

**Jucarlos Rufino de Freitas**

**Aplicação da análise logística/binomial com enfoque no Índice de Qualidade do Sono de Pittsburgh, em algumas comunidades do Sertão do Pajeú/PE.**

**Recife/2018**



**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO**  
**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA**

**Aplicação da análise logística/binomial com enfoque no Índice de Qualidade do Sono de Pittsburgh, em algumas comunidades do Sertão do Pajeú/PE.**

Dissertação julgada adequada para obtenção do título de Mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 22/02/2018 pela comissão examinadora

**Área de concentração: Biometria e Estatística Aplicada**

**Orientador: Prof. Dr. Moacyr Cunha Filho**

Recife/2018

Dados Internacionais de Catalogação na Publicação (CIP)  
Sistema Integrado de Bibliotecas da UFRPE  
Biblioteca Central, Recife-PE, Brasil

F866a Freitas, Jucarlos Rufino de.  
Aplicação da análise logística/binomial com enfoque no Índice de Qualidade do Sono de Pittsburgh, em algumas comunidades do Sertão do Pajeú/PE / Jucarlos Rufino de Freitas. – Recife, 2018.  
79 f.: il.

Orientador(a): Moacyr Cunha Filho.  
Dissertação (Mestrado) Universidade Federal Rural de Pernambuco, Pós-Graduação em Biometria e Estatística Aplicada, Recife, 2018.  
Inclui referências, apêndices e anexos.

1. Qualidade do sono. 2. PSQI 3. Análise multivariada  
4. Predição I. Cunha Filho, Moacyr, orient. II. Título.

CDD 310

**UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO  
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA**

**Aplicação da análise logística/binomial com enfoque no Índice de Qualidade do Sono de Pittsburgh, em algumas comunidades do Sertão do Pajeú/PE.**

Jucarlos Rufino de Freitas

Dissertação julgada adequada para obtenção do título de Mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 22/02/2018 pela comissão examinadora

Orientador:

---

**Prof. Dr. Moacyr Cunha Filho**  
Orientador

Banca examinadora:

---

**Prof. Dr. Frank Sinatra Gomes da  
Silva**  
Universidade Federal Rural de  
Pernambuco - UFRPE

---

**Prof. Dr. Lucian Bogdan Bejan**  
Universidade Federal Rural de  
Pernambuco - UFRPE - Examinador  
Externo

*Dedico este trabalho aos meus pais  
João Rufino e Veronice Santana.*

# Agradecimentos

Agradeço a Deus, por ser presente na minha vida, guiando meus caminhos e me proporcionando SEMPRE oportunidades grandiosas, mostrando que somos mais fortes tendo fé. E assim temos certeza que felizes são aqueles que realmente acreditam, eles nunca estarão sozinhos.

Agradeço à minha família, meu pai João Rufino, minha mãe Veronice Santana e meus irmãos e meu mais novo sobrinho João Lucas, por estar sempre presente ao longo da jornada. Aos meus primos por me apoiar nessa jornada e me fazer rir, nos momentos difíceis.

Ao meu orientador Professor Doutor Moacyr Cunha Filho, pela amizade, os ensinamentos que me passou de forma muito generosa, o companheirismo, a paciência, a dedicação e a disponibilidades durante o desenvolvimento da pesquisa.

Aos meus amigos do departamento: Glauce Guerra, Yarah Lopes, Patrícia Medeiros, Sarah Lúcia, David Avellaneda, Iloane Lima, Erivaldo Neto, Lêda Valéria, Wesley Lima, André Luiz e, todos os demais que compõem nosso departamento por dividirmos arduamente nossos fracassos e sucessos oriundos das pesquisas desenvolvidas.

Aos meus amigos da Associação dos Engenheiros Agrícola de Pernambuco (AEAPE) e clube da SUDENE, que entenderam o meu distanciamento, em muitos momentos, o meu muito obrigado por compreenderem as minhas necessidades; em outros momentos a companhia de vocês, Marcos Neto, Luiz Phelipe, João Lucas, Igor Leonardo, Lucas Oliveira, Eduardo Santiago, Rejanilton Santiago, Lidiane Araújo, Aline Andrade, Daniel Conrado e Joicy Carla, foram/é de suma importância no meu desenvolvimento pessoal.

Sem deixar de lado os amigos que vêm me acompanhando no decorrer de todo o caminho precedente para essa chegada, Luciano Matheus, Amanda Novaes, France Oliveira, Ruben Vivaldi, Roniery Lucena, Rômulo, Professora Cláudia e Professora Cristina pelo apoio nos momentos mais difíceis com quem compartilharam as alegrias e inquietações.

Agradeço a todo o corpo docente do DEINFO pelo compartilhamento dos seus conhecimentos e em particular aos professores Antônio Samuel, Tiago Alessandro, Guilherme Rocha, Frank Sinatra e Cristiane Rocha.

Ao secretário do programa Marco Antônio dos Santos pela sua dedicação e competência administrativa em seu trabalho.

À Capes, pelo concessão de apoio financeiro.

À UFRPE, que me proporcionou cursar o mestrado em Biometria e Estatística Aplicada.

*“Se o sono não tem função vital absoluta,  
então é o maior erro que o processo  
evolucionário já cometeu.”  
(A. Rechtschaffen, 1971)*



## Resumo

O sono pode ser definido como “um período de repouso para o corpo e a mente, durante o qual a volição e a consciência estão em inatividade parcial ou completa; também descrito como um estado comportamental, com postura imóvel e sensibilidade diminuída, mas prontamente reversível a estímulos externos”. Nesse contexto, o sono é um fator essencial para a qualidade de vida, proporcionando uma melhoria no bem estar amplo humano. Além disso, é importante para a recuperação da saúde em situação de doenças. Diante disso, um estudo do tipo corte transversal foi realizado no mês de Abril de 2017 com o objetivo de avaliar e modelar a qualidade do sono (QS) nas comunidades de Poço do Serrote, Catolé, Três Irmãos e Poldrinho pertencentes ao Sertão do Pajeú/PE, incorporando o questionário de avaliação do índice de Qualidade do Sono de Pittsburgh (IQSP) e as relações sociodemográficas. Além disso, foi de interesse também caracterizar o perfil geral dos moradores e identificar fatores associados à qualidade do sono. A metodologia utilizada consistirá na caracterização da pesquisa, na caracterização dos participantes, nos critérios de inclusão e exclusão e a modelagem logística binomial. Dos 120 moradores contatados, 20 recusaram a participação e 27 tiveram dados incompletos, resumindo-se a 73 questionários analisados. Os resultados mostraram que 98,7% dos moradores referiam qualidade subjetiva do sono como boa ou muito boa; entretanto, os seguintes problemas relacionados ao sono destacaram-se por sua elevada frequência: levantar-se para ir ao banheiro (56,3%); acordar no meio da noite ou muito cedo pela manhã (54,8%) em menos de 1x/semana. A análise multivariada indicou que, além da variável “Cochilo”, as variáveis “Gênero”, “Estado civil” e “Idade” também foram encontradas como fatores explicativos para qualidade de sono, com uma predição de aproximadamente 80%.

**Palavras-chaves:** Qualidade do sono, PSQI, análise multivariada, predição.

# Abstract

Sleep can be defined as “a period of rest for the body and mind, during which volition and consciousness are in abeyance and bodily functions are partially suspended; also described as a behavioral state, with characteristic immobile posture and diminished but readily reversible sensitivity to external stimuli”. In this context, sleep is an essential factor for quality of life, providing an improvement in human well-being. In addition, it is important for the recovery of health in disease situation. Therefore, a cross-sectional study was carried out in April of 2017 with the objective of evaluate and model the sleep quality (QS) in the communities of Poço do Serrote, Catolé, Três Irmãos, and Poldrinho belonging to the Sertão do Pajeú / PE, incorporating the Pittsburgh Sleep Quality Index (IQSP) questionnaire and sociodemographic relationships. In addition, it was also of interest to characterize the general profile of the residents and identify factors associated with sleep quality. The methodology used will consist of the characterization of the research, the characterization of the participants, the inclusion and exclusion criteria and the binomial logistic modeling. Of the 120 residents contacted, 20 denied participation and 27 had data incomplete, summarizing the 73 questionnaires analyzed. The results showed that 98.7% of the residents referred subjective quality sleep as good or very good; however, the following problems related to sleep were noted for their high frequency: getting up to go to the bathroom (56.3%); wake in the middle of the night or very early in the morning (54.8 %) in less than 1x/week. The multivariate analysis indicated that, in addition to the Cochilo variable, the variables ”Gender”, ”Marital status” and ”Age” were also found as explanatory factors for sleep quality, with a prediction of approximately 80%.

**Key-words:** Quality of sleep, PSQI, multivariate analysis, prediction.

# Lista de Figuras

Figura 1 – Municípios do Sertão Pajeú/2014 - PE. Fonte: Blog TV Web Sertão. . .	4
Figura 2 – População territorial do Sertão do Pajeú - PE/2010. Fonte: IBGE (2010). . .	5
Figura 3 – População territorial de Serra Talhada - PE/2010. Fonte: IBGE (2010). . .	5
Figura 4 – Memória episódica humana. “Texto traduzido”(VORSTER; BORN, 2015). . . . .	6
Figura 5 – Fases do sono. Fonte: BedTime (2017). . . . .	7
Figura 6 – Funções de regressão logística . . . . .	13
Figura 7 – A Curva de Característica de Operação de Receptor (Curva ROC) . . . .	24
Figura 8 – Localização Geográfica do município de Serra Talhada–PE . . . . .	27
Figura 9 – Comunidade Poço do Serrote. Sertão do Pajeú, PE, Brasil, 2017. . . . .	27
Figura 10 – Comunidade Poldrinho. Sertão do Pajeú, PE, Brasil, 2017. . . . .	27
Figura 11 – Comunidade Catolé. Sertão do Pajeú, PE, Brasil, 2017. . . . .	28
Figura 12 – Comunidade Três Irmãos. Sertão do Pajeú, PE, Brasil, 2017. . . . .	28
Figura 13 – Histograma idade. Sertão do Pajeú, PE, Brasil, 2017. . . . .	32
Figura 14 – Variáveis sociodemográficas. Sertão do Pajeú, PE, Brasil, 2017. . . . .	36
Figura 15 – Idade versus probabilidade predita da QS. Sertão do Pajeú, PE, Brasil, 2017. . . . .	41
Figura 16 – Valores de h versus valores ajustados do modelo QS. . . . .	42
Figura 17 – Distância de Cook referente ao modelo QS. . . . .	42
Figura 18 – Resíduo componente do desvio versus o índice para o modelo QS. . . .	42
Figura 19 – Quantis da Normal para distribuição Binomial com envelope para QS. . .	42
Figura 20 – Curva ROC do modelo QS. Sertão do Pajeú, PE, Brasil, 2017. . . . .	43

## Lista de tabelas

Tabela 1 – Notação para tabela de contingência . . . . .	9
Tabela 2 – Comparação entre dois grupos . . . . .	11
Tabela 3 – Matriz de Confusão . . . . .	25
Tabela 4 – Caracterização da amostra, de acordo com as respostas dos dados sociodemográficos. Sertão do Pajeú, PE, Brasil, 2017. . . . .	32
Tabela 5 – Frequência e porcentagem da ocorrência dos fatores que contribuíram para o componente transtornos do sono. Sertão do Pajeú, PE, Brasil, 2017. . . . .	33
Tabela 6 – Frequências e porcentagens da variável resposta do IQSP agrupada de forma dicotômica. Sertão do Pajeú, PE, Brasil, 2017. . . . .	33
Tabela 7 – Frequência e porcentagem das respostas dos moradores na questão de número 10 do IQSP. Sertão do Pajeú, PE, Brasil, 2017. . . . .	34
Tabela 8 – Frequências e porcentagens em relação aos sete componentes do IQSP. Sertão do Pajeú, PE, Brasil, 2017. . . . .	35
Tabela 9 – Frequências e porcentagens da variável resposta do IQSP. Sertão do Pajeú, PE, Brasil, 2017. . . . .	35
Tabela 10 – Frequências e porcentagens da variável resposta do IQSP agrupada de forma dicotômica. Sertão do Pajeú, PE, Brasil, 2017. . . . .	35
Tabela 11 – Distribuição de frequência entre a variável QS (Y) e as variáveis sociodemográficas (X's). Sertão do Pajeú, PE, Brasil, 2017. . . . .	37
Tabela 12 – Componentes do IQSP de acordo com os dois grupos de qualidade do sono. Sertão do Pajeú, PE, Brasil, 2017. . . . .	38
Tabela 13 – Estimativas dos parâmetros e teste de significância das variáveis selecionadas do modelo de regressão logístico multivariado final para explicar a QS. Sertão do Pajeú, PE, Brasil, 2017. . . . .	39
Tabela 14 – Intervalos de 95% de confiança para os parâmetros do modelo final, com base na estatística de Wald. Sertão do Pajeú, PE, Brasil, 2017. . . . .	40
Tabela 15 – Intervalos de confiança para as razões de chance (Odds ratio – OR) do modelo final, fixando o nível de confiança de 95%. Sertão do Pajeú, PE, Brasil, 2017. . . . .	40

## Lista de abreviaturas e siglas

IDH	Índice de Desenvolvimento Humano
BSh	Clima das estepes quentes de baixa latitude e altitude
SOL	Sono de Onda Lenta
MLP	Memória de Longo Prazo
REM	Movimento Rápido dos Olhos
NREM	Movimento Não Rápido dos Olhos
DP	Desvio Padrão
RV	Razão de Verossimilhança
IQSP	Índice de Qualidade do Sono de Pittsburgh
VR	Variável Resposta
VP	Variável Preditora
VA	Variável Aleatória
SQR	Soma dos Quadrados Residuais
VAD	Variável Aleatória Discreta
OR	Odds Ratio
FDP	Função de Distribuição de Probabilidade
MLG	Modelo Linear Generalizado
$H_0$	Hipótese Nula
$H_1$	Hipótese Alternativa
gl	grau de liberdade
Curva ROC	Curva Característica de Operação do Receptor
AAS	Amostra Aleatória Simples

# Sumário

<b>1</b>		<b>ii</b>
<b>2</b>	<b>Introdução</b>	<b>1</b>
<b>3</b>	<b>Objetivos</b>	<b>3</b>
3.1	Objetivo Geral	3
3.2	Objetivos Específicos	3
<b>4</b>	<b>Revisão de Literatura</b>	<b>4</b>
4.1	Sertão do Pajeú	4
4.2	Sono	5
4.2.1	Fases do sono	7
4.3	Análise Bivariada	8
4.3.1	Teste Qui-quadrado	9
4.3.2	Teste Exato de Fisher	10
4.4	Regressão logística simples	12
4.4.1	Estimação dos parâmetros do modelo	12
4.4.2	Interpretação dos parâmetros do modelo	15
4.4.3	Estimativa dos desvios padrão	16
4.4.4	Inferência em um modelo logístico simples	17
4.4.4.1	Teste de Wald	17
4.4.5	Intervalos de Confiança	18
4.4.5.1	Intervalo de Confiança para os parâmetros	18
4.4.5.2	Intervalo de Confiança para a OR	18
4.5	Regressão logística múltipla	19
4.6	Seleção de Variáveis	19
4.7	Análise de resíduos e técnicas de diagnóstico	20
4.7.1	Pontos de Alavanca	21
4.7.2	Resíduo de Pearson, Deviance e Resíduo de Pearson estudentizados	21
4.7.3	Distância de Cook	23
4.8	Predição	24
<b>5</b>	<b>Material e Método</b>	<b>26</b>
5.1	Caracterização da pesquisa	26
5.2	Localização da área	26

5.2.1	Comunidades de Serra Grande . . . . .	27
5.3	Caracterização dos participantes . . . . .	28
5.4	Critérios de inclusão e exclusão . . . . .	29
5.4.1	Critérios de inclusão . . . . .	29
5.4.2	Critério de exclusão . . . . .	29
5.5	Modelo logístico binomial . . . . .	29
6	Resultados e Discussão . . . . .	32
6.1	Análise bivariada – QS . . . . .	36
6.2	Análise multivariada . . . . .	38
6.2.1	Modelo logístico - QS . . . . .	39
6.2.2	Análise de resíduos e técnicas de diagnóstico - QS . . . . .	41
6.2.3	Predição - QS . . . . .	43
7	Conclusão . . . . .	44
Referências Bibliográficas . . . . .		45
Apêndices . . . . .		49
APÊNDICE A TERMO DE CONFIDENCIALIDADE . . . . .		50
APÊNDICE B FICHA DE AVALIAÇÃO . . . . .		51
APÊNDICE C SELEÇÃO DE VARIÁVEIS . . . . .		53
Anexos . . . . .		54
ANEXO A ÍNDICE DE QUALIDADE DO SONO DE PITTSBURGH . . . . .		55
ANEXO B INSTRUÇÕES DE PONTUAÇÃO PARA O ÍNDICE DE QUALIDADE DO SONO DE PITTSBURGH . . . . .		60
ANEXO C CURVA ROC E A INTERPRETAÇÃO DE SEUS ÍNDICES . . . . .		64

# 1 Introdução

O sono pode ser definido como “um período de repouso para o corpo e a mente, durante o qual a volição e a consciência estão em inatividade parcial ou completa; também descrito como um estado comportamental, com postura imóvel e sensibilidade diminuída, mas prontamente reversível a estímulos externos” (DORLAND, 2011).

Segundo Friedman e Schnurr (1995) dizem que o sono é o “desencadear deliberado de uma alteração ou redução do estado consciente, que dura muito, em média 8 horas resultando, geralmente, em sensação de energia física, psíquica e intelectual restabelecida”. Embora a literatura apresente várias definições para o sono, no geral as definições propostas pelos diferentes autores se, complementam.

Em diversas situações o sono é importante para a recuperação da saúde em situação de doença, com sua privação pode afetar a regeneração celular assim como a total recuperação da função imunitária (FONTAINE; BRIGGS; POPE-SMITH, 2001). Adicionalmente, o sono é fundamental na consolidação da memória, na visão binocular, na termorregulação, na conservação e restauração da energia (VELLUTI; REIMÃO, 1996), e na restauração do metabolismo energético cerebral (FERRARA; GENNARO, 2001). Devido a essas importantes funções, as perturbações do sono podem acarretar alterações significativas no funcionamento cognitivo, físico, ocupacional e social do indivíduo, comprometendo consideravelmente a qualidade de vida.

