

**UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE ESTATÍSTICA**

TELMO CRISTIANO GOMES DA SILVA

**Fatores explicativos da mortalidade por
tuberculose em adultos no Nordeste**

**João Pessoa
Dezembro/2011**

TELMO CRISTIANO GOMES DA SILVA

Fatores explicativos da mortalidade por tuberculose em adultos no Nordeste

Trabalho de conclusão de curso apresentado ao Curso de Graduação em Estatística da Universidade Federal da Paraíba, como requisito parcial para obtenção do título de Bacharel. Áreas de concentração: Demografia e Estatística.

Orientador: Prof. Neir Antunes Paes

**João Pessoa
Dezembro/2011**

TELMO CRISTIANO GOMES DA SILVA

Fatores explicativos da mortalidade por tuberculose em adultos no Nordeste

Trabalho de conclusão de curso apresentado ao Curso de Graduação em Estatística da Universidade Federal da Paraíba, como requisito parcial para obtenção do título de Bacharel. Áreas de concentração: Demografia e Estatística.

Aprovada em ____/____/____.

BANCA EXAMINADORA

Prof. Neir Antunes Paes
Universidade Federal da Paraíba

Prof. Antônio Marcos Moreira
Universidade Federal da Paraíba

Prof. Jozemar Pereira dos Santos
Universidade Federal da Paraíba

João Pessoa
Dezembro/2011

*Para minha família por sempre
acreditar em mim.*

AGRADECIMENTOS

- Aos meus pais e irmãs.
- Ao professor Neir Antunes Paes, por me orientar.
- Ao professor Antônio Marcos Moreira por ter aceitado o convite para participar da banca examinadora.
- Ao professor Jozemar Pereira dos Santos por ter aceitado o convite para participar da banca examinadora.
- A todos meus colegas de curso pela amizade.

RESUMO

Uma das componentes fundamentais da demografia é a mortalidade. Os índices de mortalidade subsidiam gestores públicos em saúde a identificar como e onde investir na saúde de uma população. A mortalidade por Tuberculose permanece como uma das preocupações em regiões de menos desenvolvimento econômico, como o Nordeste brasileiro. Assim, teve-se como objetivo verificar a relação da mortalidade por Tuberculose com fatores explicativos na citada região. As fontes dos dados foram o Ministério do Trabalho e Emprego, Ministério da Saúde e Sistema FIRJAN. Os passos metodológicos consistem em: a) estudar a qualidade das estatísticas de óbitos, b) aplicar o método estatístico de análise de regressão linear para verificar o grau de relação que os fatores explicativos escolhidos para o estudo têm com a mortalidade por Tuberculose. Concluiu-se que alguns fatores socioeconômicos, ocupação e doenças como HIV e Pneumonia evidenciaram uma relação significativa com a mortalidade por Tuberculose no Nordeste brasileiro.

Palavras-chave: Tuberculose, Mortalidade, Regressão Linear, Nordeste.

ABSTRACT

One of the key components of demography is mortality. Mortality rates subsidize public health managers to identify how and where to invest in the health of a population. Mortality from Tuberculosis remains a major concern in regions with less economic development, as the Brazilian Northeast. Thus, the main objective was to investigate the relationship of mortality from tuberculosis and explanatory factors cited in the region. The sources were the Ministry of Labor and Employment, Ministry of Health and FIRJAN System. The methodological steps are: a) study the accuracy of death statistics, b) apply the statistical method of linear regression analysis to assess the degree of relationship that the explanatory factors chosen for the study have to mortality from Tuberculosis. It was concluded that some factors as socioeconomic, occupation and diseases as HIV showed a significant relationship with mortality from Tuberculosis in the Brazilian Northeast.

Keywords: Tuberculosis, Mortality, Linear Regression, Northeast.

Lista de Figuras

Figura 4.1: Quadro-Resumo das variáveis componentes do IFDM por Área de Desenvolvimento.	20
Figura 5.1: Representação de uma Regressão Linear Simples.....	26
Figura 6.1: Taxa de mortalidade por 100.000 hab. para Tuberculose, AIDS e Pneumonia para ambos os sexos - Mesorregiões do Nordeste 2008.	46
Figura 6.2: Índice FIRJAN de Saúde, Emprego/Renda e Educação para ambos os sexos – Mesorregiões do Nordeste 2007.	47
Figura 6.3: Gráficos para as análises de resíduos das variáveis de ocupação para o sexo masculino – mesorregiões do Nordeste, 2008.	55
Figura 6.4: Gráficos para as análises de resíduos das variáveis de ocupação para o sexo feminino - mesorregiões do Nordeste, 2008.	55
Figura 6.5: Gráficos para as análises de resíduos dos modelos gerados através da regressão múltipla para ambos os sexos – mesorregiões do Nordeste, 2008.	59
Figura A.1: Gráficos para os homens.	64
Figura A.2: Gráficos para as mulheres.....	64

Lista de Tabelas

Tabela 2.1: Série histórica da Taxa de Incidência de Tuberculose no Nordeste e Estados (1990 a 2009).....	15
Tabela 5.1: ANOVA	29
Tabela 5.2: Análise de Variância para Significância de uma Regressão Múltipla.	33
Tabela 6.1: Cobertura de óbitos pelo Método de Brass por sexo para as mesorregiões do Nordeste, 2008.	39
Tabela 6.1: Cobertura de óbitos pelo Método de Brass por sexo para as mesorregiões do Nordeste, 2008.....	40
Tabela 6.2: Coeficientes de distribuição (β) da regressão linear de Ledermann para ambos os sexos.	41
Tabela 6.3: Proporções para as principais causas de morte para o sexo masculino para mesorregiões do Nordeste, 2008.....	43
Tabela 6.4: Proporções para as principais causas de morte para o sexo feminino para mesorregiões do Nordeste, 2008.....	44
Tabela 6.5: Matriz de Coeficientes de Correlação de Pearson e significância de alguns indicadores de mortalidade e socioeconômicas – Nordeste 2008.	50
Tabela 6.6: Resultado do teste de normalidade para o sexo masculino	52
Tabela 6.7: Resultado do teste de normalidade para o sexo feminino.	53
Tabela 6.8: R de Pearson, R^2 e p-valor para as variáveis de Construção Civil, Comércio e Administração Pública por sexo e Extrativa Mineral para as mulheres.	54
Tabela 6.9: Modelo ajustado para o sexo masculino.	57
Tabela 6.10: Modelo ajustado para o sexo feminino.	58
Tabela A: Base de Dados para o sexo masculino.	65
Tabela B: Base de Dados para o sexo feminino.	66
Tabela C.	67

