser contínua, diferenciável, monótona e deverá fazer com que os resultados obtidos estejam dentro do intervalo [0,1]. Das funções de ligação mais utilizadas nos modelos logísticos encontram-se as funções logit, complemento log-log e probit (Barnett,2008).

A função logit é definida da seguinte maneira:

$$\log\left(\frac{p_i}{1-p_i}\right) = \eta_i$$

$$p_i = \frac{e^{\eta_i}}{1+e^{\eta_i}} = \frac{e^{x_i^T \beta}}{1+e^{x_i^T \beta}}$$

A função logit é o logaritmo do quociente entre a probabilidade de Y dar em sucesso e a probabilidade de Y não dar sucesso, este quociente é de odds (chance).

A função complemento log-log:

$$\log(-\log(1-p_i)) = \eta_i$$

$$p_i = 1 - e^{-\eta_i} = 1 - e^{-x_i^T \beta}$$

E a função probit por

$$\Phi^{-1}(p_i) = \eta_i$$

$$p_i = \Phi(\eta_i) = \int_{-\infty}^{\eta_i} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

A interpretação dos coeficientes nos modelos de regressão logística baseia-se na razão de chances (*odds ratio*), ela dada pelo quociente entre a *odds* do acontecimento acontecer (Y=1) nos indivíduos com x=0 e a *odds* do mesmo acontecer (Y=1) nos indivíduos com x=1.

Nas variáveis contínuas, o aumento de uma unidade na variável explicativa j (mantendo as restantes variáveis constantes) traduz-se na multiplicação da *odds ratio* (OR) num fator igual a $\exp(\beta_j)$, para um aumento de k unidades na mesma variável, a multiplicação será de $\exp(k*\beta_j)$.

No caso de uma variável qualitativa, seja X_j , supondo que ela tem duas categorias, $\exp(\beta_j)$ corresponde à razão de chances da resposta para as duas categorias.

Se $\exp(\beta_j) > 1$, ou seja, $\beta_j > 0$, para um variável j contínua, significa que as chances de sucesso aumentam à medida que a variável aumenta. Se a variável j for qualitativa, significa que, para uma determinada categoria, a chance de sucesso é superior relativamente a outra categoria de referência.

Se $\exp(\beta_j) < 1$, ou seja, $\beta_j < 0$, para um variável j contínua, significa que a chance de sucesso diminui à medida que a variável aumenta. Se a variável j for qualitativa, significa que, para uma determinada categoria, a chance de sucesso é inferior relativamente a outra categoria.

Para a seleção de variáveis que constituem o melhor modelo logístico usaremos o método stepwise com base nos valores de prova do teste de razão de verossimilhança de Wilks entre modelos que contêm ou não variáveis, ou mesmo as suas interações. O grau de importância de uma variável terá a ver com o valor de prova deste mesmo teste entre modelos que a incluem e que a excluem. Quanto menor for este valor for maior será a importância da variável em causa. Como a importância de uma variável não tem necessariamente a ver com o facto de ela ser ou não estatisticamente significativa, há que impor um limite máximo do valor de prova de modo a atrair variáveis importantes à entrada (Turkman e Silva, 2000).

Critério de informação de Akaike (AIC)

O AIC estima a quantidade relativa de informações perdidas por um determinado modelo, ou seja, quanto menos informações um modelo perde, maior é a qualidade do mesmo e menor é o valor seu AIC. Dadas duas hipóteses (modelos estatísticos) de igual poder explicativo para determinado fenómeno, a hipótese mais simples é a que tem maior chance de estar correta. O AIC leva em conta e penaliza a complexidade dos modelos e tende a favorecer modelos mais simples. O AIC é prático na escolha do melhor modelo, mesmo sendo pobres ele seleciona aquele estimado para ser o melhor (Burnham e Anderson, 2002)

O critério de Akaike é definido por (Akaike,1973)

$$AIC = -2 \log(L) + 2K$$

Onde L é a função de máxima verossimilhança do modelo e K o número de parâmetros a estimar.

Teste de razão de verossimilhanças (LRT)

O teste de razão de verossimilhanças pode ser utilizado para comparar a qualidade do ajuste de dois modelos estatísticos encaixados.

Considerando M_0 um modelo com k_0 parâmetros e M_1 um modelo com k_1 ($k_1 < k_0$) encaixado com M_0 , a comparação dos modelos usando a função de verossimilhança é baseada na seguinte expressão:

$$LRT = -2 \ln \left(\frac{\text{verossimilhança do modelo } M_1}{\text{verossimilhança do modelo } M_0} \right)$$

Que segue uma distribuição Qui-quadrado com $k_0 - k_1$ graus de liberdade.

A análise dos desvios de dois modelos encaixados M_0 e M_1 , é uma medida de diferença do ajuste dos modelos, que pode ser entendida como o ganho de ajuste decorrente da inclusão de $k_0 - k_1$ parâmetros ao modelo mais simples.

O valor de prova (p-valor) pode ser obtido calculando $P(X_{k_0-k_1}^2 > ET \mid H_0)$, ou seja, a probabilidade de distribuição qui-quadrado com $k_0 - k_1$ graus de liberdade ser superior à estatística de teste, assumindo a hipótese nula como verdadeira(McCullagh e Nelder, 1990).

2 Análise dos dados dos inquéritos sobre o HIV/SIDA

2.1 Análise Estatística na forma descritiva

Tabela 4 Informação acerca das características dos respondentes presentes na base de dados.

	Frequência absoluta(relativa)
Ano da entrevista	
2009	3788(23,8%)
2011	7056(44,4%)
2015	5061(31,8%)
Sexo	
Feminino	10613(66,7)
Masculino	5292(33,3%)
Idade	
15-19 anos	8681(54,6%)
20-24 anos	7224(45,4%)
Área de residência	
Urbano	8117(51%)
Rural	7788(49%)
Nível de educação	
Sem escolaridade	1666(10,5%)
Primário	7828(49,2%)
Secundário	6125(38,5%)
Superior	286(1,8%)
Estatuto socioeconómico	
Muito pobre	1761(11,1%)
Pobre	2138(13,4%)
Médio	2503(15,7%)
Rica	3622(22,8%)
Muito rica	5881(37%)

Dos dados recolhidos, 44,4% deles são do ano de 2011, dois terços são de respondentes do sexo feminino, quanto á idade e á área de residência a distribuição das duas categorias encontra-se próxima do 50/50. No nível de educação, quase 90% finalizaram o ensino primário ou secundário, por fim, verifica-se que à medida que se sobe no estatuto socioeconómico mais inquiridos se encontra

Tabela 5 Informação acerca dos comportamentos e conhecimento acerca do HIV.