Algumas privações do sono são bastante comuns, a citar: os despertares noturnos, a redução da duração e o aumento da latência do sono, despertar precoce e redução da eficiência do sono (LUCCHESI et al., 2005), entre outros sintomas.

Embora, a literatura não apresente um número fixo para o total de horas de sono que uma pessoa necessite, alguns estudos indicam que este valor deve ser em torno de sete e a oito horas (TEIXEIRA, 2008). Vale ressaltar que isto foi corroborado pelo professor Shawn Youngstedt, Faculdade de Enfermagem e Inovação em Saúde da ASU Phoenix - Estados Unidos, o qual determinou que “as taxas de mortalidade e de morbidade são mais baixas com uma noite de sono de sete horas” (NORONHA, 2014).

Segundo Fontaine, Briggs e Pope-Smith (2001), nem todos os indivíduos adultos precisem do mesmo número de horas de sono, especialistas acreditam que menos de 7 horas de sono por noite, numa base contínua, pode ter consequências negativas para o



corpo e para o cérebro. Por outro lado, há estudos indicando muitas horas de sono (em média >8 horas) pode causar problemas a saúde.

Nesse contexto, Buysse et al. (1998) desenvolveram o índice de Qualidade do Sono de Pittsburgh (IQSP), para criar uma medida padronizada projetada para coletar informações consistentes sobre a natureza subjetiva dos hábitos de sono das pessoas e fornecer um índice claro que ambos os clínicos e pacientes possam utilizar (BUYSSE et al., 1989; MOLLAYEVA et al., 2016; HUNSLEY; MASH, 2008). O IQSP ganhou popularidade como uma medida que poderia ser usada na pesquisa que olha como o sono pôde ser associado com desordens do sono, depressão e desordem bipolar.

O IQSP é validado e amplamente utilizado em inúmeros Países, tendo versões em português, espanhol, holandês, Frances, norueguês, sueco, hebraico e mandarim (BUYSSE et al., 1989). Apresentou uma sensibilidade de 89,6% e uma especificidade de 86,5%. Quando traduzido para português mantém alta sensibilidade, 80%, mas a sua especificidade diminui para 68,8% (BARBA et al., 2011).

A partir do projeto Potencialidade do uso da água, manejo florestal e suas implicações na qualidade de vida das mulheres, adultas e idosas de comunidades do Sertão do Pajeú/PE, em 2014, surgiram dúvidas relacionadas a possíveis fatores que podiam contribuir diretamente na qualidade de vidas dos moradores(as) envolvidos(as), entre elas a QS. Deste modo, as conseqüências dos distúrbios do sono são fortemente relacionadas à qualidade de vida das pessoas que sofrem desse mal (MÜLLER; GUIMARÃES, 2007).

Em geral, os pesquisadores em QS têm em mente perguntas de relação causal ou aquelas que visam conceituar ou descrever a ocorrência da má (ou boa) QS e seus fatores associativos. Diante desta perspectiva este trabalho, objetivou-se avaliar e modelar a QS em algumas comunidades (Poço do Serrote, Catolé, Três Irmãos e Poldrinho) do Sertão do Pajeú/PE, incorporando o questionário de avaliação IQSP e as relações sociodemográficas. Através da modelagem logística binomial, podemos analisar os fatores (ou variáveis) de risco ou proteção associada à boa QS.

## 2 Objetivos

### 2.1 Objetivo Geral

Avaliar e modelar a QS nas comunidades de Poço do Serrote, Catolé, Três Irmãos e Poldrinho pertencentes ao Sertão do Pajeú/PE, incorporando o questionário de avaliação IQSP e as relações sociodemográficas.

### 2.2 Objetivos Específicos

- Realizar uma revisão bibliográfica sobre o sono e o IQSP;
- Caracterizar o perfil geral dos moradores com boa e má QS nas comunidades;
- Avaliar o padrão de sono nas comunidades do Sertão do Pajeú/PE e os fatores associados à QS;
- Investigar a associação entre as pontuações IQSP e as características demográficas.

## 3 Revisão de Literatura

Neste capítulo serão abordados o Sertão do Pajeú, a QS e suas respectivas fases. Na seção 3.3 apresentaremos a análise bivariada com o objetivo de explorar as relações (similaridades) entre a QS e as variáveis sociodemográficas. Em seguida nas seções 3.4 e 3.5 serão abordadas a modelagem logística simples e múltiplas para dados binários, denotando “fracasso” ou “sucesso” na ocorrência de um evento de interesse. Após a formulação do modelo iremos apresentar na Seção 3.6 o método de seleção variável para o ajuste do modelo. Formulado e ajustado o modelo iremos apresentar na seção 3.7 técnicas de resíduos e diagnósticos para a qualidade do ajuste. E por fim na seção 3.8 iremos definir uma ferramenta fundamental para medir e especificar problemas na avaliação de testes de diagnósticos.

### 3.1 Sertão do Pajeú

O Território Sertão Do Pajeú - PE abrange uma área de 13.350,30  $Km^2$  e é composto por 20 municípios (Figura 1): Afogados da Ingazeira, Brejinho, Calumbi, Carnaíba, Flores, Igaraci, Ingazeira, Itapetim, Mirandiba, Quixaba, Santa Cruz da Baixa Verde, Santa Terezinha, São José do Belmonte, São José do Egito, Serra Talhada, Sertânia, Solidão, Tabira, Triunfo e Tuparetama.

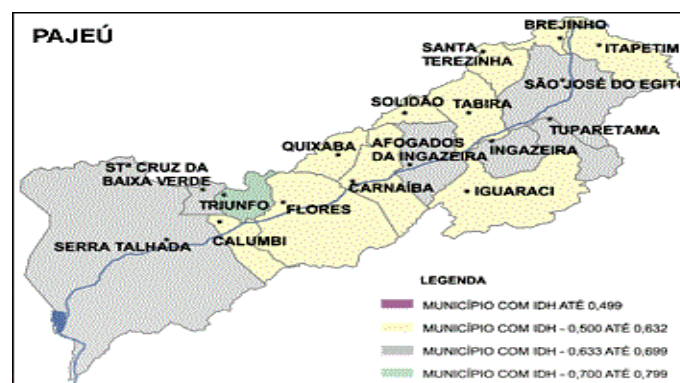


Figura 1 – Municípios do Sertão Pajeú/2014 - PE. Fonte: Blog TV Web Sertão.

A população territorial é de 395.315 habitantes, dos quais 38,89% (número de habitantes = 153.723) vivem na zona rural (Figura 2). É composta por aproximadamente 33.804 agricultores familiares, 1.810 famílias assentadas, 16 comunidades quilombolas e 1

terras indígenas (Mirandiba). Apresenta um Índice de Desenvolvimento Humano (IDH) médio de 0,65 e clima semiárido, quente e seco, na maioria de seu território.

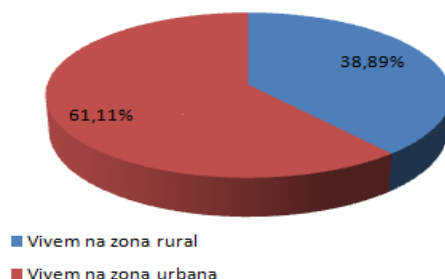


Figura 2 – População territorial do Sertão do Pajeú - PE/2010. Fonte: IBGE (2010).

Dos vinte municípios envolvendo o Sertão do Pajeú/PE o mais populoso segundo a estimativa de 2016, é o município de Serra Talhada (IBGE, 2016) abrangendo uma área de 2.980,007  $Km^2$ . A população territorial estimada é de 84.970, dos quais 22,66% (número de habitantes = 17.957) vivem na zona rural (Figura 3) e 52,25% da população residente é do gênero feminino. É composta por 4.221 agricultores familiares, 757 famílias assentadas e apresenta um IDH de 0,661.

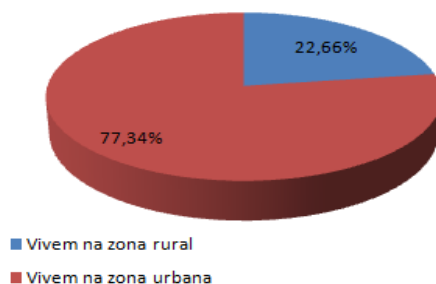


Figura 3 – População territorial de Serra Talhada - PE/2010. Fonte: IBGE (2010).

O município de Serra Talhada está localizado em latitude 07°59'31" Sul, longitude 38°17'54" Oeste, na Mesorregião do Sertão Pernambucano, Microregião do Pajeú, a uma altitude de 429 metros. O clima do município é semiárido, do tipo Bsh. Apresenta precipitação pluviométrica de 686 mm/ano e temperatura média anual de aproximadamente 24 °C.

## 3.2 Sono

O sono é uma função essencial, pois assim como não se pode sobreviver sem respirar, também não há como subsistir sem dormir. O mesmo desempenha um papel fundamental,

pois garante a recuperação física e psíquica e, não suprime apenas o estado de vigília, como também repara todas as suas funções.

Várias funções são atribuídas ao sono. A hipótese mais simples é a de que o sono se destina à recuperação do organismo de um possível débito energético estabelecido durante a vigília. Além dessa hipótese, outras funções são atribuídas, especialmente:

- **a conservação de energia**→ quando a temperatura corporal baixa durante o sono e reduz o consumo de oxigênio, sendo uma espécie de poupança de energia para conservá-la depois das tarefas realizadas durante o dia;
- **a inatividade adaptativa**→ quando o sono não deve ser visto como uma função de vulnerabilidade que persistiu no processo evolutivo por conta de seus aspectos fisiológicos desconhecidos, mas sim por aumentar o estado de eficiência de um organismo;
- **o processamento cerebral e consolidação da memória**→ quando a memória episódica recente que inicialmente armazenada no hipocampo durante o sono é reativada e distribuída para as áreas corticais associativas, onde se torna memória de longa duração (Figura 4);

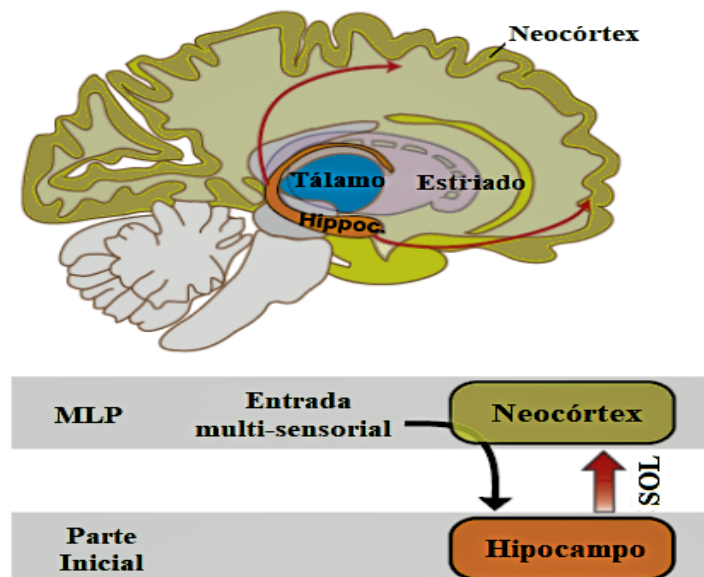


Figura 4 – Memória episódica humana. “Texto traduzido”(VORSTER; BORN, 2015).

- **a restauração**→ quando o sono é importante para reparar, rearmar ou reconstruir o que foi gasto durante a vigília. Essa função é de suma importância para eliminar produtos de degradação da atividade neuronal que se acumula durante a vigília.

Dessa forma, a QS pode ser avaliada de um ponto de vista objetivo referente a dificuldade de uma pessoa adormecer e permanecer em estado de sono ou subjetivo referente a sensação de descanso após acordar do sono.

Basicamente podemos dividir o sono em duas fases: REM (Movimento Rápido dos Olhos) e NREM (Movimento Não Rápido dos Olhos). O estado do NREM corresponde a 75% do período do sono, sendo dividido em quatro fases.

### 3.2.1 Fases do sono

As fases do sono são divididas em cinco estágios, sendo que cada estágio apresenta um característica específica (Figura 5). Os cinco estágios formam um ciclo, durando cerca de 90 - 110 minutos cada um. Normalmente, o número de ciclos varia entre quatro a cinco ciclos por noite.



Figura 5 – Fases do sono. Fonte: BedTime (2017).

Neste tempo o cérebro age de forma diferente em cada estágio. Em algumas etapas, o corpo se movimenta (NREM), enquanto em outras, permanece completamente imóvel (REM). Algumas das etapas estão diretamente ligadas com a saúde da memória, concentração e até problemas com estresse. Os estágios são divididos em:

- Estágio 1 (Transição da vigília para o sono leve): Ocupa de 4% a 5% da noite e nele ocorre a liberação da melatonina, hormônio que induz à sonolência. A temperatura corporal começa a cair, os músculos relaxam e os olhos se movem lentamente de um lado para o outro;

- Estágio 2: Ocupa de 45% a 55% do tempo, com duração média de 10 a 25 minutos. Neste estágio há uma redução no ritmo cardíaco, queda de temperatura corporal e relaxamento muscular;
- Estágio 3: Ocupa de 4% a 6% do tempo, o corpo funciona lentamente. Durante este estágio, conhecido como sono profundo ou sono lento, o cérebro torna-se menos sensível aos estímulos externos, tornando difícil o despertar.
- Estágio 4: Ocupa de 12% a 15% do tempo, com duração de cerca de 40 minutos. Neste estágio há queda de pressão sanguínea, a respiração torna-se mais lenta, o aumento de sangue nos músculos, a energia é restaurada, ocorre crescimento e reparo de tecido e hormônios são liberados. A diferença entre o estágio 3 e o 4 é que sono está em um nível ainda mais profundo ;
- REM: Ocupa de 20% a 25% do tempo, com 5 a 30 minutos de duração, a cada 90 minutos. Nesta fase os movimentos oculares rápidos, há um Aumento do fluxo sanguíneo cerebral, a ausência de reflexos espinais e quase do tônus muscular (exceto pelo controle da respiração) o corpo fica imóvel e relaxado, enquanto os músculos são desligados. Além, de está associado a atividades psíquicas, como o sonho.

A prática de exercícios físicos é importante para prevenir doenças, promover saúde e qualidade de vida, melhorar a autoestima e a QV, trazer resultados estéticos e até prolongar a expectativa de vida. Entretanto, o exercícios físicos em excesso podem trazer prejuízo, segundo especialistas.

Analogamente, o esforço físico elevado no trabalho é um fator de risco para distúrbios do sono, embora com baixa intensidade é considerado um fator de proteção para qualidade de vida (MARTINS; MELLO; TUFIK, 2001).

### 3.3 Análise Bivariada

O objetivo da análise bivariada é explorar as relações (similaridades) entre duas variáveis aleatórias X e Y. De modo geral, a quantificação do grau de associação entre duas variáveis é feita pelo coeficiente de associação ou correlação. Essas são medidas que descrevem, por meio de um único número, a associação (ou dependência) entre duas variáveis qualitativas.

Nesta seção iremos apresentar dois testes utilizando tabelas de dupla entradas (ou contingência) para verificar a associação entre duas variáveis.

### 3.3.1 Teste Qui-quadrado

O Teste Qui-quadrado é um teste de hipótese não paramétrico<sup>1</sup> que se destina encontrar um valor de dispersão para duas variáveis nominais, e avaliar a associação existente entre variáveis qualitativas. Contudo, se faz necessário verificar se foram satisfeitas as suposições do teste.

#### Suposições:

- Nível de Mensuração em escala nominal ao menos.
- $N > 25$  e frequências esperadas  $> 5$  quando ocorre o caso  $2 \times 2$  ( $N$  é o tamanho da amostra);
- Se  $k > 2$ , o número de caselas com frequência esperada inferior a 5 deve ser menos de 20% do total de células ( $k$  é o número de colunas).

Ao realizar um teste de hipóteses, primeiramente definem-se as hipóteses nula ( $H_0$ ) e alternativa ( $H_1$ ) e fixa-se o nível de significância do teste ( $\alpha$ ).

$$\begin{cases} H_0 : \text{As frequências observadas } (n_{ij}) \text{ não são diferentes das frequências esperadas} \\ H_1 : \text{As frequências observadas são diferentes das frequências esperadas} \end{cases} \quad (3.1)$$

Antes de definir a fórmula geral para a medida de associação, vamos introduzir, na Tabela 1, uma notação geral para tabelas de dupla entradas.

Tabela 1 – Notação para tabela de contingência

X \ Y	$B_1$	$B_2$	$\dots$	$B_j$	Total
$A_1$	$n_{11}$	$n_{12}$	$\dots$	$n_{1j}$	$n_{1.}$
$A_2$	$n_{21}$	$n_{22}$	$\dots$	$n_{2j}$	$n_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$A_i$	$n_{i1}$	$n_{i2}$	$\dots$	$n_{ij}$	$n_{i.}$
Total	$n_{.1}$	$n_{.2}$	$\dots$	$n_{.j}$	$n_{..}$

Considere duas variáveis aleatórias qualitativas  $X$  e  $Y$ , classificadas em  $i$  categorias  $A_1, A_2, \dots, A_i$  para  $X$  e  $j$  categorias  $B_1, B_2, \dots, B_j$  para  $Y$ , sendo composta por:

- $n_{ij}$  número de elementos pertencentes à  $i$ -ésima categoria de  $X$  e  $j$ -ésima categoria de  $Y$ ;

<sup>1</sup> Teste não paramétrico: não depende de parâmetros populacionais média  $\mu$  e variância  $\sigma^2$ .



- $n_i$ . número de elementos pertencentes à  $i$ -ésima categoria de X;
- $n_j$ . número de elementos pertencentes à  $j$ -ésima categoria de Y;
- $n_{..}$ . número total de elementos.

