



**UNIVERSIDADE ABERTA**

**Análise da Mortalidade Materna na Província de Nampula e Fatores  
Associados. Uma aplicação do modelo logístico**

**Aboo Evaristo Camacho**

**Dissertação de Mestrado**

**Mestrado em Bioestatística e Biometria**

**Orientadora:** Professora Doutora. Maria do Rosário Olaia Duarte Ramos

**Coorientadora:** Professora Doutora Teresa Paula Costa Azinheira Oliveira

## **Condições de Utilização do Trabalho por Terceiros**

Análise da Mortalidade Materna na Província de Nampula e Fatores Associados. Uma aplicação do modelo logístico © 2025 Aboo Evaristo Camacho, está licenciado sob CC BY-NC-ND 4.0. Para ver uma cópia completa desta licença, visite <https://creativecommons.org/licenses/by-nc-sa/4.0/>

## **Agradecimento**

À minha mãe, esposa e filhos pelo encorajamento, carinho e apoio incondicional em todos os momentos.

Às minhas Orientadoras Professora Doutora Maria do Rosário Olaia Duarte Ramos e a Professora Doutora Teresa Paula Costa Azinheira Oliveira, pela orientação e apoio, bem como o seu suporte incondicional e paciência durante o desenvolvimento do trabalho.

## **Dedicatória**

**À minha mãe, esposa e filhos**



## DECLARAÇÃO DE INTEGRIDADE

### STATEMENT OF INTEGRITY

Declaro ter atuado com integridade na elaboração da presente dissertação/tese. Confirmando que em todo o trabalho conducente à sua elaboração não recorri à prática de plágio ou a qualquer outra forma de falsificação de resultados.

Mais declaro que tomei conhecimento integral do Regulamento Disciplinar da Universidade Aberta, publicado no Diário da República, 2.ª série, n.º 215, de 6 de novembro de 2013.

I hereby declare having conducted my thesis with integrity. I confirm that I have not used plagiarism or any form of falsification of results in the process of the thesis elaboration.

I further declare that I have fully acknowledged Disciplinary Regulations of the Universidade Aberta (regulation published in the official journal Diário da República, 2.ª série, N.º 215, de 6 de novembro de 2013).

Universidade Aberta, 04 de março de 2025

Nome completo/Full name: Aboo Evaristo Camacho

Assinatura/Signature:

---

manuscrita ou digital / handwritten or digital

Análise da Mortalidade Materna na Província de Nampula e Fatores Associados. Uma aplicação do modelo logístico

## **Resumo**

Nas últimas décadas do último século, em todo o mundo, foram registados avanços significativos e louváveis na área científica e tecnológica, principalmente nas áreas de atendimento a mulher e criança. Por essa razão, tornou-se praticamente inadmissível que o processo de reprodução cause sequelas às mulheres assim como os seus bebés, levando-os à morte por qualquer fator ou causa.

Em Moçambique, a redução do rácio da mortalidade materna constitui um dos pilares dos objetivos do desenvolvimento sustentável no seu número 3.1. O mesmo figura dentre os indicadores de saúde mais utilizados para identificar os riscos à saúde de grupos populacionais considerados de risco (mulheres grávidas e crianças).

Para a análise dos dados foi aplicada a regressão logística como técnica estatística utilizada na relação entre variáveis. Procurou-se encontrar a influência ou relação das variáveis: número de Consultas Pré-Natais, idade gestacional, número de gestações, número de filhos, causa básica e causa direta de morte em relação à variável explicada Evitabilidade, tendo em conta que para o comité de auditoria toda a morte materna deve ser classificada como Evitável e Não evitável. Foram analisados 173 processo de óbitos maternos, dos quais se notou que 82% era evitável e cerca de 18% não evitável.

O modelo final selecionado tem como preditores significativos as variáveis, “Fez consultas pré-natal e número de consultas feitas”. Em relação os fatores que influenciaram nas mortes maternas foi possível verificar que a maioria (97 mortes) foi por fatores pessoais ou familiares, fatores institucionais 32 mortes (inclui a falta de equipamentos/consumíveis na unidade sanitária, medicamentos, falta de sangue e noutros casos falta de pessoal capacitado para dar resposta as emergências obstétricas), o terceiro grupo de fatores que mais influencia nas mortes foi o logístico com 25 mortes.

Dos fatores acima descritos, tem se que a maioria das mortes podia ser evitável ao nível das unidades sanitária.

**Palavras Chaves:** Mortalidade materna, fatores associados, Regressão Logística Múltipla

Analysis of Maternal Mortality in Nampula Province and Associated Factors. An application of the logistic model

## **Abstract**

In the last decades of the last century, throughout the world, significant and commendable advances were recorded in the scientific and technological area, mainly in the areas of care for women and children. For this reason, it has become practically unacceptable that the reproduction process causes consequences to women as well as their babies, leading to their death due to any factor or cause.

In Mozambique, reducing the maternal mortality ratio is one of the pillars of the sustainable development goals in number 3.1. It is among the health indicators most used to identify health risks for population groups considered at risk (pregnant women and children).

For data analysis, logistic regression was applied as a statistical technique used in the relationship between variables. We sought to find the influence or relationship of the variables: number of Prenatal Consultations, gestational age, number of pregnancies, number of children, underlying cause and direct cause of death in relation to the explained variable Preventability, considering that for the audit committee all maternal deaths must be classified as Preventable and Non-preventable. 173 maternal death cases were analyzed, of which it was noted that 82% were preventable and approximately 18% were non-preventable.

The final model selected has as significant predictors the variables, “Had prenatal consultations and number of consultations made”. Regarding the factors that influenced maternal deaths, it was possible to verify that the majority (97 deaths) were due to personal or family factors, institutional factors 32 deaths (including the lack of equipment/consumables in the health unit, medicines, lack of blood and in other cases lack of trained personnel to respond to obstetric emergencies), the third group of factors that most influenced deaths was logistical with 25 deaths.

From the factors described above, it is clear that most deaths could have been prevented at the level of health units.

**Keywords:** Maternal mortality, associated factors, Multiple Logistic Regression

## Índice

Lista de tabelas .....	11
Lista de gráficos.....	12
Lista de figuras .....	12
Abreviaturas.....	13
CAPÍTULO I: INTRODUÇÃO.....	14
1. Introdução.....	14
1.1. Objetivos.....	15
➤ Objetivo Geral .....	15
➤ Objetivos específicos.....	15
CAPÍTULO II: MORTALIDADE MATERNA .....	17
2. Mortalidade Materna .....	17
2.1. Alguns conceitos .....	17
2.2. Antecedentes e Situação .....	18
2.3. Magnitude da Mortalidade Materna .....	19
2.4. A mortalidade materna e seu impacto .....	20
2.5. Mortalidade Materna no Mundo.....	21
2.6. Mortalidade Materna em África .....	21
2.7. Mortalidade Materna em Moçambique .....	23
2.8. Auditoria ou Investigação do Óbito Materno.....	23
CAPÍTULO III: DADOS E MÉTODOS.....	26
3. Dados e Métodos .....	26
3.1. População .....	26
3.2. Amostra .....	26
3.3. Análise de dados.....	27

CAPÍTULO IV: REGRESSÃO LÓGISTICA.....	28
4. Regressão Logística.....	28
4.1. Regressão logística univariada ou simples .....	29
4.1.1. Transformação Logit .....	31
4.1.2. Estimação dos parâmetros do modelo .....	31
4.1.3. Ajustamento do Modelo .....	34
4.1.4. Métodos iterativos para estimação dos parâmetros que maximizam a função de verossimilhança.....	36
4.1.5. Teste de significância dos Coeficientes.....	37
4.1.6. Intervalo de Confiança para os coeficientes .....	41
4.2. Regressão logística múltipla.....	42
4.2.1. Estimação dos parâmetros .....	43
4.2.2. Qualidade do ajustamento do modelo de regressão logística múltipla.....	44
4.2.3. Teste de significância do modelo .....	47
4.2.4. Intervalo de confiança .....	48
4.3. Diagnóstico do Modelo .....	48
4.3.1. Resíduos de Pearson.....	48
4.3.2. Deviance.....	49
4.3.3. O Teste de Hosmer-Lemeshow .....	50
4.3.4. Matriz de classificação .....	51
4.4. Curva ROC .....	53
CAPÍTULO 5: RESULTADOS .....	57
5. Aplicação do Modelo aos dados.....	57
5.1. Aplicação do modelo de Regressão Logística a mortalidade materna .....	57
5.1.1. Descrição dos dados .....	57
5.2. Análise dos dados usando a regressão logística .....	64

6. DISCUSSÃO DOS RESULTADOS .....	78
7. CONCLUSÃO.....	80
8. BIBLIOGRAFIA .....	82

## Lista de tabelas

Tabela 1: Tabela de Contingência ou Matriz confusão .....	52
Tabela 2: Resumo dos casos Processados .....	57
Tabela 3: Estatísticas .....	57
Tabela 4: Grupo da Causa Básica * Evitabilidade da Morte Crosstabulation.....	63
Tabela 5: Possíveis Fatores Associados * Evitabilidade da Morte Crosstabulation .....	63
Tabela 6: Codificação de Variável Dependente .....	64
Tabela 8: Codificações de variáveis categóricas .....	65
Tabela 9: História da Iteração <sup>a,b,c</sup> .....	65
Tabela 10: Tabela de Classificação <sup>a,b</sup> .....	66
Tabela 11: Variáveis na Equação .....	66
Tabela 12: Variáveis não presentes na equação <sup>a</sup> .....	67
Tabela 13: Testes Omnibus de Coeficientes do Modelo .....	68
Tabela 14: Resumo do modelo .....	68
Tabela 15: Teste de Hosmer and Lemeshow .....	69
Tabela 16: Tabela de Contingência do Teste de Hosmer & Lemeshow .....	69
Tabela 17: Tabela de Classificação <sup>a</sup> .....	70
Tabela 18: Variáveis na Equação .....	70
Tabela 19: Matriz de Correlação .....	72
Tabela 20: Medidas Preditivas .....	72
Tabela 21: Resumo do processamento de caso .....	75
Tabela 22: Área sobre a Curva .....	76
Tabela 23: Área sob a curva apenas com variáveis acima da linha de referência .....	77

## **Lista de gráficos**

Gráfico 1: Mortes maternas por faixa etária.....	59
Gráfico 2: Mortes maternas por nº de consultas feitas .....	60
Gráfico 3: Evitabilidade da Morte .....	61
Gráfico 4: Grupo de Causa Básica de Morte.....	62
Gráfico 5: Ponto de Corte.....	73
Gráfico 6: Distribuição dos grupos de probabilidade observados e preditas .....	74
Gráfico 7: Curva ROC com todas as variáveis.....	75
Gráfico 8: Curva ROC com variáveis cujo valor esta acima do limiar .....	76
Gráfico 9: Curva ROC Geral do Modelo.....	77

## **Lista de figuras**

Figura 1: Modelo de três atrasos .....	25
Figura 2: Função Logística.....	30
Figura 3: Método de Newton-Raphson .....	37

## **Abreviaturas**

**AUC** – Area Under the Curv/ Área sobre a curva

**CID** – Classificação Internacional de Doença

**CPN** – Consulta Pré-Natal

**Dhis2** – Distrital health Information System 2

**FN** - Falsos Negativos

**FP** - Falsos Positivos

**HIV** – Vírus de Imunodeficiência Humana

**ICF** - International Coaching Federation/ Federação Internacional de Coaching

**IDS** – Inquérito Demográfico de Saúde

**INE** – Instituto Nacional de Estatística

**ODM** – Objetivos do Desenvolvimento do Milénio

**OMS** – Organização Mundial

**OPAS** – Organização Pan-Africana de Saúde

**NV** – Nados Vivos

**PNUD** - Programa das Nações Unidas para o Desenvolvimento/ United Nations Development Programme

**RMM** – Rácio de Mortalidade Materna

**ROC** - Receiver Operating Characteristic/ Característica Operacional do Receptor

**SIDA** – Síndrome de Imunodeficiência Adquirida

**SISMA** – Sistema de Informação para Saúde & Monitoria Avaliação

**UNFPA** - United Nations Population Fund/ Fundo de População das Nações Unidas

**UNICEF** - United Nations Children's Fund/ Fundo das Nações Unidas para a Infância

**VN** - Verdadeiros Negativos

**VP** - Verdadeiros Positivos

**WHO** – World Health Organization

## **CAPÍTULO I: INTRODUÇÃO**

### **1. Introdução**

Um dos Objetivos de Desenvolvimento Sustentável estabelece a redução global do rácio de mortalidade materna para menos de 70 mortes por 100.000 nascimentos vivos até 2030.

O número de óbitos maternos que ocorrem em Moçambique tem registado um ligeiro aumento, tornando-se assim numa das principais preocupações do governo moçambicano. Diante destes fatos, o Ministério da Saúde vem conduzindo uma série de medidas que visam melhorar a qualidade de assistência e atenção à saúde das mulheres e crianças. Várias evidências vêm apontando as estatísticas sobre a mortalidade materna em Moçambique, não só, mas também além-fronteiras como um dos melhores indicadores da saúde da população feminina, não obstante, é também uma das melhores ferramentas de gestão de políticas públicas que concorrem para a diminuição destes índices, constituindo em todo o caso um dos mais importante desafio de saúde pública.

Uma análise deste indicador no contexto internacional indica que foi a Europa, no século XIX, a primeira a preocupar-se com a Mortalidade Materna, iniciando ações que visavam melhorar a qualidade de assistência à saúde da mulher. Durante todo século XX várias ações de prevenção foram implementadas para a redução dos níveis de mortalidade materna. Viana, Novaes, & Calderon (2011). Nas últimas décadas do último século, em todo o mundo foram registados avanços significativos na área científica e tecnológica, principalmente nas áreas de atendimento a mulher e criança. Razão pela qual, tornou-se praticamente inadmissível que o processo de reprodução cause sequelas às mulheres assim como os seus bebés, levando-os à morte por qualquer fator ou causa.

Apesar de todos os esforços realizados e avanços notáveis, ainda no século XXI as mulheres morrem durante o ciclo gravídico-puerperal, com 90% destes óbitos de causas evitáveis e 99% ocorrendo em países em desenvolvimento Viana, Novaes, & Calderon (2011).

Moçambique, ao assinar a Declaração dos Objetivos do Desenvolvimento do Milénio (ODM), comprometeu-se em reduzir a Taxa de Mortalidade Infantil assim como taxa de mortalidade materna dos 408 por 100.000 nascimentos vivos registados em 2003 para 350 em 2010 e 250 em 2015 (ODM 5). Segundo os relatórios nacionais apresentados, esse objetivo ainda não foi

alcançado, dado que cerca de 10 mulheres morrem por dia e 3.840 mulheres ao ano por causas e complicações relacionadas a gravidez e parto possivelmente evitáveis. Medicusmundi (2011).

Devido aos altos rácios, a mortalidade materna tem sido apontada como o melhor indicador da saúde da população feminina e, conseqüentemente, a melhor ferramenta de gestão de políticas públicas voltadas para diminuição destes valores, pois as estatísticas deste indicador revelam uma preocupação, constituindo assim, um importante desafio de saúde pública que Moçambique enfrenta na atualidade, não obstante, é considerado uma das mais graves violações dos direitos humanos, ao mesmo tempo, avaliando a qualidade do serviço de saúde prestado à população, não só, mas também, correlacionando-o com o desenvolvimento socioeconómico. Trata-se de uma problemática eminentemente encontrada nos países em desenvolvimento, sendo na sua maioria evitáveis.

Diante da situação, dada a magnitude e impacto social que a mortalidade materna apresenta no seio da comunidade em Geral e em particular na mulher, e pelo fato de ser considerada um dos principais problemas de Saúde Pública, suscitou o interesse pelo estudo do mesmo, de modo a se apurar quais são os fatores contribuintes/determinantes, suas causas e possíveis propostas de redução dos seus níveis.

### **1.1. Objetivos**

#### **➤ Objetivo Geral**

- Analisar os óbitos maternos ocorridos ao longo do ano de 2020 na Província de Nampula.

#### **➤ Objetivos específicos**

- Avaliar a Evitabilidade das mortes
- Calcular a probabilidade de ocorrência do óbito materno.
- Identificar as principais causas de morte materna,
- Identificar os fatores associados que contribuíram para a ocorrência das Mortes Maternas na Província de Nampula;

A presente dissertação encontra-se organizada em capítulos, além da presente introdução. No Capítulo II, temos apresentamos o enquadramento do problema da mortalidade materna no Mundo, com enfoque para África e Moçambique em particular. Neste também estão apresentados os antecedentes e a magnitude.

No capítulo III estão apresentados os dados e métodos fazendo-se uma breve e explicação da metodologia adotada, a população de estudo bem como a sua amostra, são apresentadas as variáveis selecionadas, os procedimentos estatísticos utilizados para análise dos dados.

No capítulo IV uma apresentação de conteúdos estatísticos, isto é, a regressão logística, abordando alguns conceitos, estimação dos parâmetros, ajuste do modelo, os testes de significância, diagnósticos do modelo e a curva ROC, o que permitiu aprofundar conhecimentos teóricos relevantes sobre a análise dos dados do estudo.

No capítulo V, são apresentados os resultados numa abordagem prática de aplicação do modelo aos dados, aliando a teoria apresentada nos capítulos anteriores para obtenção dos resultados do estudo.

Por fim, é apresentada a discussão dos dados, principais conclusões do estudo, assim como algumas limitações e sugestões para a melhoria da situação da mortalidade materna.

## CAPÍTULO II: MORTALIDADE MATERNA

### 2. Mortalidade Materna

#### 2.1. Alguns conceitos

A estatística sobre a mortalidade tem sido apontadas como uma das principais fontes de informação dos sistemas de saúde no mundo inteiro e, em muitos países, constituem o tipo mais confiável de dados sobre a saúde OMS & OPAS, (2009).

Para melhor classificação é importante distinguir os conceitos de:

- **Morte Materna**

A Organização Mundial de Saúde (OMS), em 1994, na sua 10.<sup>a</sup> revisão da Classificação Internacional de Doenças (CID-10), definiu a morte materna como: “a morte de uma mulher durante a gestação ou num período de 42 dias após o término da gravidez, devida a qualquer causa relacionada com ou agravada pela gravidez ou por medidas tomadas em relação a ela, porém não devida a causas acidentais ou incidentais”

Esta definição foi adotada e aceite pelas associações de ginecologia e obstetrícia internacionais, destacando-se a *International Federation of Gynecology and Obstetrics* (IFGO) e a Federação Brasileira das Associações de Ginecologia e Obstetrícia. (FEBRASGO) Viana, Novaes, & Calderon, (2011)

- **Causas de morte**

Esta é definida como toda a doença ou afeção mórbida ou lesão que produziu ou contribuiu para a morte e as circunstâncias do acidente ou violação que produziram quaisquer de tais lesões. Esta não inclui sintomas e modos de morrer, tais como a insuficiência cardíaca ou respiratória OMS; OPAS (2009). Podendo estas ser codificados usando a CID-10. A causa de morte pode ser subdividida em causa básica e direta.

- **Causa Básica de morte**

É a doença ou lesão que iniciou a cadeia de acontecimentos patológicos que conduziram diretamente a morte ou as circunstâncias do acidente ou violação que produziu a lesão fatal OMS & OPAS, (2009).

- **Causa Direta de morte**

É apontada como a doença ou lesão que esteve presente nos momentos finais da vida do paciente ou no momento da morte.

## **2.2. Antecedentes e Situação**

As mortes resultantes da gravidez, parto e pós-parto são consideradas potenciais ameaças para as mulheres nas faixas etárias reprodutivas ou mulheres em idade fértil (15-49 anos em Moçambique). O preço que a maternidade assume sobre a vida e a saúde das mulheres, e, por conseguinte, nas suas famílias e comunidades, torna-se verdadeiramente trágico, uma vez que a mortalidade materna é, na sua maioria, evitável. Esforços tendentes a redução da mortalidade das mulheres tem sido, assim, uma das áreas de grande preocupação em praticamente todo o mundo, o que faz com que os governos de todo o mundo estabeleçam metas de tempo para a alcançar. National Institute of Public Cooperation and Child Development, (2015).

Do ponto de vista estatístico o indicador a avaliar é o rácio ou razão de mortalidade materna, pois segundo Figueiredo, (2010), este indicador estima a frequência de óbitos que ocorreram em mulheres atribuídos a causas ligadas à gravidez, ao parto e ao puerpério, relativamente ao total de nascimentos vivos. Por conseguinte, tem-se que o número de nascimentos vivos é adotado como uma aproximação ao total de mulheres que estiveram grávidas.