Sumário

1 INTRODUÇÃO	12
2 REVISÃO DA LITERATURA	14
2.1 DOENÇAS INFECCIOSAS E PARASITARIAS	15
2.1.1 TUBERCULOSE	16
2.1.1.1 FORMA DE CONTÁGIO, SINTOMAS E TRATAMENTO	16
3 OBJETIVO	18
3.1 GERAL	18
3.2 ESPECÍFICOS	18
4 FONTES DOS DADOS	19
4.1 ÍNDICE FIRJAN DE DESENVOLVIMENTO MUNICIPAL (IFDM)	19
5 METODOLOGIA	21
5.1 A CLASSIFICAÇÃO ESTATÍSTICA INTERNACIONAL DE DOENÇAS ..	21
5.2 QUALIDADES DOS ÓBITOS	22
5.2.1 COBERTURA DOS ÓBITOS.....	22
5.2.2 CAUSAS MAL DEFINIDAS	23
5.3 INDICADORES DE MORTALIDADE.....	24
5.3.1 PROPORÇÃO DE ÓBITOS.....	24
5.3.2 TAXA DE MORTALIDADE PADRONIZADA	24
5.3.3 INDICADORES SOCIOECONÔMICOS.....	25
5.4 ANÁLISE DE REGRESSÃO LINEAR	25
5.4.1 REGRESSÃO LINEAR SIMPLES	26
5.4.2 REGRESSÃO LINEAR MÚLTIPLA.....	30
5.4.3 PRINCIPAIS TESTES DE HIPÓTESE DA REGRESSÃO MÚLTIPLA	31
5.4.3.1 TESTE DE SIGNIFICÂNCIA	31
5.4.3.2 TESTE PARA β_j	33
5.5 VERIFICAÇÃO DAS SUPOSIÇÕES DO MODELO DE REGRESSÃO	34
5.5.1 ANÁLISE DE RESÍDUOS	34
5.6 MULTICOLINEARIDADE.....	35
5.6.1 MULTICOLINEARIDADE PERFEITA	36
5.6.2 AUSÊNCIA DE MULTICOLINEARIDADE	37
5.6.3 ALTO GRAU DE MULTICOLINEARIDADE.....	37

6 RESULTADOS E DISCUSSÕES	39
6.1 COBERTURA DOS ÓBITOS.....	39
6.2 DISTRIBUIÇÃO DOS ÓBITOS MAL DEFINIDOS	40
6.3 INDICADORES DE MORTALIDADE.....	42
6.3.1 PROPORÇÕES	42
6.3.2 TAXAS DE MORTALIDADE.....	45
6.4 INDICADORES SOCIOECONÔMICOS.....	47
6.5 ANÁLISES DE REGRESSÃO	48
6.5.1 COLINEARIDADE E MULTICOLINEARIDADE DOS DADOS	48
6.5.2 TRANSFORMAÇÕES DOS ÍNDICES DO FIRJAN	51
6.5.3 TESTE DE NORMALIDADE NOS RESÍDUOS DAS VARIÁVEIS CORRELACIONADAS.....	51
6.5.4 ANÁLISE DAS VARIÁVEIS OCUPACIONAIS ATRAVÉS DE REGRESSÃO SIMPLES.....	53
6.5.5 ANÁLISE DE RESÍDUOS DA TUBERCULOSE EM RELAÇÃO ÀS VARIÁVEIS OCUPACIONAIS.....	54
6.5.6 MODELO DE REGRESSÃO LINEAR MÚLTIPLA	56
6.5.7 ANÁLISE DE RESÍDUOS	58
7 CONSIDERAÇÕES FINAIS	60
REFERÊNCIAS	62
APÊNDICE A	64
ANEXO A	65
ANEXO B	66
ANEXO C	67

1 INTRODUÇÃO

A mortalidade é uma das componentes da demografia que contribui muito para se verificar o quanto uma população está se modificando no que se diz respeito ao seu aumento ou decréscimo. É também através dos índices de mortalidade que os gestores públicos em saúde podem tomar decisões acerca de como investir na saúde de uma população.

A causa evitável ou reduzível de mortalidade é uma causa que poderia ser evitada ou prevenida. Um dos meios para a diminuição dos níveis de mortalidade por uma determinada causa de morte evitável é a produção de estatísticas de saúde para se conhecer a participação e importância das causas de morte que estão ocorrendo em uma determinada região e, assim, tomar providências segundo as políticas de saúde.

A estatística é uma ferramenta importante para obtenção ou predição desses índices ou indicadores de mortalidade facilitando, assim, a tomada de decisões dos administradores públicos quanto à forma de distribuir os recursos econômicos nas diversas áreas da gestão pública entre elas a da saúde.

No mundo, uma doença evitável que sofreu um incremento em sua incidência foi a Tuberculose. Em março de 1993, a tuberculose foi declarada pela Organização Mundial da Saúde (OMS) como "emergência mundial de saúde", sendo diversos os fatores atribuídos a este aumento, tais como: surgimento do HIV/AIDS, resistência às drogas, piora das condições socioeconômicas e sucateamento dos programas de controle da tuberculose (OMS, 2002).

A taxa de incidência de tuberculose no Brasil em 1981 foi cerca de 63,40/100.000 hab. (HIJAR, 1992). Em 2000 foi de 47,81/100.000 hab., passando para 36,81/100.000 hab. em 2008 (SINAN, 2010). A análise por regiões fisiográficas mostra, por exemplo, importantes diferenças entre os coeficientes na Região Nordeste de 37,19/100.000 hab. contra a Região Centro Oeste de 22,38/100.000 em 2008 (SINAN, 2010). A região Nordeste quando comparada com o Brasil encontra-se em um nível um pouco mais elevado.

A identificação dos fatores de risco associados à tuberculose é de extrema importância para uma adequada intervenção em termos de saúde pública. O papel dos fatores socioeconômicos na determinação direta da doença ou intermediando outros

fatores deve ser identificado para um melhor entendimento da cadeia de eventos que levam à morte por Tuberculose.