	Frequência absoluta(relativa)
Idade do 1º ato sexual	
Ainda não teve relações sexuais	2990(18,8%)
8-14 anos	3269(20,6%)
15-19 anos	7942(49,9%)
20-24 anos	270(1,7%)
Não sabe	1434(9%)
Já ouviu falar do HIV	
Não	470(3%)
Sim	15435(97%)
Teve alguma DST no último ano	
Não	15133(95,1%)
Sim	617(3,9%)
Não sabe	155(1%)
Nº de parceiros sexuais no último ano, excluindo esposo/a	
0	9708(61%)
1	4782(30,1%)
2 ou mais	1094(6,9%)
Não sabe	321(2%)
Teste do HIV	
Não	9148(57,5%)
Sim	6744(42,4%)

No que concerne os comportamentos e conhecimentos dos inquiridos, cerca de 50% deles tiveram a sua primeira relação sexual entre os 15 e 19 anos contra 20,6% que a tiveram mais cedo e 1,7% mais tarde, 18,8% ainda não tinham tido relações sexuais á data da entrevista, 97% dos inquiridos já ouviram falar do HIV/SIDA e 3,9% tiveram uma DST no ano que antecedeu a entrevista. Continuando nos comportamentos, 61% dos respondentes não tiveram relações extraconjugais, cerca de 30% tiveram 1 parceiro/a sexual para além do esposo/a e quase 7% tiveram pelo menos 2, quanto ao teste do HIV, 42,4% dos inquiridos o fizeram.

Tabela 6 Tendências dos comportamentos dos jovens moçambicanos entre 2009 e 2015.

	2009	2011	2015	P		2009	2011	2015	P
	Mulheres					Homens			
Idade mediana no início sexual	16	15	16	0.0009		15	16	15	0.0014
Rural	15	15	15	0.2779		15	16	15	<0.0001
Urbano	16	16	16	<0.0001		16	16	15	0.6971
15-19 anos	15	15	15	0.6752		15	15	15	< 0.0001
20-24 anos	16	16	16	0.0002		16	16	16	0.7492
n (15-24 anos)	2137	5533	2944			1652	1523	2117	
Proporção de respondentes que nunca tiveram sexo	38,8	43,6	39,4	0,001		29,2	28,8	29,5	0,91
Rural	47,1	55,5	44,9	0,0005		37,7	37,3	42	0,21
Urbano	32,7	35,5	36,6	0,313		24,3	21,3	22,6	0,424
15-19 anos	45,7	53,6	50,9	0,003		38	38,8	40	0,658
20-24 anos	9	5,4	4	0,057		6,6	2,8	5,7	0,074
n (nunca casaram,15-24 anos)	779	2394	1234			1241	1140	1615	
Proporção de respondentes com idades entre 15 e 24 anos que não fizeram sexo nos últimos 12 meses	11,4	3	8,4	<0,0001		10,8	4,3	7,2	<0,0001
Rural	15,1	2,3	9,3	<0,0001		12	4,5	6,3	0,0011
Urbano	10,2	3,3	7,9	<0,0001		10,2	4,2	7,4	0,0008
15-19 anos	11,2	3	7,2	<0,0001		11,2	5,6	8,5	0,0044
20-24 anos	11,7	3	10	<0,0001		10,1	2,7	5,4	0,0001
n (nunca casaram,15-24 anos)	779	2394	1234			1241	1140	1615	

Os valores de prova que se verifica na parte da mediana da idade do primeiro ato sexual, foram obtidos através do teste de Kruskal-Wallis e os restantes pelo teste de comparação de proporções.

No que concerne a idade mediana do primeiro ato sexual, primeiro verifica-se que ela se situa entre os 15 e os 16 anos, segundo, que existem resultados contraditórios do teste Kruskal-Wallis onde idades medianas nos três períodos de estudo em duas categorias diferentes podem dar valores de prova do teste antagónicos, como se pode ver nas faixas etárias.

A análise da proporção de respondentes que nunca tiveram sexo dentro daqueles que nunca casaram, permite constatar que ela é menor no sexo masculino, nas zonas urbanas e nos indivíduos que têm entre 20 e 24 anos. Relativamente aos valores de prova, no sexo masculino, pode-se dizer que a referida proporção não evoluiu ao longo dos anos. O mesmo não se pode dizer no sexo feminino, onde à exceção das áreas urbanas, pelo teste de proporções, rejeita-se a hipótese de que as proporções sejam iguais.

Tabela 7 Tendências no conhecimento dos jovens moçambicanos acerca do HIV/SIDA ao longo dos anos do estudo e divididos pelo sexo dos respondentes e por área de residência e categoria etária.

	2009	2011	2015	P	2009	2011	2015	P
	Mulheres				Homens			
Ter apenas um parceiro sexual	75,3	73,5	69,7	<0,0001	76,3	89,1	73,5	<0,0001
Rural	73,1	68,3	63,6	<0,0001	73,6	88	66,8	<0,0001
Urbano	77,1	79,5	75	<0,0001	78,3	90,1	77,4	<0,0001
15-19 anos	72,7	70,3	67,8	<0,0001	73,1	87,6	71,1	<0,0001
20-24 anos	77,7	77,3	71,6	<0,0001	80,5	91	75,8	<0,0001
n (15-24 anos)	1574	3974	1922		1236	1351	1485	
Uso de preservativos	77,5	63,6	63,7	<0,0001	77,7	82	71,7	<0,0001
Rural	70,2	57,8	54,1	<0,0001	74,2	79,7	65,2	<0,0001
Urbano	83,8	70,3	72,2	<0,0001	80,3	84,4	76,1	<0,0001
15-19 anos	77,8	61,1	62,2	<0,0001	76,7	81,2	70,7	<0,0001
20-24 anos	77,3	66,6	65,3	<0,0001	79,1	83,2	73	<0,0001
n (15-24 anos)	1620	3440	1758		1260	1244	1455	
Conhecimento do HIV/AIDS	98	97,8	93,8	<0,0001	98,4	99,6	95,9	<0,0001
Rural	96,7	96,5	91,2	<0,0001	97	99,5	92,6	<0,0001
Urbano	99,1	99,4	96,1	<0,0001	99,4	99,7	98,2	<0,0001
15-19 anos	97,8	97,5	91,2	<0,0001	97,4	99,3	94,1	<0,0001
20-24 anos	98,2	98,1	96,5	<0,0001	99,6	100	98,2	<0,0001
n (15-24 anos)	2093	5410	2740		1625	1517	2030	

Pelos valores de prova, obtidos através do teste de comparação de proporções, que aparecem na tabela, pode ver-se que em todos os casos é inferior a 0,0001, isto é, em nenhum dos casos, a proporção de respostas certas é considerado igual em todos os anos do estudo, e ao ver pelos seus valores, a tendência é para que esta proporção diminua ao longo do tempo.