Esta razão é estimada pela seguinte fórmula:

$$\text{RMM} = \frac{\text{Óbitos Maternos}}{\text{Nascimentos vivos}} * 100.000$$

Uma morte materna geralmente traz consigo repercussões negativas graves no seio da família, na sua comunidade e não só, mas também na sociedade em geral, pois a mulher (mãe) é geralmente, o suporte moral, social e económico da família e da comunidade. Em termos comparativos, os países desenvolvidos, na sua maioria, são caracterizados por apresentar taxas de mortalidade materna baixas, enquanto os países em desenvolvimento geralmente apresentam taxas de mortalidade materna elevadas. Instituto Nacional de Estatística -INE, (2009).

Segundo Viana, Novaes, & Calderon, (2011), a maioria dos óbitos maternos é possivelmente evitável, principalmente quando as gestantes têm acesso a uma boa assistência sanitária que inclui o pré-natal, perinatal e pós-natal imediato.

A mortalidade materna é um indicador que reflete claramente a qualidade de vida de uma região, em particular ou especial aos cuidados prestados e à assistência à saúde da mulher. Fatores sociais como idade, raça, estado civil, escolaridade e padrão socioeconómico são geralmente descritos que demonstram existir população mais vulnerável e com maior risco de complicações.

É deveras importante estudar os fatores relacionados a mortalidade materna, pois a sua análise faz-se essencial e permite a formulação de políticas de gestão e planeamento, tendo em conta que o número de mortes maternas de um país é considerado um excelente indicador da sua realidade social, estando inversamente relacionado ao nível de desenvolvimento humano, traduzida pelo nível socioeconómico, a qualidade da assistência, a iniquidade entre os géneros e as estratégias de promoção da saúde pública. Ferraz & Bordignon, (2012).

### **2.3.Magnitude da Mortalidade Materna**

Os óbitos por gravidez e nascimento constituem potenciais ameaças a saúde das mulheres em idade fértil (as consideradas nas faixas etárias reprodutivas). O cenário que a maternidade insegura assume a vida e a saúde das mulheres, nas suas famílias ou nas suas comunidades, torna-se realmente trágico, por ser principalmente evitável. A redução da mortalidade das mulheres tem sido, portanto, uma área de grande preocupação e os governos de todo o mundo estabeleceram metas de tempo para alcançá-la. O óbito materno é um importante indicador do alcance dos serviços clínicos eficazes de saúde para os pobres e, por sua vez, atua como uma das medidas compostas para avaliar o progresso do país.

Segundo UNFPA, (2004) *apud* Nhatave, (2006), estima-se que na África Subsariana, a mulher durante a sua vida corre risco de 1 em 16 de morrer durante a gravidez e o parto, comparado com o risco de 1 em 2800 nos países desenvolvidos. Em cada minuto, alguma parte do mundo, uma mulher morre por complicações de gravidez ou de parto, totalizando 529.000 mortes cada ano, deixando para trás crianças que serão mais suscetíveis de morrer por ficarem órfãs de mãe, 95% destas mortes ocorrem nos países em desenvolvimento (África e Ásia). Somente 4% das mortes maternas ocorrem na América Latina e Caraíbas e menos de 1% nas regiões mais desenvolvidas do mundo.

Em 2000, o número estimado de mortes maternas no mundo foi de 529.000 e 95% dessas mortes ocorreram na África e na Ásia. Enquanto as mulheres em países desenvolvidos têm somente uma chance em 2.800 de morrer no parto e uma chance de 1 em 8.700 em alguns países, mulheres na África têm uma chance em 20. Em vários países, o risco de vida é maior que 1 em cada 10. Hunt & Mesquita, (2007)

Para cada mulher que morrer por complicações obstétricas, aproximadamente mais 30 sofrem lesões, infecções e incapacidades. Em 1999, por exemplo, a OMS estimou que mais de 2 milhões de mulheres que vivem em países em desenvolvimento permanecem sem tratamento para fistula obstétrica, uma lesão devastadora do parto. Não há uma única causa de morte e incapacidade para homens entre 15 e 44 anos que esteja próximo da magnitude da morte materna e da incapacidade. Hunt & Mesquita, (2007)

Dada a complexidade, a morte materna é possivelmente evitável e por tanto, socialmente inaceitável. A probabilidade de evitá-la oscila entre 90–95%, e está diretamente relacionada à qualidade da assistência recebida pela mulher durante a gestação, pré-parto, parto e pós-parto Borges, (2012).

#### **2.4.A mortalidade materna e seu impacto**

A morte materna, geralmente traz consigo consequências enormes no seio da família e geralmente representa uma grande perda na sua comunidade com repercussões ao longo das gerações, pois a mulher representa o suporte moral, social e muitas das vezes o suporte econômico da família e da comunidade, pois realiza trabalhos várias tarefas e trabalhos, que contribui para o aumento do rendimento ou redução das despesas nas famílias. Nhatave, (2006).

A morte materna traz consigo inúmeras consequências, sendo que para cada morte outras 20 mulheres sofrem de sequelas graves e deficiências relacionadas com a gravidez ou parto, somando ao final de um ano, equivalente a cerca de 10 milhões de mulheres que sobrevivem deste evento e experimentam uma complicação de saúde negativa, na sua maioria fístulas obstétricas. União Africana, (2013).

Com a ocorrência de uma morte materna, as crianças são as que mais sofrem, ficando vulneráveis com uma probabilidade de 3 a 10 vezes mais de morrer no período de 02 anos após a morte da mãe. O risco de morte para uma criança menor de 05 anos duplica se a sua mãe faleceu durante o parto. Cerca de 20% de doenças nas crianças menores de 05 anos está diretamente associada a fracas

condições de saúde materna e reprodutiva, nutrição, e qualidade de cuidados obstétricos e infantis. Nhatave, (2006).

O 5.º Objetivo do Desenvolvimento do Milénio preconizava melhorar a saúde da mulher, tendo com a principal meta n.º 6 reduzir em três quartos entre 1990 e 2015 a Taxa de mortalidade materna, facto que não foi alcançado por Moçambique. IDHS; PNUD; PUC MINAS, (2004).

## **2.5.Mortalidade Materna no Mundo**

Estima-se que um total de 13,6 milhões de mulheres tenham morrido nos 25 anos entre 1990 e 2015 devido a causas maternas. No entanto, ao longo desse tempo, o mundo tem registado progressos na redução da mortalidade materna, tendo registado um decréscimo global de 44%, passando de 385 em 1990 para 216 em 2015. Esta redução traduz-se numa diminuição de mais de 43% no número anual estimado de mortes maternas, de 532.000 em 1990 para 303.000 em 2015, e uma redução para metade do risco de vida global aproximado de uma morte materna de 1 em 73 para 1 em 180. De forma geral em todo o mundo, o rácio de mortalidade materna diminuiu em média 3,0% ao ano entre 2005 e 2015, mais do dobro da média anual estimada do declínio de 1,2 entre 1990 e 2000. WHO; UNICEF; UNFPA; World Bank Group, (2015).

Globalmente, as principais causas de mortes maternas são a hemorragia severa (24%), as causas indiretas (20%), a infeção (15%), o aborto inseguro (12%) e outras causas diretas (8%). Em Moçambique, as estatísticas existentes sobre as causas da mortalidade materna são das mortes maternas nos hospitais WHO; UNICEF; UNFPA; World Bank Group, (2015).

## **2.6.Mortalidade Materna em África**

Na África a mortalidade materna é inaceitavelmente elevada, com cerca de 295.000 mulheres que morreram durante a gravidez, parto ou no pós-parto em 2017. A grande maioria destas mortes (94%) ocorreu em Países com baixos recursos e a maioria poderia ter sido prevenidas ou evitáveis. WHO (2019).

Entre 1990 e 2005, os Objetivos de Desenvolvimento do Milénio 4 e 5 foram formulados para diminuir os rácios de mortalidade materna e infantil em três quartos e em dois terços respetivamente, o que não foi alcançado. Numa avaliação feita em 2008 apontava para morte de cerca de 358.000 mulheres ao nível mundial, vítimas de complicações obstétricas, isto é, resultantes

da gestação ou do parto, estimando-se que entre estes casos, 99% foram registados em países em desenvolvimento com maior destaque para a África Subsariana (87%), atingindo uma razão de mortalidade materna de 640 óbitos por 100.000 Nados Vivos. OMS, (2008).

Para Séraphin, et al., (2014), alguns países no Norte de África diminuíram tremendamente as taxas de mortalidade materna infantil. No entanto, a mortalidade materna entre as mulheres em nações de baixo rendimento continua 15 vezes maior do que a dos países do primeiro mundo. África Subsariana também relatou progressos.

Na África, mais de 60% dos Países têm rácios de mortalidade materna de mais de 300 mortes em cada 100.000 nascimentos vivos. Entre os países com altos rácios destacam-se o Chade e a Somália com cerca de mais de 1000 óbitos maternos em cada 100.000, a Nigéria (40.000), a República Democrática do Congo (15.000), Sudão (10.000), a Etiópia (9.000) e a República Unida da Tanzânia (8.500) e cada uma destas Nações contribui com cerca de 3 a 14% do rácio da mortalidade materna global. Ministério da Saúde, Instituto Nacional de Estatística e ICF International (2013)

Apesar destes altos rácios, alguns Países registam níveis considerados baixos, isto é, entre 20-99 mortes maternas em cada 100.000 nados vivos, destacando-se a Tunísia com 56, Egipto (66), Maurícias (60), São Tomé e Príncipe (70), Cabo-Verde (79) e Argélia (81). Os Países com rácios moderados (isto é, entre 100–299 mortes maternas em cada 100.000 nados vivos) incluem o Botsuana (160), Djibouti (200), Namíbia (200), Gabão (230), Eritreia (240) e Guiné Equatorial (UNIÃO AFRICANA, 2013). Moçambique regista um rácio de mortalidade materna situado em 480 por 100.000, Ministério da Saúde, Instituto Nacional de Estatística e ICF International (2013).

Nem tudo vai mal em África, pois destaca-se o sucesso pela realização ou alcance do objetivo do desenvolvimento do milénio (ODM 5) pela Guiné Equatorial que reduziu a taxa de mortalidade materna em 81%, passando dos 1200 para 240 em cada 100.000 nascimentos vivos. O alcance desse objetivo faz com que a Guiné Equatorial se situe entre os 10 países no mundo inteiro a alcançar esta meta. Existem vários Países na África com tendências consideradas promissoras no alcance deste objetivo, são eles o Egipto, São Tomé e Príncipe, Madagáscar e a Eritreia. União Africana (2013).

## **2.7.Mortalidade Materna em Moçambique**

Em Moçambique, a saúde materna é considerada uma das prioridades do Governo para a área de saúde, correspondendo a cerca de 80% do peso das atividades no setor, fazendo com que um importante desafio de saúde pública que Moçambique enfrenta atualmente seja a mortalidade materna. São notáveis os investimentos realizados no setor por forma a melhorar a atenção a mulher. Não obstante, apesar dos esforços ainda persistem grandes desafios no que a mortalidade materna diz respeito, principalmente na sua redução, pois o nível continua elevado, ao se situar em cerca de 408 mortes maternas em cada 100 mil nascidos vivos em 2003 segundo o Inquérito Nacional Demográfico de Saúde, (2003 e 2011) Instituto Nacional de Estatística -INE (2009)

Em Moçambique as cinco principais causas de mortalidade materna são: A rotura uterina, a hemorragia pós-parto, a pré-eclâmpsia severa/eclâmpsia, o HIV/SIDA e a Sépsis. São considerados como principais fatores que contribuem para as mortes maternas, são apontadas como: o uso insuficiente dos serviços de saúde, fraco uso das consultas pré-natais; baixa percentagem de partos institucionais e serviços pós-natais, as normas culturais desfavoráveis e perceções negativas sobre os serviços de saúde; os casamentos e gravidez precoces e a baixa cobertura do uso dos serviços de planeamento familiar. Centro de Investigação em Saúde Pública (2013)

Em Moçambique, os valores da razão da mortalidade materna considerados na literatura para a década de 90 variam de 910 a 1500 por 100.000 nascimentos vivos. World Health Organization (2013).

Segundo o Centro de Investigação em Saúde Pública (2013) os diferenciais da Mortalidade Materna por Província variam de acordo com a especificidade do local e sua cultura, sendo as Províncias da Região Norte as que apresentam maiores índices.

## **2.8.Auditoria ou Investigação do Óbito Materno**

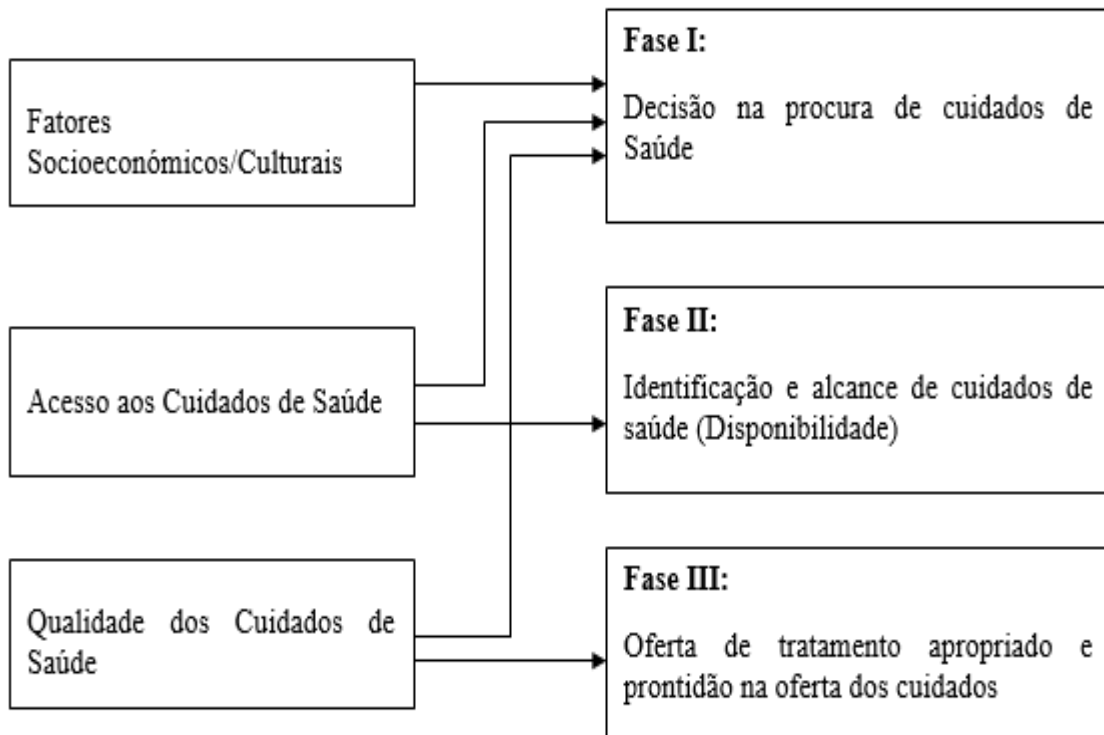
Após a ocorrência de um óbito materno, a Unidade Sanitária por onde o mesmo ocorreu, deve no prazo de 24h preencher a ficha de notificação em triplicado e enviar ao nível imediatamente superior. No prazo de até 07 dias o nível imediatamente superior (Comité Distrital) após a receção deve convocar o comité institucional do mesmo nível para investigar as circunstâncias em que o mesmo ocorreu, desde o momento da admissão da mulher, incluindo as intervenções nela feitas até

ao óbito. A seguir envia o relatório ao nível imediatamente superior (Comité Provincial) que por sua vez convoca os seus membros para uma discussão mais detalhada do óbito.

O comité ao analisar o óbito, deve iniciar pelo perfil socioeconómico da falecida, se esta fez ou não alguma consulta pré-natal e a assistência prestada nesta consulta, assistência prestada durante o parto, aborto e puerpério, o preenchimento correto da história clínica, bem como as intervenções feitas em todos os momentos, verificando assim se era um óbito materno declarado ou não, se a morte era evitável ou não, e quais seriam os fatores de Evitabilidade e as medidas de prevenção cabíveis.

Durante o processo de investigação das mortes foram identificadas várias barreiras do lado da procura de cuidados por parte das mulheres grávidas as quais foram denominadas por Fase I e II que refletem os atrasos na decisão de procurar cuidados, na identificação e acesso a um centro de saúde ou unidade hospitalar. Essas barreiras impedem as mulheres nos países em desenvolvimento de utilizar e aceder aos serviços de parto. Knight, Self, & Kennedy (2013).

Por tanto, foram resumidos em três revisões através do diagrama abaixo:



**Figura 1: Modelo de três atrasos**

**Fonte: Knight, Self, & Kennedy (2013)**

Segundo a ficha de auditoria de mortes maternas, podem ser encontrados os seguintes fatores em cada uma das fases:

- **Fatores Socioeconômicos/Culturais** são influenciados pelas situações pessoais ou familiares ao nível da comunidade, também chamados de 1ª demora, tais como: atraso da mulher na procura dos cuidados de saúde e recuso de tratamento na US.
- **Fatores de acessibilidade as Unidades Sanitárias**, também chamados de logísticos, podem ser: falta de transporte de casa para a Unidade Sanitária, Falta de transporte e/ou comunicação entre as Unidades Sanitárias.
- **Qualidade de cuidados:** Também designados por fatores Institucionais na Unidade Sanitária, incluem: falta de equipamento/consumíveis, falta de medicamento e falta de pessoal capacitado.

## **CAPÍTULO III: DADOS E MÉTODOS**

### **3. Dados e Métodos**

Para aos objectivos foram analisados os Processos de mulheres grávidas que deram entrada na maternidade e que tiveram alguma complicação obstétrica, resultando em óbitos (óbitos maternos discutidos pelo Comité Provincial de auditoria das Mortes Maternas e Neonatais) e as que não resultaram em óbitos durante o ano 2020, com dados individualizados, para se apurar quais foram os principais causas e fatores que contribuíram para a ocorrência do óbito. Os dados serão analisados utilizando o pacote estatístico SPSS versão 26, auxiliado com o pacote Jamovi versão 2.3.28 com aplicação do modelo logístico.

Para tal serão consideradas como variáveis de estudo:

- a) Número de consultas pré-natais (ante natal),
- b) A idade gestacional,
- c) A idade da mulher,
- d) Número de gestações,
- e) Número de filhos,
- f) Causa básica,
- g) Causa direta,
- h) Desfecho (Óbito/Não Óbito)
- i) Evitabilidade (Evitável/Não Evitável)

#### **3.1. População**

Ao longo do ano de 2020 a Província registou 10703 complicações obstétricas (diretas e indiretas) resultando em 205 óbitos.

#### **3.2. Amostra**

Para a realização do estudo serão tomados em consideração duas amostras:

- Das mulheres grávidas que deram entrada na maternidade e tiveram alguma complicação obstétrica cujo desfecho foi óbito em 2020 (205 óbitos).

Durante o ano de 2020 houve um registo de cerca de 10703 complicações obstétricas de mulheres grávidas que deram entradas nas maternidades da Província de Nampula, tendo resultado em 205 óbitos.

### **3.3.Análise de dados**

Os dados recolhidos serão analisados, utilizando o pacote estatístico SPSS (Statistical Package for Social Science) Versão 26, auxiliado ao pacote de análise de dados Jamovi versão 2.3.28. Terão como base a utilização do modelo de regressão logística.

## CAPÍTULO IV: REGRESSÃO LÓGISTICA

### 4. Regressão Logística

Os métodos de regressão tornaram-se um componente de qualquer análise de dados, tendo como principal foco a descrição da relação entre uma variável de resposta e uma ou mais variáveis explicativas. Relações essas que, em muitas vezes, a variável de resultado é discreta, assumindo dois ou mais valores possíveis. A regressão logística é fundamental nas últimas décadas do século XX, o método padrão de análise e a sua aplicação vem ganhando maior visibilidade. Inicialmente foi concebido e aplicado em pesquisas epidemiológicas, tendo ganho a sua aceitação, de tal forma que atualmente, o método é comumente empregue em muitos campos de atuação humana, como pesquisas biomédicas, negócios e finanças, criminologia, ecologia, engenharia, política de saúde, linguística e biologia da vida selvagem.