Poucos, no entanto, são os estudos epidemiológicos desenvolvidos no Brasil em que tenham sido investigados tais fatores de risco para a tuberculose, particularmente no Nordeste.

Sendo assim, justificam-se estudos com tal propósito, ou seja, que investiguem a associação dos óbitos por Tuberculose com seus principais fatores desencadeantes.

2 REVISÃO DA LITERATURA

A Tuberculose é classificada pela organização mundial de saúde, como uma doença infecciosa de amplitude mundial. O Brasil ocupa a 15ª posição entre os 22 países que abrigam cerca de 80% de todos os casos de Tuberculose no mundo (WORLD HEALTH ORGANIZATION, 2002). O vírus da imunodeficiência humana (HIV) tem contribuído para o incremento da epidemia da Tuberculose, dificultando o seu controle no mundo, principalmente em populações de baixa renda, nas quais sua associação com o HIV representa a principal causa de morte (COCK, 2006).

As duas doenças (principalmente a Tuberculose) estão concentradas em áreas de pobreza, onde existem recursos mínimos para o diagnóstico, tratamento e controle da infecção e os serviços de saúde pública não atendem às necessidades para o controle das epidemias, apesar das grandes conquistas no tratamento tanto da Tuberculose como da AIDS (BARTLETT, 2007; OMS, 2002). A coinfeção pelo bacilo de Koch e pelo HIV pode elevar em 25 vezes o risco de desenvolver a Tuberculose (DALEY & DIPERRI, 1992; HIJJAR, 1994).

Nos países em desenvolvimento, o impacto da coinfeção pela Tuberculose e pelo HIV é ainda mais alarmante. O número de coinfectados na América Latina praticamente se quadruplica quando se toma como referência o Canadá e os Estados Unidos (SUDRE, 1991; WORLD HEALTH ORGANIZATION, 1993).

Estudos indicaram que a probabilidade de um indivíduo infectado pelo HIV desenvolver a infecção Tuberculosa após a exposição ao bacilo é mais alta que a de um indivíduo não infectado (PABLOS-MENDES, 1997).

A região do Nordeste brasileiro concentrou 7,3% de todos os casos de AIDS registrados no Brasil entre os anos de 1993-95. Entre esse período e o início da epidemia, apresentou um crescimento de 1.100% em sua taxa de incidência (MINISTÉRIO DA SAÚDE, 1995). Essa região ocupou na mesma época o segundo lugar em número de casos de Tuberculose no País (HIJJAR, 1994).

Diante desse quadro, em 1998, o Ministério da Saúde (MS) do Brasil declarou a Tuberculose como prioridade nacional. Experiências em outros países mostraram que mesmo em condições socioeconômicas adversas, um programa de controle da Tuberculose (PCT) bem estruturado pode modificar um cenário semelhante ao encontrado no Brasil.

A predisposição que um indivíduo apresenta para desenvolver a doença deve-se à interação de fatores genéticos e ambientais. A identificação dos fatores associados à ocorrência da Tuberculose é uma meta importante para que sejam traçadas medidas eficazes de controle.

A Tabela 2.1 traz a série histórica da incidência de Tuberculose para o Nordeste e seus Estados no período de 1990 a 2009 para fins comparativos e de evolução.

Tabela 2.1: Série histórica da Taxa de Incidência de Tuberculose no Nordeste e Estados (1990 a 2009).

UF e Região	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Maranhão	81,60	72,37	68,41	64,45	64,68	69,87	61,66	62,64	57,10	36,30	52,00	46,02	46,96	44,66	44,89	45,22	41,13	38,90	33,62	33,34
Piauí	64,33	63,09	67,50	70,44	68,03	69,47	52,60	49,82	51,46	48,06	43,65	40,65	38,06	35,40	37,37	36,18	32,67	32,23	28,37	26,49
Ceará	74,05	66,19	59,52	68,75	68,19	67,04	58,17	53,31	51,84	55,84	46,04	46,97	46,94	50,46	49,03	49,36	42,90	40,94	43,36	44,35
Rio Grande do Norte	57,17	50,09	55,74	56,12	56,34	48,72	47,92	34,61	47,25	45,06	64,68	36,98	37,86	39,06	39,99	36,06	32,76	28,63	30,13	30,95
Paraíba	44,67	39,55	41,77	50,36	48,31	42,52	40,23	37,82	46,13	67,54	38,79	32,78	32,90	33,71	34,41	33,76	27,35	26,71	27,57	27,64
Pernambuco	53,02	52,11	53,78	62,69	61,36	70,07	58,25	53,48	53,84	53,95	48,09	47,58	50,01	52,79	54,19	52,69	47,83	46,14	46,06	44,96
Alagoas	51,10	56,00	54,54	51,14	47,97	42,64	44,13	41,53	38,09	43,90	40,85	39,94	39,69	40,99	40,13	41,71	37,40	37,34	36,93	36,79
Sergipe	45,56	41,63	41,23	42,21	40,92	42,24	39,41	36,45	36,44	38,13	30,82	23,88	24,76	28,11	25,80	34,35	29,69	24,49	28,71	28,07
Bahia	60,67	64,05	58,72	74,67	77,04	68,78	62,96	61,41	61,12	58,44	49,58	55,36	47,02	51,01	49,62	48,12	43,94	38,78	36,92	38,86
Nordeste	61,53	59,37	57,37	64,92	64,79	63,29	56,18	52,97	53,31	55,55	47,60	45,99	44,14	46,14	45,88	45,39	40,65	37,88	37,19	37,58

Fonte: Sinan/SVS/MS atualizado em 03/09/2010.

2.1 DOENÇAS INFECCIOSAS E PARASITARIAS

A Tuberculose faz parte do grupo de causas das Doenças Infecciosas e Parasitárias. Infecção é a penetração, multiplicação e/ou desenvolvimento de um agente infeccioso em determinado hospedeiro; doença infecciosa reflete as conseqüências das lesões causadas pelo agente e pela resposta do hospedeiro manifestada por sintomas e sinais e por alterações fisiológicas, bioquímicas e histopatológicas.