Ao testar a associação entre duas variáveis numa determinada população, em geral não se conhece a frequência esperada ( $n_{ij}^*$ ), no entanto, é possível estimá-la. Sob a  $H_0$  (4.1) de que as variáveis X e Y não sejam associadas, temos que

$$\frac{n_{i1}}{n_{.1}} = \frac{n_{i2}}{n_{.2}} = \dots = \frac{n_{ij}}{n_{.j}} \quad \text{para } i=1,2,\dots$$

ou ainda

$$\frac{n_{ij}}{n_{.j}} = \frac{n_{i.}}{n_{..}} \quad \text{para } i=1,\dots \text{ e } j=1,\dots$$

de onde se deduz, que a frequência esperada é dada por

$$n_{ij}^* = \frac{n_{i.}n_{.j}}{n_{..}} \quad \text{para } i=1,\dots \text{ e } j=1,\dots$$

Para testar as hipóteses (4.1), a estatística teste é dada por

$$\chi^2 = \sum \frac{(n_{ij} - n_{ij}^*)^2}{n_{ij}^*},$$

sendo necessário uma regra de decisão para rejeitar ou não  $H_0$ .

### Regra de decisão:

Rejeita-se  $H_0$  se  $p\text{-valor} < \alpha$  ou  $\chi^2 > \chi_{(k-1),(1-\alpha)}^2$ . Obtemos os valores críticos de  $\chi_{(k-1),(1-\alpha)}^2$  através da consulta a Tábua C – Siegel.

### 3.3.2 Teste Exato de Fisher

Esta técnica foi proposta na década de 1930, quase simultaneamente por R. A. Fisher, J. O. Irwin e F. Yates, em publicação separadas, mas é mais conhecida pelo nome do primeiro autor. É chamada de teste “exato” porque calcula a probabilidade exata de se obter, ao acaso, os resultados observados nas caselas ou resultados ainda mais extremos, para os mesmos totais das margens.

O Teste Exato de Fisher é a alternativa para tabelas 2x2 quando não se pode usar o teste  $\chi^2$  para valor esperado  $< 5$  ou o número total de indivíduos estudados é menor do que 25, sendo necessário verificar se foram satisfeitas as suposições do teste.

**Suposições:**

- Amostras aleatórias e independentes;
- Duas classes mutuamente exclusivas;
- Nível de Mensuração em escala nominal ao menos.

Ao realizar um teste de hipóteses, primeiramente definem-se as hipóteses  $H_0$  e  $H_1$  e fixa-se o valor do  $\alpha$

$$\begin{cases} H_0 : P_I = P_{II} \\ H_1 : P_I \neq P_{II} \end{cases} \quad (3.2)$$

em que  $P_I$  e  $P_{II}$  são as probabilidades do evento I e II.

Seja I e II duas amostras aleatórias e independentes, agrupadas em duas classes – e + (Tabela 2).

Tabela 2 – Comparação entre dois grupos

	+	-	
I	A	B	A+B
II	C	D	C+D
	A+C	B+D	N

Para testar as hipóteses (4.2), a estatística teste é dada por

$$P = \frac{(A+B)!(C+D)!(A+C)!(B+D)!}{N!A!B!C!D!} \quad (3.3)$$

em que P calcula a probabilidade de ocorrência, ao acaso, das frequências observadas quando  $H_0$  for verdadeira.

Deve-se ressaltar que os números obtidos nos fatoriais em (4.3) podem resultar em valores grande, dificultando o calculo. Uma forma de contornar esse resultado é aplicar a função logaritmica em (4.3) e, em seguida, aplicar a função exponencial no valor obtido para chegar ao valor de P.

**Regra de decisão:**

Rejeita-se  $H_0$  se  $P < \alpha$ . (teste unilateral) ou

Rejeita-se  $H_0$  se  $P^* = 2P < \alpha$ .( teste bilateral)

### 3.4 Regressão logística simples

Os modelos lineares generalizados (MLG) foram formulados para unificar diversos modelos estatístico, incluindo a regressão logística (NELDER; BAKER, 1972), usado para avaliar a relação entre uma ou mais variáveis preditoras (VP) e uma variável resposta (VR). A regressão logística é útil para situações nas quais deseja-se estar apto a prever a presença ou ausência (1 ou 0) de uma característica ou resultado, baseado num conjunto de VP (FIGUEIRA, 2006).

Seja  $Y = VR$  uma variável aleatória discreta (VAD) seguindo distribuição Bernoulli, que é da classe da família exponencial cujo o parâmetro  $\pi$  denota a probabilidade de “sucesso” no valor de  $x$ . O modelo de regressão logística tem forma linear, sendo considerada a função de ligação logit ( $g(x)$ ),

$$Y_i = g(x) + \epsilon_i = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) + \epsilon_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

em que

- 

$$\pi(x) = \frac{\exp(g(x))}{1 + \exp(g(x))} \quad (3.4)$$

$\pi(x)$  denota a probabilidade de “sucesso” no valor de  $x$ ;

- $\beta_0$  é o parâmetro fixo;
- $x_i$  são as VP;
- $\beta_1$  representa a variação esperada no logaritmo da chance por unidade de variação na variável  $x_1$ .
- $\epsilon_i$  é o erro experimental com distribuição normal com média ( $\mu$ ) e variância ( $\sigma^2$ ).

Em (4.4), é necessário que  $\beta_1 < 0$  para que  $\pi$  seja crescente e que  $\beta_0 > 0$  para que  $\pi$  seja decrescente (Figura 6). Quando  $x \rightarrow \infty$ ,  $\pi(x) \rightarrow 0$  quando  $\beta_1 < 0$  e  $\pi(x) \rightarrow 1$  quando  $\beta_0 > 0$ . Caso  $\beta_1 = 0$ , o lado direito da equação (4.4) se torna constante e  $Y$  é independente de  $X$ .

#### 3.4.1 Estimação dos parâmetros do modelo

Na estimação dos parâmetros do modelo  $(\beta_0, \beta_1)$  utilizaremos o método da máxima verossimilhança (RYAN, 2008). Fornecendo-nos valores para os parâmetros desconhecidos que maximizam a probabilidade de se obter determinado conjunto de dados.

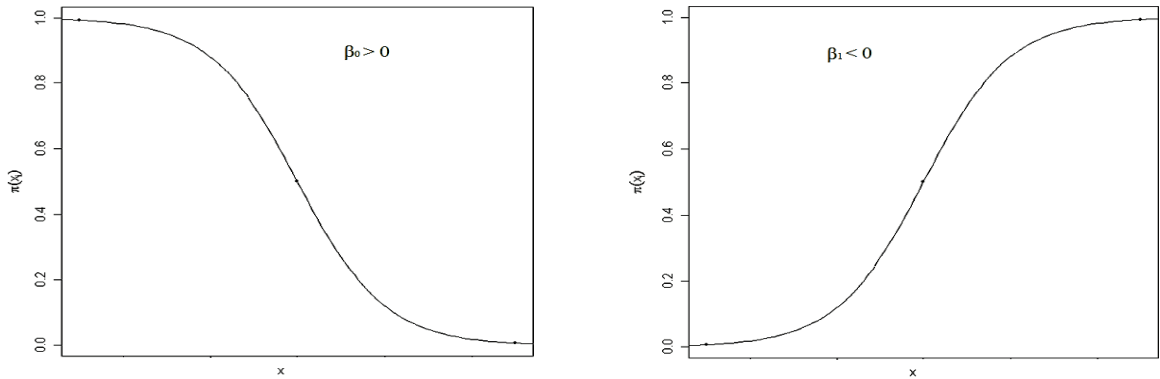


Figura 6 – Funções de regressão logística

O método da máxima verossimilhança é ilustrado a seguir. Suponha uma amostra aleatória  $n$  independente  $(x_i, y_i)$ ,  $i=1,2,\dots,n$  com função densidade de probabilidade (FDP) definida por:

$$f(y_i, \pi_i) = \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \quad (3.5)$$

em que,

- $y_i$  é a VR;
- $n$  é o tamanho da amostra.

Assumindo que  $(x_0, y_0), \dots, (x_n, y_n)$  são independentes, a função de verossimilhança é da seguinte forma

$$\begin{aligned} \mathbf{L}(\beta) &= \prod_{i=1}^n f(y_i, \pi_i) = \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i)^{1-y_i} \\ &= \prod_{i=1}^n \pi_i^{y_i} (1 - \pi_i) (1 - \pi_i)^{-y_i} \\ &= \prod_{i=1}^n \frac{\pi_i^{y_i} (1 - \pi_i)}{(1 - \pi_i)^{y_i}} \\ &= \prod_{i=1}^n \left( \frac{\pi_i}{1 - \pi_i} \right)^{y_i} (1 - \pi_i). \end{aligned} \quad (3.6)$$

Aplicando o logaritmo neperiano ( $\ln$ ) em 4.6, temos

$$\log(\mathbf{L}(\beta)) = \log \left\{ \prod_{i=1}^n \left( \frac{\pi_i}{1 - \pi_i} \right)^{y_i} (1 - \pi_i) \right\} = \sum_{i=1}^n \log \left\{ \left( \frac{\pi_i}{1 - \pi_i} \right)^{y_i} (1 - \pi_i) \right\} \quad (3.7)$$

Detalhando  $\log\left(\frac{\pi_i}{1-\pi_i}\right)$  e considerando que,

$$\pi_i = \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}, \text{ temos}$$

$$\begin{aligned} \log\left(\frac{\pi_i}{1-\pi_i}\right) &= \log\left(\frac{\frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}}{1 - \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}}\right) = \\ &= \log\left(\frac{\frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}}{\frac{1 + \exp(\beta_0 + \beta_1 x_i) - \exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}}\right) = \log\left(\frac{\frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}}{1}\right) = \\ &= \log(\exp(\beta_0 + \beta_1 x_i)) = \beta_0 + \beta_1 x_i. \end{aligned}$$

Dessa forma, a expressão (4.7), pode ser reescrita como

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

Portanto,

$$\log(L(\beta)) = \sum_{i=1}^n y_i (\beta_0 + \beta_1 x_i) - \sum_{i=1}^n \log(1 + \exp(\beta_0 + \beta_1 x_i)) \quad (3.8)$$

O valor de  $\beta$  que maximiza  $\log(L(\beta))$  é encontrado após derivar-se  $\log(L(\beta))$  em relação aos parâmetros  $(\beta_0, \beta_1)$ :

$$\frac{\partial \log(L(\beta))}{\partial(\beta_0)} = \sum_{i=1}^n \left[ y_i - \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)} \right] \quad (3.9)$$

$$\frac{\partial \log(L(\beta))}{\partial(\beta_1)} = \sum_{i=1}^n \left[ y_i x_i - \frac{\exp(\beta_0 + \beta_1 x_i) x_i}{1 + \exp(\beta_0 + \beta_1 x_i)} \right]. \quad (3.10)$$

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

$$\begin{cases} \sum_{i=1}^n (y_i - \pi_i) = 0; \\ \sum_{i=1}^n x_i (y_i - \pi_i) = 0. \end{cases} \quad (3.11)$$

Tais equações são não-lineares nos parâmetros e requerem o emprego de processo iterativo na sua solução, por exemplo, Newton-Raphson (GOURIEROUX; MONFORT, 1995).

### 3.4.2 Interpretação dos parâmetros do modelo

A interpretação dos parâmetros do modelo é obtida pela razão da probabilidade de sucesso ( $\pi$ ) com a probabilidade de fracasso ( $1-\pi$ ), usando a função odds ratio (OR) ou razão de chances. Essa função é definida por:

$$\begin{aligned} g(x) &= \frac{\pi(x)}{[1 - \pi(x)]} = \frac{\frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}}{1 - \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}} = \\ &= \frac{\frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)}}{\frac{1}{1 + \exp(\beta_0 + \beta_1 x_i)}} = \exp(\beta_0 + \beta_1 x_i). \end{aligned}$$

Assim, ao tomarmos dois valores distintos da VP,  $x_j$  e  $x_{j+1}$ , obtemos

$$OR = \frac{g(x_{j+1})}{g(x_j)} = \frac{\exp(\beta_0 + \beta_1 x_{j+1})}{\exp(\beta_0 + \beta_1 x_j)}. \quad (3.12)$$

Aplicando a função logaritmo ( $\log$ ) em 4.12, temos

$$\log(OR) = \log \left[ \frac{g(x_{j+1})}{g(x_j)} \right] = \log [g(x_{j+1})] - \log [g(x_j)]$$

$$= \beta_0 + \beta_1 x_{j+1} - \beta_0 - \beta_1 x_j = \beta_1 (x_{j+1} - x_j). \quad (3.13)$$

Assumindo que  $x_{j+1} - x_j = 1$  unidade, então

$$\log(OR) = \log(\exp(\beta_1)) = \beta_1. \quad (3.14)$$

Dessa forma, temos a relação dos resultados de ocorrência entre os indivíduos  $x_{j+1}$  e  $x_j$ , fazendo, portanto, algumas análises:

$$(\text{Fator de risco}) \quad \beta_1 > 0 \Rightarrow OR > 1 \Rightarrow \pi(x_{j+1}) > \pi(x_j) \quad (3.15)$$

$$(\text{Fator de proteção}) \quad \beta_1 < 0 \Rightarrow OR < 1 \Rightarrow \pi(x_{j+1}) < \pi(x_j). \quad (3.16)$$

### 3.4.3 Estimativa dos desvios padrão

A estimativa dos desvios padrão (DP) no modelo logístico é obtida a partir da matriz de informação de Fisher. Assim sendo, podemos obter a matriz de informação de Fisher  $I(\hat{\beta})$  a partir dos dados, da seguinte forma,

$$I(\hat{\beta}) = \begin{bmatrix} 1 & 1 \\ x_1 & x_2 \end{bmatrix} \begin{bmatrix} \hat{\pi}_1(1 - \hat{\pi}_1) & 0 \\ 0 & \hat{\pi}_2(1 - \hat{\pi}_2) \end{bmatrix} \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \end{bmatrix} = X' W X. \quad (3.17)$$

Aplicando a inversa em 4.17, temos

$$\hat{\Sigma} = I^{-1}(\hat{\beta}) = \begin{bmatrix} \sigma^2(\hat{\beta}_1) & Cov(\hat{\beta}_1, \hat{\beta}_2) \\ Cov(\hat{\beta}_2, \hat{\beta}_1) & \sigma^2(\hat{\beta}_2) \end{bmatrix}$$

em que

- $\hat{\beta}_0$  é o estimador do parâmetro fixo;
- $\hat{\beta}_1$  representa o estimador da variação esperada no logaritmo da chance por unidade de variação na variável  $x_1$ ;
- $\hat{\Sigma}$  é o estimador da matriz de covariâncias dos estimadores  $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)$ .

Com base no estimador da matriz de covariâncias  $\widehat{\Sigma}$  o desvio padrão de  $\widehat{\beta}$  é definido por:

$$\widehat{DP}(\widehat{\beta}_1) = \sqrt{\widehat{\sigma}^2(\widehat{\beta}_1)}; \quad (3.18)$$

$$\widehat{DP}(\widehat{\beta}_2) = \sqrt{\widehat{\sigma}^2(\widehat{\beta}_2)}. \quad (3.19)$$

### 3.4.4 Inferência em um modelo logístico simples

Após estimar os parâmetros, temos interesse de verificar se a VP é significativa no modelo. Assim, iremos aplicar testes de hipóteses para determinar se a VP no modelo é significativa em relação a VR.

#### 3.4.4.1 Teste de Wald

O teste de Wald é um teste de hipóteses para avaliar o modelo de regressão logística como um todo, tem como finalidade aferir o grau de significância de cada coeficiente da equação logística ( $\widehat{\beta}_1$ ), inclusive a constante ( $\widehat{\beta}_0$ ).

**Hipóteses:**

$$\begin{cases} H_0 : \beta_1 = 0 \\ H_1 : \beta_1 \neq 0 \end{cases} \quad (3.20)$$

Para testar esta hipótese (4.20), a estatística teste é dada por

$$W_j = \frac{\widehat{\beta}_1}{\widehat{DP}(\widehat{\beta}_1)}$$

em que

- $\widehat{DP}(\widehat{\beta}_1)$  é o estimador do desvio padrão de  $\widehat{\beta}_1$ ;
- $W_j$  avalia se o parâmetro é estatisticamente significativo no modelo.

O p-valor é definido como

$$p - \text{valor} = P[|Z| > |W_j||H_0] = 2P[Z > |W_j||H_0],$$



onde  $Z$  denota uma VA com distribuição normal padrão. Se  $p\text{-valor} < \alpha$ , o coeficiente  $\hat{\beta}_1$  é significativo.

Tendo em vista que o teste de Wald frequentemente não rejeita  $H_0$  (valor-p  $> 5\%$ ) quando a variável de fato é significativa. Jr e Donner (1977) sugerem a aplicação do teste da razão de verossimilhança (RV) para testar se realmente a variável não é significativa quando o teste de Wald não rejeita a  $H_0$ .

### 3.4.5 Intervalos de Confiança

#### 3.4.5.1 Intervalo de Confiança para os parâmetros

A base da construção dos intervalos de confiança para os parâmetros é a mesma teoria estatística que usamos para os testes de significância do modelo. O intervalo de confiança de  $100(1-\alpha)\%$  para o parâmetro  $\beta_j$  é

$$IC(\beta_j, 1 - \alpha) = [\hat{\beta}_j - z_{1-\alpha/2}DP(\hat{\beta}_j); \hat{\beta}_j + z_{1-\alpha/2}DP(\hat{\beta}_j)]$$

em que

- $j$  assume valores 0 ou 1;
- $z_{1-\alpha/2}$  é o ponto da normal padrão correspondente a  $100(1-\alpha)\%$ .