A regressão logística é usada principalmente para modelar uma variável binária (0,1) com base em uma ou mais outras variáveis, chamadas de preditores. A variável binária modelada é geralmente chamada de variável de resposta ou variável dependente. Hilbe (2015)

Segundo Garson (2014), a regressão logística concentra-se nos seguintes termos:

- **Odds:** é uma razão formada pela probabilidade de um evento ocorrer dividida pela probabilidade de que o evento não ocorra, é geralmente a probabilidade de obter um “1” dividida pela probabilidade de obter um “0”. Ou seja, na regressão logística binária, “1” é previsto e “0” é geralmente a categoria de referência. No caso da multinomial, qualquer valor mais baixo pode ser previsto e o valor codificado mais alto é geralmente a categoria de referência;
- **Odds ratio:** As razões de chances são a principal medida de tamanho do efeito para a regressão logística;
- **Log odds:** é o coeficiente previsto pela regressão logística também chamado de “*logit*”. Que é o logaritmo natural das chances de a variável dependente igualar algum valor (por exemplo, 1 em vez de 0 na regressão logística binária). O log das probabilidades é igual ao log natural da probabilidade de o evento ocorrer dividido pela probabilidade do evento não ocorrer.

- **Estimativas de parâmetros:** estes são os coeficientes de regressão logística para as variáveis independentes e a constante em uma equação de regressão logística. As estimativas de parâmetros estão no lado direito da equação de regressão logística e os *logits* estão no lado esquerdo.

A regressão logística é também considerada uma abordagem parcialmente não paramétrica, não exigindo suposições sobre o comportamento probabilístico dos dados em estudo, não só, mas também pelo fato deste apresentar pelo menos duas razões para a utilização de variáveis-resposta dicotômicas, que segundo Hosmer & Lemeshow (2000) podem ser:

- a) Do ponto de vista matemático, este modelo é extremamente flexível e fácil de ser utilizado;
- b) Permite uma interpretação de resultados bastante simples, rica e direta. Não só, como também, dá a possibilidade ao pesquisador de estimar de forma direta a probabilidade de uma observação pertencer a determinada classe, tornando possível testar a significância de um grande número de variáveis que contribuem mais para a separabilidade entre classes, fazendo de igual modo suposições sobre os comportamentos probabilísticos das variáveis independentes.

Dada a facilidade de utilização deste método, foi ganhando inúmera popularidade a sua aplicação em análise de dados e segundo Hilbe, (2015), este está presente em quase todos os pacotes estatísticos de análise de dados e de uso geral, sendo considerado uma das rotinas mais importantes de se usar em áreas como a saúde, estatística médica, área financeiras, etc.

Ainda segundo Hilbe, (2015), de uma forma geral, a regressão logística é normalmente usada em pesquisas devido a três principais propósitos:

- Prever a probabilidade de uma variável de resultado ou resposta;
- Categorizar resultados ou previsão;
- Estimar o risco associados a cada variável preditora do modelo.

Este modelo de regressão, permite o uso de um método para calcular ou prever a probabilidade de um evento específico, não só, mas também, quando se tem uma variável dependente.

#### **4.1. Regressão logística univariada ou simples**

Tomando em consideração que  $x_1$  seja a variável explicativa, teremos então o  $y_1$ , como o número de ocorrências de um determinado evento, cujas observações podem ser representadas por  $i = 1, 2, \dots, n$ . No caso se assume que a variável resposta tem uma distribuição binomial. Na sua forma simples, a regressão logística, admitindo que o valor médio condicional, pode ser expresso como uma equação linear em “x” é expressa por:

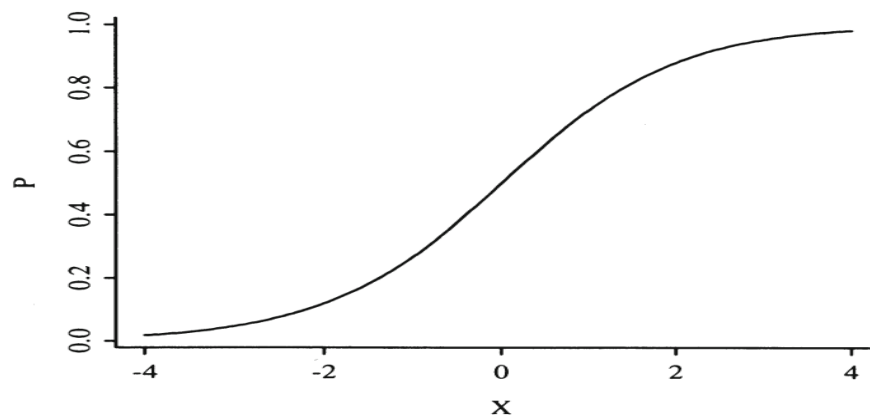
$$E[Y|X] = \beta_0 + \beta_1 x \quad (1)$$

Por forma a simplificar a notação é importante considerar que  $\pi(x) = E[Y|X = x]$ . Quando a variável resposta toma 2 valores distintos, assume-se que:

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \quad (2)$$

Uma vez que a variável binária  $Y$  é a  $Prob\{Y = 1\}$ , tal modelo, por definição, não pode ajustar os dados em toda a faixa dos preditores, pois um modelo puramente linear  $E\{Y|X\} = Prob\{Y = 1|X\} = X\beta$  só pode permitir que a  $Prob\{Y = 1|X\}$  exceda 1. Harrell, Jr. (2001)

A função do modelo pode ser expressa pelo seguinte gráfico:



**Figura 2: Função Logística**

**Fonte:** Frank E. Harrell, Jr (2001)

#### 4.1.1. Transformação Logit

Ao modelo logístico pode se aplicar uma transformação fulcral, denominada por **logit** cujo objetivo é linearização o modelo através do logaritmo. Esta transformação define-se:

$$g(x) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \beta_0 + \beta_1 x \quad (3)$$

$$g(x) = \ln\left(\frac{\frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}}{1 - \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}}\right) = \ln\left(\frac{\frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}}{\frac{1}{1 + e^{\beta_0 + \beta_1 x}}}\right)$$

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

Segundo Hosmer & Lemeshow, (2000) no modelo de regressão logística, a transformação assume uma grande importância, por possui várias propriedades do modelo de regressão linear, tais como:

- A função logit, é linear nos parâmetros;
- Pode ser contínua;
- Os seus valores podem variar em  $\mathbb{R}$ .

Geralmente a transformação do logit de  $\pi(x)$ , é chamada de *Odds*, e é expressa em  $\frac{\pi(x)}{1 - \pi(x)}$

#### 4.1.2. Estimação dos parâmetros do modelo

No modelo de regressão linear, assume-se que uma observação da variável de resposta pode ser expressa como  $Y_x = E[Y|X = x] + \varepsilon_x$ , onde  $\varepsilon$  é chamado de erro e representa o desvio de uma observação da média condicional. A suposição mais comum é que  $\varepsilon$  obedece a uma distribuição normal com média zero e alguma variância constante entre os níveis da variável independente. Hosmer & Lemeshow, (2000).

Segue-se que a distribuição condicional da variável de resposta dado  $x$  será normal com média  $E[Y|X]$  e uma variância constante. Para o caso da variável de resposta dado  $x$  como  $y_x = \pi(x) + \varepsilon$ ,

O intervalo de valores para uma variável de resultado contínua é de  $-\infty$  a  $+\infty$ , e esse mesmo intervalo se aplica a qualquer estatística calculada, como a média. Em contraste, uma variável de

resultado binária como é o caso na regressão logística é restrita a ser 0 ou 1 e qualquer estatística de resumo resultante deve ser 0 ou 1 ou ter um valor entre esses dois extremos. Assim, um resultado binário leva a estimativas de proporções com uma característica particular. Tai & Machin, (2014).

Para Agresti, (2019) em uma variável de resposta binária  $Y$ , modela-se a probabilidade de sucesso,  $P(Y=1)$ . No caso de variável explicativa única, denota-se isso por  $\pi(x)$  para enfatizar que o valor de  $P(Y=1)$  depende do valor  $x$  dessa variável. Assumindo-se que as observações são variáveis binomiais independentes com o parâmetro  $\pi(x)$ , que por sua vez varia conforme o valor de  $x$ , para tal, a quantidade  $\varepsilon$  pode assumir um dos dois valores possíveis:

$$a) Y = 1 \rightarrow \varepsilon = 1 - \pi(x) \text{ com probabilidade } \pi(x) \quad (4)$$

onde:

- $Y$  é o resultado binário (por exemplo:  $Y = 1$  para o “sucesso” e  $Y = 0$  para o “fracasso”)
- $\pi(x) = P[Y = 1|X = x]$  é a probabilidade predita do sucesso dada a entrada " $x$ "
- $\varepsilon = 1 - \pi(x)$  é o erro residual quando o verdadeiro resultado é  $Y = 1$ ,

Assim:

$$\pi(x) = P[Y = 1|X = x] = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

$$b) Y = 0 \rightarrow \varepsilon = -\pi(x) \text{ com probabilidade } 1 - \pi(x), \text{ onde} \quad (5)$$

$$1 - \pi(x) = P[Y = 0|X = x] = \frac{1}{1 + e^{\beta_0 + \beta_1 x}}$$

A estimação dos parâmetros do modelo é feita utilizando o método de máxima verosimilhança, fornecendo em geral valores para os parâmetros  $\beta_0$  e  $\beta_1$  que em princípio são desconhecidos, de modo a maximizar a probabilidade de se obter determinado conjunto de valores. Deste modo a função de verosimilhança será expressa por:

$$P[Y_1 = y_1, \dots, y_n | \beta_0, \beta_1] = \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i} = \prod_{i=1}^n \left( \frac{\pi_i}{1 - \pi_i} \right)^{y_i} (1 - \pi_i)^{1-y_i} \quad (6)$$

Aplicando o logaritmo aos dois membros da equação anterior tem-se:

$$L(\beta_0, \beta_1 | (x_i, y_i)) = \sum_{i=1}^n y_i(\beta_0 + \beta_1 x_i) - \sum_{i=1}^n \ln(1 + e^{\beta_0 + \beta_1 x_i}) \quad (7)$$

Na estimação por máxima verosimilhança os parâmetros  $\beta_0$  e  $\beta_1$  são os valores de estimadores  $\hat{\beta}_0$  e  $\hat{\beta}_1$  que irão maximizar o logaritmo da função de verosimilhança. Deste modo para maximizar a função deve-se derivar em relação aos parâmetros do modelo, daí que se obtém a seguinte expressão:

$$\frac{\partial}{\partial \beta_0} L(\beta_0, \beta_1 | (x_i, y_i)) = \sum_{i=1}^n y_i - \sum_{i=1}^n \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \quad (8)$$

$$\frac{\partial}{\partial \beta_1} L(\beta_0, \beta_1 | (x_i, y_i)) = \sum_{i=1}^n y_i x_i - \sum_{i=1}^n x_i \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \quad (9)$$

A seguir deve-se igualar a zero e substituir  $\beta_0$  e  $\beta_1$  por  $\hat{\beta}_0$  e  $\hat{\beta}_1$ , obtendo-se

$$\begin{aligned} \sum_{i=1}^n y_i - \sum_{i=1}^n \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}} &= 0 \\ \sum_{i=1}^n y_i x_i - \sum_{i=1}^n x_i \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}} &= 0 \end{aligned} \quad (10)$$

Por forma a resolver as equações acima, é necessário recorrer aos chamados métodos iterativos numéricos, cujos resultados decorrentes da aplicação deste método devem ser incluídos na matriz denominada informação de Fisher, através da seguinte expressão:

$$I(\hat{\beta}) = \begin{bmatrix} \sum_{i=1}^n \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{(1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i})^2} & \sum_{i=1}^n x_i \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{(1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i})^2} \\ \sum_{i=1}^n x_i \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{(1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i})^2} & \sum_{i=1}^n x_i^2 \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{(1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i})^2} \end{bmatrix} \quad (11)$$

Uma vez obtidas a estimativa dos parâmetros do modelo é possível determinar as probabilidades estimadas através da seguinte expressão:

$$\pi_i = \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}} \quad (12)$$

### 4.1.3. Ajustamento do Modelo

Considerando uma amostra de  $n$  observações independentes do par com características  $(X = x, Y = y)$  ou  $(x_i, y_i)$ ,  $i = 1, 2, \dots, n$ , onde  $y_i$  denota o valor de uma variável binária e  $x_i$  é o valor da variável independente e supondo-se que a variável de resposta seja codificada como 0 ou 1, representando a ausência ou a presença desta mesma característica, respetivamente. De modo a ajustar o modelo de regressão logística da equação (2) a um conjunto de dados exige que se estimem os valores de  $\beta_0$  e  $\beta_1$ , que são os parâmetros desconhecidos, através do método da máxima verossimilhança (Maximum Likelihood Estimation) Hosmer & Lemeshow, (2000). Para estimar estes parâmetros desconhecidos, o método mais usado é o dos mínimos quadrados. Neste método escolhe-se os valores de  $\beta_0$  e  $\beta_1$  que minimizam a soma dos desvios quadrados dos valores observados de  $Y$  a partir dos valores previstos com base no modelo.

Os estimadores de  $\beta$  são valores que maximizam o  $l(\beta)$ . Embora não haja uma solução fechada, recomendar-se o uso de métodos de algoritmo de Newton-Raphson de modo a encontrar os parâmetros ótimos.

Como dito anteriormente, se  $Y$  está codificado entre 0 e 1, a expressão para  $\pi(x)$  na equação (12), atribui para valores arbitrários de  $\beta_0$  e  $\beta_1$  o vetor do parâmetro a probabilidade condicional de que  $Y = 1$ , dado  $x$ . Assim se  $P[Y = 1|X]$ . Dado que  $1 - \pi(x)$  na equação (5), dá a probabilidade condicional de que  $Y = 0$ , dado  $x$ ,  $P[Y = 0|x]$ . Para os pares  $(x_i, y_i)$ . Assim:

$$\pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \quad (13)$$

Assumindo que as observações são independentes, a função de verossimilhança pode ser obtida como o produto dos termos dados na expressão (13), assim:

$$l(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \quad (14)$$

Aplicando o  $\log$  da função de verosimilhança, tem-se:

$$l(\beta) = \ln[L(\beta)] = \left[ \prod_{i=1}^n \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i} \right] \quad (15)$$

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

Supondo-se que  $\pi(x_i) = 1$  e  $1 - \pi(x_i) = 2$ , teremos:

$$\begin{aligned} l(\beta) &= \sum_{i=1}^n \left[ y_i(\beta_0 + \beta_1 x_i) + \ln \left[ \frac{1}{1 + e^{\beta_0 + \beta_1 x_i}} \right] \right] \quad (16) \\ &= \sum_{i=1}^n [y_i(\beta_0 + \beta_1 x_i) + \ln(1) - \ln(1 + e^{\beta_0 + \beta_1 x_i})] \\ &\quad \sum_{i=1}^n [y_i(\beta_0 + \beta_1 x_i) - \ln(1 + e^{\beta_0 + \beta_1 x_i})] \\ &\quad \sum_{i=1}^n [y_i(\beta_0 + y_i \beta_1 x_i) - \ln(1 + e^{\beta_0 + \beta_1 x_i})] \end{aligned}$$

Para se ter os valores de  $\beta$  que maximizam a função em (15), deve-se diferenciar  $L(\beta)$  em relação aos parâmetros  $\beta_0$  e  $\beta_1$ .

$$\frac{\partial \ln[L(\beta)]}{\partial \beta_0} = \sum_{i=1}^n \left[ y_i - \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \right] = \sum_{i=1}^n [y_i - \pi(x_i)] \quad (17)$$

$$\frac{\partial \ln[L(\beta)]}{\partial \beta_0} = \sum_{i=1}^n \left[ y_i x_i - x_i \frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}} \right] = \sum_{i=1}^n x_i [y_i - \pi(x_i)] \quad (18)$$

Após a derivação, temos de igual o resultado a zero, assim, teremos:

$$\sum [y_i - \pi(x_i)] = 0 \quad (19)$$

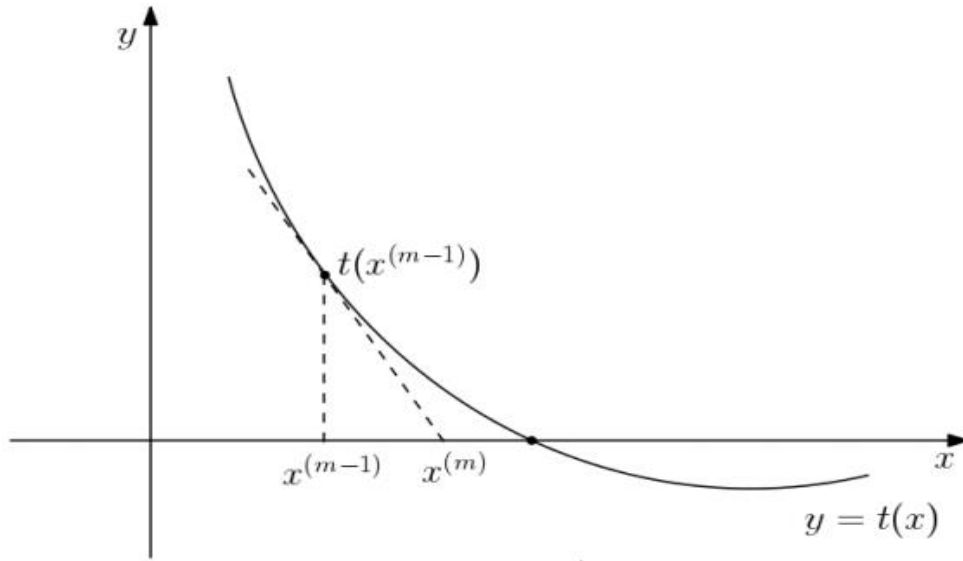
e

$$\sum x_i [y_i - \pi(x_i)] = 0 \quad (20)$$

#### 4.1.4. Métodos iterativos para estimação dos parâmetros que maximizam a função de verosimilhança

Métodos iterativos são um conjunto de procedimento que visam estimar a solução de um determinado problema, no qual se recorre às iterações sucessivos que geram uma sequência de aproximações de tal solução. O método comumente utilizado para a estimação de raízes é o de **Newton-Raphson**. Santos, (2017)

Este método é melhor visualizado na forma gráfica como abaixo demonstrada.



**Figura 3: Método de Newton-Raphson**

*Fonte:* Santos, (2017)

Os resultados obtidos através da aplicação do método representado na figura 1, são incluídos na matriz denominada informação de Fisher, que apresenta a seguinte forma:

$$I(\hat{\beta}) = \begin{bmatrix} \sum_{i=1}^n \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{(1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i})^2} & \sum_{i=1}^n x_i \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{(1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i})^2} \\ \sum_{i=1}^n x_i \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{(1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i})^2} & \sum_{i=1}^n x_i^2 \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{(1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i})^2} \end{bmatrix} \quad (20)$$

Uma vez obtidas a estimativa dos parâmetros do modelo é possível calcular as probabilidades estimadas através da equação (2):

$$\pi_i = \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}$$

Na função logística, os parâmetros  $\hat{\beta}_0$  e  $\hat{\beta}_1$  têm significados semelhantes

#### 4.1.5. Teste de significância dos Coeficientes

Após a estimação dos coeficientes do modelo ajustado é importante fazer uma avaliação da significância das variáveis no modelo, o que implica geralmente a formulação hipótese do teste

para determinar se as variáveis independentes no modelo estão significativamente relacionadas à variável de resultado. Para testar a significância do coeficiente de uma variável em um modelo é importante verificar se o modelo que inclui a variável em questão nos diz mais sobre a variável de resultado (ou resposta) do que um modelo que não inclui essa variável. Hosmer & Lemeshow, (2000).

Assim, comparam-se os valores observados da variável de resposta aos valores previstos por cada um dos dois modelos (o primeiro com a variável e o segundo sem a variável em alusão).