As doenças infecciosas e parasitárias podem ser causadas pelos seguintes mecanismos: invasão e destruição dos tecidos por ação mecânica, por reação inflamatória ou por ação de substâncias líticas (lisinas); ação de toxinas específicas, elaboradas pelos germes infectantes ou parasitos, capazes de causar danos locais e/ou à distância nas células dos hospedeiros; indução de reação de hipersensibilidade com resposta imune do hospedeiro capaz de produzir lesões em suas próprias células e tecidos (MINISTÉRIO DA SAÚDE, 2005).

2.1.1 TUBERCULOSE

2.1.1.1 FORMA DE CONTÁGIO, SINTOMAS E TRATAMENTO

A Tuberculose é uma doença infecciosa (que pode passar de uma pessoa para outra), transmitida por um microorganismo que é um tipo especial de bactéria, chamada *Mycobacterium Tuberculosis* ou bacilo de Koch, em homenagem ao seu descobridor, o alemão chamado Robert Koch. Popularmente, ela é conhecida como “tísica”, “mancha no pulmão” ou “pulmão fraco”. É uma doença que atinge os pulmões, mas também pode atingir os rins, ossos, pleura, meninges, gânglios e outros órgãos, nesses casos a doença recebe o nome de Tuberculose Extrapulmonar.

Quando uma pessoa com Tuberculose Pulmonar tosse, fala ou espirra, ela expelle gotículas contaminadas com o bacilo. A pessoa que estiver próxima, ao respirar, pode inspirar o bacilo para o seu pulmão. Quando isso ocorre, dizemos que o bacilo foi transmitido de uma pessoa para outra. O contato direto e permanente com o paciente em ambiente fechado, com pouca ventilação e ausência de luz solar, representam uma maior chance de infecção.

Não se adquire Tuberculose bebendo no copo ou utilizando o mesmo talher do paciente, desde que bem lavados. Pessoas com as defesas naturais comprometidas, como os infectados pelo vírus da AIDS, têm uma possibilidade maior de contaminação, por isso, devem estar mais atentas aos sintomas da Tuberculose.

O tratamento dos doentes é a melhor forma de se evitar a disseminação da Tuberculose, pois quanto mais rápido for feito o diagnóstico, menor será a chance de disseminar a doença para outras pessoas. Quando há um doente em casa, as condições da habitação são importantes para diminuir o contágio. Manter a casa limpa, ventilada e

deixar o sol entrar. Para os portadores do bacilo, uma boa alimentação diminui o risco de adoecer e aumenta a possibilidade de cura. Alimentar-se bem é fundamental.

A vacinação com BCG (bacilo de Calmette-Guérin), aplicada em crianças recém-nascidas, não protege contra o contágio, mas evita formas graves da doença.

Os principais sintomas da tuberculose são tosse com catarro (por mais de 15 dias), febre (mais comumente ao entardecer), suores noturnos, falta de apetite, emagrecimento, indisposição e mal-estar.

O tratamento é feito com a combinação de diferentes tipos de medicamentos. A duração é de seis meses e pode curar praticamente todos os casos. Já no primeiro mês, a pessoa sente-se muito melhor, mas é imprescindível não interromper o tratamento, pois os sintomas voltam e o bacilo se tornará resistente aos medicamentos utilizados. Por isso, o tratamento não pode ser interrompido.

3 OBJETIVO

Este trabalho por sua natureza articula as áreas do conhecimento da Estatística, Demografia e Epidemiologia. Traçam-se, assim, os seguintes objetivos:

3.1 GERAL

Verificar a associação da mortalidade por Tuberculose com fatores explicativos para a população adulta por sexo das mesorregiões do Nordeste no ano de 2008.

3.2 ESPECÍFICOS

- a. Construir indicadores de mortalidade por causa básica;
- b. Aplicar a análise de regressão linear.

4 FONTES DOS DADOS

Os dados de óbitos e população foram obtidos através do Ministério da Saúde pelo site do DataSus (2008). Os dados populacionais são estimativas para o ano de 2008. As variáveis socioeconômicas referentes à ocupação foram extraídas do site do Ministério do Trabalho e Emprego (2008). Os indicadores de Emprego/Renda, Educação e Saúde foram retirados do banco de informações do FIRJAN (2010). Os softwares usados para se obter os resultados deste trabalho foram o Statistical Package for the Social Sciences (SPSS) versão 18 e o Microsoft Office Excel 2007.

Descrição das variáveis referentes à ocupação e as causas básicas: AIDS e Pneumonia:

- a. Extrativa Mineral: são os trabalhadores com idade adulta ocupados na extração de minérios para ambos os sexos;
- b. Construção Civil: número de pessoas por sexo que trabalham na construção civil;
- c. Comércio: número de trabalhadores para ambos os sexos que trabalham no comércio;
- d. Administração Pública: funcionários de instituições públicas por sexo;
- e. AIDS: doença causada pelo vírus da imunodeficiência humana (HIV) classificada no grupo de Doenças Infecciosas e Parasitárias da CID-10;
- f. Pneumonia: infecção ou inflamação dos pulmões que se encontra no grupo das Doenças Respiratórias da CID-10.

4.1 ÍNDICE FIRJAN DE DESENVOLVIMENTO MUNICIPAL (IFDM)

O (IFDM) nasceu em resposta à necessidade de se monitorar anualmente o desenvolvimento socioeconômico de uma região, considerando as diferentes realidades de sua menor divisão federativa: o município (FIRJAN, 2010).

O IFDM Emprego/Renda acompanha a movimentação e as características do mercado formal de trabalho, com base nos dados disponibilizados pelo Ministério do Trabalho e Emprego. Como o próprio nome sugere, o indicador trabalha com dois subgrupos, ambos com o mesmo peso (50% do total) no indicador final: emprego