#### 3.4.5.2 Intervalo de Confiança para a OR

Sejam os limites do intervalo de confiança para  $\beta_1$

$$\begin{cases} \beta_I = \hat{\beta}_1 - z_{1-\alpha/2}DP(\hat{\beta}_1) \\ \beta_S = \hat{\beta}_1 + z_{1-\alpha/2}DP(\hat{\beta}_1) \end{cases}$$

em que

- $\beta_I$  é o limite inferior;
- $\beta_S$  é o limite superior;

O intervalo de confiança para a OR é

$$IC(OR, 1 - \alpha) = [e^{\beta_I}; e^{\beta_S}].$$

### 3.5 Regressão logística múltipla

Na secção 4.4 introduzimos a regressão logística simples, ou seja, para o caso onde temos uma única VP. Considere agora o caso onde se tem um conjunto  $p$  de VP's expresso pelo vetor  $X^T = (X_1, X_2, X_3, \dots, X_p)$  e  $\beta = (\beta_1, \beta_2, \beta_3, \dots, \beta_p)^T$  o vetor de parâmetros desconhecidos (JR; LEMESHOW; STURDIVANT, 2013). No modelo de regressão múltipla, a probabilidade de sucesso é dada por

$$E(Y|x) = \pi(x) = \frac{\exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi})}{1 + \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi})}$$

em que

- $\beta_0$  é o parâmetro fixo;
- $\beta_j$  representa a variação esperada no logaritmo da chance por unidade de variação na variável  $x_j$  com  $j=1,2,\dots, p$ .

Assim, a função de ligação  $g(\cdot)$  toma a forma

$$g(x) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi}$$

e o modelo logístico múltiplo é definido por

$$Y_i = g(x) + \epsilon_i.$$

De forma análoga a estimativa dos parâmetros, a inferência e o intervalos de confiança no caso múltiplo pode ser visto como uma extensão do caso simples, onde tem-se agora em vez de um preditor  $X$ , um conjunto com  $p$  preditores.

### 3.6 Seleção de Variáveis

Uma vez definido o conjunto de VP a ser inserida num modelo, estamos interessados em encontrar um modelo reduzido que inclua apenas as VP's mais importantes para explicar a probabilidade de sucesso  $\pi(x)$ . Assim, em seguida, será discutido o método de seleção Backward para a seleção do modelo.

**Seleção Backward:**

O método, baseia-se num algoritmo que constrói iterativamente um modelo incorporando inicialmente todas as variáveis e depois, por etapas, cada uma podendo ser ou não eliminada.

**Procedimento:**

Iniciamos o procedimento pelo modelo

$$Y_j = \beta_0 + \beta_1 x_1 + \beta_j x_j \quad (j = 2, \dots, p)$$

**1. Testamos**

$$\begin{cases} H_0 : \beta_j = 0 \\ H_1 : \beta_j \neq 0 \end{cases} \quad j = 2, \dots, p$$

Seja  $\alpha^*$  o maior nível de significância descritivo dentre os  $q$  testes RV. Se  $\alpha^* > \alpha$ , a variável correspondente sai do modelo. Vamos supor que  $x_1$  tenha saído do modelo ( $\alpha_{x_1}^* > \alpha$ ). Então, ajustamos o modelo

$$Y_j = \beta_0 + \beta_2 x_2 + \beta_j x_j \quad (j = 3, \dots, p)$$

**2. Testamos**

$$\begin{cases} H_0 : \beta_j = 0 \\ H_1 : \beta_j \neq 0 \end{cases} \quad j = 3, \dots, p$$

Seja  $\alpha^*$  o maior nível de significância descritivo dentre os  $(q - 1)$  testes RV. Se  $\alpha^* > \alpha$ , a variável correspondente sai do modelo. Repetimos o procedimento até que ocorra  $\alpha^* < \alpha$ .

Dessa forma, o pesquisador precisa definir o nível de significância ( $\alpha$ ), muitos adotam 0,05. Segundo Lee e Koval (1997), a escolha  $\alpha = 0,05$  é muito rigorosa excluindo variáveis importantes no modelo. Assim, é indicado escolher um valor de  $\alpha$  entre 0,15 e 0,20.

### 3.7 Análise de resíduos e técnicas de diagnóstico

Após a formulação e ajuste do modelo de regressão logística, iremos apresentar nesta seção: a análise dos resíduos e as medidas de análise de diagnóstico da qualidade do ajuste.

### 3.7.1 Pontos de Alavanca

A análise de diagnóstico na modelagem estatística serve para avaliar se uma observação tem influência sobre as estimativas dos parâmetros e na estatística de ajuste. Quando uma observação se encaixa mal no modelo, alguns autores tendem a excluir tal observação e construir um novo modelo para as observações restantes forçando um ajuste perfeito. Porém, se a observação apresenta um ponto de grande alavancagem e for considerada influente não se pode remover da análise pois causará mudanças apreciáveis nas estimativas dos  $\beta$ 's (CORDEIRO; NETO, 2004).

Na regressão logística, os pontos de alavancagem podem ser detectados pelos elementos  $h_{ii}$  da matriz chapéu generalizada, definida por

$$\hat{H} = \hat{W}^{1/2} X (X' \hat{W} X)^{-1} X' \hat{W}^{1/2}$$

em que  $\hat{W}$  é a matriz diagonal  $I \times I$   $\hat{w}_i = n_i \hat{\pi}_i (1 - \hat{\pi}_i)$  para as  $n_i$  observações correspondente ao  $i$ -ésimo preditor.

Desta forma, para o modelo logístico, os pontos de alavanca ( $h_i$ ) são dados por

$$h_i = \hat{w}_i (1, x_i) (X' W X)^{-1} (1, x_i)' = w_i x b_i$$

em que  $b_i = (1, x_i) (X' W X)^{-1} (1, x_i)'$ .

Como a matriz  $\hat{H}$  é simétrica e idempotente, tem-se que:

- $0 \leq h_i \leq 1$ ;
- $tr(H) = \sum_{i=1}^n h_i = p + 1$  onde  $tr(H)$  é o traço da matriz H;

Deste modo, os pontos de alavanca medem o quão distante a observação  $x_i$  está das demais  $n-1$  observações no espaço definido pelas VP do modelo. Hoaglin e Welsch (1978) sugerem usar  $h_{ii} > \frac{2p}{n}$  para indicar os pontos de alavancagem.

### 3.7.2 Resíduo de Pearson, Deviance e Resíduo de Pearson estudentizados

A análise residual na modelagem estatística está relacionada com a qualidade do ajuste, comparando os valores observados com os ajustados. Além, de verificar a existência de pontos discrepantes que interfiram no ajuste do modelo.

Dentre os tipos de resíduos mais comuns nos MLG's, encontra-se o resíduo de Pearson, que é também o mais simples, sendo definido por

$$e_i = \frac{y_i - \hat{\mu}_i}{[\widehat{\text{var}}(Y_i)]^{1/2}} = \frac{y_i - n_i \hat{\pi}_i}{[n_i \hat{\pi}_i (1 - \hat{\pi}_i)]^{1/2}} \quad (3.21)$$

em que

- $Y_i \sim \text{Bin}(\pi, n_i)$  correspondente a  $i$ -ésima ( $i=1, \dots, N$ ) observação da VR;
- $\hat{\pi}_i$  é a estimativa do modelo  $P(Y=1)$ ;
- $n_i \hat{\pi}_i$  é o número ajustado de sucessos.

Uma desvantagem desse tipo de resíduo (4.21) é que ele é geralmente assimétrico para distribuições não normais.

A estatística de Pearson para testar o ajuste do modelo satisfaz

$$X^2 = \sum_{i=1}^N e_i^2$$

em que  $X^2$  é uma componente da estatística de Pearson generalizada, segundo Demétrio e Cordeiro (2007) e  $e_i \sim N(0, 1)$  para  $n$  grande.

Uma vez que  $\pi_i$  é estimado por  $\hat{\pi}_i$  e os valores de  $\hat{\pi}_i$  depende de  $y_i$ , contudo,  $(y_i - \hat{\mu}_i)$  tendem a ser menores que  $(y_i - \mu_i)$ . Se  $X^2$  tem  $gl = v$ ,  $X^2 = \sum_i e_i^2$  é assintoticamente comparável à soma dos quadrados de  $v$  independente da variável aleatória  $N(0,1)$ . Assim, quando o modelo é válido,  $\frac{E(\sum_i e_i^2)}{N} \approx \frac{v}{N} < 1$ .

Porém, o resíduo de Pearson estudentizados é ligeiramente maior em valor absoluto e aproximadamente  $N(0,1)$ , quando o modelo é válido. Haja vista que o ajuste usa a alavancagem de uma matriz de chapéu estimada ( $\hat{h}_i$ ), definido por

$$r_i = \frac{e_i}{\sqrt{1 - \hat{h}_i}} = \frac{y_i - \hat{\mu}_i}{\sqrt{\pi_i (1 - \hat{\pi}_i) (1 - \hat{h}_i)}} \quad (3.22)$$

em que  $\hat{h}_i$  é o  $i$ -ésimo elemento da diagonal da matriz de projeção ortogonal  $H = W^{1/2} X (X' W X)^{-1} X' W^{1/2}$ .

Em geral se faz o uso de gráfico de resíduo de Pearson estudentizados para localizar valores absolutos maiores que  $[-2; 2]$  fornecendo evidências de pontos aberrantes ocorrendo a falta de ajuste.

Quando um MLG se encaixa mal de acordo com um teste geral de bondade de ajuste, o exame dos resíduos destaca onde o ajuste é pobre. Dessa forma, o resíduo muito utilizado dentro da metodologia dos MLG's, é o resíduo componente do desvio. Seja

$$D(y, \hat{\mu}) = \sum_i \frac{2}{\phi_i} \left[ y_i(\tilde{\theta}_i - \hat{\theta}_i) - b(\tilde{\theta}_i) + b(\hat{\theta}_i) \right] = \sum_i d_i$$

em que

- $D(y, \hat{\mu})$  é a deviance que considera a proporção do desvio na variável resposta que explica o modelo.;
- $\tilde{\theta}_i$  o estimador do modelo ajustado;
- $\hat{\theta}_i$  o estimador do modelo saturado;
- $b(\cdot)$  é a função que depende do parâmetros  $(\tilde{\theta}_i, \hat{\theta}_i)$ ;
- e  $\phi$  é o parâmetro de dispersão.

Em resumo, o desvio residual para a observação  $i$  é

$$\sqrt{d_i} \times \text{sinal}(y_i - \hat{\mu}_i)$$

onde  $d_i$  também pode ser definido por

$$d_i = 2 \left( y_i \log \left( \frac{y_i}{n_i \hat{\pi}_i} \right) + (n_i - y_i) \log \left( \frac{n_i - y_i}{n_i - n_i \hat{\pi}_i} \right) \right).$$

Segundo Collet (2003) a distribuição assintótica da deviance é dada por

$$D \sim X_{n-p}^2$$

em que  $p$  é o número de parâmetros estimados no modelo.

### 3.7.3 Distância de Cook

A distância de Cook tem como objetivo medir a influência de cada observação nos parâmetros do modelo (COOK, 1979). Essa medida é calculada através da diferença entre  $\hat{\beta}$  e  $\hat{\beta}_{(-j)}$ , que representam, respectivamente, as estimativas de máxima verossimilhança calculadas usando todos os I padrões de covariáveis e excluindo as observações com padrão  $i$  e ainda padronizada pela matriz da covariância de  $\hat{\beta}$ . Pregibon (1981), mostrou através de uma aproximação linear, que a distância de Cook, para a regressão logística é dada por

$$\Delta\hat{\beta}_j = (\hat{\beta} - \hat{\beta}_{(-j)})'(X'WX)(\hat{\beta} - \hat{\beta}_{(-j)})$$

onde uma observação é dita influente se satisfazer  $\Delta\hat{\beta}_j > 1$ .

### 3.8 Predição

A Curva de Característica de Operação de Receptor (Curva ROC) é uma ferramenta fundamental para medir e especificar problemas na avaliação de testes de diagnósticos (MARGOTTO, 2010), em que traçamos a taxa verdadeira positiva (Sensibilidade) em função da taxa de falsos positivos (1-Especificidade) em diferentes pontos de corte de um parâmetro (Figura 7).

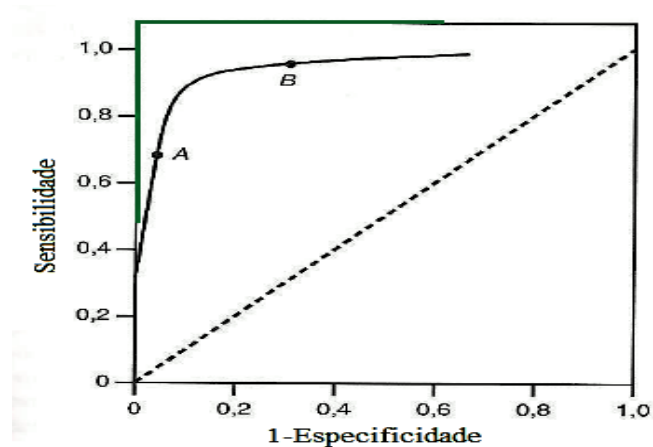


Figura 7 – A Curva de Característica de Operação de Receptor (Curva ROC)

Dessa forma, cada ponto na Curva ROC representa um par sensibilidade ou especificidade correspondente a um determinado limiar de decisão. Assim, a área sob a curva ROC (acurácia) é uma medida de quão bem um parâmetro pode distinguir entre dois grupos de diagnóstico (boa QS ou ruim). Já a linha diagonal pontilhada representa um teste que é positivo ou negativo, aleatoriamente.

A medida que um critério para o teste positivo torna-se mais rigoroso, por exemplo, o ponto “A” da curva (Figura 7) ao movimentar-se para baixo e para a esquerda a sensibilidade diminui e especificidade aumenta. Se adotar um critério menos evidente, para identificar os positivos, por exemplo, o ponto “B” da curva ao movimentar-se para cima e para a direita a sensibilidade aumenta e especificidade diminui.

A escolha do ponto de corte na Curva ROC deve ser baseada em uma combinação da sensibilidade e 1-especificidade que mais se aproxima do canto superior esquerdo. Na prática, curvas consideradas boas estarão entre a linha diagonal (pontilhada) e a

linha perfeita (verde), onde quanto maior a distância da linha diagonal, maior poder de discriminação (Figura 7).

Após a formulação e ajuste do modelo, e determinado o ponto de corte, é importante avaliar o poder de discriminação do modelo. Para essa avaliação, métricas foram criadas a partir da matriz de confusão (Tabela 3).

Tabela 3 – Matriz de Confusão

		Valor observado		
		Y=1	Y=0	
Valor predito	Y=1	Verdadeiro positivo (VP)	Falso positivo (FP)	VP + FP
	Y=0	Falso negativo (FN)	Verdadeiro negativo (VN)	FN + VN
Total		VP + FN	FP + VN	N

Com base na Tabela 3, pode-se definir cinco métricas usadas para avaliar o poder de discriminação do modelo:

- **Sensibilidade:** é a proporção dos VP, ou seja, avalia a capacidade do modelo classificar um indivíduo  $\hat{Y} = 1$  dado que realmente é  $Y = 1$

$$SENS = \frac{VP}{VP + FN}.$$

- **Especificidade:** é a proporção dos VN, ou seja, avalia a capacidade do modelo classificar um indivíduo  $\hat{Y} = 0$  dado que realmente é  $Y = 0$

$$ESPEC = \frac{VN}{FP + VN}.$$

- **Valor preditivo positivo:** é a proporção de VP em relação a todas as predições positivas

$$VP+ = \frac{VP}{VP + FP}.$$

- **Valor preditivo negativo:** é a proporção de VN em relação a todas predições negativas

$$VP- = \frac{VN}{FN + VN}.$$

- **Acurácia:** refere-se ao grau em que o teste ou uma estimativa baseada em um teste é capaz de determinar o verdadeiro valor do que está sendo medido.

$$ACC = \frac{VP + VN}{N}.$$



## 4 Material e Método

Neste capítulo será abordada a caracterização da pesquisa, abrangendo o tipo de pesquisa escolhida, a caracterização dos participantes, os critérios de inclusão e exclusão e a modelagem logística binomial.

### 4.1 Caracterização da pesquisa

O presente trabalho está vinculado a Universidade Federal Rural de Pernambuco (UFRPE), com parceria Instituto Federal de Pernambuco (IFPE) e Universidade de Pernambuco (UPE). O mesmo faz parte do projeto de pesquisa intitulado: “Potencialidade do uso da água, manejo florestal e suas implicações na qualidade de vida das mulheres, adultas e idosas de comunidades do Sertão do Pajeú/PE”. aprovado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)- nº463077/2014-8. Trata-se de um estudo do tipo corte transversal, descritivo, conceitual e analítico. Realizado no Departamento de Estatística e Informática (DEINFO), ao Programa de Pós-Graduação em Biometria e Estatística Aplicada (PPGBEA) da UFRPE.

### 4.2 Localização da área

O espaço selecionado para o estudo da QS foi realizado junto a comunidades do Sertão do Pajeú-PE, localizado no município de Serra Talhada, situado na microrregião do Pajeú, no Sertão do Estado de Pernambuco.

A área de estudo está localizada nas comunidades de Serra Grande (Figura 8) - Poço do Serrote (Figura 9), Poldrinho (Figura 10), Catolé (Figura 11) e Três Irmãos (Figura 12), sendo este próximo à Serra da Lagartixa, pertencente ao limite municipal entre as cidades de Serra Talhada e Floresta, região do alto Sertão do Pajeú, ambientes semiáridos do Estado de Pernambuco, de coordenadas geográficas de  $-38^{\circ}23'55''51$  longitude Oeste e  $-8^{\circ}07'06''72$  latitude Sul.