Os resultados da tabela abaixo, de certa forma, corroboram as conclusões da tabela anterior. Para as três perguntas, verifica-se que em 2015, último ano do estudo, a proporção de respostas corretas é inferior do que em 2009, ano do início do estudo, o que pode levar a pensar que o conhecimento dos jovens moçambicanos relativamente ao HIV/SIDA esteja a diminuir.

Tabela 8 Tendências do conhecimento dos jovens moçambicanos sobre o HIV/SIDA ao longo dos anos do estudo.

Pergunta com resposta correta	2009	2011	2015
	n (%)	n (%)	n (%)
Ter apenas 1 parceiro sexual	2810(74,2)	5325(75,5)	3407(67,3)
Usar preservativos ao fazer sexo	2880(76)	4684(66,4)	3213(63,5)
Conhecimento abrangente de HIV/SIDA	3718(98,1)	6927(98,2)	4790(94,6)

Tabela 9 Tabela cruzada das proporções de jovens moçambicanos a utilizarem preservativos com o mais recente parceiro sexual.

Uso de preservativos no último ano com o mais recente parceiro/a	Não	Sim
Sexo		
Masculino	43,9	29,7
Feminino	56,1	15,9
Ano de entrevista		
2009	54,3	20,6
2011	52,4	17,7
2015	49,7	24,2
Idade		
15-19 anos	40,6	19,3
20-24 anos	65,7	21,8
Área de residência		
Urbano	40,6	32,1
Rural	63,9	8,3
Nível educacional		
Sem escolaridade	74,9	2,6
Primário	60	10,2
Secundário	37	36,3
Superior	22	65,7
Estatuto socioeconómico		
Muito pobre	70,6	3,3
Pobre	69,3	4,6
Médio	63,6	7,5
Rica	54,9	17,2
Muito rica	33,4	38,8
Nº de parceiros/as, excluindo esposo/a		
0	53,7	4,3
1	53,6	46,4
2 ou mais	44,5	55,5

Em 2015, a proporção de inquiridos que usaram preservativo na mais recente relação sexual é superior do que nos restantes anos de estudo, embora ainda seja bem inferior aos que responderam negativamente a esta questão, o uso de preservativo é bem superior nos meios urbanos do que nos rurais. À medida que se sobe quer no nível educacional quer no estatuto socioeconómico, a proporção de inquiridos que usaram preservativo aumenta, tal se verifica no número de parceiros sexuais à exceção do esposo/a. De referir, que os resultados desta tabela têm em conta os valores omissos.

Tabela 10 Tabela cruzada com as proporções de jovens moçambicanos a terem feito o teste de HIV.

Teste de HIV	Não	Sim
Sexo		
Masculino	74,8	25,1
Feminino	48,9	51
Ano da entrevista		
2009	65,4	34,3
2011	57,9	42,1
2015	51,1	48,8
Área de residência		
Urbano	52,3	47,6
Rural	67	37
Nível educacional		
Sem escolaridade	65,7	34,2
Primário	62,6	37,3
Secundário	50	50
Superior	31,5	68,5
Estatuto socioeconómico		
Muito pobre	70,9	29,1
Pobre	70,4	29,5
Médio	64,1	35,8
Rica	50,7	49,2
Muito rica	50,2	49,7
Uso de preservativo no último ato como mais recente parceiro		
Não	52,1	47,8
Sim	47,3	52,7
Idade do 1º ato sexual		
Ainda não teve relações	91,5	8,4
8-14 anos	55,2	44,6
15-19 anos	49,3	50,7
20-24 anos	43,7	56,3
Tem conhecimento do HIV		
Não	93,8	6,2
Sim	56,4	43,5

A proporção de inquiridos do sexo feminino que fizeram o teste de HIV é duas vezes superior ao verificado no sexo oposto, essa mesma proporção aumenta ao longo dos anos do estudo e à medida que se sobe no nível de escolaridade e no estatuto socioeconómico. 47,6% dos respondentes que vivem nas zonas urbanas fizeram o teste do HIV contra 37% nas zonas rurais, no que toca á idade do 1º ato sexual, observa-se

que quanto mais tarde ela acontece maior é proporção de indivíduos que fizeram o tal teste. Nos respondentes que admitiram não conhecerem o HIV/SIDA, apenas 6,2% fizeram o teste para a mesma doença, mas mesmo naqueles que a conhecem a maioria não fez o teste, por fim, no uso de preservativo, quer numa categoria quer noutra, a proporção de inquiridos a terem feito o teste de HIV anda próximo dos 50%, sendo superior nos que responderam afirmativamente ao uso de preservativo na mais recente relação sexual.

Tabela 11 Tabela cruzada com as proporções de jovens moçambicanos que já ouviram falar do HIV/SIDA.