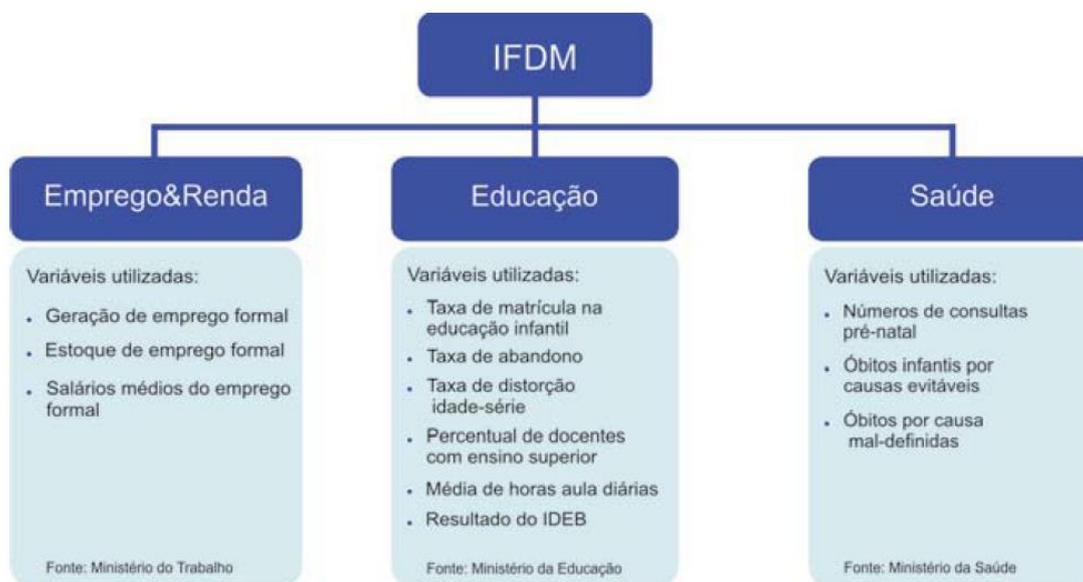
formal (postos de trabalho gerados) e renda (remuneração média mensal do trabalhador formal).

O IFDM Educação foi idealizado para captar tanto a oferta como a qualidade da educação infantil e do ensino fundamental nos municípios brasileiros, de acordo com as competências constitucionais dos municípios (FIRJAN, 2010).

Segundo a OMS, a atenção básica “é o primeiro nível de contato dos indivíduos, da família e da comunidade com o sistema nacional de saúde pelo qual os cuidados de saúde são levados o mais proximamente possível aos lugares onde pessoas vivem e trabalham, e constituem o primeiro elemento de um continuado processo de assistência à saúde (FIRJAN, 2010)”. Partindo da premissa de que o nível primário de atendimento à população é prioritário e deve existir em todos os municípios brasileiros o IFDM adotou as variáveis Atendimento Pré-Natal, Óbitos Mal Definidos e Óbitos Infantis por causas evitáveis por se constituírem fatores reconhecidamente de atenção básica à Saúde.

A Figura 4.1 mostra os índices com as respectivas variáveis utilizadas na sua formação.

Figura 4.1: Quadro-Resumo das variáveis componentes do IFDM por Área de Desenvolvimento.



Fonte: Sistema FIRJAN, 2010.

5 METODOLOGIA

5.1 A CLASSIFICAÇÃO ESTATÍSTICA INTERNACIONAL DE DOENÇAS

A Classificação Estatística Internacional de Doenças e Problemas Relacionados com a Saúde, que também é designada pela sigla CID (em inglês: International Statistical Classification of Diseases and Related Health Problems - ICD) fornece códigos relativos à classificação de doenças e de uma grande variedade de sinais, sintomas, aspectos anormais, queixas, circunstâncias sociais e causas externas para ferimentos ou doenças. A cada estado de saúde é atribuída uma categoria única à qual corresponde um código, que contém até 6 caracteres. Tais categorias podem incluir um conjunto de doenças semelhantes.

A CID é publicada pela Organização Mundial de Saúde (OMS, 1996) e é usada globalmente para estatísticas de morbidade e de mortalidade, sistemas de reembolso e de decisões automáticas de suporte em medicina. O sistema foi desenhado para permitir e promover a comparação internacional da coleção, processamento, classificação e apresentação do tipo de estatísticas de saúde.

A CID é revista periodicamente e encontra-se na sua décima edição. A CID-10, como é conhecida, foi desenvolvida em 1992 para registrar as estatísticas de mortalidade. Atualizações anuais (menores) e tri-anuais (maiores) são publicadas pela OMS.

A CID-10 é dividida nos seguintes capítulos:

- I. Algumas doenças infecciosas e parasitárias;
I.I. Tuberculose;
- II. Neoplasias (tumores);
- III. Doenças sangue órgãos hematopoéticos e alguns transtornos imunitários;
- IV. Doenças endócrinas nutricionais e metabólicas;
- V. Transtornos mentais e comportamentais;
- VI. Doenças do sistema nervoso;
- VII. Doenças do olho e anexos;
- VIII. Doenças do ouvido e da apófise mastóide;
- IX. Doenças do aparelho circulatório;
- X. Doenças do aparelho respiratório;
- XI. Doenças do aparelho digestivo;
- XII. Doenças da pele e do tecido subcutâneo;
- XIII. Doenças sistêmicas osteomuscular e tecido conjuntivo;

- XIV. Doenças do aparelho geniturinário;
- XV. Gravidez parto e puerpério;
- XVI. Algumas afecções originadas no período perinatal;
- XVII. Malformações congênitas deformidades e anomalias cromossômicas;
- XVIII. Sintomas, sinais e achados anormais de exames clínicos e de laboratório, não classificados em outra parte;
- XIX. Lesões, envenenamento e algumas outras consequências causas externas;
- XX. Causas externas de morbidade e mortalidade.

5.2 QUALIDADES DOS ÓBITOS

A qualidade dos dados foi abordada através dos sub-registros de óbitos e das Causas Mal Definidas constantes no Capítulo XVIII da CID-10.

Após as devidas correções nas causas básicas (Tuberculose, HIV e Pneumonia) de óbitos de interesse foram calculados os seguintes indicadores: Taxa de Mortalidade Padronizada para a população em idade adulta – (TMPPIA) e Proporção de Óbitos (percentual). As taxas foram calculadas para as 42 mesorregiões do Nordeste. Também foram desagregadas por sexo, tendo a população do Nordeste do ano de 2008 como a padrão.

A causa denominada de “Outros” é a soma das demais causas.

5.2.1 COBERTURA DOS ÓBITOS

Para avaliar a cobertura dos dados de óbitos do ano de 2008 das mesorregiões do Nordeste, fez-se uso da técnica de estimação do sub-registro de óbitos, denominada de Equação do Balanço de Crescimento ou Técnica de Brass (Brass, 1975). Esta técnica baseia-se em supostos que quando satisfeitos permitem obter uma estimativa consistente. A seguir é feita uma breve exposição metodológica da técnica.