#### a) Teste da razão de verosimilhança

O teste da razão de verosimilhança é usado para avaliar a significância de cada um dos parâmetros  $\beta$  do modelo, através da comparação das diferenças entre o **log** da função de verosimilhança do modelo com todas as variáveis e o **log** da função de verosimilhança sem a variável. Sendo expressa por:

$$G = -2 \ln \left[ \frac{\text{Verosimilhança do modelo com a variável (ajustado)}}{\text{Verosimilhança do modelo sem a variável (saturado)}} \right] \quad (21)$$

ou

$$G = -2 \sum_{i=1}^n \left[ y_i \ln \left( \frac{\hat{\pi}_i}{y_i} \right) + (1 - y_i) \ln \left( \frac{1 - \hat{\pi}_i}{1 - y_i} \right) \right] \quad (22)$$

Onde:

$$\hat{\pi}_i = \hat{\pi}(x_i)$$

Para situações em que os valores das variáveis do modelo assumem os resultados 0 e 1, a probabilidade do modelo sem a variável é 1. No caso em que temos  $\hat{\pi}_i = y_i$ , a probabilidade é dada por:

$$l(\text{Verosimilhança do modelo sem a variável}) = \prod_{i=1}^n y_i^{y_i} * (1 - y_i)^{(1-y_i)} = 1 \quad (23)$$

Olhando para a equação (21), o deviance é dado por:

$$G = -2 \ln(\text{likelihood do modelo ajustado})$$

No modelo simples em que a variável independente não está no modelo é fácil a estimativa de máxima verosimilhança de  $\beta_0$  é o  $\ln(n_1/n_0)$  onde  $n_1 = \sum y_i$  e  $n_0 = \sum(1 - y_i)$  e o valor do preditor considerado constante é  $n_1/n_0$ . Assim, o valor do G é obtido através de:

$$G = -2 \ln \left[ \frac{\left(\frac{n_1}{n}\right)^{n_1} \left(\frac{n_0}{n}\right)^{n_0}}{\prod_{i=1}^n \hat{\pi}_i^{y_i} (1 - \hat{\pi}_i)^{(1-y_i)}} \right] \quad (24)$$

ou

$$G = -2 \left\{ \sum_{i=1}^n [y_i \ln(\hat{\pi}_i) + (1 - y_i) \ln(1 - \hat{\pi}_i)] - [n_1 \ln(n_1) + n_0 \ln(n_0) - n \ln(n)] \right\} \quad (25)$$

A estatística G é chamada de **deviance**, pois ela desempenha o mesmo papel que a **SQRes** como na regressão linear, permitindo a comparação entre os valores observados e preditos.

Tomando em consideração que a hipótese de que  $\beta_1 = 0$ , a estatística de G apresenta distribuição assintótica de qui-quadrado com grau de liberdade 1, tem-se:

$$H_0: \beta_1 = b_1$$

$$H_1: \beta_1 \neq b_1, \text{ para } b_1 \in \mathbb{R}$$

Assim, rejeita-se a  $H_0$  se  $\mathbb{P}[\chi^2 > G] < \alpha$

## b) Teste de Wald

O teste de Wald é baseado na distribuição assintótica de  $\beta$  e é uma generalização do teste  $t$  de Student, utilizado também para avaliar se um parâmetro é estatisticamente significativo. A estatística deste teste é obtida através da razão do coeficiente pelo seu respectivo erro padrão.

Segundo Hosmer & Lemeshow, (2000), Este teste é obtido comparando-se a estimativa de máxima verosimilhança do parâmetro  $\hat{\beta}_1$ , com uma estimativa de seu erro padrão. A razão resultante, sob a hipótese de que  $\beta_1 = 0$ , seguirá uma distribuição normal padrão. Em alguns casos, costuma não rejeitar a hipótese nula quando esta deveria ser rejeitada. Examinando a eficiência do teste de Wald

recomenda-se que o teste da razão de verossimilhança seja usado quando há dúvidas de que o teste de Wald tenha falhado.

Portanto ele deve avaliar também as hipóteses em:

$$\text{Hipóteses: } \begin{cases} H_0: \beta_1 = \beta_2 = \dots \beta_n = 0 \\ H_1: \text{Pelo menos um } \beta \neq 0 \end{cases}$$

Não só, olhando para as hipóteses, a estatística do teste, para avaliar se o parâmetro  $\mathbf{b}$  é igual a zero, deve assim ser especificada.

$$W = \frac{\hat{\beta}}{\sqrt{\text{var}(\hat{\beta})}} \sim N(0,1) \quad (26)$$

ou

$$w_j = \frac{\hat{\beta}}{\text{var}(\hat{\beta}_j)} \bigcap_{\text{sob } H_0} \chi^2_{(1)} \quad (27)$$

### c) Teste de Score

A estatística do Teste de Score é dada pela seguinte expressão:

$$TS = \frac{\sum_{i=1}^n x_i (y_i - \bar{y})}{[\bar{y}(1 - \bar{y}) \sum_{i=1}^n (x_i - \bar{x})^2]^{1/2}} \quad (28)$$

Onde  $\bar{y} = \bar{\pi}$ , é a proporção de sucesso

Neste teste é importante considerar as seguintes hipóteses

$$\text{Hipóteses: } \begin{cases} H_0: \beta_1 = \beta_2 = \dots \beta_n = 0 \\ H_1: \text{Pelo menos um } \beta \neq 0 \end{cases}$$

#### 4.1.6. Intervalo de Confiança para os coeficientes

Segundo Hosmer & Lemeshow, (2000), para testar a significância do modelo, é importante calcular e interpretar os intervalos de confiança para os estimadores  $\beta_0$  e  $\beta_1$ , baseados nos respectivos testes de Wald.

Assim, o intervalo a  $100(1 - \alpha)\%$  de confiança para  $\beta_0$  é dado por:

$$\hat{\beta}_1 \pm Z_{1-\frac{\alpha}{2}} \widehat{SE}(\hat{\beta}_1) \quad (29)$$

e para o intercepto temos

$$\hat{\beta}_0 \pm Z_{1-\frac{\alpha}{2}} \widehat{SE}(\hat{\beta}_0) \quad (30)$$

O Intervalo baseado no estimador do desvio padrão de  $\hat{\beta}_0$ , onde  $Z_{1-\alpha/2}$  é o quantil 100  $(1 - \alpha/2)\%$ , da distribuição normal padrão.

Uma vez que o **logit** é a parte linear do modelo de regressão logística e, como tal, é mais parecido com a linha ajustada em um modelo de regressão linear, o seu estimador é dado por:

$$\hat{g}(x) = \hat{\beta}_0 + \hat{\beta}_1 x \quad (31)$$

Assim, o estimador da variância do estimador do **logit** requer a obtenção da variância de uma soma, como parte linear do modelo de regressão logístico através da seguinte expressão:

$$\widehat{Var} [\hat{g}(x)] = \widehat{Var} (\hat{\beta}_0) + x^2 \widehat{Var} (\hat{\beta}_1) + 2x \widehat{Cov} (\hat{\beta}_0, \hat{\beta}_1) \quad (32)$$

Os valores de  $\widehat{Cov} (\hat{\beta}_0, \hat{\beta}_1)$ , são geralmente obtidos pelo método de máxima verosimilhança, através da matriz de variância e covariâncias. Assim deve se ter em conta a seguinte expressão:

$$\hat{g}(x) \pm z_{1-\alpha/2} \widehat{SE}[\hat{g}(x)] \quad (33)$$

onde  $\widehat{SE}[\hat{g}(x)]$  é a raiz quadrada positiva do estimador de variância

A variância de uma soma é igual à soma da variância de cada termo e duas vezes a covariância de cada par possível de termos formados a partir dos componentes da soma. Os pontos finais de um intervalo de confiança baseado em Wald de 100  $(1 - \alpha)\%$  para o **logit**. Neste caso, tem-se:

Por tanto, o estimador do **logit** e seu intervalo de confiança fornecem uma base para o estimador do valor ajustado, neste caso a probabilidade logística e seu intervalo de confiança associado.

Os limites do intervalo de confiança baseado em Wald de 100  $(1 - \alpha)\%$  para o valor ajustado são:

$$\frac{e^{\hat{g}(x) \pm z_{1-\alpha/2} \widehat{SE}[\hat{g}(x)]}}{1 + e^{\hat{g}(x) \pm z_{1-\alpha/2} \widehat{SE}[\hat{g}(x)]}} \quad (34)$$

Para o **Odds** considera-se o seguinte intervalo e confiança, sendo dados  $\beta_0$  e  $\beta_1$  como limites inferior e superior do intervalo  $(IC_{\beta_1 1-\alpha})$ :

$$IC_{OddsRatio\ 1-\alpha} = [e^{\beta_l}, e^{\beta_s}] \quad (35)$$

## 4.2. Regressão logística múltipla

A maior parte dos estudos clínicos e epidemiológicos de respostas binárias são particularmente focalizados nos efeitos potenciais de vários preditores, quando estes são categóricos e em número reduzido. No entanto, para um número maior de preditores potenciais e / ou quando alguns são medições contínuas, os métodos de regressão têm uma série de vantagens. Vittinghoff *et al*, (2005)

Este modelo é particularmente usado para situações em que temos uma variável resposta com mais de uma variável explicativa. Para a representação destas variáveis em forma de vetores no  $k$  é dada pela seguinte expressão:  $X = (X_1, X_2, X_3, \dots; X_k)$ . Neste modelo, a representação da probabilidade condicional é expressa pela seguinte definição  $P = \left(\frac{1}{X}\right) = \pi X$  (SILVA, 2016)

A função de ligação (**logit**) é expressa por:

$$E[X] = \pi(X) = \frac{e^{g(x)}}{1 + e^{g(x)}} \quad (36)$$

e

$$g(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots, \beta_k X_k \quad (37)$$

Segundo Vittinghoff *et al*, (2005) a função **logit** é também designada de parte sistemática do modelo e pode ser dada por:

$$\log \left[ \frac{P(X_1, X_2, X_3, \dots; X_K)}{1 - P(X_1, X_2, X_3, \dots; X_K)} \right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots, \beta_k X_k \quad (38)$$

A expressão em (3) pode ser escrita em termos de probabilidade de resultado da seguinte a partir da seguinte forma:

$$P(X_1, X_2, X_3, \dots; X_K) = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots, \beta_k X_k)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots, \beta_k X_k)} \quad (39)$$

#### 4.2.1. Estimação dos parâmetros

A estimativa dos parâmetros do modelo, obtém-se através dos componentes do vetor  $\beta = (\beta_0, \beta_1, \dots, \beta_p)$  que correspondem aos parâmetros do modelo utilizando o método de máxima verossimilhança. Tendo em conta que a função de verossimilhança é a probabilidade dos dados observados como uma função dos parâmetros  $\beta$ , tem-se que  $y$  é codificado em 0 e 1, tal que  $\pi(x) = P(Y = 1|x)$  e  $1 - \pi(x) = P(Y = 0|x)$ . Quando  $y_i = 1$  a contribuição para a função de verossimilhança é igual a  $\pi(x_i)$  e quando  $y_i = 0$  a função de verossimilhança será  $1 - \pi(x)$ . Tomando em conta que o par  $(x_i, y_i)$  é  $\pi(x_i)^{y_i} * [1 - \pi(x_i)]^{1-y_i}$ .

Como de pressuposto, as observações são independentes, a função de verossimilhança é obtida pelo produto de cada par  $(x_i, y_i)$ .

Assim:

$$L(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} * [1 - \pi(x_i)]^{1-y_i} \quad (40)$$

Uma vez sabido que o objetivo é estimar o  $\beta$  que maximiza a equação acima, o **log** da verossimilhança será definido por:

$$L(\beta) = \ln(L(\beta)) = \sum_{i=1}^n \{y_i * \ln(\pi(x)) + (1 - y_i) * \ln(1 - \pi(x))\} \quad (41)$$

ou

$$L(\beta_0, \beta_1, \dots, \beta_p | (x_i, m_i, y_i)) = \sum_{i=1}^n [y_i g(x) - \ln(1 + e^{g(x)})] \quad (42)$$

Derivando a função em (14), igualando a zero e substituindo pelos estimadores dos parâmetros, tem-se a seguinte equação:

$$\sum_{i=1}^n y_i(1 + e^{g(x)}) - \sum_{i=1}^n e^{g(x)} = 0 \quad (43)$$

e

$$\sum_{i=1}^n x_i y_i(1 + e^{g(x)}) - \sum_{i=1}^n x_i e^{g(x)} = 0 \quad (44)$$

As equações descritas acima são não lineares em  $\beta$ , que são obtidas diferenciando a função de verosimilhança em relação aos coeficientes  $p+1$ , por isso, necessitam de outros métodos especiais iterativos de estimação.

#### 4.2.2. Qualidade do ajustamento do modelo de regressão logística múltipla

Assim como no caso univariado, para testar o ajustamento do modelo requer que se obtenham estimativas de vetor  $\beta' = (\beta_0, \beta_1, \dots, \beta_n)$ , sendo necessário formular as hipóteses:

- $H_0$ : O modelo ajusta-se aos dados
- $H_1$ : O modelo não se ajusta aos dados

O método de estimação usado é o de máxima verosimilhança, quase idêntico ao aplicado no caso do modelo univariado, diferindo no caso em que  $\pi(X)$  é definido como se apresenta na equação (34). Portanto, deve-se considerar  $p + 1$  equações de verosimilhança que são obtidas diferenciando a função de verosimilhança em relação aos coeficientes  $p+1$ .

Este método estima os coeficientes que maximizam a probabilidade de se achar a realização da variável dependente (Marôco, 2018). A variável independente  $Y$  toma apenas dois valores (0 e 1), sendo cada um deles uma variável aleatória dicotômica de tal forma que  $Y_j \sim B(1, \pi)$  assim:

$$P(Y = y_j) = \pi^{y_j}(1 - \pi)^{1-y_j} \quad (45)$$

Tomando em conta que todas as observações fossem independentes a função de verosimilhança que dá a probabilidade conjunta será dada por:

$$L = P(Y = y_1) \times P(Y = y_2) \times \dots \times P(Y = y_n) = \prod_{j=1}^n \pi_j^{y_j} (1 - \pi_j)^{1-y_j} = \quad (46)$$

$$= \prod_{j=1}^n \left( \frac{e^{\beta_0 + \beta_1 X_{1j} + \dots + \beta_p X_{pj}}}{1 + e^{\beta_0 + \beta_1 X_{1j} + \dots + \beta_p X_{pj}}} \right)^{y_j} * \left( \frac{1}{1 + e^{\beta_0 + \beta_1 X_{1j} + \dots + \beta_p X_{pj}}} \right)^{1-y_j}$$

As equações de verosimilhança resultantes podem ser expressas como segue:

$$\ell(\beta) = \log L(\beta) = \sum_{i=1}^n [Y_i \log P(Y_i = 1|X_i) + (1 - Y_i) \log(1 - P(Y_i = 1|X_i))] \quad (47)$$

Substituindo  $P(Y_i = 1|X_i)$  pela função logística:

$$\ell(\beta) = \sum_{i=1}^n [Y_i(X_i\beta) - \log(1 + e^{X_i\beta})] \quad (48)$$

Derivando  $\ell(\beta)$  em relação a  $\beta$  e igualando a zero, tem-se:

$$\frac{\partial \ell(\beta)}{\partial \beta} = \sum_{i=1}^n X_i (Y_i - P(Y_i = 1|X_i)) = 0 \quad (48)$$

Se considerarmos  $\hat{\beta}$  como a solução para essas equações, assim, os valores ajustados para o modelo de regressão logística múltipla são  $\hat{\pi}(x_i)$ , o valor de a expressão na equação (10) calculada usando  $\hat{\beta}$  e  $X_i$ .

Pode também estimar-se as variâncias e covariâncias dos coeficientes através da razão de máxima verosimilhança, que são obtidos a partir da matriz das segundas derivadas parciais da função de verosimilhança, cuja forma é:

$$\frac{\partial^2 L(\beta)}{\partial \beta_j^2} = - \sum_{i=1}^n x_{ij}^2 \pi_i (1 - \pi_i) \quad (49)$$

e

$$\frac{\partial^2 L(\beta)}{\partial \beta_j \partial \beta_l} = - \sum_{i=1}^n x_{ij} x_{il} \pi_i (1 - \pi_i) \quad (50)$$

para  $j, l = 0, 1, 2, \dots, p$  onde  $\pi_i$  denota  $\pi(x_i)$ . Considerando os pontos na matriz  $(p+1) \times (p+1)$  contendo o negativo dos termos dados nas equações acima (21) e (22) como  $I(\beta)$ . Esta matriz é denominada por matriz de informação observada. As variâncias e covariâncias dos coeficientes estimados são obtidas a partir da inversa desta matriz que denotamos como  $Var(P) = I^{-1}(\beta)$ .

Os estimadores das variâncias e covariâncias, que serão denotados por  $\widehat{Var}(\hat{\beta})$ , são obtidos avaliando  $Var(\beta)$  em  $\hat{\beta}$ . Usando  $\widehat{Var}(\hat{\beta}_j)$  e  $\widehat{Cov}(\hat{\beta}_j, \hat{\beta}_l)$ ,  $j, l = 0, 1, 2, \dots, p$ , para denotar os valores nesta matriz.

Na maioria das vezes, existe sempre a oportunidade de usar apenas os erros padrão estimados dos coeficientes estimados, que são denotados como:

$$\widehat{SE}(\hat{\beta}_j) = [\widehat{Var}(\hat{\beta}_j)]^{1/2} \quad (51)$$

para  $j, l = 0, 1, 2, \dots, p$ , será usada a notação no desenvolvimento de métodos para teste de coeficiente em estimativa de intervalo de confiança.

A formulação da matriz será útil no ajuste do modelo e avaliação do mesmo se  $\hat{I}(\hat{\beta}) = \mathbf{X}'\mathbf{V}\mathbf{X}$  onde  $\mathbf{X}$  é uma matriz diagonal  $n$  por  $n$  com elemento geral  $\hat{\pi}_i(1 - \hat{\pi}_i)$ . Por tanto a matriz é:

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \quad (52)$$

e a matriz  $\mathbf{V}$  é

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

### 4.2.3. Teste de significância do modelo

Para testar a significância dos parâmetros na regressão logística múltipla, seguem o mesmo princípio que uma regressão linear, sendo necessário comparar os valores observados com os valores preditos da variável resposta. Essa comparação é baseada na função de verosimilhança e é amplamente conhecida como teste da razão de verosimilhança. A Estatística do teste é dada por:

$$D = -2 * \ln \left[ \frac{\text{verosimilhança do modelo ajustado}}{\text{verosimilhança do modelo saturado}} \right] \quad (53)$$

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

ou seja

$$D = -2 \ln(L_S) + 2 \ln(L_C) \quad (54)$$

Onde  $\hat{\pi}_i = \hat{\pi}(x_i)$ ,  $L_S$  é a verosimilhança do modelo sem a co-variável e  $L_C$  é a verosimilhança do modelo com a co-variável.

Para avaliar a significância de uma variável explicativa, deve-se comparar os valores de  $D$  com e sem a variável em que questão e verificar se esta é ou não significativa.

Neste modelo, importa também saber se existe pelo menos uma variável que seja significativa para o modelo. Tomando em consideração a hipótese nula ( $H_0$ ), o  $n$  coeficientes são iguais a zero, de tal forma que a estatística  $D$  tem a distribuição Qui-Quadrado com  $n$  graus de liberdade. Para tal o  $L_C$  com as  $n$  variáveis explicativas e  $L_S$  apenas com o intercepto.

Outro teste usado é o **Wald**, que como objetivo testar a significância dos coeficientes do modelo, se este é diferente de zero. Serve também para avaliar se uma determinada variável apresenta uma relação estatisticamente significativa em relação a variável dependente.

Assim deve-se considerar as seguintes hipóteses:

$$H_0: \beta_1 = 0 \quad e \quad H_0: \beta_1 \neq 0, j = 0, \dots, p \quad (55)$$

A estatística de teste é dada por:

$$W_j = \frac{\hat{\beta}}{\hat{\sigma}(\hat{\beta}_j)} \quad (56)$$

ou

$$W = \hat{\beta}^T (X^T V X) \hat{\beta} \quad (57)$$

A partir desta, olhando para a distribuição Qui-quadrado com  $p + 1$  graus de liberdade sob a hipótese de que cada um dos coeficientes ( $p + 1$ ) seja igual a zero.

#### 4.2.4. Intervalo de confiança

Os intervalos de confiança são obtidos tendo em conta o valor do respetivo teste de Wald, e são dados por IC  $100 (1 - \alpha)\%$  para o parâmetro  $\beta_j$ :

$$IC(\beta_j, 1 - \alpha) = [\hat{\beta}_j \pm z_{1-\alpha} \sigma(\hat{\beta}_j)] \quad (58)$$

O intervalo de confiança para o **Odds Ratio**, é obtido de forma semelhante a do modelo univariado e é dado por:

$$IC_{OddsRatio\ 1-\alpha} = [e^{\beta_l}, e^{\beta_s}] \quad (59)$$

### 4.3. Diagnóstico do Modelo

Para a validação da qualidade de qualquer modelo de regressão é necessário ter em conta a análise dos resíduos. De modo a avaliar as distâncias entre os valores observados e estimados. É importante considerar dois tipos de resíduos possíveis usados: resíduos de **Pearson** e da **Deviance**. Hosmer & Lemeshow, (2000).