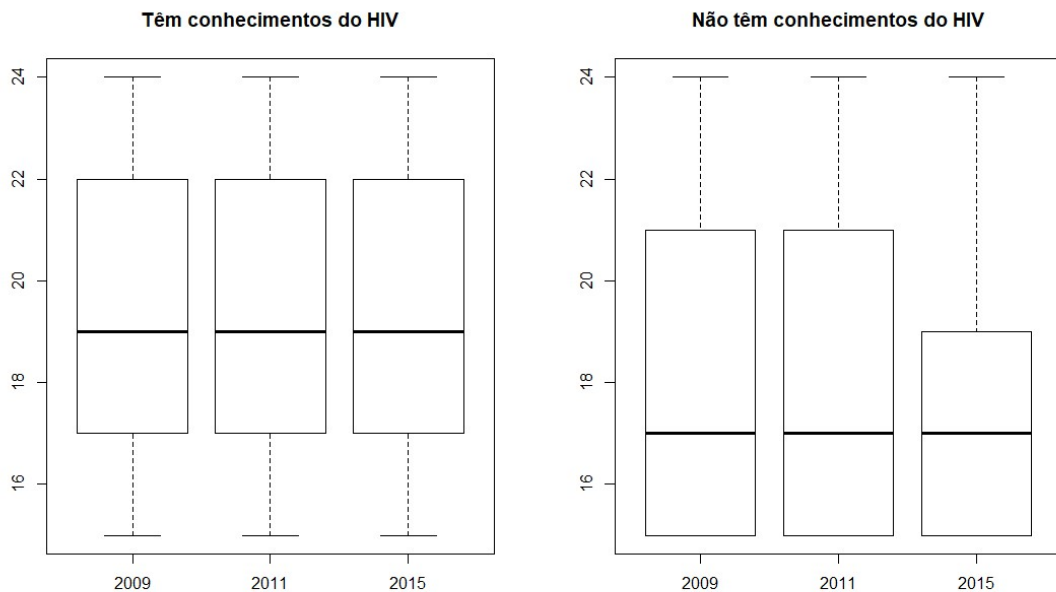
Tem conhecimento do HIV	Não	Sim
Sexo		
Masculino	2,3	97,7
Feminino	3,3	96,7
Ano da entrevista		
2009	1,8	98,2
2011	1,8	98,2
2015	5,4	94,6
Idade		
15-19 anos	3,9	96,1
20-24 anos	1,8	98,2
Área de residência		
Urbano	1,4	98,6
Rural	4,6	95,4
Nível educacional		
Sem escolaridade	7,7	92,3
Primário	3,9	96,1
Secundário	0,6	99,4
Superior	0,3	99,7
Estatuto socioeconómico		
Muito pobre	5,1	94,9
Pobre	5,7	94,3
Média	4,8	95,2
Rica	2,3	97,7
Muito rica	1	99
Teste HIV		
Não	4,8	95,2
Sim	0,4	99,6
Idade 1º ato sexual		
Ainda não teve relações sexuais	7,1	92,9
8-14 anos	2,1	97,9
15-19 anos	1,4	98,6
20-24 anos	1,5	98,5

Tabela 11 (continuação) Tabela cruzada com as proporções de jovens moçambicanos que já ouviram falar do HIV/SIDA.

Tem conhecimento do HIV	Não	Sim
Nº de parceiros/as sexuais, excluindo esposo/a		
0	4	96
1	1,2	98,8
2 ou mais	0,5	99,5
Teve uma DST no último ano		
Não	3,1	96,9
Sim	0,3	99,7
Uso de preservativo no último ato sexual		
Não	2,3	97,7
Sim	0,5	99,5

A proporção de respondentes a conhecerem o HIV é inferior no ano 2015 (94,6%) do que nos restantes anos (98,2%), é maior nas zonas urbanas (98,6% contra 95,4% nas rurais), nos que fizeram o teste de HIV (99,6% contra 95,2% nos que não a fizeram), nos que tiveram uma DST (99,7% contra 96,9% nos não tiveram) e nos que utilizaram o preservativo na mais recente relação sexual (99,5% contra 97,7% nos não utilizaram). Esta mesma proporção aumenta à medida que se sobe no nível de escolaridade, de estatuto socioeconómico e à medida que aumenta o número de parceiros sexuais.

Figura 1 Distribuição etária dos respondentes que têm conhecimentos acerca do HIV e os que não têm ao longo dos anos.

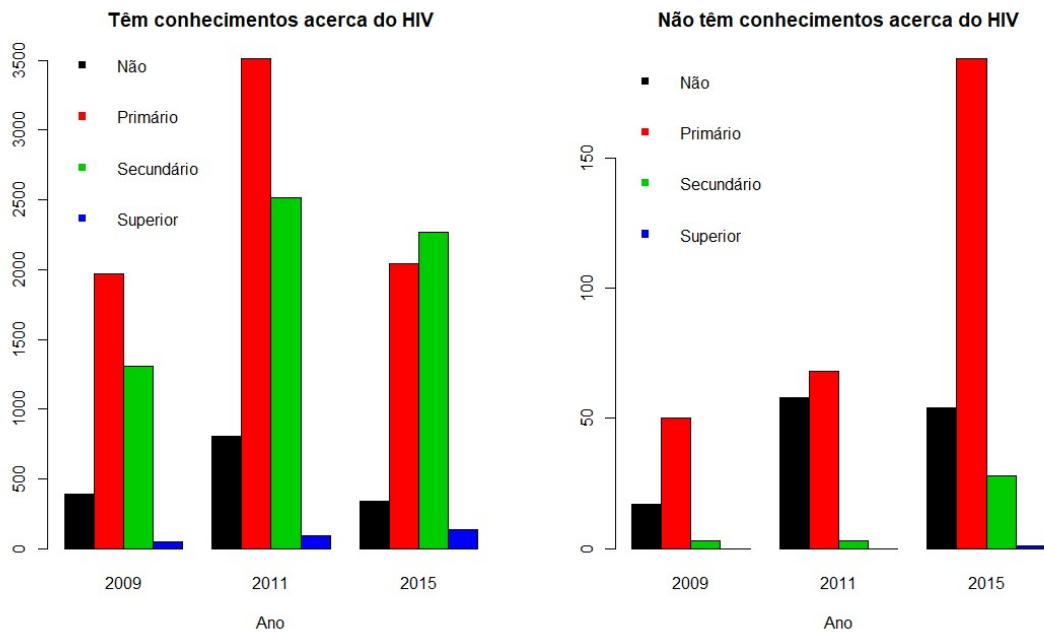


Na parte dos respondentes que têm conhecimentos acerca do HIV a distribuição etária é exatamente a mesma nos três anos do estudo, com uma mediana igual a 19 anos e os primeiros e terceiros quartis a serem de 17 e 22 anos, respetivamente.

No outro lado, vemos que a mediana dos respondentes que afirmaram não terem ouvido falar do HIV é, nos três anos, de 17 anos. De resto, observa-se uma maior concentração da distribuição etária no ano de 2015 onde 75% das pessoas que não têm conhecimentos acerca do HIV têm uma idade menor ou igual a 19 anos, quando nos outros anos é menor ou igual a 21 anos.