Brass demonstrou que em uma população estável, a equação de balanço do crescimento pode ser expressa como:

$$\frac{N(a)}{N(a+)} = r + f \left[\frac{D(a+)}{N(a+)} \right]$$

onde,

$N(a)$ = população na idade a ;

$N(a+)$ = população com idade igual ou superior a idade a ;

$D(a+)$ = óbitos de pessoas com idade igual ou superior a idade a ;

r = taxa de crescimento da população;

f = fator de correção dos óbitos;

$N(a)/N(a+)$ = taxa de nascimentos da geração de idade a .

Os fatores de correção para as mesorregiões do Nordeste foram estimados por Paes e Marinho (2009) para o ano de 2005. Os resultados dos fatores de correção para este trabalho foram atualizados para o ano de 2008.

5.2.2 CAUSAS MAL DEFINIDAS

As causas Mal Definidas são os registros dos óbitos sem o diagnóstico pelo médico da causa da morte de determinado indivíduo. Então, para se obter resultados mais fidedignos foi necessário que de alguma forma esses óbitos Mal Definidos fossem alocados para os óbitos Definidos, como a Tuberculose, que é o foco deste trabalho.

Para a construção da base de dados foi utilizado o método desenvolvido por Ledermann (1955), e explorado por Vallin (1987) que consiste em uma regressão linear simples onde a variável dependente representa a população de uma determinada causa e a variável independente representa a população da causa mal definida. Então, estimou-se o parâmetro β (coeficiente de declividade) e distribuiu-se as proporções do grupo denominado de *Sintomas, sinais e achados anormais de exames clínicos e de laboratório, não classificados em outra parte* (causa mal definida), para os demais grupos de causas, obtendo-se uma Base de Dados mais próxima da realidade.

As proporções calculadas para os Estados a partir da redistribuição das causas mal definidas foram, por sua vez distribuídas uniformemente para todas as mesorregiões dos Estados correspondentes, ou seja, se uma causa como as Doenças Infecciosas e Parasitárias obteve um percentual de 17%, para determinado Estado, todas as mesorregiões desse Estado receberam a mesma proporção de Mal Definidas para este grupo de causa em questão.

Como as causas específicas Tuberculose e AIDS, são subgrupos das Doenças Infecciosas e a Pneumonia das Doenças Respiratórias dividiu-se os valores dos β encontrados para os grupos pelo total das causas que formam o subgrupo. Por exemplo, o primeiro subgrupo denominado Doenças Infecciosas e Parasitárias é formado por sete causas, então, o resultado do β para este grupo foi dividido por sete encontrando, assim, os β correspondentes para a Tuberculose e a AIDS.

5.3 INDICADORES DE MORTALIDADE

5.3.1 PROPORÇÃO DE ÓBITOS

Calculou-se a distribuição percentual para cada grupo de causas, causa e subcausa básicas de óbitos, utilizando-se como numerador o número de óbitos do grupo de causas, causa ou subcausa e no denominador o total geral de óbitos por todos os grupos de causas.

5.3.2 TAXA DE MORTALIDADE PADRONIZADA

Este método é necessário para tornar comparáveis as mesorregiões do Nordeste, uma vez que elas se apresentam com diferentes distribuições etárias. A população utilizada como padrão foi a estrutura etária da população do Nordeste, referente ao ano de 2008.

Para que se possa calcular a taxa de mortalidade padronizada (TX_p) faz-se necessário calcular a taxa de mortalidade observada (TX_o) que é o resultado do quociente entre o número de óbitos residentes observados e a população residente para um determinado ano específico e região. Em seguida faz-se o cálculo dos óbitos esperados (O_e) que é o produto da taxa de mortalidade observada pelo valor da população padrão.

A taxa de mortalidade padronizada (TX_p) é o quociente entre o número de óbitos esperados (O_e) e a população residente para um determinado ano específico e região. Ela é expressa por 100.000 habitantes.

Para a obtenção das taxas e das proporções foram calculadas as médias do triênio (2007, 2008 e 2009) para evitar eventuais flutuações dos óbitos.

5.3.3 INDICADORES SOCIOECONÔMICOS

Emprego/Renda, Educação e Saúde constituem as três esferas contempladas pelo FIRJAN (2010), todas com peso igual no cálculo para determinação do índice de desenvolvimento dos municípios brasileiros. O índice varia de 0 a 1, sendo que, quanto mais próximo de 1, maior será o nível de desenvolvimento da localidade, o que permite a comparação entre municípios ao longo do tempo. Devido à atribuição de notas de corte fixas, essa metodologia possibilita determinar com precisão se a melhora relativa ocorrida em determinado município decorre da adoção de políticas específicas, ou se o resultado obtido é apenas reflexo da queda dos demais municípios (FIRJAN, 2010).

5.4 ANÁLISE DE REGRESSÃO LINEAR

A literatura pesquisada para descrever as análises de regressão linear foi baseada nas referências: (SEARLE, 1987; SEBER, 1977).

Ao se estudar um fenômeno, deve-se coletar dados (informações) sobre ele. A partir destes dados, analisá-los e obter conclusões.

Para a análise desses dados foi usada a regressão linear. Através dos recursos matemáticos e estatísticos oferecidos pela Análise de Regressão pode-se encontrar alguma função que estime o comportamento do conjunto de dados em relação a uma variável de interesse neste caso a tuberculose.

Análise de regressão é uma técnica de modelagem utilizada para analisar a relação entre uma variável dependente (Y) e uma ou mais variáveis independentes $X_1, X_2, X_3, \dots, X_n$. O objetivo dessa técnica é identificar (estimar) uma função que descreve, o mais próximo possível, a relação entre essas variáveis e assim poder prever o valor que a variável dependente (Y) irá assumir para um determinado valor da variável independente X.

O modelo de regressão pode ser escrito genericamente como:

$$Y = f(X_1, X_2, X_3, \dots, X_n) + \varepsilon,$$

onde o termo ε representa uma perturbação aleatória na função, ou o erro da aproximação. O número de variáveis independentes varia de uma aplicação para outra, quando se tem apenas uma variável independente chama-se Modelo de Regressão

Simples, quando se tem mais de uma variável independente chama-se de Modelo de Regressão Múltipla. A forma da função $f(\cdot)$ também varia, podendo ser representada por um modelo linear, polinomial ou até mesmo uma função não linear.