#### 4.3.1. Resíduos de Pearson

Os resíduos de Pearson são definidos pela seguinte expressão:

$$r(y_i, \hat{\pi}) = r_j = \frac{y_i - \hat{\pi}}{\sqrt{\hat{\pi}_j(1 - \hat{\pi})}}, j = 1, 2, \dots, n \quad (60)$$

A estatística de teste de Pearson é designada por estatística de Qui- Quadrado de Pearson e é calculada da seguinte forma:

$$\chi^2 = \sum_{j=1}^n r(y_i, \hat{\pi})^2 \quad (61)$$

onde:

$$\hat{\pi}_i = \frac{e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}}{1 + e^{\hat{\beta}_0 + \hat{\beta}_1 x_i}} \quad (62)$$

Para distribuição Qui-Quadrado com  $n - (p + 1)$  graus de liberdade, com  $p$  covariáveis, para testar a  $H_0$ , deve considerar-se o modelo ajustado como sendo adequado, isto é,

$$H_0 : O \text{ modelo ajustado é adequado}$$

#### 4.3.2. Deviance

Os resíduos do Deviance ou desvios podem ser padronizados de igual forma como em Modelos de Regressão Lineares. Este mede a diferença entre o modelo atual ou final em relação ao modelo saturado.

Assim os resíduos Deviance são definidos como:

$$d(y_i, \hat{\pi}_j) = \pm \left\{ 2 \left[ y_j \ln \left( \frac{y_j}{m_j \hat{\pi}_j} \right) + (m_j - y_j) \ln \left( \frac{(m_j - y_j)}{m_j (1 - \hat{\pi}_j)} \right) \right] \right\}^{1/2} \quad (63)$$

Nesta equação, o sinal + ou -, é o mesmo sinal que o de  $(y_j - m_j y_j)$ . Para os padrões de covariáveis com  $y_j = 0$ , o resíduo de desvio é:

$$d(y_i, \hat{\pi}_j) = -\sqrt{2m_j |\ln (1 - \hat{\pi}_j)|} \quad (64)$$

Para  $y_j = m$ , o deviance é:

$$d(y_i, \hat{\pi}_j) = \sqrt{2m_j |\ln (\hat{\pi}_j)|} \quad (65)$$

A estatística resumida baseada nos resíduos de Deviance é a Deviance (Hosmer & Lemeshow, 2000) e é dada por:

$$D = \sum_{j=1}^j d(y_i, \hat{\pi}_j)^2 \quad (66)$$

### 4.3.3. O Teste de Hosmer-Lemeshow

Hosmer & Lemeshow, (2000) desenvolveram uma série de resíduos que podem ser usados para o modelo logístico binário. Eles também desenvolveram um teste de qualidade de ajuste que foi aceite pela maioria dos estatísticos como uma das melhores maneiras de avaliar o ajuste de modelos logísticos HILBE, (2015)

O teste de Hosmer & Lemeshow tem o objetivo de testar a qualidade do ajuste, isto é, comprovar se o modelo obtido pode explicar adequadamente os dados observados. Baseia-se na divisão dos dados em função das probabilidades previstas. Para o primeiro grupo tem-se elementos cuja probabilidade prevista seja inferior a 10%, neste usa-se o  $g= 10$  e contém  $n'_1 = n/10$  elementos com as menores probabilidades estimadas e o último grupo contém  $n'_{10} = n/10$  com as maiores probabilidades estimadas. Como no segundo grupo, o uso dos elementos do grupo  $g = 10$  resulta em pontos de valores definidos nos valores de

$\frac{k}{10}, k = 1, 2, \dots, 9$  contêm todos os sujeitos com probabilidades estimadas entre pontos de corte adjacentes, isto é, valores compreendidos entre 10% e 20%. Este processo é realizado até que se forme o 10º grupo.

A estatística do teste é dada por:

$$\hat{C} = \sum_{k=1}^g \frac{(O_k - n'_k \bar{\pi}_k)^2}{n'_k \bar{\pi}_k (1 - \bar{\pi}_k)} \quad (67)$$

Onde  $n'_k$  é o número total de objetos ou elementos no grupo  $k^o$  grupo,  $C_k$  denota o número de padrões de covariáveis no  $k$ -ésimo decil.

$$O_k = \sum_{j=1}^{c_1} y_j \quad (68)$$

é o número de casos registados no  $k$ -ésimo decil e

$$\bar{\pi}_k = \sum_{j=1}^{c_1} \frac{m_j \hat{\pi}_j}{n'_k} \quad (69)$$

é a média das probabilidades estimadas

Noutras circunstâncias, usa-se como alternativa o denominador da equação (67), que é obtido considerando o  $O_k$  como a soma das variáveis aleatórias independentes não identicamente distribuídas. Assim, é necessário padronizar a diferença quadrada entre os valores esperados e os observados, que é dada por:

$$\sum_{j=1}^{c_1} m_j \hat{\pi}_j (1 - \hat{\pi}_j) = n'_k \bar{\pi}_k (1 - \bar{\pi}_k) + \sum_{j=1}^{c_1} (\hat{\pi}_j - \bar{\pi}_k)^2 \quad (70)$$

#### 4.3.4. Matriz de classificação

Uma forma simples e intuitiva de resumir os resultados do ajuste de um modelo de regressão logística é por meio de uma tabela ou matriz de classificação ou confusão que é um dos testes de ajuste originais usados na regressão logística e originalmente aplicado na Análise Discriminante. Esta tabela é produzida pela classificação cruzada da variável de resposta com respostas previstas que são derivadas do modelo ajustado. O valor usual para o ponto de corte é  $c = 0,5$ ; no entanto,  $c = 0,5$  nem sempre é o valor ideal Myers, Montgomery, & Vining, (2010).

Segundo Hosmer & Lemeshow, (2000), nesta abordagem, as probabilidades estimadas são usadas para prever se um elemento pertence ao grupo. Presumivelmente, se o modelo prediz a participação do elemento no grupo com precisão de acordo com algum critério, então acredita-se que isso forneça evidências de que o modelo se ajusta.

Suponde que  $P(Y = 1) = \theta_1$  e  $X \sim N(0,1)$  pertencente ao grupo com  $Y = 0$  e  $X \sim N(\mu, 1)$  no grupo com  $Y = 1$ . Assim, teremos:

$$\beta_0 = \ln \left[ \frac{\theta_1}{1 - \theta_1} \right] - \frac{\mu^2}{2} \quad (71)$$

### a) Elementos da Matriz de classificação ou matriz confusão

Os elementos a ter em conta numa Matriz de Classificação são:

- **Verdadeiros Positivos (VP):** São os casos em que o modelo previu corretamente os casos "Positivo" sendo que a classe real também é "Positivo".
- **Falsos Positivos (FP):** São os casos em que o modelo previu incorretamente os casos "Positivo", sendo que a classe real é "Negativo". Também são conhecidos como "erros do tipo I" ou "falsos alarmes".
- **Verdadeiros Negativos (VN):** São os casos em que o modelo previu corretamente os casos "Negativo" sendo que a classe real também é "Negativo".
- **Falsos Negativos (FN):** São os casos em que o modelo previu incorretamente os casos "Negativo", sendo que a classe real é "Positivo". Também são conhecidos como "erros do tipo II" ou "omissões".

Para se entender melhor, é importante olhar para a representação visual de uma matriz, segundo a tabela abaixo:

*Tabela 1: Tabela de classificação ou matriz confusão*

	Valor Observado	
	Real Positivo	Real Negativo
Positivo Estimado	VP	FN
Negativo Estimado	FN	VP

Tendo em conta os valores dos elementos constantes da matriz, é possível calcular várias métricas de avaliação do ajuste de modelo, tais como:

- **Precisão:** É a proporção das predições corretas, considerando todos os elementos da amostra, sem, no entanto, ter em conta o que é positivo e o que é negativo e é dada pela seguinte expressão:

$$Precisão = \frac{VP + VN}{P + N} \quad (74)$$

Onde  $P$  indica o número total de eventos e  $N$  o número total de não eventos.

- **Sensibilidade:** é a medida usada para verificar corretamente os resultados positivos previstos do número total de resultados positivos, isto é, a proporção de verdadeiros positivos, avaliando a capacidade do modelo em classificar um indivíduo como positivo, dado que é realmente positivo, dada através de:

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (75)$$

Uma alta *sensibilidade* indica que o modelo está sendo eficaz em prever os elementos positivos, minimizando o número de falsos negativos. Por outro lado, uma baixa sensibilidade indica que o modelo está prevendo muitos casos positivos e tem uma tendência a prever mais falsos negativos.

- **Especificidade:** é a capacidade de um modelo prever corretamente os verdadeiros negativos (VN) das previsões totais de negativos (VN + FP), dado que é realmente negativo.

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (76)$$

Geralmente uma alta *especificidade* indica que o modelo tem uma boa capacidade de evitar falsos alarmes, ou seja, ele é capaz de prever os elementos verdadeiramente negativos sem classificá-las erroneamente como positivas. Por outro lado, uma baixa especificidade sugere nos que o modelo está prevendo muitos falsos positivos, o que pode ser problemático, dependendo do contexto.

Em resumo, a tabela de classificação é mais apropriada quando a classificação é um objetivo declarado da análise; caso contrário, deve apenas complementar métodos mais rigorosos de avaliação de adequação. Hosmer & Lemeshow, (2000)

#### 4.4. Curva ROC

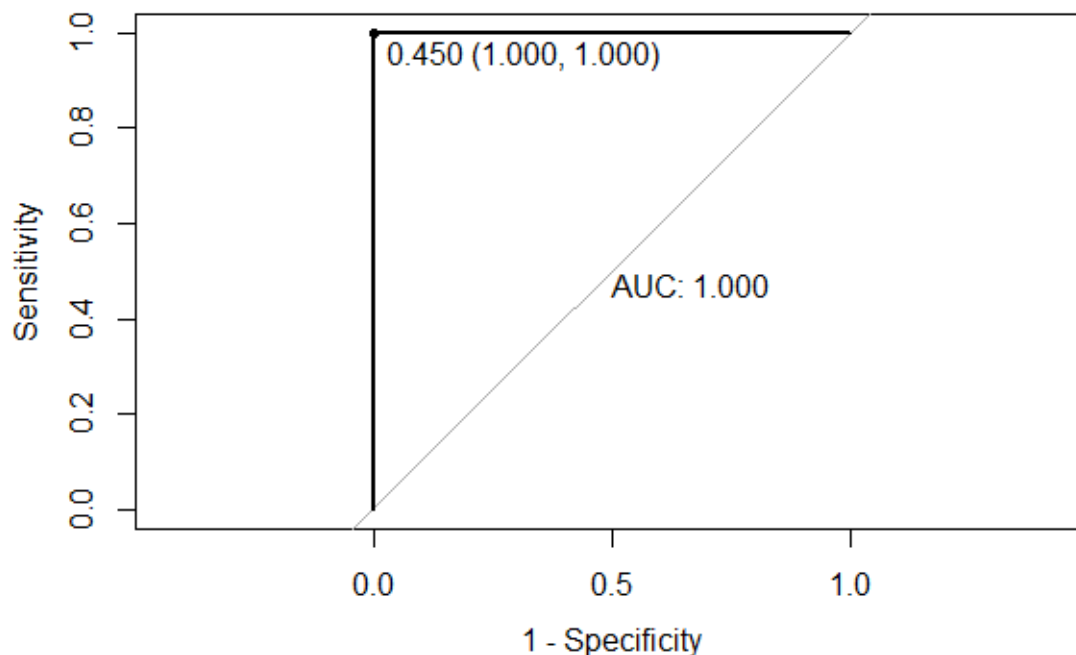
A Curva ROC (Receiver Operating Characteristic) é uma ferramenta gráfica usada para avaliar o desempenho de um modelo logístico binária à medida que o limiar de decisão varia ao classificar

os casos. A curva ROC mostra a relação entre a sensibilidade e a especificidade para diferentes valores de limiar. Ela é obtida traçando a sensibilidade no eixo vertical (eixo Y) e o inverso da especificidade no eixo horizontal (eixo X) para diferentes valores de limiar de decisão. Assim a cada ponto na curva corresponde a um limiar diferente. Quanto mais próxima a curva estiver do canto superior esquerdo do gráfico, melhor será o desempenho do modelo. Myers, Montgomery, & Vining, (2010).

Uma outra forma de analisar a curva ROC é analisando a área sob a Curva ROC (AUC-ROC), que quantifica a qualidade global do modelo, isto é, quanto maior o valor da AUC-ROC, melhor é o desempenho do modelo.

Hosmer & Lemeshow, (2002) sugerem que a área sob a curva ROC é um guia geral de quão bem o modelo discrimina os dados, com os seguintes critérios de avaliação:

- **ROC = 0.5**, indica sem discriminação, isto é, o modelo apresenta um desempenho aleatório,
- **0.7 < ROC < 0.8**, desempenho aceitável,
- **0.8 < ROC < 0.9**, excelente desempenho,
- **ROC = 1**, indica um desempenho perfeito



**Figure 1: Curva ROC, baseado em dados hipotéticos**

**Fonte:** Adaptada pelo autor (Dados hipotéticos)

De forma resumida, a Curva ROC é uma ferramenta visual que permite avaliar o desempenho ou a eficácia geral de um modelo de regressão logística binária e comparar diferentes modelos. A forma e a posição da curva em relação à linha diagonal de referência e o valor da AUC-ROC (área da curva ROC), são importantes indicadores do desempenho do modelo em termos do equilíbrio entre Taxa de Especificidade e a Taxa de Sensibilidade.

## **PARTE II: APRESENTAÇÃO DOS RESULTADOS**

## CAPÍTULO 5: RESULTADOS

### 5. Aplicação do Modelo aos dados

#### 5.1. Aplicação do modelo de Regressão Logística a mortalidade materna

##### 5.1.1. Descrição dos dados

Dos dados recolhidos e utilizados, correspondem a cerca de 84% devido a indisponibilidade de processos de auditorias previamente definidos para a amostra.

**Tabela 2: Resumo dos casos Processados**

Casos não ponderados <sup>a</sup>		N	Percentagem (%)
Casos Seleccionados	Incluídos na Análise	168	97,1
	Casos Perdidos	5	2,9
	Total	173	100,0
Casos não seleccionados		0	0
Total		173	173

*a. Se o peso estiver em vigor, consulte a tabela de classificação para o número total de casos.*

Da tabela acima temos o resumo dos casos usados para a análise de onde pode-se ver que dos 173 óbitos, 97.1% foram seleccionados para a análise e 2.9% são os casos considerados omissos.

#### Caixa 1: Comando para obtenção da tabela resumo das Estatísticas

GET

```
FILE='D:\Abo2\UAB.pt\Dissertação\Base de Dados dissertação_M.sav'.
```

```
DATASET NAME DataSet1 WINDOW=FRONT.
```

```
FREQUENCIES VARIABLES=Idade_Mulher NrConsultas_Ante_Natal IdadeGestacional Gesta  
Partos
```

```
/FORMAT=NOTABLE
```

```
/NTILES=4
```

```
/STATISTICS=STDDEV VARIANCE RANGE MINIMUM MAXIMUM SEMEAN MEAN
```

```
MEDIAN MODE SUM SKEWNESS SESKEW
```

```
KURTOSIS SEKURT
```

```
/ORDER=ANALYSIS.
```

**Tabela 3: Estatística Descritiva**

		Idade da Mulher	Número de Consultas Ante Natais feitas	Idade Gestacional da Mulher	Número de gestações	Número de Partos
N	Valid	171	172	173	171	171
	Missing	2	1	0	2	2
Mean		26,88	2,20	33,17	4,60	3,58
Std. Error of Mean		,596	,116	,566	,255	,223
Median		27,00	2,00	37,00	4,00	3,00
Mode		27	1	39	1	1
Std. Deviation		7,797	1,524	7,442	3,330	2,918
Variance		60,786	2,323	55,377	11,089	8,515
Skewness		,126	,433	-1,264	1,041	,737
Std. Error of Skewness		,186	,185	,185	,186	,186
Kurtosis		-,753	-1,090	1,050	,752	-,113
Std. Error of Kurtosis		,369	,368	,367	,369	,369
Range		38	5	37	15	12
Minimum		11	0	5	0	0
Maximum		49	5	42	15	12
Sum		4597	378	5739	786	613
Percentiles	25	20,00	1,00	28,00	2,00	1,00
	50	27,00	2,00	37,00	4,00	3,00
	75	33,00	4,00	39,00	6,00	5,00

Do total de processos esperados (N=205) para fazer parte da amostra, foi possível obter 173 dos 205 processos correspondente a 84% de disponibilidade. Assim, das variáveis quantitativas analisadas, para o caso da idade pode notar-se que foram inclusos na análise 171 observações e 2 omissas, destas mulheres que tiveram alguma complicação e tenha resultado em óbitos tinham uma idade compreendida entre os 11 a 49 anos, com uma idade média de 26 anos. Infelizmente notou-se na Província o registo de 38 (22%) óbitos maternos em adolescentes (sendo 2 gestantes com 11 anos de idade). De forma isolada foi analisada a idade e assim, em termos de idade, revela um outro problema que merece de todo uma atenção especial que é o casamento e gravidezes precoces pois segundo dados disponíveis do SISMA/Dhis2 (Sistema de Informação, Monitoria e Avaliação) de 2016 a 2020 foram atendidas nas unidades sanitárias da Província cerca de 400 mil adolescentes grávidas.

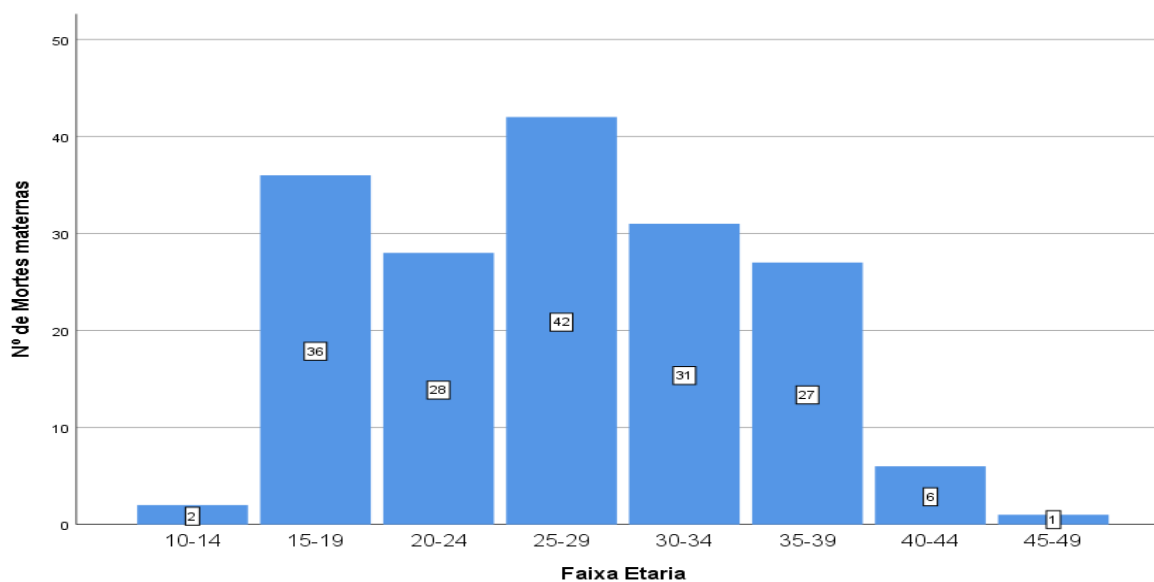
Em relação a idade gestacional nota-se que a menor foi de 5 semanas e a gestante com a idade gestacional mais alta teve 42 semanas. Olhando para o número de gestações, dentre as mulheres

que faleceram existiam aquelas que estavam grávidas pela primeira vez (primigestas) e por outro lado mulheres com 15 gestações.

No que diz respeito ao número de Consultas Pré-Natais ou ante-natal feitas, das 172 mulheres com informação sobre a realização da consulta (1 caso omissos), nota-se em princípio fraca aderência as mesmas, dado que existiram mulheres que faleceram sem ter feito pelo menos uma (cerca de 9%), 37% fizeram apenas uma consulta. Olhando para o recomendado pela OMS, que no mínimo a mulher deve fazer 4 consultas de seguimento durante a sua gestação, nota-se o não cumprimento por parte das mulheres, pois apenas 27% das mulheres teve pelo menos 4 consultas. A variável idade gestacional não apresenta dados omissos, pelo que foram validados todos os 173 casos, sendo a que apresenta uma média maior em relação as outras.

No que concerne ao número de partos por mulheres que faleceram, pode ver-se que, 13.3% das mulheres faleceram sem que tivesse ocorrido o parto, isto é, faleceram grávidas, 18.5% tiveram apenas um parto anterior. No entanto existe um grupo (7%) de mulheres que podem ser consideradas múltíparas, com número de partos variando entre 09 a 12, constituindo um grande risco para a saúde das mulheres. Ainda desse conjunto existiram 1.7% das mulheres sem informação sobre o número de partos, constituindo em princípio um erro de registo.