Será interessante verificar os números do conhecimento acerca do HIV ao longo dos anos pelos diferentes níveis de escolaridade.

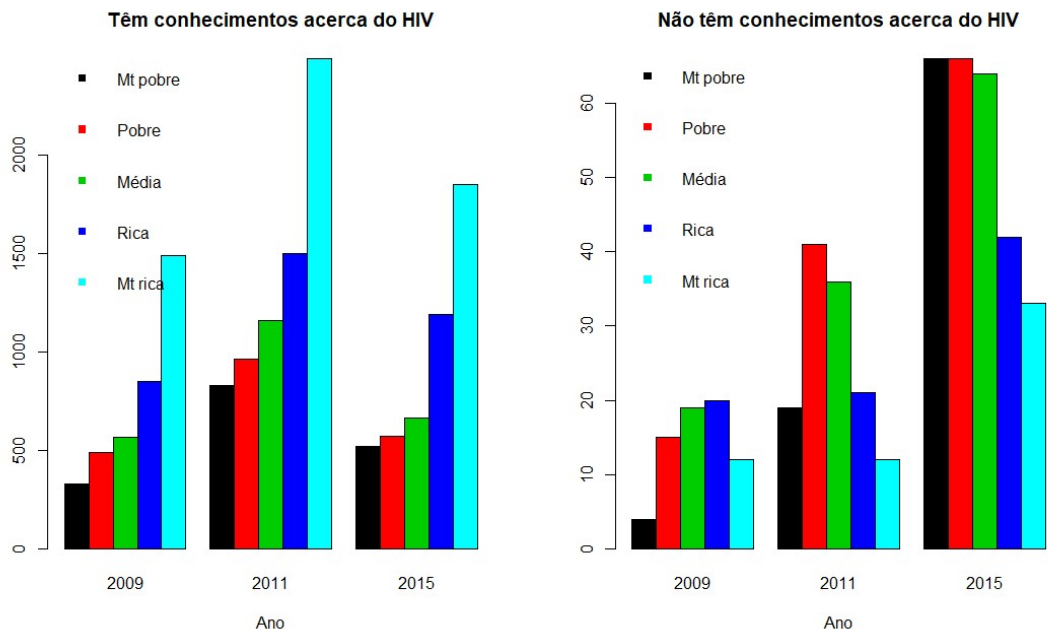
Figura 2 Evolução da frequência absoluta do conhecimento do HIV ao longo dos anos pelos níveis de escolaridade.



Na categoria dos respondentes de têm conhecimentos acerca do HIV, em 2015 o número de inquiridos com a escolaridade primária é quase idêntico ao verificado no ano 2009, o mesmo acontece nos que não frequentaram a escola, pelo contrário o número de inquiridos com o ensino secundário aumentou de 2009 para 2015, ultrapassando mesmo a barra do número relativo aos inquiridos com o ensino primário.

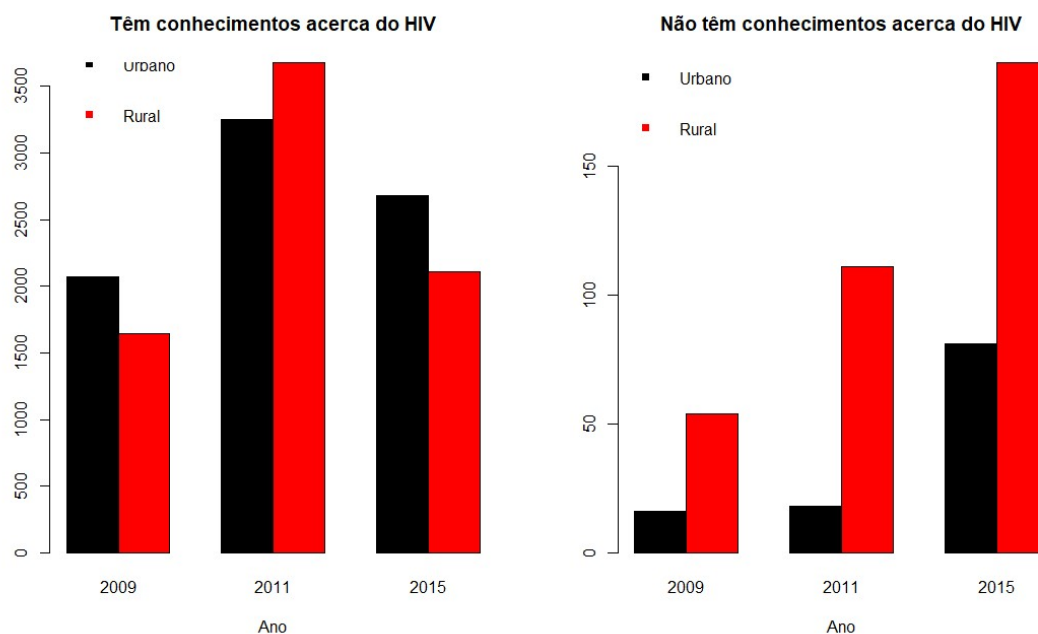
Na categoria oposta, a predominância está nos inquiridos que tem o ensino primário e nos que não frequentaram a escola, o que pode explicar mediana inferior no gráfico 1, mas o mais relevante até acaba por ser a evolução do número de respondentes a afirmarem não conhecerem o HIV ao longo dos anos, na categoria dos que têm o ensino primário ele passou de cerca de 50 em 2009 para cerca de 200 em 2015, ou seja, 4 vezes mais, o mesmo se verifica nos outros níveis de escolaridade.

Figura 3 Evolução da frequência absoluta do conhecimento acerca do HIV ao longo dos anos pelos diferentes estatutos socioeconômicos.