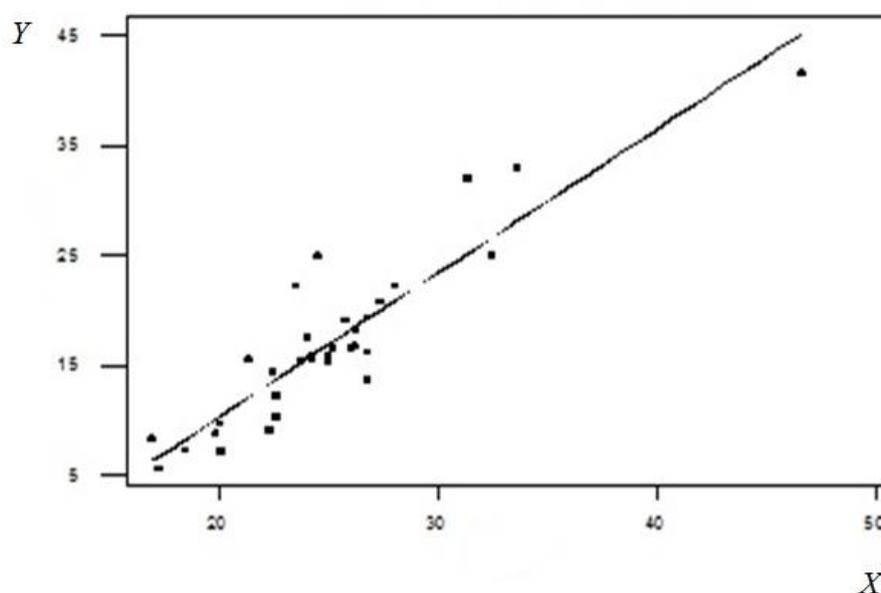
5.4.1 REGRESSÃO LINEAR SIMPLES

Este modelo é utilizado quando existe uma relação linear entre a variável dependente (Y_i) e a variável independente (X_i) (neste caso apenas uma). A função que expressa esse modelo será dada pela forma a seguir:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon, \quad (1)$$

A seguir é obtida uma reta de regressão, então, o primeiro passo na sua interpretação é verificar o sinal β_1 . Se for positivo como no gráfico da Figura 5.1, indica que, quanto maior o valor de X_i , maior o valor de Y_i ; se negativo, indica quanto maior o valor de X_i , menor o valor de Y_i .

Figura 5.1: Representação de uma Regressão Linear Simples



Verifica-se pelo mesmo que nem todos os pontos tocam a reta, e essa diferença é o erro (ε), que pode ter sido ocasionado por um erro de leitura dos dados;

Mas supõe-se que em média esses erros tendem a se anular, ou seja:

$$E(\varepsilon_i) = 0$$

Uma vez escolhido o modelo de regressão, deve-se estimar seus parâmetros, neste caso os coeficientes da equação da reta, β_0, β_1 . Isso pode ser feito a partir da aplicação do Método dos Mínimos Quadrados.

Tirando a média sobre a equação acima, tem-se:

$$\bar{Y} = \beta_0 + \beta_1 \bar{X} \quad (2)$$

uma vez que a média dos erros é zero.

Subtraindo as duas equações tem-se:

$$Y_i - \bar{Y} = (\beta_0 - \beta_0) + (\beta_1)(X_i - \bar{X}) + \varepsilon_i$$

Chamando de y e x as diferenças centradas nas médias, $(Y - \bar{Y})$ e $(X - \bar{X})$ respectivamente, tem-se que:

$$y_i = \beta_1 x_i + \varepsilon_i$$

ou ainda,

$$\varepsilon_i = y_i - \beta_1 x_i$$

Fazendo a soma dos quadrados dos erros,

$$\sum (\varepsilon_i)^2 = \sum (y_i - \beta_1 x_i)^2$$

$$\sum (\varepsilon_i)^2 = \sum y_i^2 - \sum 2\beta_1 x_i y_i + \sum \beta_1^2 x_i^2$$

como β_1 é uma constante,

$$\sum (\varepsilon_i)^2 = \sum y_i^2 - 2\beta_1 \sum x_i y_i + \beta_1^2 \sum x_i^2$$

Como o objetivo é estimar uma equação que minimize os erros, deve-se então derivar a equação acima em relação à β_1 e igualar a zero. E como não se tem os verdadeiros valores e sim uma amostra, ou seja, o valor a ser determinado é um estimador do verdadeiro valor populacional, a nova nomenclatura para β_1 será $\hat{\beta}_1$. Com isso tem-se:

$$0 = -2 \sum x_i y_i + 2 \hat{\beta}_1 \sum x_i^2$$

Que pode ser reescrita como:

$$\hat{\beta}_1 = \frac{\sum x_i y_i}{\sum x_i^2} \quad (3)$$

E o estimador $\hat{\beta}_0$, pode ser calculado a partir de:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \quad (4)$$

Sendo que a equação de estimativa será dada por:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X} \quad (5)$$

Pode-se perceber que as diferenças $(Y - \hat{Y})$ são relativamente pequenas. Uma análise mais cuidadosa pode ser feita através da aplicação de testes estatísticos, nesse caso ANOVA (teste de variância) e teste t-Student.

A ANOVA é mostrada na Tabela 5.1.