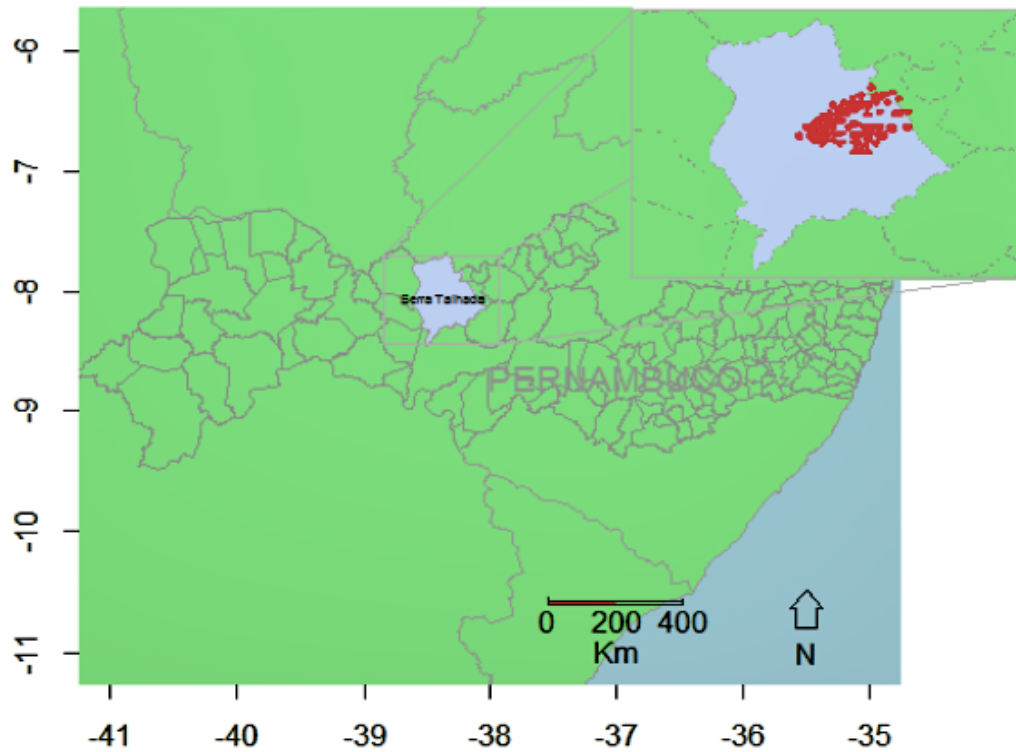


Figura 8 – Localização Geográfica do município de Serra Talhada–PE

#### 4.2.1 Comunidades de Serra Grande

Os dados foram coletados nas comunidades de Serra Grande em parceria com os extensionistas da Universidade de Pernambuco (UPE), do Instituto Federal de Pernambuco (IFPE) e da Universidade Federal Rural de Pernambuco conforme as figuras abaixo.



Figura 9 – Comunidade Poço do Serrote. Sertão do Pajeú, PE, Brasil, 2017.



Figura 10 – Comunidade Poldrinho. Sertão do Pajeú, PE, Brasil, 2017.



Figura 11 – Comunidade Catolé.  
Sertão do Pajeú, PE,  
Brasil, 2017.



Figura 12 – Comunidade Três  
Irmãos. Sertão do  
Pajeú, PE, Brasil,  
2017.

Dentre as comunidades envolvidas, a comunidade Catolé é um dos mais importantes núcleos produtivos da cidade de Serra Talhada, contando com expressiva concentração populacional e, proporcionalmente, com alta produtividade na agricultura familiar local e na agropecuária, localizada a 45 km do município de Serra Talhada, possui uma área total de 740 habitantes.

Em ambas as comunidades as atividades são definidas com base no gênero. Segundo os moradores(as), o gênero masculino trabalham nas atividades consideradas mais “pesadas”, onde estão envolvidos na implementação da agricultura e pastoreio do gado. Já o Feminino é responsável pelo cuidado dos caprinos e ovinos, trabalham na agricultura realizando pequenos tratamentos culturais e colheita, além das atividades domésticas.

### 4.3 Caracterização dos participantes

Antes da aplicação dos questionários, os moradores(as) selecionados leram o Termo de Confidencialidade (Apêndice A), onde comprometemo-nos a manter em anonimato, sob sigilo absoluto, durante e após o término do estudo, todos os dados que identifiquem o sujeito da pesquisa, usando apenas para divulgação dos dados inerentes ao desenvolvimento do estudo. Para obtenção dos dados foram aplicados a Ficha de Avaliação (Apêndice B) e o instrumento de pesquisa IQSP (Anexo A).

A Ficha de Avaliação é composta por questões objetivas de fácil compreensão, elaborado pelo próprio pesquisador, com finalidade de classificar a amostra de acordo com dados sócio-demográficos. O IQSP é um instrumento que avalia a qualidade subjetiva

do padrão do sono referente aos seus hábitos de sono durante o mês passado. Possui 10 perguntas, sendo que as questões de 1 a 4 são subjetivas, enquanto que as de 5 a 10 são objetivas, havendo espaço para registro de comentários, se caso de necessidade. As questões estão distribuídas em 7 componentes: 1) qualidade subjetiva, 2) latência, 3) duração, 4) eficiência habitual do sono, 5) distúrbios do sono, 6) uso de medicações para dormir e 7) disfunção durante o dia (TOMFOHR et al., 2013; COLE et al., 2006). Cada um dos componentes recebe uma pontuação que varia de 0 a 3. Ao final, as 7 componentes são somadas fornecendo o IQSP, que pode variar de 0 a 21 . Pontuação de 0-4 indica boa QS, de 5-10 indicam QS ruim e  $>10$  indicam distúrbio do sono.

Para compor a amostra foi utilizado o processo de amostragem aleatória simples (AAS), onde participaram desta pesquisa 73 moradores(as) dos quais 17 eram da comunidade de Poço do Serrote, 23 da comunidade Catolé, 17 da comunidade Três Irmãos e 16 da comunidade Poldrinho.

## 4.4 Critérios de inclusão e exclusão

### 4.4.1 Critérios de inclusão

- As questões devem ser respondidas em relação a mês passado;
- Ser morador de uma das comunidades.

### 4.4.2 Critério de exclusão

- Ausência durante qualquer etapa da realização do estudo.

## 4.5 Modelo logístico binomial

A fim de modelar a QS nas comunidades de Poço do Serrote, Catolé, Três Irmãos e Poldrinho pertencentes ao Sertão do Pajeú/PE, foi proposto o modelo de regressão logística binomial. Dessa forma temos que os dados foram organizados como “Boa” QS, respostas correspondente a 0-4, e “Ruim”, respostas correspondentes  $\geq 5$ . A proposta se deu ao fato da pesquisa apresentar um pequeno percentual entre algumas respostas, inviabilizando o uso dos dados de forma individual, sendo necessário o agrupamento. A VR foi assim definida:  $Y=1$ , se os moradores têm uma boa QS, e  $Y=0$  caso contrário (Anexo B).

O questionário aplicado (Apêndice B), em alguns momentos, deixou os moradores(as) constrangidos em responder algumas perguntas, devido a sua natureza pessoal. O

problema de trabalhar com os voluntários que responderam a todas as perguntas é o viés das estimativas, pois desconsideram os casos em que os voluntários não responderam todas as questões.

Foi realizada uma análise descritiva e a apresentação das variáveis mensuradas foi feita através de tabelas ou gráficos. Para análise bivariada, foi aplicado o Teste Qui-Quadrado ou exato de Fisher, quando necessário, com o objetivo de encontrar a dispersão entre duas variáveis categóricas. Todas as conclusões foram tomadas ao nível de significância de 5%.

A VR se comporta seguindo a distribuição Bernoulli ( $Y \sim \text{Ber}(\pi)$ ), ou seja, uma variável binária, que é da classe da família exponencial. Assim, os modelos ajustados são uma aplicação do modelo de regressão logística, sendo considerada a função de ligação logit:

$$Y_i = g(x) + \epsilon_i = \frac{\log(\pi(x))}{1-\pi(x)} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \epsilon_i$$

$$E(Y|x) = \pi(x) = \frac{e^{g(x)}}{1+e^{g(x)}}$$

onde  $Y_i \sim B(\pi, n_i)$  corresponde a  $i$ -ésima observação da variável resposta, com suas respectivas variáveis explicativas  $X = (X_{1i}, X_{2i}, \dots, X_{pi})$  e  $\pi$  é a chance de sucesso da  $i$ -ésima variável, (AGRESTI, 2007).

Após a formulação e ajuste do modelo  $\alpha = [0,15; 0,20]$  (Apêndice C), se faz necessário observar se existem falhas no ajuste do modelo proposto, por exemplo, multicolinearidade e heterocedasticidade. Dessa forma, foram realizadas a análise dos resíduos e as medidas de análise de diagnóstico da qualidade do ajuste.

Após estudar a bondade do ajuste, se faz necessário observar o poder de predição do modelo. Essa predição é mensurada a partir do gráfico da Curva ROC, representando a taxa das estimativas verdadeiras que são positivas versus a taxa das estimativas falsas que são positivas.

Uma vez que o modelo proposto seja validado, a interpretação do modelo se dá em função da Razão de Chance. Segundo Agresti (2007), ela pode ser entendida como uma medida de associação entre tabelas de contingência que considera a probabilidade do sucesso  $\pi$ , variando entre 0 e  $\infty$ . Temos que a OR é definida por:

$$g(x) = \frac{\pi(x)}{[1-\pi(x)]} = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}} = e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}$$

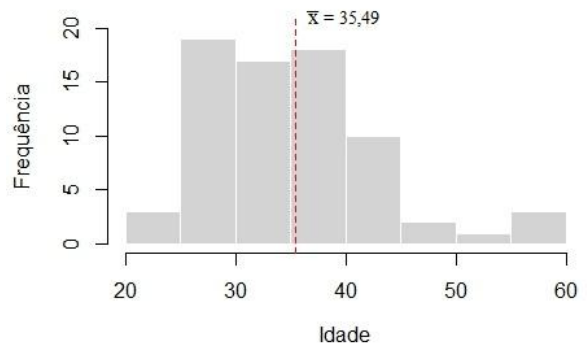
Para a análise estatística dos dados, foram utilizados os softwares estatísticos Microsoft Office Excel 2007, RStudio (RStudio Team, 2015) e IBM SPSS Statistics 20 (SPSS, 2011).

## 5 Resultados e Discussão

Dos 120 moradores contatados, 20 recusaram a participação e 27 tiveram dados incompletos, resumindo-se a 73 questionários analisados. A amostra caracterizou-se por discreto predomínio do sexo feminino (63,0 %). A idade variou de 20 a 60 anos, com média 35,49 anos (DP = 7,734) e mediana de 35 (Figura 13). Foram mais frequentes os moradores solteiros (63,0%), a maioria alegaram ser preto (78,1 %) e recebiam uma renda familiar de 1 salário mínimo (91,8 %), conforme a Tabela 4.

Tabela 4 – Caracterização da amostra, de acordo com as respostas dos dados sociodemográficos. Sertão do Pajeú, PE, Brasil, 2017.

Dados	Frequência	%
<b>Gênero</b>		
Masculino	27	37,0
Feminino	46	63,0
<b>Estado civil</b>		
Solteiro(a)	20	27,4
Casado(a)	53	72,6
<b>Cor/Raça</b>		
Branco	9	12,3
Preto	57	78,1
Pardo	7	9,6
<b>Renda Familiar</b>		
1 salário mínimo	67	91,8
2 salários mínimos	5	6,8
3 salários mínimos	1	1,4



A Tabela 5 refere-se à análise dos fatores que contribuíram para os componentes transtornos do sono, destacando-se dentre eles a necessidade de levantar-se para ir ao banheiro (56,3%) e o despertar no meio da noite ou de manhã muito cedo (54,8%) por menos de uma vez por semana. Diferente dos resultados encontrados por Mansano-Schlosser e Ceolim (2012), com valores de 83,6% e 47,9%, respectivamente, para três vezes por semana ou mais.

Tabela 5 – Frequência e porcentagem da ocorrência dos fatores que contribuíram para o componente transtornos do sono. Sertão do Pajeú, PE, Brasil, 2017.

Questões	0		1		2		3	
	N	%	N	%	N	%	N	%
5a) Não conseguiu adormecer em até 30 min.	51	69,9	18	24,7	4	5,5	0	0,0
5b) Acordou no meio da noit./de manhã cedo	26	35,6	40	54,8	0	0,0	7	9,6
5c) Levantou para ir ao banheiro	25	34,2	41	56,3	0	0,0	7	9,6
5d) Teve dificuldade para respirar	64	87,7	3	4,1	4	5,5	2	2,7
5e) Tossiu ou roncou muito alto	62	84,9	8	11,0	0	0,0	3	4,1
5f) Sentiu muito frio	65	89,0	8	11,0	0	0,0	0	0,0
5g) Sentiu muito calor	52	78,1	21	28,8	0	0,0	0	0,0
5h) Teve sonhos ruins ou pesadelos	57	78,1	11	15,1	1	1,4	4	5,5
5i) Teve dor	55	75,3	15	20,5	0	0,0	3	4,1

0 = Nenhuma vez; 1 = Menos de 1x/semana; 2 = 1 ou 2x/semana;  
3 = 3x/semana ou mais.

Segundo a National Sleep Foundation, adormecer em até 30 minutos e acordar no máximo uma vez de madrugada são alguns dos parâmetros para um sono de qualidade. Analisando esse dois parâmetros em nosso estudo pode-se observar na Tabela 6 que 90,6 % moradores informaram que adormecem em menos de 30 min e 90,4% acordam no máximo 1x/semana ou nenhuma (Tabela 5 - Questão 5.b) corroborando com a National Sleep Foundation.

Tabela 6 – Frequências e porcentagens da variável resposta do IQSP agrupada de forma dicotômica. Sertão do Pajeú, PE, Brasil, 2017.

Questão 2	Frequência	%
≤ 30 min	66	90,6
>30 min	7	9,4
Total	73	100

A Tabela 7 expõe a frequência e porcentagem das respostas dos moradores na questão de número 10 do IQSP. No que se refere à QS, 52% relataram que o sono é uma necessidade e que o ato de cochilar após as refeições também é uma necessidade (70,9%). Nos resultados obtidos por Barichello et al. (2009), ocorreu um inversão nas respostas, sendo o ato de cochilar um prazer (89,1%). Vale ressaltar que a ocorrência de cochilos, pode ser associadas à sonolência diurna, foi relatada por 55% dos moradores. Em contrapartida, Floyd (2002) afirma que cochilos longos tendem a ser prejudiciais, associando-se a maiores períodos de latência e despertares noturnos mais freqüentes.



Tabela 7 – Frequência e porcentagem das respostas dos moradores na questão de número 10 do IQSP. Sertão do Pajeú, PE, Brasil, 2017.

Questões	Frequência	%
<b>Para você o sono é</b>		
Um prazer	27	37,0
Uma necessidade	38	52,0
Hábito	8	11,0
<b>Você cochila</b>		
Sim	48	65,8
Não	25	34,2
<b>Caso sim, intencionalmente</b>		
Sim	16	21,93
Não	32	43,87
<b>Para você cochilar é</b>		
Um prazer	14	29,1
Uma necessidade	34	70,9

A consistência interna do IQSP revelou-se baixa (alfa de Cronbach = 0,451) para essa amostra, resultado corroborado pelo trabalho de Araújo et al. (2008), que encontrou um valor de 0,438. Estudos têm revelado que um valor baixo de alfa pode significar apenas número pequeno de perguntas e de respondentes (VIEIRA, 2000). Devido a esse resultado optou-se por analisar cada componente separadamente, em função das variáveis do estudo.

Ao avaliar o componente Qualidade subjetiva do sono, observou-se que 98,7% dos moradores classificaram-na como muito boa ou boa. Quanto à Latência do sono, 74,0% referiam conciliar o sono em menos de 30 minutos; entretanto, 24,7% demoravam mais de 30 minutos, menos de uma vez por semana (Tabela 5 - Questão 5.a). A Duração do sono era superior a sete horas por noite para 78,1% dos moradores. A Eficiência do sono, segundo 75,3% dos moradores, era superior a 85%. Quanto ao Uso de medicação para dormir, prescrita ou não pelo médico, 87,7% dos moradores informaram que não utilizavam. A ocorrência de Sonolência diurna e distúrbios durante o dia foram negados por 72,6% dos moradores; destaca-se, no entanto, que 26,0% relataram um pequena indisposição ou falta de entusiasmo para realizar suas atividades diárias (Anexo A - Questão 9), são apresentados na Tabela 8.

Tabela 8 – Frequências e porcentagens em relação aos sete componentes do IQSP. Sertão do Pajeú, PE, Brasil, 2017.

Componentes	0		1		2		3	
	N	%	N	%	N	%	N	%
1) Qualidade subjetiva do sono	14	19,2	58	79,5	1	1,4	0	0,0
2) Latência do sono	37	50,7	17	23,3	12	16,4	7	9,6
3) Duração do sono	57	78,1	8	11,0	6	8,2	2	2,7
4) Eficiência habitual do sono	55	75,3	7	9,6	5	6,8	6	8,2
5) Distúrbios do sono	6	8,2	62	84,9	5	6,8	0	0,0
6) Uso de medicação para dormir	64	87,7	9	12,3	0	0,0	0	0,0
7) Sonolência diurna e distúrbios durante o dia	53	72,6	20	27,4	0	0,0	0	0,0

Verificou-se que 68,5% dos moradores apresentavam uma boa QS (Tabela 9), de acordo com a classificação obtida no IQSP, resultado 5% inferior foi observado por, Rahe et al. (2015), com 63,5% para boa QS. Entretanto, nas respostas ao componente qualidade subjetiva do sono (Tabela 8), 98,7% de todos os moradores responderam que essa era boa (1) ou muito boa (0).

Tabela 9 – Frequências e porcentagens da variável resposta do IQSP. Sertão do Pajeú, PE, Brasil, 2017.

Pontuação	Frequência	%
0 a 4	50	68,5
5 a 10	21	28,8
>10	2	2,7
Total	73	100

A partir da Tabela 9, pode-se notar que pontuação >10 (distúrbios do sono) apresentou um baixo percentual (2,7%) inviabilizando o uso dos dados de forma individual, sendo necessário o agrupamento com a categoria 5 a 10 (QS Ruim) obtendo-se a Tabela 10.

Tabela 10 – Frequências e porcentagens da variável resposta do IQSP agrupada de forma dicotômica. Sertão do Pajeú, PE, Brasil, 2017.

Pontuação	Frequência	%
0 a 4	50	68,5
$\geq 5$	23	31,5
Total	73	100

Avaliando as variáveis socioeconômicas (Figura 14), verificou-se que as respostas eram 100% sim ou 100% não para diferentes perguntas. Deste modo, não permite estabelecer relações entre as variáveis socioeconômicas com a VR (Análise Bivariada), isto é, determinar

se as diferenças entre a distribuição de duas variáveis são estatisticamente significativas. Além, de serem eliminadas na primeira seleção de variáveis.