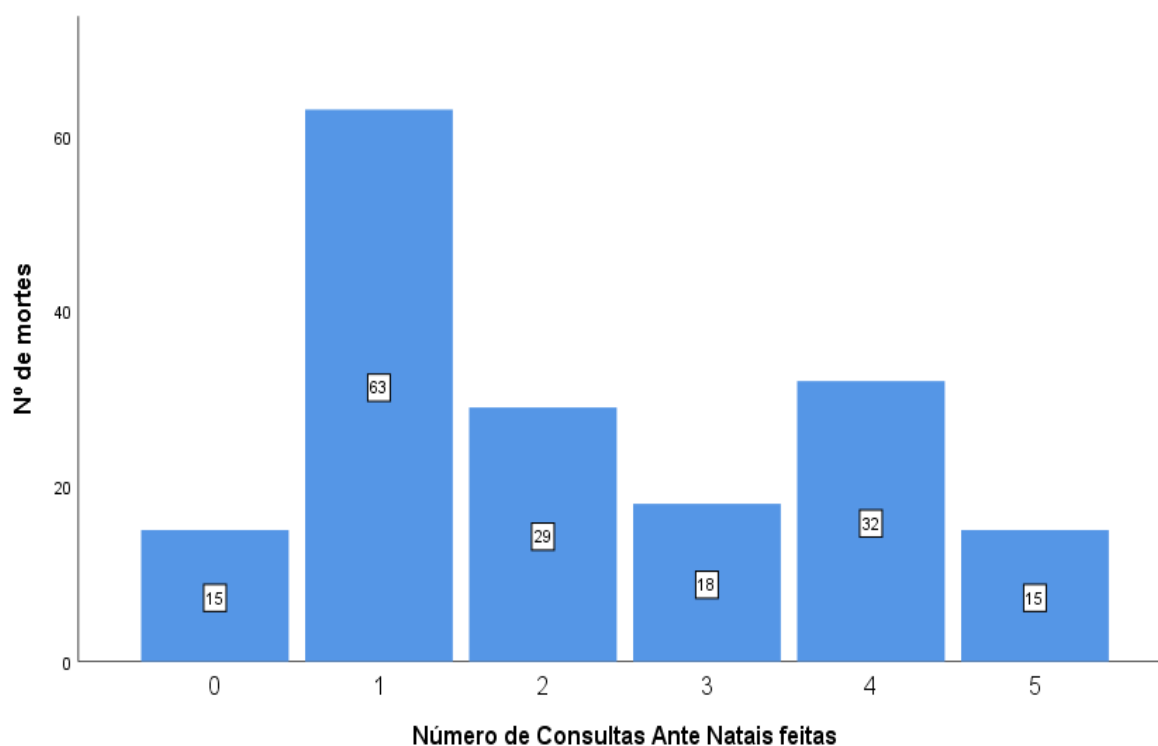
### **Gráfico 1: Mortes maternas por faixa etária**



**Fonte:** SPS (Elaborado pelo autor – baseado nos dados do estudo)

Das 173 mortes analisadas em termos de padrão de morte, caracterizou-se por um maior número de mortes em mulheres nas faixas etárias de 25-29 anos com cerca de 42 mortes e um rácio de mortalidade estimado em 18/100.000 Nascimentos vivos, seguidas das faixas etária dos 15-19 anos com 36 mortes um rácio de 15/100.000 Nascimentos vivos e 30-34 anos com 30 mortes e um rácio de 13/100000 Nascimentos vivos. Assim, é notório a ocorrência de mortes em adolescentes dos 10 aos 19 anos (38 mortes).

**Gráfico 2: Mortes maternas por nº de consultas feitas**



**Fonte:** SPS (Elaborado pelo autor – baseado nos dados do estudo)

De acordo com os dados do gráfico acima, pode notar-se que 9% das mulheres que morreram não tiveram pelo menos uma consulta pré-natal, 36% tiveram apenas uma consulta. Existindo cerca de 9% que fizeram 5 consultas.

## Caixa 2: Comando SPSS para obtenção dos gráficos da Evitabilidade e causa básica

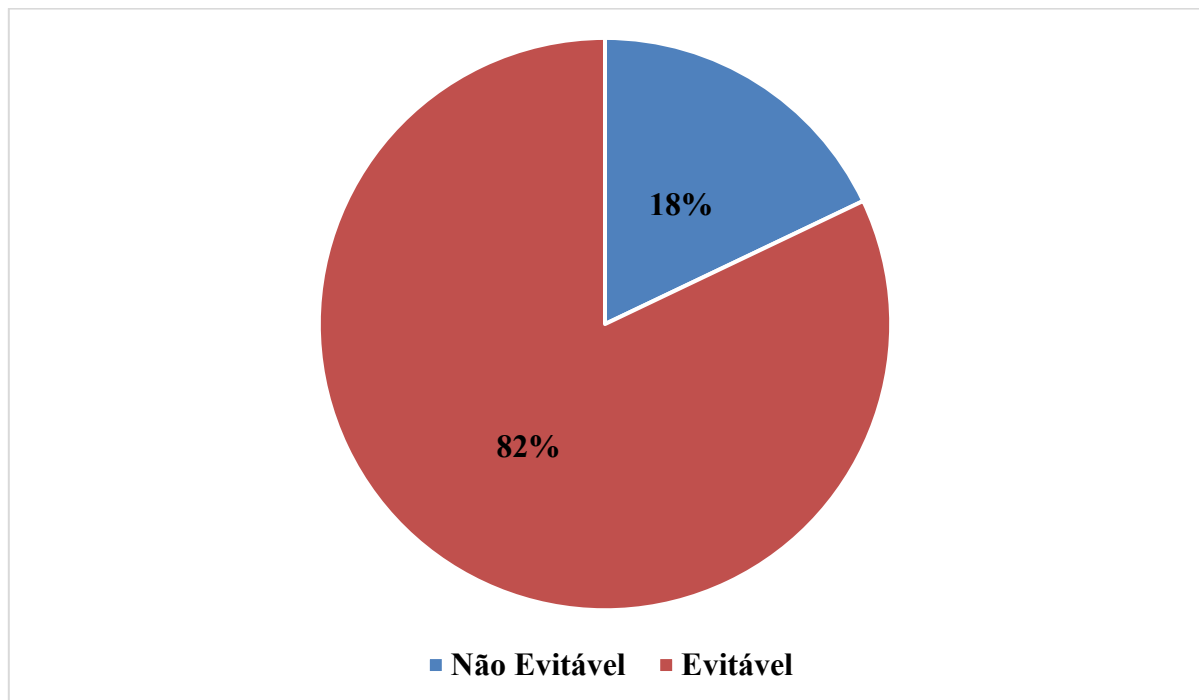
```
FREQUENCIES VARIABLES=Evitabilidade Grupo_de_CB
```

```
/FORMAT=NOTABLE
```

```
/PIECHART PERCENT
```

```
/ORDER=ANALYSIS.
```

**Gráfico 3: Evitabilidade da Morte**

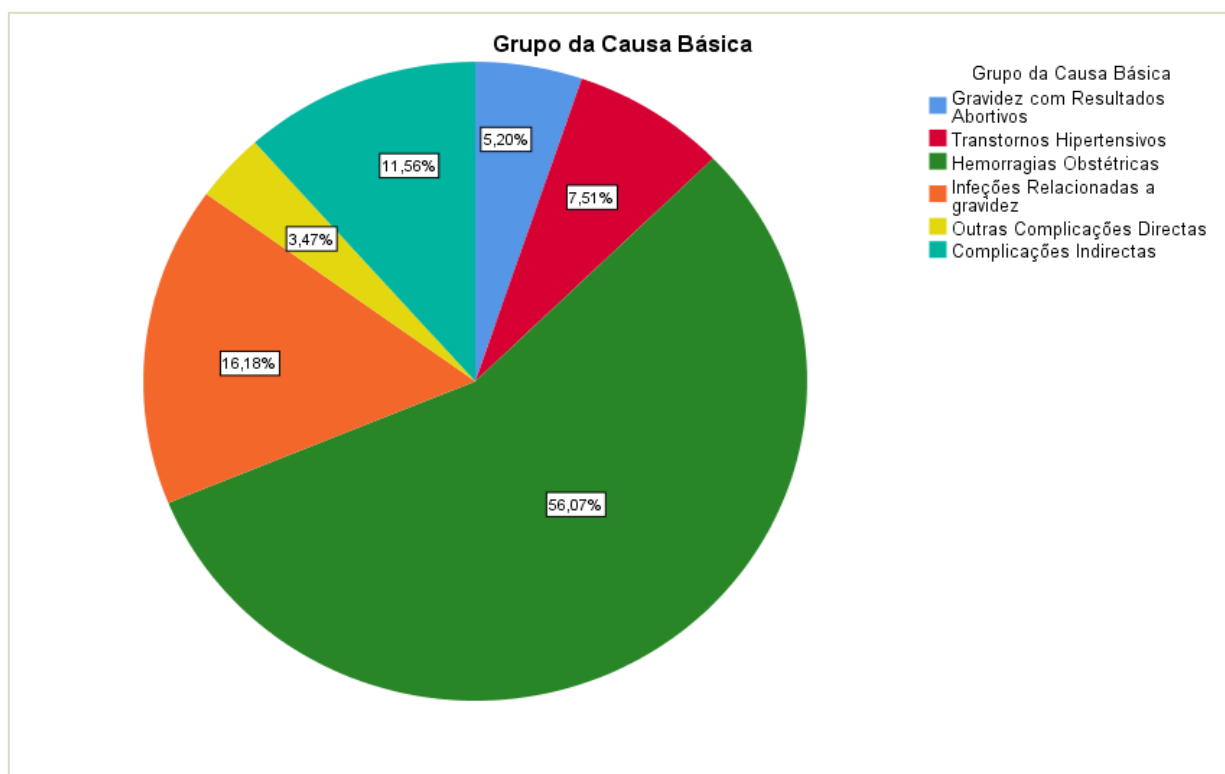


**Fonte:** SPS (Elaborado pelo autor – baseado nos dados do estudo)

A Evitabilidade, consiste em avaliar se as circunstâncias em ocorreu o óbito, se eram possíveis ou não de evitar. Assim ela pode ser evitável ou não evitável (possivelmente evitável). Tendo em conta que o propósito da investigação de óbito materno não é responsabilização ou culpabilização dos provedores ou serviços, mas sim, prevenir a ocorrência de mortes futuras por causas similares, os comitês definem os critérios que utilizarão para estabelecer a Evitabilidade do óbito materno, considerando não só as intervenções clínicas adotadas mas também vários fatores como: Fatores da Comunidade (pessoais/familiares), Institucionais (que tem haver com o provedor, e a instituição como um todo) (Ministério da Saúde, Brasil, 2009)

Dos óbitos registados ao longo do período 82% eram possivelmente evitáveis, existindo certos fatores que possam ter ocorrido por detrás do mesmo e que condicionaram a ocorrência destes. A maioria das complicações que essas mulheres experimentam durante a gravidez e o trabalho de parto levando-as a morte podem ser evitadas com recursos e precauções conhecidas ou disponíveis. (Evans, Jeemon, Stryko, Toole, & Catts, 2018)

**Gráfico 4: Grupo de Causa Básica de Morte**



**Fonte:** SPS (Elaborado pelo autor – baseado nos dados do estudo)

Dos óbitos registados, consta que a maior causa de mortalidade são as hemorragias obstétricas com cerca de 56.1% dos casos, seguida das infeções relacionadas a gravidez, que podem ter ocorrido durante o trabalho de partos com cerca de 16.2%. Segundo (Evans, Jeemon, Stryko, Toole, & Catts, 2018), nos Países em desenvolvimento, 99% da mortalidade em mulheres grávidas ocorre devido a Infeção, doenças hipertensivas, sangramento, sendo consideradas como as principais causas com cerca de 75% das complicações maternas. O risco dessas complicações é maior em gestantes com HIV e malária (outras complicações).

**Tabela 4: Grupo da Causa Básica \* Evitabilidade da Morte Crosstabulation**

		Evitabilidade da Morte		Total
		Não Evitável	Evitável	
Grupo da Causa Básica	Gravidez com Resultados Abortivos	2	7	9
	Transtornos Hipertensivos	2	11	13
	Hemorragias Obstétricas	13	84	97
	Infeções Relacionadas a gravidez	4	24	28
	Outras Complicações Directas	2	4	6
	Complicações Indiretas	8	12	20
<b>Total</b>		<b>31</b>	<b>142</b>	<b>173</b>

**Fonte:** SPS (Elaborado pelo autor – baseado nos dados do estudo)

Do cruzamento da Evitabilidade e o grupo de causas, tem-se que em todas as causas o maior número de morte foi evitável ou possivelmente evitável, sendo a maior causa de morte foi a hemorragias obstétricas com maior número de óbitos evitável, seguida de Infeções relacionadas a gravidez, parto/parto

**Tabela 5: Possíveis Fatores Associados \* Evitabilidade da Morte Crosstabulation**

		Evitabilidade da Morte		Total
		Não Evitável	Evitável	
Possíveis Fatores Associados	Pessoais ou Familiares	6	37	43
	Lógicos	0	13	13
	Institucionais	1	31	32
	Pessoais e Logísticos	6	19	25
	Pessoais e Institucionais	6	18	24
	Sem Informação	6	15	21
<b>Total</b>		<b>25</b>	<b>133</b>	<b>158</b>

**Fonte:** SPS (Elaborado pelo autor – baseado nos dados do estudo)

Da associação da Evitabilidade da morte e possíveis fatores associados tem-se que, 28% das mortes maternas evitáveis ou possivelmente evitáveis foram relacionados aos fatores pessoais ou familiares que inclui o atraso da mulher na procura de cuidados de saúde e em outras situações até recusa das mulheres em ser referida para outra unidade hospitalar ou Unidade Sanitária, seguida de fatores institucionais com 21% de mortes evitáveis que incluiu a falta de equipamento médico, falta de medicamento e noutros casos erros de procedimento clínicos durante a intervenção do profissional de saúde.

Nota se na tabela que 13 casos foram considerados omissos por não terem sido registados algum fator relacionado.