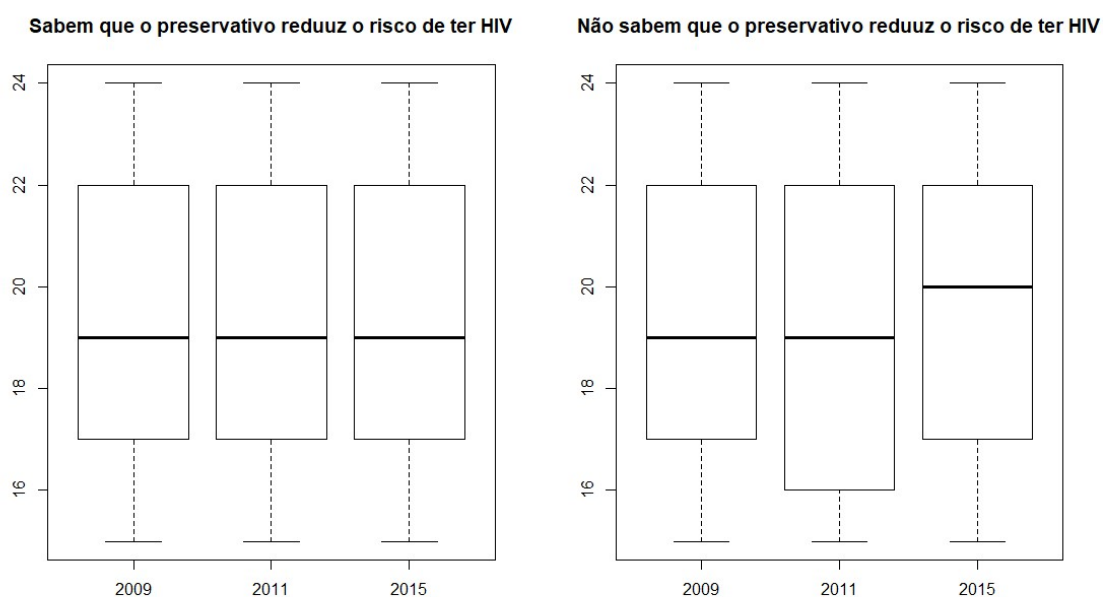
Se no lado dos inquiridos que responderam afirmativamente, em todos os anos, vê-se que quanto mais alto é o estatuto maior é a frequência absoluta, no outro lado o mesmo verifica-se uma evolução ao longo dos anos, se em 2009 a predominância estava nas classes médias no ano de 2015 ela está nas classes mais baixas. Outra coisa que é verificada na categoria dos inquiridos que afirmaram não conhecerem o HIV é o aumento significativo desse número em todos os níveis socioeconômicos.

Figura 4 Evolução da frequência absoluta, ao longo dos anos, dos inquiridos que têm e os que não têm conhecimentos acerca do HIV nas diferentes áreas residência.



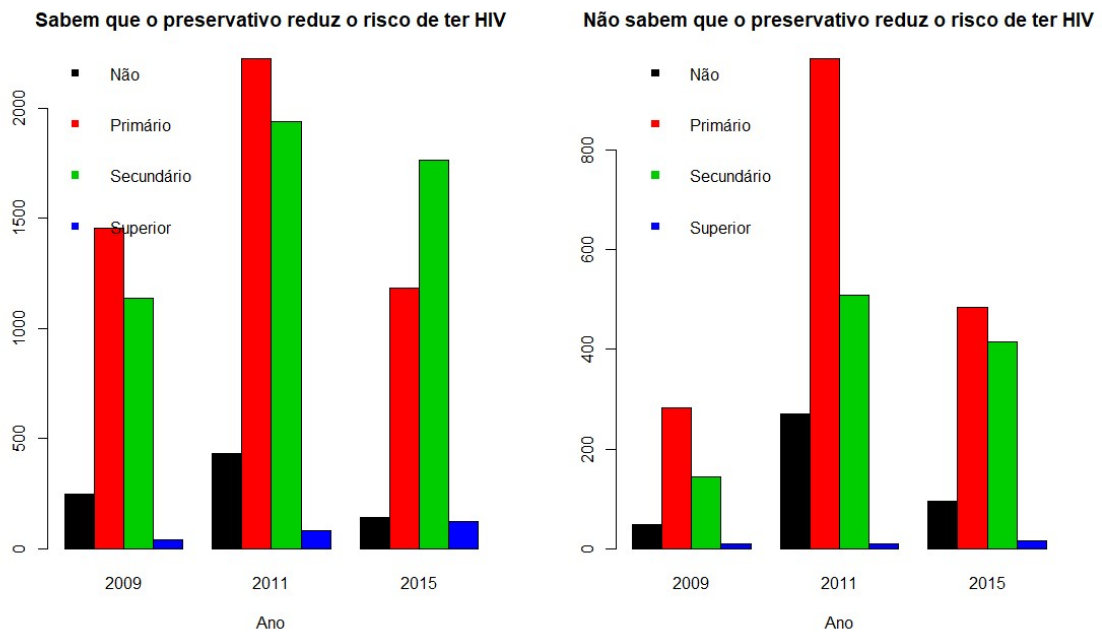
O número de inquiridos a não terem conhecido o HIV conhece um aumento significativo quer eles vivam em zonas rurais ou urbanas. Outra coisa que se verifica é que em qualquer dos anos a maioria dos respondentes que não conhecem o HIV vive nas zonas rurais, no contrário, á exceção de 2011, a maioria dos que conhecem vivem nas zonas urbanas de Moçambique.

Figura 5 Distribuição etária dos respondentes que sabem e os que não sabem que o preservativo o risco de ter HIV nos diferentes anos do estudo.



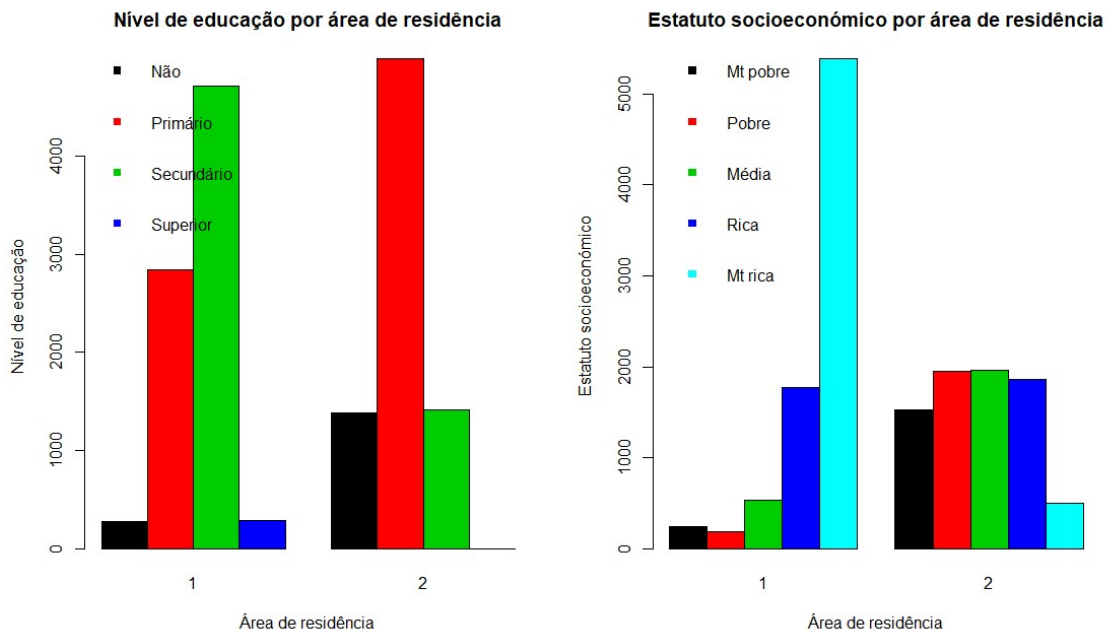
No lado dos que responderem de o preservativo reduz o risco de ter HIV, a distribuição etária é a mesma nos três anos do estudo com a mediana a situar-se nos 19 anos e os primeiros e terceiros quartis nos 17 e 22 anos, respetivamente. No outro lado, nota para o aumento da mediana no ano de 2015 em que ela passou de 19 para 20 anos.