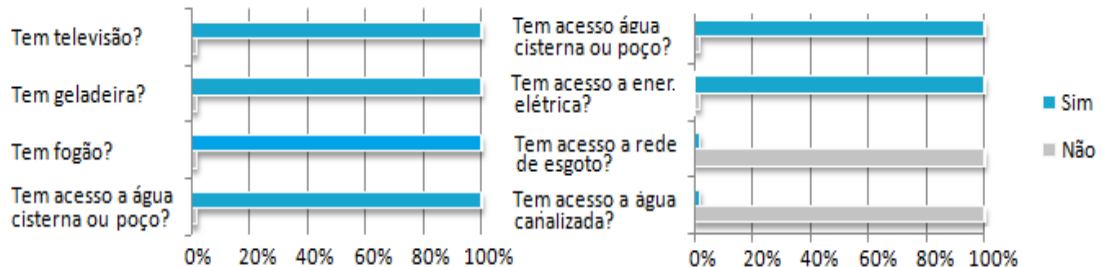


Figura 14 – Variáveis sociodemográficas. Sertão do Pajeú, PE, Brasil, 2017.

## 5.1 Análise bivariada – QS

A fim de verificar se existe associação entre a variável resposta QS e as variáveis explicativas inicialmente foram realizadas análises bivariadas. Essa por sua vez tem como objetivo aplicar o teste  $\chi^2$  de independência de Pearson ou exato de Fisher, quando necessário, em tabelas de contingência,  $2 \times 2$ , comparando as distribuições das variáveis explicativas nas duas definições das variáveis respostas (QS ruim e Boa).

Observou-se na Tabela 11 associação estatisticamente significativa entre a QS e Gênero ( $p=0,019$ ), sendo o número de mulheres (72,0%) estatisticamente superior no grupo boa QS. Pascotto e Santos (2013), encontraram resultados distintos, em relação às mesmas variáveis; entre a QS e Idade ( $p=0,030$ ), havendo uma maior proporção na categoria acima da mediana (65,2%), apresentando uma QS ruim. Reclamações de sono constituem uma das dificuldades mais comuns que enfrentam adultos de meia idade e idosos (KING et al., 1997).

Tabela 11 – Distribuição de frequência entre a variável QS (Y) e as variáveis sociodemográficas (X's). Sertão do Pajeú, PE, Brasil, 2017.

X's \ Y	Ruim		Boa		Valor-p
	N	%	N	%	
<b>Gênero</b>					
Masculino	13	56,5	14	28,0	
Feminino	10	43,5	36	72,0	0,019
<b>Idade</b>					
<Mediana(36)	8	34,8	31	62,0	
≥Mediana(36)	15	65,2	19	38,0	0,030
<b>Estado civil</b>					
Solteiro(a)	8	34,8	12	24,0	
Casado(a)	15	65,2	38	76,0	0,337
<b>Cor/Raça</b>					
Branca	4	17,4	5	10,0	
Preta	16	69,6	41	82,0	
Parda	3	13,0	4	8,0	0,441
<b>Renda familiar</b>					
1 salário mínimo	21	91,3	46	92,0	
2 salários mínimos	2	8,7	3	6,0	
3 salários mínimos	0	0,0	1	2,0	0,761
<b>Cochila</b>					
Não	12	52,2	14	28,0	
Sim	11	47,8	36	72,0	0,048

Ao analisar a QS de acordo com o estado civil, observou-se a não associação estatisticamente significativa ( $p = 0,337$ ), havendo maiores prevalências na categoria casado para ruim e boa QS (65,2 % e 76,0%). Pesquisa realizada por Martini et al. (2012) apresentou resultados similares ao nosso estudo, em que não houve associação significativa entre as variáveis; e entre QS e Cochila ( $p = 0,048$ ), em que se constata que a maior parte desses moradores(as) investigados (72,0%) possuem uma boa qualidade do sono (Risco Relativo (RR): 1.972; IC 95% 1,0160 – 3,8278), em que há um maior risco do grupo QS ruim ao grupo boa QS. Cardoso et al. (2009) apresentaram resultados similares (RR) aos de nossa pesquisa, em seu estudo realizado com estudantes de Medicina (RR: 1,108; IC 95% 1,0047 – 1,2240).

Dentre as variáveis que foram comparadas, ou seja, cor/raça e renda família mensal per capita, nenhuma delas apresentou diferença significativa entre os dois grupos. Rossi et al. (2017), obteve resultados similares. No grupo QS ruim, 91,3% apresentavam uma renda familiar de 1 salário mínimo e 65,2% eram casados, enquanto no grupo QS boa, 76,0% eram solteiros.

Em relação à Tabela 12, os componentes que mais interferiram na qualidade de sono ruim foram a latência ( $p < 0,001$ ), duração do sono ( $p = <0,001$ ) e eficiência habitual do sono ( $p = <0,001$ ). Segundo Pascotto e Santos (2013), com relação à avaliação da qualidade do sono em estudantes de ciências da saúde, constatou que o único componente que não interferiu na QS ruim foi a eficiência habitual do sono ( $p = <0,142$ ).

Tabela 12 – Componentes do IQSP de acordo com os dois grupos de qualidade do sono. Sertão do Pajeú, PE, Brasil, 2017.

Componentes do PSQI	Ruim		Boa		Valor-p
	N	%	N	%	
<b>Qualidade subjetiva do sono</b>					
Muito boa/Boa	50	100,0	22	95,7	
Ruim/Muito Ruim	-	-	1	4,3	0,357
<b>Latência do sono</b>					
≤ 15 min / 16 – 30 min	44	88,0	10	43,5	
≤ 31 – 60 min / > 60 min	6	12,0	13	56,5	<0,001
<b>Duração do sono</b>					
> 7h / 6 – 7h	50	100,0	15		65,2
5 – 6h / < 5h	-	-	8	34,8	<0,001
<b>Eficiência habitual do sono</b>					
> 85% / 75% – 84%	50	100,0	12	52,2	
65% – 74% / < 65%	11	47,8	-	-	<0,001
<b>Distúrbios do sono</b>					
Nenhum/ Menos de 1 vez/semana	48	96,0	20	87,0	
1 a 2 vezes/semana/3 ou mais vezes/semana	2	4,0	3	13,0	0,317
<b>Uso de medicação para dormir</b>					
Nenhum/ Menos de 1 vez/semana	23	100,0	50	100,0	
1 a 2 vezes/semana/3 ou mais vezes/semana	-	-	-	-	-
<b>Sonolência diurna e distúrbios durante o dia</b>					
Nenhuma/Pequena	23	100,0	50	100,0	
Moderada/Muita	-	-	-	-	-

## 5.2 Análise multivariada

Já conhecemos a associação entre a VR com as VP individualmente. A fim de buscar uma função que consiga explicar a VR baseando-se nas VP conjuntamente, foi proposto o modelo logístico multivariado que é um caso particular dos MLG, em que VR assume valores, 0 para QS ruim ( $\geq 5$ ) e 1 para Boa QS (0 a 4), caracterizando essa variável como Bernoulli. Uma vez definido o conjunto de VP a ser inserida num modelo (??), estamos interessados em encontrar um modelo reduzido que inclua apenas as VP mais importantes para explicar a probabilidade de sucesso  $\pi(X)$ .

### 5.2.1 Modelo logístico - QS

A Tabela 13 a seguir apresenta as estimativas dos parâmetros e teste de significância das variáveis selecionadas do modelo de regressão logístico multivariado final para explicar a QS. De acordo com a Tabela 13, com relação a variável gênero, pode-se notar que, fixada as demais variáveis, a chance do morador ter uma Boa QS é 3 vezes maior para o gênero masculino. Já em relação variável estado civil, pode-se notar que, fixada as demais variáveis, a chance do morador ter uma Boa QS é quase 6 vezes maior para classe solteiro neste estudo. O morador que relatou cochilar após o almoço, reduziu a chance em 73,6% de não ter uma boa QS, com relação aos que apresentam dependência. A OR para cada ano de incremento na idade é 0.90, ou seja, a cada ano de idade (20 - 58 anos), há uma redução de 10% no risco para boa QS nesta população.

Tabela 13 – Estimativas dos parâmetros e teste de significância das variáveis selecionadas do modelo de regressão logístico multivariado final para explicar a QS. Sertão do Pajeú, PE, Brasil, 2017.

<b>Coefficientes</b>					
	<b>Estimativa</b>	<b>Erro padrão</b>	<b>Wald</b>	<b>Valor-p</b>	<b>OR</b>
(Intercepto)	2,904	1,475	3,875	0,049 *	18,25
Gênero(1)	1,163	0,582	3,987	0,046 *	3,199
Idade	-0.099	0,043	5,172	0,023 *	0,90
Estado civil(1)	1,791	0,766	5,467	0,019 *	5,99
Cochila(1)	-1,334	0,621	4,605	0,032 *	0,264
---					
Signif. codes:	0 *** 0,001	0,001 ** 0,01	0,01 * 0,05	0,05 . 0,1	
Null devia.	Resid. devia.	AIC	Núm. inter.	Disper. gaus.	(1)
90,97	74,63	84,63	4	1	prim. categ.

Estudos evidenciam que existe uma tendência ao aumento da prevalência de insônia com o envelhecimento, sendo que a remissão desse quadro é pouco provável com o avanço da idade (PHILLIPS; ANCOLI-ISRAEL, 2001). Acredita-se então, que a necessidade de sono pode não mudar com a idade, mas a capacidade de dormir o necessário diminui (NEIKRUG; ANCOLI-ISRAEL, 2010). Assim, este conceito geral sobre mudanças do hipnograma com o envelhecimento deve ser aceito com ressalvas (FERNANDES, 2006).

Na Tabela 14, são apresentados os intervalos de 95% de confiança para os parâmetros do modelo final, com base na estatística de Wald. Podemos notar que o valor zero não está contido nos intervalos de confiança abaixo, ou seja, nós não conseguimos rejeitar a hipótese nula de que um coeficiente de regressão particular é zero, dado que os outros preditores estão no modelo.

Tabela 14 – Intervalos de 95% de confiança para os parâmetros do modelo final, com base na estatística de Wald. Sertão do Pajeú, PE, Brasil, 2017.

	<b>IC 95%</b>	
	<i>Lim<sub>inf</sub></i>	<i>Lim<sub>sup</sub></i>
Intercepto	0,013	5,795
Gênero	0,021	2,304
Idade	-0,183	-0,013
Estado civil	0,289	3,291
Cochila	-2,551	-0,115

*Lim<sub>inf</sub>* = Limite inferior;  
*Lim<sub>sup</sub>* = Limite superior

A Tabela 15 a seguir apresenta os intervalos de confiança para a OR no modelo final. Nesta pode-se notar que, o intervalo de confiança não incluiu o 1, a variável idade e cochila é um fator de risco para Boa QS nesta população.

Tabela 15 – Intervalos de confiança para as razões de chance (Odds ratio – OR) do modelo final, fixando o nível de confiança de 95%. Sertão do Pajeú, PE, Brasil, 2017.

	<b>IC 95%</b>	
	<i>Lim<sub>inf</sub></i>	<i>Lim<sub>sup</sub></i>
Gênero	1,022	10,018
Idade	0,832	0,986
Estado civil	1,336	26,884
Cochila	0,078	0,891

*Lim<sub>inf</sub>* = Limite inferior;  
*Lim<sub>sup</sub>* = Limite superior

Um dos pressupostos do modelo logístico é que só podemos inserir uma variável contínua no modelo se houver evidência de linearidade, ou seja, de que para cada incremento desta variável, corresponda um decréscimo ou acréscimo na probabilidade predita do evento. Na Figura 15 pode-se verificar que existe uma tendência linear decrescente da probabilidade predita da QS pelo modelo na medida em que aumenta a idade. Portanto, a idade pode ser modelada como uma variável quantitativa contínua no modelo.

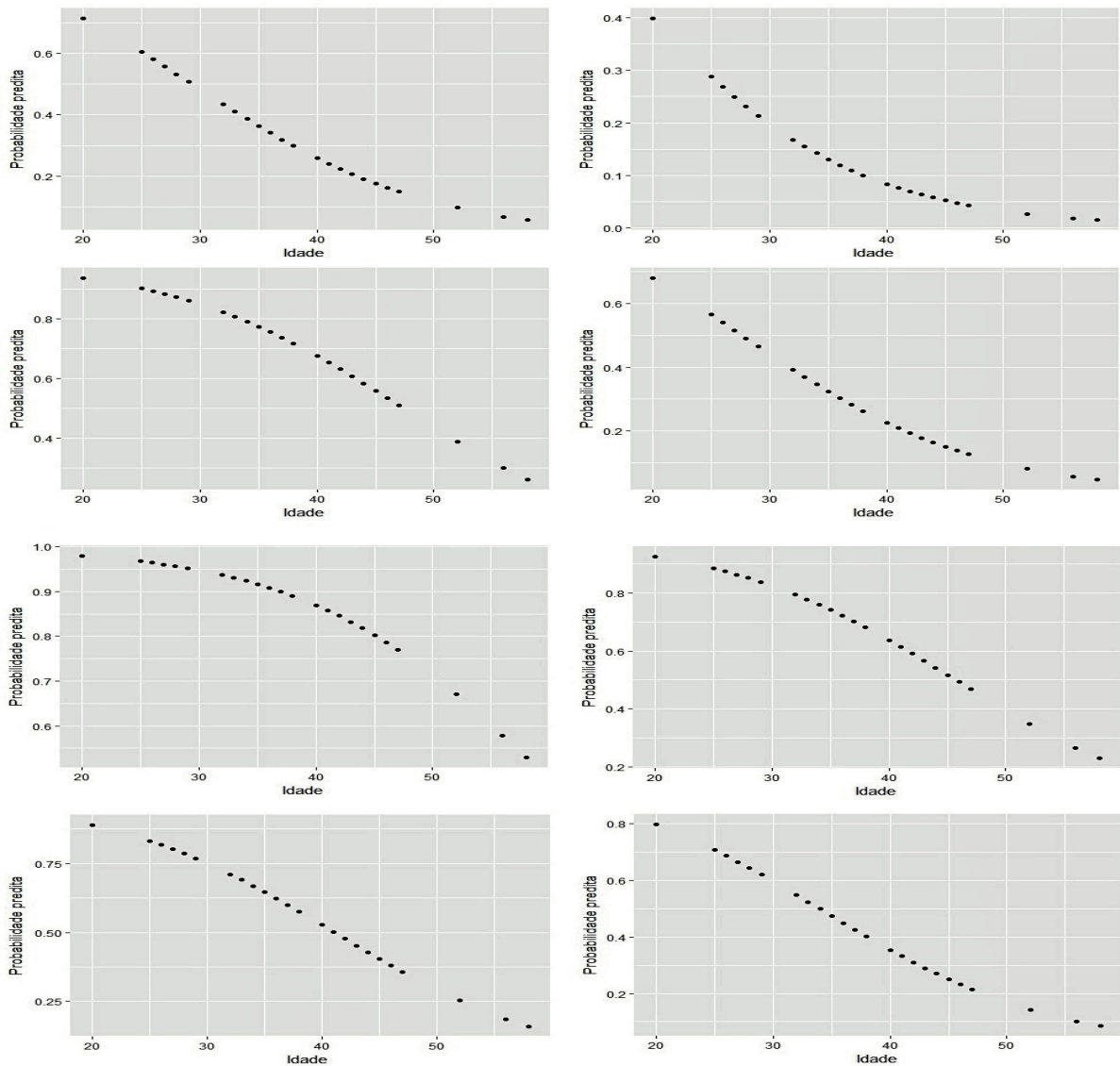


Figura 15 – Idade versus probabilidade predita da QS. Sertão do Pajeú, PE, Brasil, 2017.

### 5.2.2 Análise de resíduos e técnicas de diagnóstico - QS

Na Figura 16, os valores da medida  $h$  aparecem plotados versus os valores ajustados. Analisando este gráfico, pode-se perceber que os pontos 24, 39 e 46 se distanciam dos demais pontos, destacando-se. Por isso, temos que essas observações são consideradas outliers. Quanto ao gráfico de influência Figura 17, observa-se que todos os pontos se encontram próximos e abaixo de  $D_i < 1$ , indicando que não se tem de fato pontos aberrantes. Isto nos leva a crer que o modelo está bem ajustado.



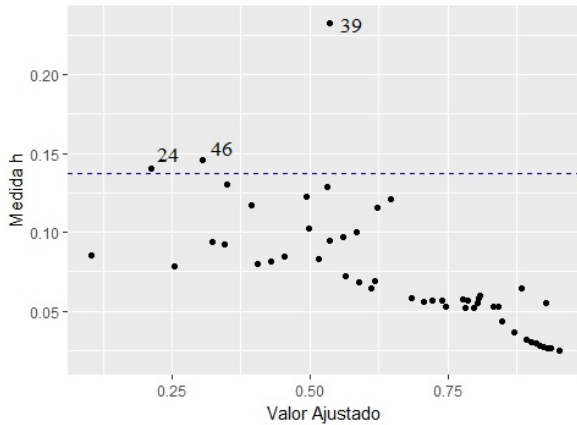


Figura 16 – Valores de  $h$  versus valores ajustados do modelo QS.

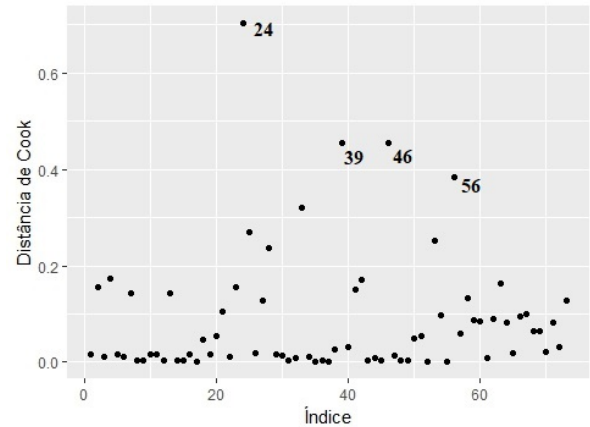


Figura 17 – Distância de Cook referente ao modelo QS.