## 5.2. Análise dos dados usando a regressão logística

### *Caixa 3: Comando SPSS para construção do modelo de regressão logístico*

```
ANALYZE → REGRESSION → BINARY LOGISTIC
```

```
LOGISTIC REGRESSION VARIABLES Evitabilidade
```

```
/METHOD=ENTER Idade_Mulher Fez_CPN NrConsultas_Ante_Natal IdadeGestacional Gesta Partos
```

```
/CONTRAST (Fez_CPN)=Indicator
```

```
/SAVE=PRED PGROUP RESID LRESID DEV
```

```
/CLASSPLOT
```

```
/PRINT=GOODFIT CORR ITER(1) CI(95)
```

```
/CRITERIA=PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

***Tabela 6: Codificação de Variável Dependente***

Valor Original	Valor Interno
Não Evitável	0
Evitável	1

Para o estudo foi considerada como variável depende a Evitabilidade do óbito que pode ser: Evitável e Não Evitável, havendo algumas literaturas que definem essa classificação como Evitável/Possivelmente Evitável.

Assim, a Codificação da Variável Dependente, nos informa quais valores foram atribuídos a variável dependente. Para os em análise foram codificados os dados com “Não Evitável” = 0 e “Evitável” = 1. Assim foram modelados os valores “1” (ou seja, a probabilidade de se Evitar a morte materna tomando em consideração vários fatores associados (ver tabela 5).

**Tabela 7: Codificações de variáveis categóricas**

		Frequência	Codificação de Parâmetros
			(1)
Fez consulta ante-natal/pré-natal	Não	74	1,000
	Sim	94	,000

**Fonte:** SPS (Elaborado pelo autor – baseado nos dados do estudo)

Na tabela acima, temos a contagem da realização da consulta cuja a codificação da variável categórica independente e as suas classes de referência foram codificadas com (0,000) as de realização de alguma consulta pré-natal. Assim para análise da variável “**Fez consulta ante-natal/pré-natal**” os seus resultados referem para a classe “*Sim*”

## **BLOCK 0: BEGINNING BLOCK/BLOCO INICIAL**

Neste bloco estão apresentadas as saídas do bloco inicial que irão explicar o quão bom o modelo funciona sem a avaliação ou inclusão do Chi-quadrado ou por outra segundo Pallant, (2007) são os resultados da análise sem nenhuma das nossas variáveis independentes usadas no modelo. Acrescenta Meyers, Gamst, & Guarino (2013) que são apresentados apenas os resultados para o modelo somente de intercepção, calculado apenas com a constante na equação, mas sem nenhuma das variáveis preditoras.

**Tabela 8: Histórico da Iteração** <sup>a,b,c</sup>

Iteração		-2 Log likelihood	Coeficientes
			Constante
Passo 0	1	159,155	1,286
	2	157,665	1,510
	3	157,658	1,526
	4	157,658	1,526

a. Constant is included in the model.

b. Initial -2 Log Likelihood: 157,658

c. Estimation terminated at iteration number 4 because parameter estimates changed by less than ,001.

A tabela "Histórico de Iterações" é geralmente usada em conjunto com técnicas de análise estatística, como regressão logística ou análise de sobrevivência, que envolvem a maximização de uma função de verossimilhança ou minimização de uma função de erro para encontrar os parâmetros do modelo que melhor se ajustam aos dados. Ela exibe informações sobre o processo iterativo

avaliando o ajuste dos parâmetros do modelo até que uma solução que atenda aos critérios de convergência seja encontrada. Nela encontramos as iterações. Do ponto de vista do nosso modelo, nota-se que a concordância ou convergência perfeita no ajuste do modelo foi alcançada a partir da 3 linha, com 3 casas decimais.

**Tabela 9: Tabela de Classificação <sup>a,b</sup>**

Observados			Preditos		
			Evitabilidade da Morte		Porcentagem Correta
			Não Evitável	Evitável	
Passo 0	Evitabilidade da Morte	Não Evitável	0	30	,0
		Evitável	0	138	100,0
	Porcentagem Geral				82,1

a. Constant is included in the model.

b. The cut value is ,500

A tabela de classificação, fornece o número de casos observados previstos pelo modelo só com a constante. Por outro lado fornece a contagem do número de casos em cada variável ou resultado binário. Com apenas o intercepto no modelo, a previsão baseia-se exclusivamente nas frequências, de tal modo que foram observados nela 138 óbitos evitáveis e 30 não evitáveis considerados corretamente classificados. Assim, considerando a não existência de alguma informação adicional, na melhor das hipóteses um óbito evitável, na sua classificação (previsão) estaria correto em 82,1% dos casos. Usando o modelo nulo em todas as observações a morte teve a probabilidade de ser evitável estimada em 0,821 obtida através de:

$$\hat{\pi}_j = \frac{e^{1,526}}{1 + e^{1,526}} = \frac{4,5997}{5,5997} = 0,821$$

Assim, de forma geral, estaríamos afirmando que o óbito materno é em todo o caso possivelmente evitável, independentemente dos fatores.

**Tabela 10: Variáveis na Equação**

		B	S.E.	Wald	df	Sig.	Exp(B)
Passo 0	Constante	1,526	,201	57,389	1	,001	4,600

Fonte: SPS (Elaborado pelo autor – baseado nos dados do estudo)

As variáveis no modelo ou na equação apenas com a constante são descritas na tabela acima. Nesta não estão apresentadas as variáveis predictoras, sendo que o único fator é o intercepto, bem como as

estimativas dos seus coeficientes, erro padrão e o teste de Wald. A razão das chances ilustrada por Exp (B) tem um valor de 4,600 revelando que a Evitabilidade é muito maior do que a Não Evitabilidade, isto é, a ocorrência do óbito evitável é 4,6 vezes maior do que a inevitável.

Assim, a probabilidade de se evitar a morte materna é tida a partir de:

$$\hat{\pi}_0 = \frac{\sum_{n=1}^n Y_i}{n} = \frac{138}{168} = 0.821429 \cong 0.821$$

A estimativa da constante do modelo é:

$$\hat{\beta}_0 = Ln\left(\frac{0.821429}{0.178571}\right) = 1,526056 \cong 1.526$$

E o erro padrão é obtidos através de:

$$\hat{\sigma}_{\beta_0} = \frac{1}{\sqrt{n\hat{\pi}_0(1 - \hat{\pi}_0)}} = \frac{1}{\sqrt{168 * 0.821 * 0.179}} = 0.20144 \cong 0.201$$

**Tabela 11: Variáveis não presentes na equação <sup>a</sup>**

		Score	df	Sig.	
Passo 0	Variáveis	Idade da Mulher	,288	1	,591
		Fez consulta Ante-Natal/Pré-Natal(1)	,102	1	,750
		Número de Consultas Ante Natais feitas	4,908	1	,027
		Idade Gestacional da Mulher	,849	1	,357
		Número de gestações	1,571	1	,210
		Número de Partos	,262	1	,608
	Estatística Geral	15,661	6	,016	

**Fonte:** SPS (Elaborado pelo autor – baseado nos dados do estudo)

As variáveis não inclusas no modelo são apresentadas na tabela acima, incluindo as estimativas da estatística da pontuação para cada uma das variáveis e para todo o modelo como um só, os graus de liberdade e o *p-value*. Tomando em consideração que para  $\alpha = 0.05$ , temos apenas as variáveis Número de Consultas Ante Natais feitas são consideradas estatisticamente significativas no modelo.

## BLOCK 1: MÉTODO = ENTER

Neste bloco, Segundo Marôco (2018), é respeitante ao modelo como um todo, isto é, o modelo completo, ao qual é incluído na análise o Chi-quadrado, como ilustrado na tabela abaixo. É também neste bloco onde o modelo (conjunto de variáveis preditoras) é testado Pallant, (2007).

**Tabela 12: Testes Omnibus de Coeficientes do Modelo**

		Chi-square	df	Sig.
Passo 1	Passo	16,497	6	,011
	Bloco	16,497	6	,011
	Modelo	16,497	6	,011

**Fonte:** SPS (Elaborado pelo autor – baseado nos dados do estudo)

O *Testes Omnibus de Coeficientes de Modelo* contém o modelo qui-quadrado, um teste estatístico da hipótese nula de que todos os coeficientes preditores são zero MEYERS, GAMST, & GUARINO (2013).

Pode ver-se da tabela que esta apresenta o valor do teste do rácio de verosimilhanças entre o modelo nulo e os modelos em cada um dos passos, bloco e modelo final. A utilização do modelo “*Enter*” indica que todos os valores são iguais dado que existe apenas um único passo, bloco e modelo. Para o modelo, o valor dado de  $X^2(6) = 16,497$ ; com  $p < 0.05$  que indica existir pelo menos uma variável independente no modelo com o poder preditivo sobre a variável dependente (Evitabilidade).

**Tabela 13: Resumo do modelo**

Passo	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	141,161 <sup>a</sup>	,094	,154

*a. Estimation terminated at iteration number 5 because parameter estimates changed by less than ,001.*

A tabela resumo do modelo, fornece informações sobre a utilidade do modelo ou o quão bom o modelo se ajusta aos dados, representada pelos 3 índices. Com todas variáveis no modelo, temos a estatística da verosimilhança -2Loglikelihood que é usado para comparar o ajuste do modelo como todas as variáveis, sendo que para o nosso caso tem um valor de 141,161. Segundo MEYERS, GAMST, & GUARINO (2013), normalmente não se interpreta essa estatística diretamente, mas é usada apenas para comparar diferentes modelos logísticos.

Em suma, o modelo apresenta um ajuste significativo tendo em conta o valor do - 2Log likelihood (141,161), embora o valor do Pseudo  $R^2$  de Cox&Senll (0,094) sugira que outras variáveis não incluídas poderiam ser necessárias para melhor explicar a variável dependente. Já o  $R^2$  de Nagelkerke indica uma explicação de cerca de 15%.

**Tabela 14: Teste de Hosmer and Lemeshow**

Passo	Chi-square	df	Sig.
1	8,913	8	,350

**Fonte:** SPS (Elaborado pelo autor – baseado nos dados do estudo)

O teste de *Hosmer & Lemeshow* avalia se as probabilidades de os valores previstos corresponderem valores observados, com o valor de  $X^2(8) = 8.913$  e  $p = 0.350$ , que não é significativo, por ser maior que 0.05, conclui-se que o modelo se ajusta aos dados, havendo uma correspondência aceitável entre as probabilidades previstas e observadas.

**Tabela 15: Tabela de Contingência do Teste de Hosmer & Lemeshow**

		Evitabilidade da Morte = Não Evitável		Evitabilidade da Morte = Evitável		Total
		Observado	Esperado	Observado	Esperado	
Passo 1	1	6	7,253	11	9,747	17
	2	6	5,624	11	11,376	17
	3	7	4,404	10	12,596	17
	4	1	3,023	16	13,977	17
	5	2	2,615	16	15,385	18
	6	4	2,300	13	14,700	17
	7	2	2,044	15	14,956	17
	8	0	1,417	17	15,583	17
	9	2	,948	15	16,052	17
	10	0	,370	14	13,630	14

**Fonte:** SPS (Elaborado pelo autor – baseado nos dados do estudo)

A Tabela de Contingência para o Teste de Hosmer & Lemeshow demonstra de forma mais detalhada o teste de Hosmer & Lemeshow. Nela os dados estão divididos em 10 grupos com base na variável de resultado. Esses grupos são definidos por taxas crescentes de Evitabilidade (chamadas de “etapas” na tabela). Para o primeiro grupo (Etapa 1 para a não evitável) representa os óbitos com menor probabilidade de ser não evitável com cerca de 6 casos. Para a mesma etapa temos um outro grupo (Evitável) em que temos cerca de 11 óbitos com menor probabilidade de ser

evitável. Os valores observados e os esperados (com base no modelo de previsão) correspondem razoavelmente bem para todas as etapas, pelo que se considera um resultado desejável.

**Tabela 16: Tabela de Classificação <sup>a</sup>**

Observados			Preditos		
			Evitabilidade da Morte		Percentagem Correta
			Não Evitável	Evitável	
Passo 1	Evitabilidade da Morte	Não Evitável	0	30	,0
		Evitável	2	136	98,6
	Percentagem Geral				81,0

*a. The cut value is ,500*

A tabela de classificação no modelo final, isto é, com todos os preditores inclusos, tem uma precisão da Evitabilidade da morte em 81,0%. Apesar de termos a especificidade do modelo praticamente nula e a sensibilidade em 98,6%, podemos considerar que a inclusão dos preditores contribuiu para uma previsão aceitável com algumas fragilidades na classificação das mortes não evitáveis. Comparadas as tabelas do modelo inicial com apenas o intercepto e a do modelo final que incluí as variáveis predictoras, podemos concluir que os valores estão muito próximos, pois houve uma perda ligeira da percentagem de acerto dos casos dos 82.1% para 81.0%, não havendo grande variação.

Este resultado poderá estar relacionado com o fato de a amostra estar desequilibrada entre os óbitos evitáveis e não evitáveis.

**Tabela 17: Variáveis na Equação**

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Passo 1 <sup>a</sup>	Idade da Mulher	0,001	0,034	0	1	0,986	1,001	0,936	1,07
	Fez consulta Ante-Natal/Pré-Natal (1)	2,004	0,727	7,591	1	0,006	7,42	1,783	30,873
	Número de Consultas Ante Natais feitas	0,952	0,299	10,158	1	0,001	2,592	1,443	4,655
	Idade Gestacional da Mulher	0,01	0,026	0,143	1	0,705	1,01	0,959	1,063
	Número de gestações	-0,222	0,141	2,489	1	0,115	0,801	0,608	1,055
	Númerode Partos	0,207	0,167	1,535	1	0,215	1,23	0,886	1,707
	Constant	-1,292	1,378	0,879	1	0,348	0,275		

*a. Variable(s) entered on step 1: Idade da Mulher, Fez consulta Ante-Natal/Pré-Natal, Número de Consultas Ante Natais feitas, Idade Gestacional da Mulher, Número de gestações, Número de Partos.*

A tabela denominada “variáveis na equação” apresenta para cada preditor o coeficiente de regressão parcial de pontuação bruta e seu erro padrão e indicam a quantidade de variação esperada no log odds quando há uma variação de uma unidade na variável preditora, com todas as outras variáveis no modelo mantidas constantes. Quando um coeficiente tem valor próximo de “0” indica que não a mudança na variável de resultado associada à variável preditora não é considerável. Quando todas as variáveis são incluídas no modelo, temos apenas duas significantes que são: Fez consulta ante-Natal/Pré-Natal (1) e Número de Consultas Ante Natais feitas. As outras variáveis não são mais significantes.

Assim, tomando em consideração o valor das variáveis na tabela acima pode-se ter a seguinte equação:

$$E(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

$$E(x) = -1.292 + 0.001 \text{Idade} + 2.004 \text{Fez}_{CPN} + 0.952 \text{Nr de Consultas} \\ + 0.10 \text{Idade Gestacional} - 0.222 \text{Nr de Gestações} + 0.207 \text{Nr de Partos}$$

Se, na regressão logística, a probabilidade de ocorrência de um evento pode ser estimada diretamente no caso da variável dependente assumir apenas dois possíveis valores (Evitável “1” e Não Evitável “0”) e haver um conjunto de variáveis independentes  $X_1, X_2, \dots, X_k$ .

Assim, o modelo de regressão logística será escrito da seguinte forma:

$$P(Y = 1) = \frac{1}{1 + e^{-g(x)}}$$

onde,

$$g(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

Logo,  $g(x) = E(x)$

**Tabela 18: Matriz de Correlação**

		Constant	Idade da Mulher	Fez consulta Ante-Natal/Pré-Natal (1)	Número de Consultas Ante Natais feitas	Idade Gestacional da Mulher	Número de gestações	Número de Partos
Passo 1	Constant	1,000	-,484	-,517	-,558	-,558	-,022	,098
	Idade da Mulher	-,484	1,000	-,036	-,018	-,033	-,233	-,075
	Fez consulta ante-Natal/Pré-Natal (1)	-,517	-,036	1,000	,810	-,051	-,005	,037
	Número de Consultas Ante Natais feitas	-,558	-,018	,810	1,000	-,031	,016	-,001
	Idade Gestacional da Mulher	-,558	-,033	-,051	-,031	1,000	,083	-,114
	Número de gestações	-,022	-,233	-,005	,016	,083	1,000	-,860
	Número de Partos	,098	-,075	,037	-,001	-,114	-,860	1,000

**Fonte:** SPS (Elaborado pelo autor – baseado nos dados do estudo)

A tabela acima que apresenta os coeficientes de correlação entre pares de variáveis do conjunto de dados, variando de -1 a 1. Ela é útil para entender as relações entre as variáveis e pode ajudar a identificar possíveis padrões, associações ou dependências entre elas. Assim:

- Um valor igual a -1 ou 1 indica uma correlação perfeita (podendo ser positiva ou negativa em função do seu valor).
- Um valor próximo de 1 indica uma correlação positiva forte (quando uma variável aumenta, a outra também tende a aumentar), para um valor próximo de -1 indica uma correlação negativa forte (quando uma variável aumenta, a outra tende a diminuir).
- Um valor próximo de 0 indica uma correlação fraca ou inexistente, isto é, as variáveis têm pouca relação linear entre si.

**Tabela 19: Medidas Preditivas**

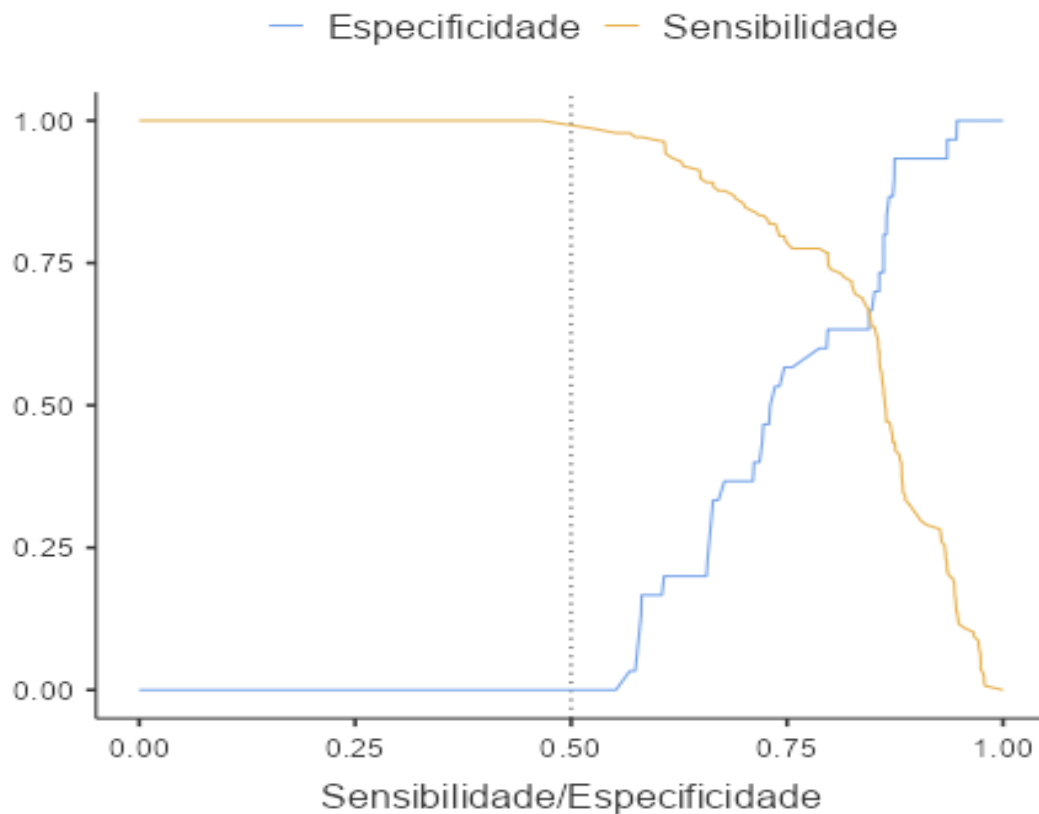
Acurácia	Especificidade	Sensibilidade	AUC
0.810	0.00	0.986	0.729

*Nota.* O valor de corte é 0.5

As medidas preditivas podem ser usadas para avaliar o desempenho de modelos de classificação, para além de que fornecem uma compreensão mais detalhada sobre as previsões do modelo em

diferentes aspectos. Pode notar-se na tabela (ver também a tabela 17) a acurácia que a proporção de previsões corretas em relação ao total de previsões do modelo dada em 81%.

**Gráfico 5: Ponto de Corte**



**Fonte: Jamovi (Elaborado pelo autor – baseado nos dados do estudo)**

Do gráfico acima pode concluir-se que o ponto de interseção entre a sensibilidade e a especificidade é muito maior que o ponto de corte, concordando com o valor da acurácia do modelo em prever de forma real o número de casos verdadeiramente positivos classificados pelo modelo. Por outra indica que se a probabilidade de um caso for evitável for de 0.81, esta será verdadeiramente classificado como evitável na amostra e se a probabilidade for menor que 0.81 será verdadeiramente classificado como não evitável.



**Caixa 4: Comando para Obtenção da Curva ROC**

Analyse → Classify → ROC Curve

ROC Idade\_Mulher Fez\_CPN NrConsultas\_Ante\_Natal IdadeGestacional Gesta Partos BY Evitabilidade (1)

/PLOT=CURVE(REFERENCE)

/PRINT= COORDINATES

/CRITERIA=CUTOFF(INCLUDE) TESTPOS(LARGE) DISTRIBUTION(FREE) CI(95)

/MISSING=EXCLUDE.

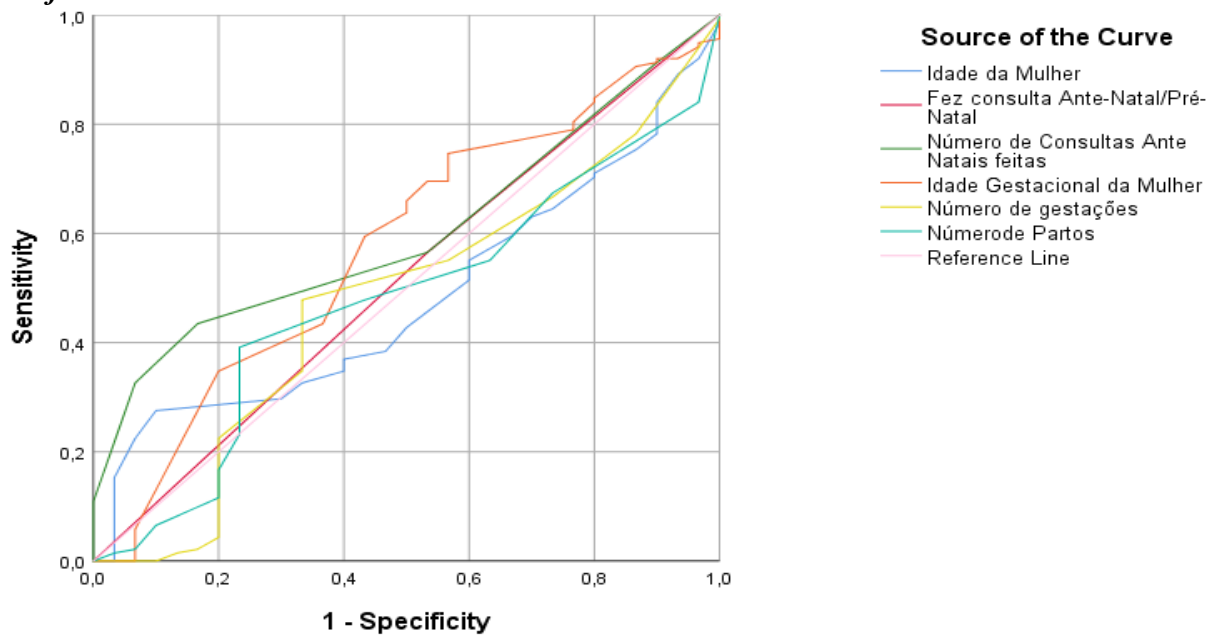
**Tabela 20: Resumo do processamento de caso**

Evitabilidade da Morte	Valid N (listwise)
Positivo <sup>a</sup>	138
Negativo	30
Em falta	5

Larger values of the test result variable(s) indicate stronger evidence for a positive actual state.

a. The positive actual state is Evitável.

**Gráfico 7: Curva ROC com todas as variáveis**



Diagonal segments are produced by ties.