Figura 6 Evolução, ao longo dos anos, da frequência absoluta, dos inquiridos que sabem e os que não sabem que o preservativo reduz o risco de ter HIV nos diferentes níveis de escolaridade.



Nas categorias dos inquiridos com o ensino primário e que não frequentaram a escola, a proporção de inquiridos que sabe que o preservativo reduz o risco de HIV diminuiu de 2009 para 2015 enquanto que essa proporção cresceu entre os inquiridos que desconhecem esse facto.

Figura 7 Relação das frequências absolutas dos níveis de educação e de estatuto socioeconómico com a área de residência (1=Urbano, 2=Rural).



Estes resultados mostram-nos uma diferença social entre as zonas urbanas e rurais de Moçambique. No que concerne o nível educacional, se nas zonas urbanas predominam os inquiridos com o ensino secundário, nas zonas rurais são inquiridos com o ensino primário que se encontram com maior frequência e ficando nas zonas rurais existem muito mais respondentes que não frequentaram a escola do que nas zonas urbanas e quanto aos inquiridos com o ensino superior quase não existem nas áreas rurais. Na relação com estatuto socioeconómico essa diferença social é mais evidente, se nas zonas urbanas predominam as classes mais elevadas nas zonas rurais são as classes médias, médias baixas que existem com maior frequência.

2.2 Análise estatística na forma de modelo de regressão logística

Como referido no módulo 1.3 sobre a metodologia utilizada no trabalho, um modelo de regressão requer uma variável resposta, ora escolha data1 caiu para a variável que diz respeito ao conhecimento ou não do respondente acerca do sida, primeiro porque ela condiz com o tema do trabalho e segundo porque ela não contém valores omissos, o que ajuda no cálculo do modelo. Esta variável é uma variável binária, e por esta razão o modelo que foi utilizado aqui foi o modelo de regressão logística.

- Idade
- Sexo
- Ano da entrevista
- Área de residência
- Nível educacional
- Estatuto socioeconómico
- Se o respondente trabalha
- Estado civil
- Se o respondente teve uma DST no ano transato
- Idade do 1º ato sexual categorizada
- Nº de parceiros sexuais no ano transato, excluindo esposo(a)
- Se o respondente fez o teste do HIV

Começou-se por um modelo contendo todas as variáveis atrás referidas, chamemos este modelo como o modelo completo, de seguida pelo método stepwise ficou-se com um modelo reduzido, ou seja, com menos variáveis explicativas.

Tabela 12 Resultados do modelo de regressão logístico.

	Beta	OR(exp(beta))	Std.Error	Valor de prova
Interceção	2,77869		0,55922	< 0,0001
Idade	0,03660	1,0372	0,02516	0,1458
Sexo (Masculino)				
Feminino	-0,73911	0,4775	0,12559	< 0,0001
Área de residência (Urbano)				
Rural	-0,30098	0,7401	0,14418	0,0368
Nível de escolaridade (Sem escolaridade)				
Primário	0,79339	2,2109	0,12384	< 0,0001
Secundário	2,44132	11,4882	0,22706	< 0,0001
Superior	2,08267	8,0258	1,03345	0,0439
Estatuto socioeconómico (Muito pobre)				
Pobre	-0,25253	0,7768	0,15295	0,0987
Médio	-0,31225	0,7318	0,15454	0,0433
Rico	-0,02104	0,9792	0,17468	0,90415
Muito rico	0,24944	1,2833	0,22450	0,26652
Estado Civil (Nunca casou)				
Casado/a	0,44901	1,5668	0,25112	0,07378
Vive com alguém	0,68617	1,9861	0,25657	0,0075
Viúvo/a	0,09205	1,0964	0,82909	0,9116
Divorciado/a	0,52205	1,6855	0,56412	0,3548
Vive sozinho/a	-0,14936	0,8613	0,30531	0,6247
Idade do 1º ato sexual (Nunca teve)				
8-14 anos	0,42617	1,5314	0,25635	0,0964
15-19 anos	0,45771	1,5804	0,24464	0,0614
20-24 anos	0,28079	1,3242	0,57988	0,6282
Nº de parceiros/as sexuais, excluindo esposo/a, no último ano (0)				
1	0,66642	1,9472	0,23348	0,0043
2 ou mais	1,07396	2,9269	0,46306	0,0204

Tabela 12 (continuação) Resultados do modelo de regressão logístico.

	Beta	OR(exp(beta))	Std.Error	Valor de prova
Teve alguma DST no último ano (Não)				
Sim	1,59653	4,9359	0,72259	0,0271
Teste HIV (Não)				
Sim	2,33103	10,2886	0,20522	< 0,0001
Ano da entrevista (2009)				
2011	0,25950	1,2963	0,16788	0,1222
2015	-1,46536	0,2310	0,14639	< 0,0001

Quanto aos resultados, no que concerne às variáveis que já estavam no modelo obtido anteriormente, chega-se mais ou menos às mesmas conclusões, ou seja, o sexo feminino e as zonas rurais têm um efeito negativo no conhecimento do HIV relativamente á categoria oposta, no sentido contrário, a frequência escolar, o facto de já ter tido relações sexuais, o facto de ter mais parceiros/as sexuais para além do companheiro/a, o ter uma DST e o facto de ter feito o teste de HIV têm efeito positivo. No que diz respeito ao ano da entrevista, se um inquirido do ano 2011 tem maior chance de responder positivamente relativamente a um do ano 2009, para outro do 2015 esta chance corresponde a nem sequer $\frac{1}{4}$ daquela que se tem em encontrar uma resposta positiva num inquirido do ano 2009, este resultado que se vê no ano 2015 pode explicar em parte o valor de negativo que se verificou no modelo de efeitos fixos, na variável do ano da entrevista em que se podia concluir que o conhecimento diminuía ao longo dos anos.