Analisando os resíduos a fim de validar o modelo final ajustado, segundo a Figura 18, os resíduos apresentam bom comportamento aleatório, concentrando-se entre os valores 0 e 1, com duas observação se distanciando dos limites. Quanto ao gráfico dos quantis da Normal para distribuição Binomial com envelope para QS (Figura 19), observa-se que a suposição de que os erros são binomiais foi validada, uma vez que nenhum resíduo estimado ultrapassou o envelope. Além disso, observa-se que o modelo conseguiu captar o comportamento estudado.

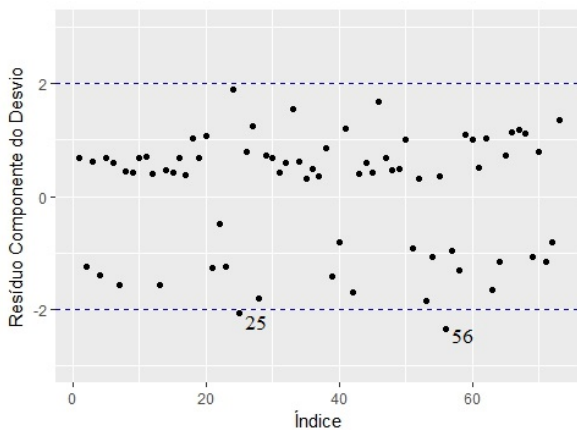


Figura 18 – Resíduo componente do desvio versus o índice para o modelo QS.

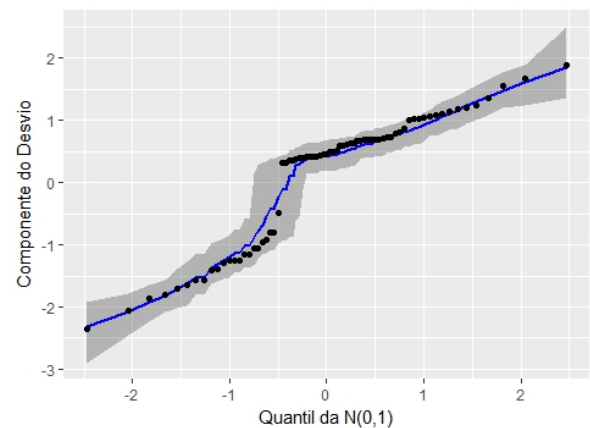


Figura 19 – Quantis da Normal para distribuição Binomial com envelope para QS.

Vale ressaltar que diferentes autores a partir de diferentes pontos de corte definem quais observações são pontos de alavanca ou de influência. Por isso, não é possível

simplesmente excluir tais observações da análise (24, 39 e 46). Segundo Cordeiro e Neto (2004) se um ponto de grande alavancagem for também influente ele não pode ser removido da análise, pois causará mudanças apreciáveis nas estimativas dos  $\beta$ 's. Desse modo, deve-se estar familiarizado com a “natureza” dos dados de maneira a conhecer qual o comportamento dos diversos elementos na população e, quando isso não ocorrer, estar preparado para fazer uma análise mais detalhada da situação para decidir o quão importante é manter ou eliminar uma observação, ou conjunto de observações de determinado padrão ou modificar o modelo de regressão utilizado, de maneira que ele considere essas características dos dados tornando-o mais “eficiente”.

Segundo Paula (2004), a razão de encontrar pontos com comportamentos atípicos em relação aos demais não implica, necessariamente, na eliminação de tais pontos. Dessa forma, este deve ser o último recurso a ser utilizado. Outras soluções seriam: mudanças na distribuição postulada para a variável resposta, inclusão, eliminação ou mesmo transformação de variáveis explicativas ou uso de métodos robustos (VENABLES; RIPLEY, 2013).

### 5.2.3 Predição - QS

Avaliando a curva ROC para o modelo de regressão logística, apresentada na Figura 20, verifica-se que a área sob a curva corresponde ao valor de 0,79, demonstrando um aceitável poder de discriminação (Anexo C) de acordo com a classificação dada por Hosmer e Lemeshow (2000). Este resultado pode ser justificado quando analisado em conjunto com o cálculo de adequação do modelo.

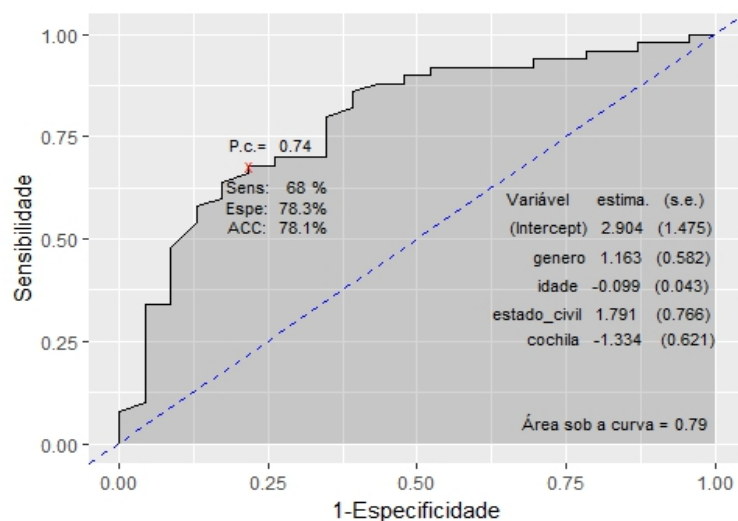


Figura 20 – Curva ROC do modelo QS. Sertão do Pajeú, PE, Brasil, 2017.

## 6 Conclusão

Visto que o desempenho dos moradores na agricultura familiar e agropecuária está fortemente relacionado com sua QS foi de fundamental importância avaliar e modelar a QS em algumas comunidades envolvidas. Deste modo, Martins, Mello e Tufik (2001) apontam que um dos fatores que contribuem para QS é o fator exercício físico, o qual facilita o sono por aumentar o gasto energético durante a vigília e restabeleça a condição adequada para um novo episódio de vigília, sendo esta uma variável fixada neste estudo.

A partir do instrumento IQSP foi observado que os moradores apresentavam sono regular e contínuo, com uma média de aproximadamente oito horas por noite. Porém com necessidade de levantar-se para ir ao banheiro (56,3%) e despertar no meio da noite ou de manhã muito cedo (54,8%) por menos de uma vez por semana. Os agricultores relataram que acordam entre 4h00 e 7h30 e dormem pouco depois das 20h00. A intensa atividade física do dia contribui para um sono profundo à noite, onde 90,4 % dos moradores relatam adormecer em menos 30 minutos.

A partir do modelo proposto, foi constatado que as variáveis, gênero, estado civil, idade e cochilar são possíveis fatores explicativos da qualidade do sono. Com relação à qualidade do sono, pode-se observar que as variáveis que merecem maior atenção são a variável idade e cochilar, uma vez que essas são as únicas variáveis que têm o poder de reduzir a qualidade do sono. Uma vez que fixada as demais, a variável idade tem o poder de reduzir a qualidade do sono em 10%. E fixada as demais, a variável cochilar tem o poder de reduzir em 73,6% a qualidade do sono.

É importante ainda ressaltar que o modelo não apresentou indícios de estar mal ajustado, sem efeito de heterocedasticidade, ou seja, a suposição das variâncias serem constantes em relação ao tempo  $t$  foi satisfeita e sem efeito de multicolinearidade, onde não foi verificada existência de uma perfeita (ou exata) relação linear entre algumas ou todas as variáveis explicativas de um modelo de regressão. Além, do modelo apresentar um aceitável poder de discriminação (79%).

Para trabalhos futuros sugere-se verificar o ajuste do modelo para diferentes cenários (Regiões) adicionando a estes novos fatores que influenciam diretamente ou indiretamente na QS. Por outro lado, encontrar fatores não identificados na literatura que possam ser possíveis fatores de risco ou proteção na população.

## Referências Bibliográficas

- AGRESTI, A. Logistic regression. **An Introduction to Categorical Data Analysis, Second Edition**, Wiley Online Library, p. 99–136, 2007.
- ARAÚJO, C. L. d. O. et al. Qualidade do sono de idosos residentes em instituição de longa permanência. [sn], 2008.
- BARBA, M. E. F. d. et al. Validation of the brazilian portuguese version of the pittsburgh sleep quality index. **Revista HCPA. Porto Alegre**, 2011.
- BARICHELO, E. et al. Qualidade do sono em pacientes submetidos à cirurgia oncológica. **Revista Latino-Americana de Enfermagem**, Universidade de São Paulo, v. 17, n. 4, 2009.
- BedTime. **Quais são as fases do sono?** 2017. (<https://www.bedtime.com.br/blog/curiosidade-do-sono/quais-sao-as-fases-do-sono/>). Acessado em 15/06/2017.
- BUYSSE, D. J. et al. The pittsburgh sleep quality index: a new instrument for psychiatric practice and research. **Psychiatry research**, Elsevier, v. 28, n. 2, p. 193–213, 1989.
- CARDOSO, H. C. et al. Avaliação da qualidade do sono em estudantes de medicina. **Rev. bras. educ. méd**, v. 33, n. 3, p. 349–355, 2009.
- COLE, J. C. et al. Validation of a 3-factor scoring model for the pittsburgh sleep quality index in older adults. **Sleep**, Oxford University Press, v. 29, n. 1, p. 112–116, 2006.
- COLLET, D. **Modelling binary data**. [S.l.]: Chapman & Hall, 2003.
- COOK, R. D. Influential observations in linear regression. **Journal of the American Statistical Association**, Taylor & Francis, v. 74, n. 365, p. 169–174, 1979.
- CORDEIRO, G. M.; NETO, E. d. A. Modelos paramétricos. **Pernambuco: UFRPE**, 2004.
- DEMÉTRIO, C.; CORDEIRO, G. Modelos lineares generalizados. **SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA**, UFSM Santa Maria, v. 12, 2007.
- DORLAND, W. A. N. **Dorland's Illustrated Medical Dictionary32: Dorland's Illustrated Medical Dictionary**. [S.l.]: Elsevier Health Sciences, 2011.
- FERNANDES, R. M. F. O sono normal. **Medicina (Ribeirao Preto. Online)**, v. 39, n. 2, p. 157–168, 2006.
- FERRARA, M.; GENNARO, L. D. How much sleep do we need? **Sleep medicine reviews**, Elsevier, v. 5, n. 2, p. 155–179, 2001.

- FIGUEIRA, C. V. Modelos de regressão logística. 2006.
- FLOYD, J. A. Sleep and aging. **Nursing Clinics of North America**, Elsevier, v. 37, n. 4, p. 719–731, 2002.
- FONTAINE, D. K.; BRIGGS, L. P.; POPE-SMITH, B. Designing humanistic critical care environments. **Critical Care Nursing Quarterly**, LWW, v. 24, n. 3, p. 21–34, 2001.
- FRIEDMAN, M. J.; SCHNURR, P. P. The relationship between trauma, post-traumatic stress disorder, and physical health. Lippincott Williams & Wilkins Publishers, 1995.
- GOURIEROUX, C.; MONFORT, A. **Statistics and econometric models**. [S.l.]: Cambridge University Press, 1995. v. 1.
- HOAGLIN, D. C.; WELSCH, R. E. The hat matrix in regression and anova. **The American Statistician**, Taylor & Francis, v. 32, n. 1, p. 17–22, 1978.
- HUNSLEY, J.; MASH, E. J. **A guide to assessments that work**. [S.l.]: Oxford University Press, 2008.
- IBGE. **Estimativas da população nos municípios brasileiros**. 2016. (<http://cod.ibge.gov.br/GEZ>). Acessado em 21/07/2017.
- JR, D. W. H.; LEMESHOW, S.; STURDIVANT, R. X. **Applied logistic regression**. [S.l.]: John Wiley & Sons, 2013. v. 398.
- JR, W. W. H.; DONNER, A. Wald's test as applied to hypotheses in logit analysis. **Journal of the American Statistical Association**, Taylor & Francis, v. 72, n. 360a, p. 851–853, 1977.
- KING, A. C. et al. Moderate-intensity exercise and self-rated quality of sleep in older adults: a randomized controlled trial. **Jama**, American Medical Association, v. 277, n. 1, p. 32–37, 1997.
- LEE, K. I.; KOVAL, J. J. Determination of the best significance level in forward stepwise logistic regression. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, v. 26, n. 2, p. 559–575, 1997.
- LUCCHESI, L. M. et al. Sleep in psychiatric disorders. **Revista Brasileira de Psiquiatria**, SciELO Brasil, v. 27, p. 27–32, 2005.
- MANSANO-SCHLOSSER, T. C.; CEOLIM, M. F. Fatores associados à qualidade do sono de idosos submetidos à quimioterapia. **Revista Latino-Americana de Enfermagem**, Universidade de São Paulo, v. 20, n. 6, 2012.
- MARGOTTO, P. R. Curva roc como fazer e interpretar no spss. **Escola Superior de Ciências da Saúde, Distrito Federal**, 2010.
- MARTINI, M. et al. Fatores associados à qualidade do sono em estudantes de fisioterapia. **Fisioterapia e Pesquisa**, SciELO Brasil, v. 19, n. 3, p. 261–267, 2012.

- MARTINS, P. J. F.; MELLO, M. T. d.; TUFIK, S. Exercício e sono. **Revista Brasileira de Medicina do Esporte**, SciELO Brasil, v. 7, n. 1, p. 28–36, 2001.
- MOLLAYEVA, T. et al. The pittsburgh sleep quality index as a screening tool for sleep dysfunction in clinical and non-clinical samples: A systematic review and meta-analysis. **Sleep medicine reviews**, Elsevier, v. 25, p. 52–73, 2016.
- MÜLLER, M. R.; GUIMARÃES, S. S. Impacto dos transtornos do sono sobre o funcionamento diário e a qualidade de vida. **Estudos de psicologia**, Pontifícia Universidade Católica de Campinas, v. 24, n. 4, p. 519–528, 2007.
- NEIKRUG, A. B.; ANCOLI-ISRAEL, S. Sleep disorders in the older adult—a mini-review. **Gerontology**, Karger Publishers, v. 56, n. 2, p. 181–189, 2010.
- NELDER, J. A.; BAKER, R. J. **Generalized linear models**. [S.l.]: Wiley Online Library, 1972.
- NORONHA, N. Afinal, quantas horas devemos dormir? **Sleep Medicine**, 2014.
- PASCOTTO, A. C.; SANTOS, B. R. M. d. Avaliação da qualidade do sono em estudantes de ciências da saúde. **Revista do Instituto de Ciências da Saúde, São Caetano do Sul**, v. 31, n. 3, p. 306–310, 2013.
- PAULA, G. A. **Modelos de regressão: com apoio computacional**. [S.l.]: IME-USP São Paulo, 2004.
- PHILLIPS, B.; ANCOLI-ISRAEL, S. Sleep disorders in the elderly. **Sleep Medicine**, Elsevier, v. 2, n. 2, p. 99–114, 2001.
- PREGIBON, D. Logistic regression diagnostics. **The Annals of Statistics**, JSTOR, p. 705–724, 1981.
- RAHE, C. et al. Associations between poor sleep quality and different measures of obesity. **Sleep medicine**, Elsevier, v. 16, n. 10, p. 1225–1228, 2015.
- ROSSI, G. R. E. et al. Avaliação da qualidade do sono e fatores associados em pacientes diabéticos tipo 2. v. 41, n. 3, p. 350–358, 2017.
- RStudio Team. **RStudio: Integrated Development Environment for R**. Boston, MA, 2015. Disponível em: <http://www.rstudio.com/>.
- RYAN, T. P. **Modern regression methods**. [S.l.]: John Wiley & Sons, 2008. v. 655.
- SILVA, K. K. M. et al. Alterações do sono e a interferência na qualidade de vida no envelhecimento. **Revista de enfermagem UFPE on line-ISSN: 1981-8963**, v. 11, n. 1, p. 422–428, 2017.
- SPSS, I. **Ibm spss statistics base 20**. Chicago, IL: SPSS Inc, 2011.
- TEIXEIRA, R. A. Devemos dormir quantas horas por dia? **Canal Instituto do Cérebro de Brasília (ICB) conhecendo melhor o seu cérebro**, 2008.

TOMFOHR, L. M. et al. Psychometric characteristics of the pittsburgh sleep quality index in english speaking non-hispanic whites and english and spanish speaking hispanics of mexican descent. **Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine**, American Academy of Sleep Medicine, v. 9, n. 1, p. 61, 2013.

VELLUTI, R.; REIMÃO, R. Fisiologia do sono. **Sono: estudos abrangentes, 2<sup>a</sup> ed.** Rio de Janeiro: Atheneu, p. 1–15, 1996.

VENABLES, W. N.; RIPLEY, B. D. **Modern applied statistics with S-PLUS.** [S.l.]: Springer Science & Business Media, 2013.

VIEIRA, S. **Como elaborar questionários.** [S.l.]: Editora Atlas SA, 2000.

VORSTER, A. P.; BORN, J. Sleep and memory in mammals, birds and invertebrates. **Neuroscience & Biobehavioral Reviews**, Elsevier, v. 50, p. 103–119, 2015.