**Fonte: SPS (Elaborado pelo autor – baseado nos dados do estudo)**

**Tabela 21: Area sobre a Curva**

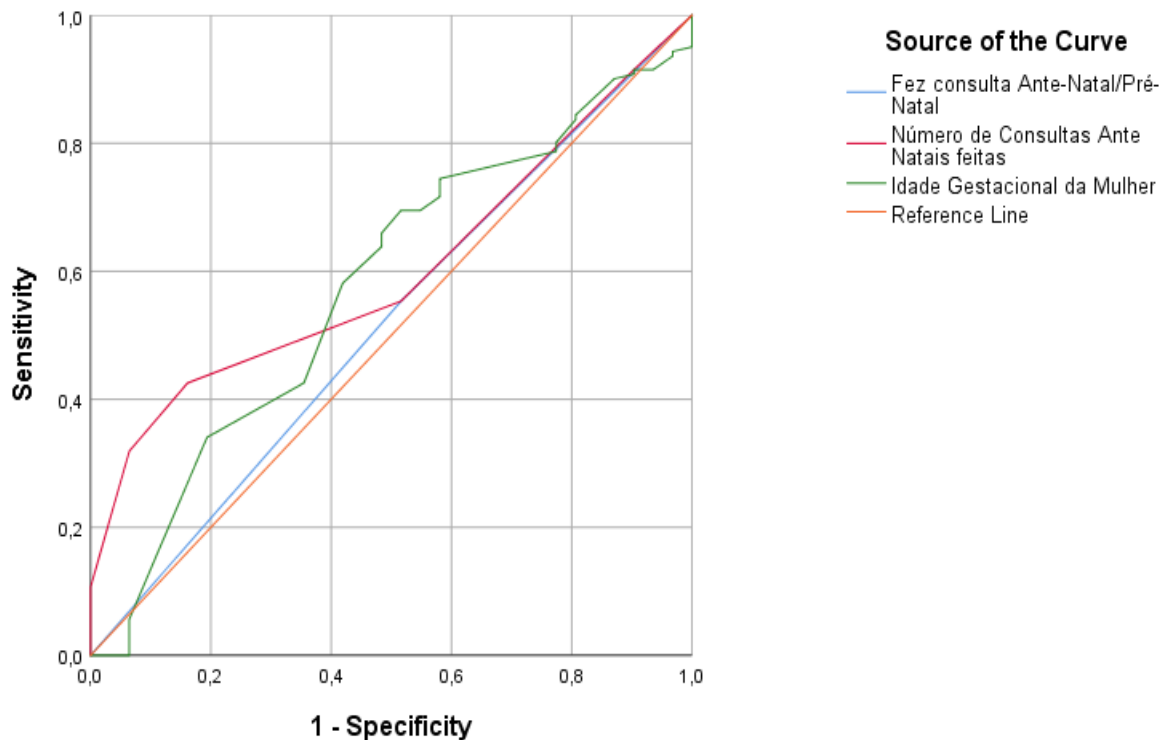
Resultado do teste da Variável(s)	Area
Idade da Mulher	,480
Fez consulta Ante-Natal/Pré-Natal	,516
Número de Consultas Ante Natais feitas	,603
Idade Gestacional da Mulher	,577
Número de gestações	,477
Númerode Partos	,476

The test result variable(s): Idade da Mulher, Fez consulta Ante-Natal/Pré-Natal, Número de Consultas Ante Natais feitas, Idade Gestacional da Mulher, Número de gestações, Númerode Partos has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

Acima temos o gráfico da curva ROC com todas as variáveis, onde podemos ver que algumas delas situam-se abaixo da linha de referência, cujos valores podem ser vistos a partir da tabela 22.

Assim será necessária criar um outro gráfico com as variáveis que apresentam valores acima da linha da referência.

**Gráfico 8: Curva ROC com variáveis cujo valor esta acima do limiar**



Diagonal segments are produced by ties.

**Fonte:** SPSS (Elaborado pelo autor – baseado nos dados do estudo)

**Tabela 22: Área sob a curva apenas com variáveis acima da linha de referência**

Test Result Variable(s)	Area	Std. Error <sup>a</sup>	Asymptotic Sig. <sup>b</sup>	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
Fez consulta ante-Natal/Pré-Natal	,519	,057	,747	,406	,631
Número de Consultas Ante Natais feitas	,600	,048	,081	,507	,694
Idade Gestacional da Mulher	,576	,056	,184	,466	,687

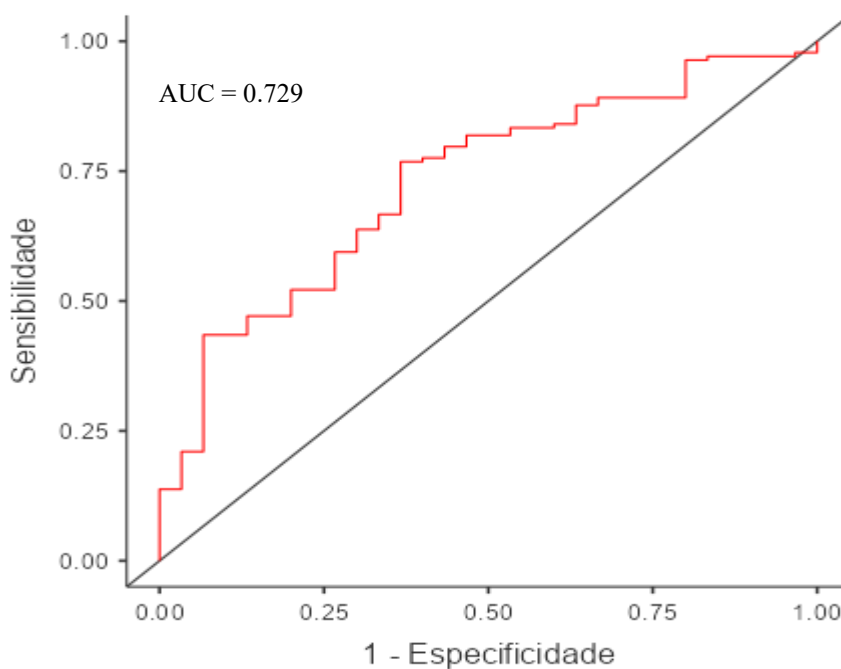
The test result variable(s): Fez consulta Ante-Natal/Pré-Natal, Número de Consultas Ante Natais feitas, Idade Gestacional da Mulher has at least one tie between the positive actual state group and the negative actual state group. Statistics may be biased.

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

Com auxílio do Jamovi, foi possível criar o gráfico da curva ROC que generaliza a avaliação da capacidade discriminatória do modelo.

**Gráfico 9: Curva ROC Geral do Modelo**



**Fonte:** Jamovi (Elaborado pelo autor – baseado nos dados do estudo)

De forma geral pode dizer-se que a área da Curva ROC é 0.729 (ver tabela 20), o que indica uma capacidade discriminatória ou desempenho moderadamente bom do modelo. O modelo é capaz de separar as classes de forma eficaz na maioria dos casos. Isto é, o modelo tem um desempenho melhor que um classificador aleatório.

## 6. DISCUSSÃO DOS RESULTADOS

Numa análise documental de relatórios anuais do sector da Saúde na Província de Nampula, revelam que nos últimos 4 anos (2020 a 2023) foram registadas cerca de 748 mortes maternas intra-hospitalares, como um rácio de mortalidade maternas estimado em cerca de 70 mortes em cada 100.000 Nascimentos Vivos, sendo uma média anual de cerca de 187 mortes anuais. No ano de 2020. Este rácio numa análise primária, se mostra satisfatório pois os dados do **IDS (1997, 2003, 2011 e 2023)**, apresentam uma taxa de mortalidade materna no País estimada em 408 por 100.000 nascimentos vivos. Por outro lado segundo, a **OMS** o rácio deve situar-se entre 10-20 mortes maternas por cada 100.000NV, pois a redução da taxa de mortalidade materna tem uma estreita relação com a redução do índice da pobreza absoluta, e este por sua vez, com redução dos problemas socioeconômicos que afetam os países em vias de desenvolvimento dos quais Moçambique faz parte.

De forma pormenorizada, foram analisadas as mortes referentes ao ano de 2020, sendo o ano que mais mortes foram registadas (205), de onde se concluiu que, foi registado o maior rácio de mortalidade materna dos últimos 5 anos com um valor estimado em 87/100000 NV, valor este considerado alto pela OMS, visto que o objectivos de Desenvolvimento Sustentável prevê que esta seja de 70/100000 NV até 2030.

Por forma a responder aos objectivos do presente trabalho foram analisados 173 disponíveis dos 205 processos esperados.

No concernente a mortalidade por faixas etárias, revelou-se um risco de mortalidade maior em mulheres dos 25 a 29 anos com 42 mortes e um rácio de 19/100000 NV, seguindo-se o rácio observado em adolescentes de idades compreendidas entre os 11 a 19 anos com 38 mortes e um rácio de morte estimado em 16/100000 NV, tendo se observado morte em adolescentes de 11 anos.

O estudo mostrou que 82% das mortes foram evitáveis ou poderiam ser evitadas, evidenciando um cenário não muito bom. Pesquisas mostraram que aproximadamente 80% das mortes maternas poderiam ser evitadas se as mulheres tivessem acesso a serviços essenciais para a maternidade e de cuidados básicos de saúde (UNICEF 2009). A situação da Evitabilidade da morte materna em Nampula apresenta o mesmo cenário citado pela UNICEF, pois o seu quadro mostrou que 55% destas mortes podiam ser evitadas. Segundo A maioria das complicações que essas mulheres

experimentam durante a gravidez e o trabalho de parto levando-as a morte podem ser evitadas com recursos e precauções conhecidas ou disponíveis. (Evans, Jeemon, Stryko, Toole, & Catts, 2018)

O número médio de consultas pré-natais por mulher foi de 2, existindo mulheres que morreram sem ter feito sequer uma consulta pré-natal (9%) e outras com 5 consultas pré-natais (9%), sendo que a maioria (36%) fez apenas uma consulta, contrastando com a recomendação da OMS, que as mulheres devem no mínimo fazer 4 consultas de seguimento antes do parto de modo a obter mais cuidados de saúde durante a gestação.

## 7. CONCLUSÃO

Este trabalho teve como objetivo analisar a mortalidade materna ao longo do ano de 2020 na Província de Nampula e identificar fatores associados bem como avaliar o peso da sua contribuição para esse desfecho. Este rácio para além de apresentar uma desigualdade social em termos assistenciais, sinalizam a deficiente organização do próprio sistema de saúde e a existência de um paradoxo em que as coberturas das consultas pré-natais revelam-se sempre altas, isto é, acima de 100% contrastando com a qualidade das mesmas, fazendo com que as gestantes cheguem até a data do parto com fraco despiste ou diagnóstico das complicações, mesmo as que podem ter feito 4 ou mais consultas.

A prevenção de mortes maternas em Nampula ainda não foi alcançada, isto porque, em função do número de mortes que parece ser reduzido se comparada as mortes por outras causas passa quase por despercebido ao nível da sociedade e esse fenómeno ainda não é tido pela maioria das pessoas como um problema de saúde que deve ser priorizado e prevenido.

Das mortes analisadas obteve-se que a maior parte delas era evitável ou possivelmente evitáveis com cerca de 82% dos casos e apenas cerca de 18% dos casos era não evitável. Em termos de padrão de morte, caracteriza-se por um maior número de mortes em mulheres nas faixas etárias de 25-29 anos com cerca de 42 mortes e um rácio de mortalidade estimado em 18/100.000 Nascimentos vivos, seguidas das faixas etária dos 15-19 anos com 36 mortes um rácio de 15/100.000 Nascimentos vivos e 30-34 anos com 30 mortes e um rácio de 13/100000 Nascimentos vivos. Assim, é notório a ocorrência de mortes em adolescentes dos 10 aos 19 anos (38 mortes).

Da análise do modelo, pode assim dizer-se que este é razoavelmente bom, com uma acurácia de 81% e com uma boa capacidade de classificar os casos de forma eficaz na sua maioria. Por outra nota-se que a Evitabilidade das mortes esta estritamente relacionada as variáveis “Fez consulta pré-natal!” e “Número de Consultas feitas” pois estas apresentam-se significativas no modelo final confirmando a norma da Organização Mundial de Saúde que afirma que quanto maior for o número de consultas ante natal (pelo menos 4) menor é a probabilidade da gestante desenvolver alguma complicação e maior é a probabilidade de se evitar a morte maternas pois esta mulher terá durante a sua gestação os cuidados necessário.

A aplicação do modelo logístico permitiu identificar as hipóteses que se levantam em relação aos principais fatores associados à mortalidade materna, como:

- **Acesso limitado aos serviços de saúde:** Onde se nota que maior parte das mortes ocorreu em zonas rurais onde as mulheres enfrentam barreiras geográficas e financeiras para ter um atendimento pré-natal bem como um parto condigno.
- **Baixa Escolaridade:** A fraca educação que esta diretamente relacionada com menor poder de informação e conhecimentos sobre os cuidados a ter durante a gravidez e menor probabilidade de procura de cuidados de saúde,
- **Idade materna:** Foi notório que existem adolescentes e mulheres acima dos 35 anos que apresentam maiores risco de complicações e morte durante a gravidez e parto,
- **Condições socioeconómicas:** A fraca disponibilidade de recursos ou a pobreza estão diretamente relacionadas a dificuldade de acesso a cuidados de saúde, principalmente para acesso a Unidades em longas distâncias devido a falta de transporte.
- **Número de consultas Pré-natais:** Mulheres com menos de 4 consltas pré-natais tiveram maior probabilidade de morte materna, oque evidencia a necessidade de um reforço no acompanhamento durante a gestação.

Este modelo apresentou limitação na estimação de alguns coeficientes, especificamente na correlação entre as variáveis qualitativas e quantitativas, obtendo-se alguns intervalos do Odds com pouca precisão.

De forma geral, para que haja uma redução da mortalidade materna na Província de Nampula, é importante que hajam esforços coordenados entre as instituições governamentais, parceiros/organizações não governamentais, sociedade covil bem como as estruturas comunitárias, como foco na equidade, sensibilização as comunidades para adesão aos cuidados de saúde, educação e acesso universal a serviços de saúde de qualidade.

Numa perspetiva futura seria interessante estudar a relação entre a ocorrência ou não do óbito materna (desfecho), olhando para o número de consultas pré-natais e o local onde esta mulher teve o parto, por forma a perceber se a determinada unidade hospitalar em função dos recursos materiais e humanos existentes justifica a ocorrência ou não do óbito e sua Evitabilidade.

## 8. BIBLIOGRAFIA

- Agresti, A. (2019). *An Introduction to Categorical Data Analysis. Third Edition*. Florida, United States: John Wiley & Sons, Inc.
- Borges, C. L. (2012). *Epidemiologia e Gestão: Análise da Mortalidade Materna em Aracaju. Dissertação de Mestrado*. Salvador da Bahia.
- Cabral, C. I. (2013). *Aplicação do Modelo de Regressão Logística num Estudo de Mercado. Projeto Mestrado em Matemática Aplicada à Economia e à Gestão*. Lisboa: Universidade de Lisboa. Faculdade de Ciências.
- Centro de Investigação em Saúde Pública. (2013). *É Urgente Reduzir A Mortalidade Materna em Moçambique. Policy Brief n°1*. Maputo: CEPSA.
- Chatterjee, S., & Hadi, A. S. (2012). *Regression Analysis by Example. 5 th Edition*. New Jersey: Wiley.
- Denis, D. J. (2019). *SPSS data analysis for univariate, bivariate, and multivariate statistics*. New Jersey: John Wiley & Sons, Inc.
- Dhingra, D. (2021, July 18). *Analytics Vidhya*. Retrieved from Analytics Vidhya: <https://www.analyticsvidhya.com/blog/2021/05/in-depth-understanding-of-confusion-matrix/>
- Evans, M., Jeemon, J., Stryko, C., Toole, L., & Catts, P. (2018, May). A Systematic Review: Infant and Maternal Mortality and Morbidity in Developing Countries. *A Systematic Review: Infant and Maternal Mortality and Morbidity in Developing Countries*, p. 54.
- Ferraz, L., & Bordignon, M. (2012). *Mortalidade Materna no Brasil: Uma Realidade Que Precisa Melhorar. Revista Baiana de Saúde Pública, artigo original de tema livre, v.36, n.2, p.527-538 abr./jun*. Baia.
- Field, A. (2009). *Descobrendo a estatística usando o SPSS. tradução Lorí Viali. – 2ª Edição*. Porto Alegre: Artmed.
- Figueiredo, Y. M. (2010). *Análise da mortalidade materna no município de Governador Valadares, 2002-2004*. Em – Rev. Min. Enferm. Brasília.

- Garson, G. D. (2014). *Logistic Regression: Binary And Multinomial*. USA: Statistical Associates Publishing.
- Gray, C. D., & Kinnear, P. R. (2012). *IBM SPSS statistics 19 made simple*. Hove and New York: Psychology Pres.
- Harrell, F. (2001). *Regression modeling strategies : with applications to linear models, logistic regression and survival analysis*. New York: Springer series in statistics.
- Hilbe, J. M. (2015). *Practical Guide to Logistic Regression*. USA: CRC Press.
- Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression, Second Edition*. USA: JOHN WILEY & SONS, INC.
- Hunt, P., & Mesquita, J. B. (2007). *Reducing Maternal Mortality: The Contribution of the Right to the Highest Attainable Standard of Health*. Uk: University of Essex, Colchester, Essex, CO4 3SQ, UK. Retrieved Dezembro 10, 2020, from <http://repository.essex.ac.uk/9719/1/reducing-maternal-mortality-contribution-right-highest-attainable-standard-health.pdf>
- IDHS; PNUD; PUC Minas. (2004). *Colecção de estudos temáticos sobre os Objectivos do Desenvolvimento*. Belo Horizonte.
- Instituto Nacional de Estatística. (2009). *Mortalidade em Moçambique Inquérito Nacional sobre Causas de Mortalidade, 2007/8*. Maputo: INE.
- Instituto Nacional de Estatística -INE. (2009). *Mortalidade Em Moçambique, Inquérito Nacional sobre Causas de Mortalidade, 2007/8 Relatório Preliminar*. Maputo: INE.
- Knight, H. E., Self, A., & Kennedy, S. H. (2013, Maio 21). *Why Are Women Dying When They Reach Hospital on Time? A Systematic Review of the 'Third Delay'*. (R. C. Young, Ed.) PLoS ONE 8(5), p. 9. doi:10.1371/journal.pone.0063846
- Landau, S., & Everitt, B. S. (2004). *A handbook of statistical analyses using SPSS*. Chapman & Hall/CRC Press LLC.
- Leech, N. L., Barrett, K. C., & Morgan, G. A. (2005). *SPSS for Intermediate Statistics: Use and Interpretation*. Second Edition. Mahwah, New Jersey 07430: Lawrence Erlbaum Associates, Inc.

- Marôco, J. (2018). *Análise Estatística com SPSS Statistics*. 7ª Edição. Pêro Pinheiro: Gráfica Manuel Barbosa & Filhos.
- Marques, H. R., Manfroi, J., DE Castilho, M. A., & Noal, M. L. (2006). *Metodologia da Pesquisa e do Trabalho Científico*. 2ª Edição. Brasília: UCDB.
- Medicusmundi. (2011). *Termos de Referência para a Realização duma Análise Sobre os fatores de risco de mortalidade materna*. Nampula.
- Meyers, L. S., Gamst, G. C., & Guarino, A. J. (2013). *Performing data analysis using IBM SPSS*. New Jersey: Wiley & Sons, Inc.
- Ministerio da Saude (MISAU), Instituto Nacional de Estatística (INE) e ICF International (ICFI). (2013). *Moçambique Inquérito Demográfico e de Saúde 2011*. Calverton, Maryland, USA: MISAU, INE e ICFI.
- Ministério da Saúde, Brasil. (2009). *Manual dos comitês de mortalidade materna*. Brasília: MS.
- Myers, R. H., Montgomery, D. C., & Vining, G. G. (2010). *Generalized linear models : with applications in engineering and the sciences*. New Jersey: John Wiley & Sons, Inc.
- National Institute of Public Cooperation and Child Development. (2015). *An Analysis of Levels and Trends in Maternal Health and Maternal Mortality Ratio in India - A Report*. New Delhi.
- Nhatave, I. (2006). *Saúde Materna em Moçambique - Revisão de Literatura*. Maputo.
- OMS; OPAS. (2009). *Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde*. Decima Revisão. (Vol. 2). (C. C. Português, Trans.) São Paulo: Editora da USP.
- Pallant, J. (2007). *SPSS Survival Manual. A Step by Step Guide to Data Analysis using SPSS for Windows*. Third. Edition. New York: Open University Press.
- Santos, R. J. (2017). *Regressão logística em dados com eventos raros*. Dissertação de Mestrado apresentada à Faculdade de Ciências da Universidade do Porto em Engenharia Matemática. Porto: FCUP.

S raphin, M. N., Ngnie-Teta, I., Ayoya, M. A., R. Khan, M., W. Striley, C., Boldon, E., & Mamadoultai bou, A. (2014, November 25). *Determinants of Institutional Delivery Among Women of Childbearing Age in Rural Haiti*. *Maternal and Child Health Journal*, 18, p. 11. doi:DOI 10.1007/s10995-014-1646-1

Silva, J. P. (2016). *Modelos de Regress o Linear e Log stica utilizando o software R*. Disserta o orientada por Prof<sup> </sup> Doutora Teresa Paula Costa Azinheira Oliveira & Prof. Doutor Am lcar Manuel do Ros rio Oliveira. Lisboa: Uab.pt. Retrieved 06 20, 2024, from <http://hdl.handle.net/10400.2/6129>

Souza, M. S. (n.d.). *Mortalidade Materna: Perfil Epidemiol gico Em Sergipe (2001- 2010)*.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics. 6th Edition*. USA: Pearson Education, Inc.

Tai, B. C., & Machin, D. (2014). *Regression methods for medical research*. Oxford, OX4 2DQ, UK: John Wiley & Sons, Ltd.

Uni o Africana. (2013). *Relat rio de Situa o sobre a Sa de Materna, Neonatal e Infantil*.

Vergara, S. C. (2004). *Projectos e Relat rios de Pesquisa em Administra o. 5<sup>a</sup> Edi o*. S o Paulo: Atlas S.A.

Viana, R. D., Novaes, M. R., & Calderon, I. M. (2011). *Mortalidade Materna - uma abordagem atualizada/Maternal Mortality - an update approach*.

Vittinghoff, E., Shiboski, S. C., Glidden, D. V., & McCulloch, C. E. (2005). *Regression methods in biostatistics : linear, logistic, survival, and repeated measures models*. USA: Springer.

WHO. (2019, September 19). Maternal mortality. WHO. *Maternal mortality*. Retrieved Janeiro 03, 2021, from <https://www.who.int/en/news-room/fact-sheets/detail/maternal-mortality>

WHO; UNICEF; UNFPA; World Bank Group; UNPD. (2015). *Trends in maternal mortality: 1900 to 2015*. Estimates by WHO, UNICEF, UNFPA, World Bank Group . Switzerland.

World Health Organization. (2013). *Making pregnancy safer in Mozambique*. *World Health Organization*. Retrieved 03 03, 2021, from <http://www.afro.who.int/index.php?>