No que diz respeito às restantes variáveis, no estatuto socioeconómico verifica-se que apenas a classe mais rica tem um efeito positivo no conhecimento do HIV relativamente á classe mais pobre. No estado civil pode-se concluir que o facto de já ter tido uma relação favorece o conhecimento do HIV.

Tabela 13 Comparação do valor AIC do modelo obtido pelo método stepwise com o modelo completo.

	AIC
Modelo stepwise	3255,807
Modelo completo	3259,295

Comparativamente ao modelo com todas as variáveis (modelo completo), o modelo obtido apresenta um valor AIC menor, o que pode indicar que ele seja o melhor modelo entre os dois.

Tabela 14 Análise de desvios entre os modelos completos e o modelo obtido pelo método stepwise.

	Desvio	Graus de liberdade	Valor de prova
Modelo completo	3197,8	15876	
Modelo stepwise	3180,3	15895	0,09401

A análise de desvio corrobora os resultados verificados na tabela anterior, com um valor de prova suficientemente alto (superior a 0,05) pode se aceitar a hipótese de que o modelo stepwise seja um melhor modelo.

Os resultados pelo modelo (tabela 9) quando comparados com os resultados que se verificaram nos trabalhos que se fizeram sobre as tendências e os determinantes acerca do conhecimento do HIV entre as mulheres no Bangladesh e relativamente aos jovens entre os 15 e os 24 anos na Nigéria que utilizaram o mesmo método no seu modelo, existem alguns pontos em comum. Nos dois modelos a chance de encontrar um respondente a ter conhecimentos acerca do HIV aumenta com a idade, com o nível de escolaridade e com o estatuto socioeconómico e ela diminui se o respondente viver numa zona rural. Onde os resultados divergem concerne a variável do ano da entrevista, se nos trabalhos realizados no Bangladesh e na Nigéria a chance de encontrar uma respondente a ter conhecimento do HIV aumenta com os anos, no caso deste trabalho uma baixa significativa no último ano em que se fizeram entrevistas.

2.3 Conclusões

Uma das conclusões que se pode retirar depois de observados os resultados dos módulos anteriores é que o nível de conhecimento aumenta com o nível de escolaridade e com o estatuto socioeconómico, no entanto ele diminui quando um inquirido vem de uma zona rural. E como vimos no Gráfico 7, nas zonas urbanas encontram-se com maior frequência inquiridos com os níveis mais altos quer em termos educacionais ou socioeconómicos enquanto que nas áreas rurais são os correspondentes níveis mais baixos que se encontram com maior frequência.

Ou seja, se o nível de conhecimento acerca do HIV diminui com o abaixamento dos níveis de escolaridade e socioeconómico e com o facto de o inquirido viver numa área rural de Moçambique. E se é nas zonas rurais que se encontram com maior frequência respondentes que com baixos níveis de escolaridade e de estatuto socioeconómico, a primeira conclusão que se pode tirar do trabalho é de que algo fazer nas áreas rurais, como por exemplo aumentar a informação sobre a doença nestas áreas.

Um resultado interessante obtido neste trabalho foi a diminuição do nível de conhecimento acerca do HIV no ano de 2015, não se sabe os fatores que levaram a esta diminuição, mas é um resultado que pode ser um sinal de alerta.

Neste trabalho seria esperado ter resultados das entrevistas do ano de 2003, mas infelizmente não se conseguiu ter estes dados, no futuro, seria interessante continuar este trabalho agora tendo já acesso aos dados das entrevistas de 2003.

3 Referências

1. Md. Tuhin Sheikh, Md. Nizam Uddin and Jahidur Rahman Khan (2017). A comprehensive analysis of trends and determinants of HIV/AIDS knowledge among the Bangladeshi women based on Bangladesh Demographic and Health Surveys, 2007-2014
2. Ayodeji B. Oginni, Sylvia B. Adebajo, Babatunde A. Ahonsi (2017). Trends and Determinants of Comprehensive Knowledge of HIV among Adolescents and Young Adults in Nigeria, 2003 - 2013
3. Observador. Moçambique. Estimativa de portadores de VIH sobe de 1,9 para 2,1 milhões. <https://observador.pt/2018/03/14/mocambique-estimativa-de-portadores-de-vih-sobe-de-19-para-21-milhoes/>
4. DW News. HIV-SIDA em Moçambique. <https://www.dw.com/pt-002/hiv-sida-em-mo%C3%A7ambique/t-17422644>
5. DW News. Moçambique: Aumenta nível de prevalência do HIV. <https://www.dw.com/pt-002/mo%C3%A7ambique-aumenta-n%C3%ADvel-de-preval%C3%Aancia-do-hiv/a-38759684>
6. INSIDA (2009). Dados sobre HIV nos Jovens, INSIDA, Moçambique 2009 <https://dhsprogram.com/pubs/pdf/HF33/HF33yp.pdf>
7. Daniel, Wayne W. (1990). Applied nonparametric statistics (em inglês). [S.l.]: PWS-Kent Publ. ISBN 9780534919764
8. Edite Manuela da G.P Fernandes (1999). Estatística Aplicada: 163-164.
- 9 A. J., Barnett, A. (2008). An Introduction to Generalized Linear Models.123-138
10. Antónia Turkman, Giovani Silva (2000). Modelos lineares generalizados - da teoria à prática.
11. Burnham, K. P.; Anderson, D. R. (2002), Model Selection and Multimodel Inference: A practical information-theoretic approach (2nd ed.). 60-65
12. Akaike, H. (1973). Information theory as an extension of the maximum likelihood principle. 267–281
13. McCullagh, P., Nelder, J. A. (1990). Generalized Linear Models. Chapman e Hall, 2ed. 35-36