## Apêndices

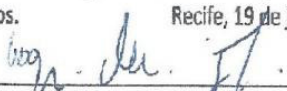



# APÊNDICE A – TERMO DE CONFIDENCIALIDADE

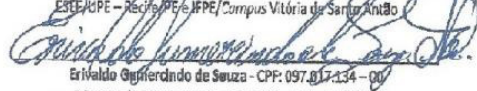
(Elaborado de acordo com a Resolução 466/2012 – CNS/CONEP)


Em referência a pesquisa intitulada: "Aplicação da análise logística/binomial com enfoque no Índice de Qualidade do Sono de Pittsburgh, nas comunidades de Poço do Serrote, Catolé, Três Irmãos e Poldrinho\_Sertão do Pajeú/PE", do mestrando do Programa em Biometria e Estatística Aplicada (PPGBEA) da Universidade Federal Rural de Pernambuco (UFRPE) Jucarlos Rufino de Freitas, CPF 101.437.544-40, RG 8.663.646 SDS, com Endereço a Rua Luziânia, nº 111, Monsenhor Fabrício - CEP 50.680-030. E-mail: jrdf1@de.ufpe.br / jucarlos123@hotmail, referente ao Projeto aprovado na chamada CNPq/MDA/SPM-PR Nº 11/2014\_Apoio à implantação e manutenção de Núcleos de Extensão em Desenvolvimento Territorial, Aprovado no Processo nº Processo 463077/2014-8 intitulado: "POTENCIALIDADE DO USO DA ÁGUA, MANEJO FLORESTAL E SUAS IMPLICAÇÕES NA QUALIDADE DE VIDA DAS MULHERES, ADULTAS E IDOSAS DE COMUNIDADES DO SERTÃO DO PAJEÚ/PERNAMBUCO"; eu, Prof. Dr. Moacyr Cunha Filho (UFRPE - CPF: 329.439.974 – 04) e minha equipe composta pela Professora Dr<sup>a</sup>. Ana Patrícia Siqueira Tavares Falcão (ESEF/UPE e IFPE – Campus Vitória - CPF: 817.664.834 – 53) e pelo(a)s estudantes Erivaldo Gumercindo de Souza (CPF: 097.017.134 – 00), Giselly de Oliveira Silva (CPF: 107.135.944 – 46), Iloane dos Santos Lima (CPF: 078.064.644 – 41), Manoel Rivelino Gomes de Oliveira (CPF: 032.176.494 – 35) e pelo Presidente do Colegiado da Cidadania do Sertão do Pajeú, respondendo pelos assentamentos de Catolé, Poço do Serrote, Poldrinho e Três Irmãos, Dr. José de Anchieta Alves de Queiroz (CPF: 547.9995.544-68) comprometemo-nos a manter em anonimato, sob sigilo absoluto, durante e após o término do estudo, todos os dados que identifiquem o sujeito da pesquisa, usando apenas para divulgação dos dados inerentes ao desenvolvimento do estudo. Assim como também nos comprometemos com a destruição, após o término da pesquisa, de todo e qualquer tipo de mídia que possa vir a identificá-lo(a)s tais como: filmagens, fotos, gravações, questionários, formulários e outros.


Recife, 19 de janeiro de 2017

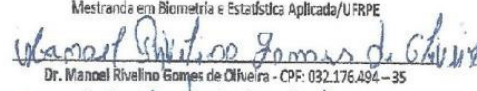
  
Prof. Dr. Moacyr Cunha Filho - CPF: 329.439.974 – 04  
UFRPE – Recife/PE

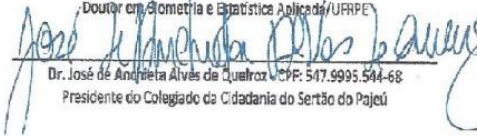
  
Prof. Dr. Ana Patrícia Siqueira Tavares Falcão - CPF: 817.664.834 – 53  
ESEF/UPE – Recife/PE e IFPE/Campus Vitória de Santo Antão

  
Erivaldo Gumercindo de Souza - CPF: 097.017.134 – 00  
Mestrando em Biometria e Estatística Aplicada/UFRPE

  
Giselly de Oliveira Silva - CPF: 107.135.944 – 46  
Mestrando em Biometria e Estatística Aplicada/UFRPE

  
Iloane dos Santos Lima - CPF: 078.064.644 – 41  
Mestrando em Biometria e Estatística Aplicada/UFRPE

  
Dr. Manoel Rivelino Gomes de Oliveira - CPF: 032.176.494 – 35  
Doutor em Biometria e Estatística Aplicada/UFRPE

  
Dr. José de Anchieta Alves de Queiroz - CPF: 547.9995.544-68  
Presidente do Colegiado da Cidadania do Sertão do Pajeú

# APÊNDICE B – FICHA DE AVALIAÇÃO

Data: \_\_/\_\_/\_\_ Pesquisador (a): \_\_\_\_\_ Hora: \_\_\_\_\_

Nome/ Iniciais: \_\_\_\_\_

**1) Gênero:**

- Masculino  
 Feminino

**2) Idade:** \_\_\_\_\_.

**3) Estado Civil:**

- Solteiro(a)  
 Casado(a)  
 Divorciado(a)  
 Viúvo(a)  
 Separado(a)

**4) Nº de pessoas que moram no domicílio:** \_\_\_\_\_.

**5) Cor\raça:**

- Branca  
 Preta  
 Parda  
 Amarela  
 Indígena  
 Não sabe / Não respondeu

**6) Renda familiar:**

- 1 salário mínimo  
 2 salários mínimos  
 3 salários mínimos  
 Não sabe/não respondeu

**7) Tem acesso a água canalizada?**

- Não  
 Sim  
 Não sabe/não respondeu

**8) Tem acesso a rede de esgoto?**

- Não  
 Sim  
 Não sabe/não respondeu

**9) Tem acesso água cisterna ou poço?**

- Não
- Sim
- Não sabe/não respondeu

**10) Tem acesso a energia elétrica?**

- Não
- Sim
- Não sabe/não respondeu

**11) Tem acesso a água cisterna ou poço?**

- Não
- Sim
- Não sabe/não respondeu

**12) Tem fogão?**

- Não
- Sim
- Não sabe/não respondeu

**13) Tem geladeira?**

- Não
- Sim
- Não sabe/não respondeu

**14) Tem televisão?**

- Não
- Sim
- Não sabe/não respondeu

# APÊNDICE C – SELEÇÃO DE VARIÁVEIS

## Variável Idade de natureza quantitativa:

A Tabela 1\* a seguir apresenta a estimativas dos parâmetros e teste de significância das variáveis selecionadas do modelo geral 1 para explicar a QS. Nesta pode-se observar que variável Renda é a primeira variável a ser eliminada do modelo apresentando um p-valor acima de 5%.

Tabela 1\*: Estimativas dos parâmetros e teste de significância das variável selecionadas do modelo geral 1 para explicar a QS. Sertão do Pajeú, PE, Brasil, 2017.

Coeficientes					
	Estimativa	Erro padrão	Wald	p-valor	Odds
(Intercepto)	22,440	40193,08	0,00	1,000	5565776180,359
Gênero(1)	1,304	0,638	4,175	0,041 *	3,684
Idade	-0,110	0,048	5,186	0,023 *	0,896
Estado civil(1)	2,354	0,856	7,566	0,006 **	10,531
Cor/raça			3,796	0,150	
Cor/raça(1)	0,490	1,231	0,158	0,691	1,632
Cor/raça(2)	1,784	1,112	2,572	,109	5,952
<b>Renda</b>			<b>0,013</b>	<b>0,993</b>	
<b>Renda(1)</b>	<b>-20,922</b>	<b>40193,082</b>	<b>0,000</b>	<b>1,000</b>	<b>0,000</b>
<b>Renda(2)</b>	<b>-21,074</b>	<b>40193,082</b>	<b>0,000</b>	<b>1,000</b>	<b>0,000</b>
Cochila(1)	-1,603	0,664	5,833	0,016 *	0,201
---					
Signif. codes:	0 *** 0,001	0,001 ** 0,01	0,01 * 0,05	0,05 . 0,1	
Null devia.	Resid. devia.	AIC	Núm. inter.	Disper. gaus.	(1)
90,97	69,03	-	20	1	prim. categ.

A Tabela 1\*\* a seguir apresenta a estimativas dos parâmetros e teste de significância das variáveis selecionadas do modelo reduzido 1 para explicar a QS. Nesta pode-se observar que variável Cor/raça é a segunda variável a ser eliminada do modelo apresentando um p-valor acima de 5%.

Tabela 1\*\*: Estimativas dos parâmetros e teste de significância das variáveis selecionadas do modelo reduzido 1 para explicar a QS. Sertão do Pajeú, PE, Brasil, 2017.

Coeficientes					
	Estimativa	Erro padrão	Wald	p-valor	Odds
(Intercepto)	1,209	1,773	0,465	0,495	3,351
Genero(1)	1,389	0,632	4,834	0,028*	4,011
Idade	-0,105	0,046	5,233	0,022*	0,901
EstadoCivil(1)	2,352	0,855	7,572	0,006**	10,510
<b>Cor/raça</b>			<b>4,216</b>	<b>0,122</b>	
<b>Cor/raça(1)</b>	<b>0,501</b>	<b>1,214</b>	<b>0,170</b>	<b>0,680</b>	<b>1,651</b>
<b>Cor/raça(2)</b>	<b>1,859</b>	<b>1,107</b>	<b>2,820</b>	<b>0,093</b>	<b>6,419</b>
Cochila(1)	-1,532	0,658	5,416	0,020*	0,216
---					
Signif. codes:	0 *** 0,001	0,001 ** 0,01	0,01 * 0,05	0,05 . 0,1	
Null devia.	Resid. devia.	AIC	Núm. inter.	Disper. gaus.	(1)
90,97	70,14	-	5	1	prim. categ.

# Anexos

# ANEXO A – ÍNDICE DE QUALIDADE DO SONO DE PITTSBURGH

## Instruções:

**As questões a seguir serão referentes aos seus hábitos de sono durante o mês passado. Suas respostas devem indicar o mais corretamente possível o que aconteceu na maioria dos dias e noites do mês passado. Por favor, responda a todas as questões:**

1) Durante o mês passado a que horas você foi deitar a noite, na maioria das vezes?

HORÁRIO DE DEITAR: \_\_\_\_\_

Comentários: \_\_\_\_\_

2) Durante o mês passado, quanto tempo (em minutos) você demorou para pegar no sono, na maioria das vezes?

QUANTOS MINUTOS DEMOROU PARA PEGAR NO SONO: \_\_\_\_\_

Comentários: \_\_\_\_\_

3) Durante o mês passado, a que horas você acordou de manhã, na maioria das vezes?

HORÁRIO DE ACORDAR: \_\_\_\_\_

Comentários: \_\_\_\_\_

4) Durante o mês passado, quantas horas de sono por noite você dormiu? (Pode ser diferente do número de horas que você ficou na cama)

HORAS DE SONO POR NOITE: \_\_\_\_\_

Comentários: \_\_\_\_\_

**Para cada uma das questões seguintes, escolha uma única resposta que você ache mais correta. Por favor responda a todas as questões.**

5) Durante o mês passado, quantas vezes você teve problema para dormir por causa de:

(a) demorar mais de 30 minutos (meia hora) para pegar no sono:

- nenhuma vez
- menos de uma vez por semana
- uma ou duas vezes por semana
- três vezes por semana ou mais

Comentários: \_\_\_\_\_

---

(b) acordar no meio da noite ou de manhã muito cedo:

- nenhuma vez
- menos de uma vez por semana
- uma ou duas vezes por semana
- três vezes por semana ou mais

Comentários: \_\_\_\_\_

---

(c) levantar-se para ir ao banheiro

- nenhuma vez
- menos de uma vez por semana
- uma ou duas vezes por semana
- três vezes por semana ou mais

Comentários: \_\_\_\_\_

---

(d) ter dificuldade para respirar:

- nenhuma vez
- menos de uma vez por semana
- uma ou duas vezes por semana
- três vezes por semana ou mais

Comentários: \_\_\_\_\_

---

(e) tossir ou roncar muito alto

- nenhuma vez
- menos de uma vez por semana
- uma ou duas vezes por semana

três vezes por semana ou mais

Comentários: \_\_\_\_\_

---

(f) sentir muito frio:

nenhuma vez

menos de uma vez por semana

uma ou duas vezes por semana

três vezes por semana ou mais

Comentários: \_\_\_\_\_

---

(g) sentir muito calor:

nenhuma vez

menos de uma vez por semana

uma ou duas vezes por semana

três vezes por semana ou mais

Comentários: \_\_\_\_\_

---

(h) ter sonhos ruins ou pesadelos:

nenhuma vez

menos de uma vez por semana

uma ou duas vezes por semana

três vezes por semana ou mais

Comentários: \_\_\_\_\_

---

(i) outras razões por favor descreva:

---

---

(j) quantas vezes você teve problemas para dormir por esta razão, durante o mês passado:

nenhuma vez

menos de uma vez por semana

uma ou duas vezes por semana



três vezes por semana ou mais

Comentários: \_\_\_\_\_

---

6) Durante o mês passado, como você classificaria a qualidade do seu sono?

muito boa

boa

ruim

muito ruim

Comentários: \_\_\_\_\_

---

7) Durante o mês passado, você tomou algum remédio para dormir, receitado pelo médico ou indicado por outra pessoa (farmacêutico, amigo, familiar) ou mesmo por sua conta?

nenhuma vez

menos de uma vez por semana

uma ou duas vezes por semana

três vezes por semana ou mais.

Quais? \_\_\_\_\_

Comentários: \_\_\_\_\_

---

8) Durante o mês passado você teve dificuldade de ficar acordado enquanto estava dirigindo, fazendo suas refeições ou participando de qualquer outra atividade social, quantas vezes isso aconteceu?

nenhuma vez

menos de uma vez por semana

uma ou duas vezes por semana

três vezes por semana ou mais.

Quais? \_\_\_\_\_

Comentários: \_\_\_\_\_

---

9) Durante o mês passado você sentiu indisposição ou falta de entusiasmo para realizar suas atividades diárias?

- nenhuma
- pequena
- moderada
- muita

Comentários: \_\_\_\_\_

---

10) Para você, o sono é:

- um prazer
- uma necessidade
- outro.

Quais? \_\_\_\_\_

Comentários: \_\_\_\_\_

---

Você cochila?

- sim
- não

Comentários: \_\_\_\_\_

---

Caso sim, você cochila intencionalmente, ou seja, porque quer cochilar?

- sim
- não

Para você, cochilar é:

- um prazer
- uma necessidade
- outro.

Quais? \_\_\_\_\_

Comentários: \_\_\_\_\_

---

# ANEXO B – INSTRUÇÕES DE PONTUAÇÃO PARA O ÍNDICE DE QUALIDADE DO SONO DE PITTSBURGH

## Componente 1: Qualidade subjetiva do sono:

Examine a questão 6 e atribua a pontuação da seguinte maneira:

Resposta	Escore
Muito boa	0
Boa	1
Ruim	2
Muito ruim	3

Pontuação do componente 1 .....

## Componente 2: Latência do sono:

- Examine a questão 2 e atribua a pontuação de a seguinte maneira:

Resposta	Escore
< ou = 15 minutos	0
16 a 30 minutos	1
31 a 60 minutos	2
> 60 minutos	3

- Examine a questão 5a e atribua a pontuação da seguinte maneira:

Resposta	Escore
Nenhuma vez	0
Menos de 1 vez/semana	1
1 a 2 vezes/semana	2
3 vezes/semana	3

- Some a pontuação da questão 2 e 5a:

Soma	Escore
0	0
1 a 2	1
3 a 4	2
5 a 6	3

Pontuação do componente 2 .....

**Componente 3: Duração do sono:**

1. Examine questão 4 e atribua a pontuação da seguinte maneira:

Resposta	Escore
> 7 horas	0
6 a 7 horas	1
5 a 6 horas	2
< 5 horas	3

Pontuação do componente 3 .....

**Componente 4: Eficiência habitual do sono:**

1. Examine a questão 2 e atribua a pontuação da seguinte maneira:

1.1. Escreva o número de horas dormidas (questão 4):

1.2. Calcule o número de horas no leito:

{horário de levantar (questão 3) – horário de deitar (questão 1)}

1.3. Calcule a eficiência do sono:

{nº de horas dormidas / nº de horas no leito} x 100 = eficiência do sono(%)

1.4. Atribua a pontuação do componente 4 da seguinte maneira:

Eficiência do sono(%)	Escore
> 85%	0
75- 84%	1
64-74%	2
< 65%	3

Pontuação do componente 4 .....

**Componente 5: Distúrbios do sono:**

1. Examine as questões de 5b a 5j e atribua a pontuação:

Resposta	Escore
Nenhuma vez	0
Menos de 1 vez/semana	1
2 a 2 vezes/semana	2
3 vezes/semana	3

2. Some a pontuação de 5b a 5j.

3. Atribua a pontuação do componente 5 da seguinte forma:

<b>Resposta</b>	<b>Escore</b>
0	0
1 a 9	1
10 a 18	2
19 a 27	3

Pontuação do componente 5 .....

**Componente 6: Uso de medicação para dormir:**

1. Examine a questão 7 e atribua a pontuação da seguinte maneira:

<b>Resposta</b>	<b>Escore</b>
Nenhuma vez	0
Menos de 1 vez/semana	1
1 a 2 vezes/semana	2
3 vezes/semana	3

Pontuação do componente 6 .....

**Componente 7: Disfunção durante o dia:**

1. Examine a questão 8 e atribua a pontuação da seguinte maneira:

<b>Resposta</b>	<b>Escore</b>
Nenhuma vez	0
Menos de 1 vez/semana	1
1 a 2 vezes/semana	2
3 vezes/semana	3

2. Examine a questão 9 e atribua a pontuação da seguinte maneira:

<b>Resposta</b>	<b>Escore</b>
Nenhuma	0
Pequena	1
Moderada	2
Muita	3

3. Some a pontuação das questões 8 e 9

4. Atribua a pontuação do componente 7 da seguinte maneira:

<b>Resposta</b>	<b>Escore</b>
0	0
1 a 2	1
3 a 4	2
5 a 7	3

Pontuação do componente 7 .....

**Os escores dos sete componentes são somados para conferir uma pontuação global do PSQI, a qual varia de 0 a 21.**

<b>Pontuação</b>	<b>Qualidade do sono</b>
0 a 4	Boa
5 a 10	Ruim
>10	Distúrbio do sono

## ANEXO C – CURVA ROC E A INTERPRETAÇÃO DE SEUS ÍNDICES

Ponto de Corte	Descrição
ROC = 0,5	Sugere sem poder discriminante
$0,7 \leq \text{ROC} < 0,8$	Aceitável poder discriminante
$0,8 \leq \text{ROC} < 0,9$	Excelente poder discriminante
$\text{ROC} \geq 0,9$	Excepcional poder discriminante

Fonte: Adaptado de Hosmer e Lemeshow (2000, p.162).