Tabela 5.1: ANOVA

Soma dos Quadrados	Graus de Liberdade (g.l.)	Quadrados Médios (QM)	Teste F
$SQE = \hat{\beta}_1^2 \sum x_i^2$	1	$SQE/g.l.$	$SQEmed/SQRmed$
$SQR = \sum (Y - \hat{Y})^2$	n-2	$SQR/g.l.$	
$SQT = \sum y_i^2$	n-1	$SQE/g.l. + SQR/g.l.$	

Obs: O grau de liberdade em relação ao SQE é devido a termos apenas uma variável independente; Em relação a SQT, os graus devem ser iguais a variância amostral, ou seja, n-1 (onde n é o número da elementos da amostra); E o grau de liberdade para SQR seria dado pela diferença entre este, ou seja n-2.

onde,

soma dos quadrados dos totais de y centrado

$$SQT = \sum y_i^2 \quad (6)$$

soma dos quadrados explicados

$$SQE = \sum \hat{Y}_i^2 = \sum \hat{\beta}_1^2 x_i^2 = \hat{\beta}_1^2 \sum x_i^2 \quad (7)$$

soma dos quadrados dos resíduos

$$SQR = \sum (Y - \hat{Y})^2 \quad (8)$$

Um outro parâmetro utilizado constantemente é o coeficiente de determinação, R^2 , que explica percentualmente a relação entre as variáveis do problema.

$$R^2 = \frac{SQE}{SQT} \quad (9)$$

5.4.2 REGRESSÃO LINEAR MÚLTIPLA

Em algumas situações mais do que uma variável independente (X_1, X_2, \dots, X_n) pode ser necessária para prever o valor da variável dependente (Y_i). O modelo matemático para esse caso é dado abaixo:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i \quad (10)$$

Que para as n observações é escrito da forma:

$$\begin{aligned} Y_1 &= \beta_0 + \beta_1 X_{11} + \beta_2 X_{21} + \dots + \beta_k X_{k1} + \varepsilon_1 \\ Y_2 &= \beta_0 + \beta_1 X_{12} + \beta_2 X_{22} + \dots + \beta_k X_{k2} + \varepsilon_2 \\ \dots & \dots \dots \dots \dots \dots \dots \dots \\ Y_n &= \beta_0 + \beta_1 X_{1n} + \beta_2 X_{2n} + \dots + \beta_k X_{kn} + \varepsilon_n \end{aligned}$$

Que forma na realidade um sistema linear, que se pode escrever na forma de matriz como:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{21} & X_{k1} \\ 1 & X_{12} & X_{22} & X_{k2} \\ \dots & \dots & \dots & \dots \\ 1 & X_{1n} & X_{2n} & X_{kn} \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \dots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{bmatrix}$$

Que escrevendo ainda em outra forma mais compacta tem-se:

$$Y = \beta X + \varepsilon \quad (11)$$

O estimador para β será dado por:

$$\hat{\beta} = (X'X)^{-1}(X'Y) \quad (12)$$

Pela equação acima, há necessidade que o produto $X'X$, tenha uma matriz inversa, o que implica na condição obrigatória que nenhuma coluna da matriz X seja combinação linear das outras.

5.4.3 PRINCIPAIS TESTES DE HIPÓTESE DA REGRESSÃO MÚLTIPLA

Agora serão apresentados os testes de hipóteses, os quais são a primeira etapa da verificação do modelo, que prossegue com a análise de resíduos.

5.4.3.1 TESTE DE SIGNIFICÂNCIA

O primeiro teste de hipótese é o da significância da regressão, que consiste em verificar se existe uma correlação linear entre a resposta e algumas das variáveis explicativas, o que corresponde testar as hipóteses:

$$H_0: \beta_1 = \dots = \beta_p = 0$$

$$H_1: \beta_j \neq 0 \text{ para pelo menos um } j. (j = 1, 2, \dots, p)$$

Para avaliar a significância de um modelo de regressão é necessário utilizar o teste da razão de verossimilhança. Para definir a estatística do teste são necessárias as seguintes somas de quadrados:

$$SQT = Y'Y - \frac{\left(\sum_{i=1}^n Y_i\right)^2}{n} \quad \{\text{soma de quadrados total}\}; \quad (13)$$

$$SQE = Y'Y - \hat{\beta}'X'Y \quad \{\text{soma de quadrados do erro}\}; \quad (14)$$

$$SQR = \hat{\beta}' X' Y - \frac{\left(\sum_{i=1}^n Y_i \right)^2}{n} \quad \{\text{soma de quadrados da regressão}\}. \quad (15)$$

As duas últimas somas de quadrados são uma decomposição da soma total, ou seja,

$$SQT = SQR + SQE. \quad (16)$$

A estatística obtida pela razão de verossimilhança é

$$F_0 = \frac{MQR}{EMQ},$$

e as fórmulas de MQR, média dos quadrados da regressão e EMQ, erro médio quadrático, são fornecidas na Tabela 5.2.

Sob H_0 , utilizando-se o teorema de Cochran, que é demonstrado por Lindgren (1976), têm-se $SQR/\sigma^2 \sim \chi_{2p}$ e $SQE/\sigma^2 \sim \chi_{2n-p-1}$. Sob a hipótese nula que especifica a adequação do modelo, F_0 tem distribuição F de Snedecor com p graus de liberdade no numerador e $n-p-1$ graus de liberdade no denominador.

A regra de decisão do teste é rejeitar H_0 se

$$F_0 > F_{(\alpha, p, n-p-1)},$$

onde $F_{(\alpha, p, n-p-1)}$ é o percentil da distribuição F de Snedecor a um nível α de significância.

Dos resultados anteriores observa-se que sempre $E(EMQ) = \sigma^2$ e que $E(MQR) = \sigma^2$ apenas quando $\beta_1 = \dots = \beta_p = 0$. A idéia principal do teste é comparar MQR com EMQ, isto é, se o quociente entre estas duas quantidades for pequeno, isto significa que a variação de Y explicada pela regressão é próxima daquela explicada pelo erro, ou seja, não é significativa.

Particionar a soma dos quadrados de Y corresponde a uma metodologia que é comumente chamada de análise de variância.

Tabela 5.2: Análise de Variância para Significância de uma Regressão Múltipla.

Fonte de Variação	Soma de Quadrados	Graus de liberdade	Quadrado Médio	F ₀
Regressão	SQR	p	$MQR = \frac{SQR}{p}$	$\frac{MQR}{EMQ}$
Resíduo	SQE	n - p - 1	$EMQ = \frac{SQE}{(n - p) - 1}$	
Total	SQT	n - 1	$QMT = \frac{SQT}{n - 1}$	

Uma outra medida muito utilizada na análise de regressão é o percentual da variação dos dados explicado pelo modelo, que é dado por

$$R^2 = \frac{MQR}{QMT}. \quad (17)$$

5.4.3.2 TESTE PARA β_j

Depois da verificação de que pelo menos um dos parâmetros é significativo, deve-se testar a significância de cada um deles, ou seja:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

Se $H_0: \beta_j = 0$ for aceita, isto determina que a covariável X_j pode ser excluída do modelo. A estatística do teste é

$$t_0 = \frac{\hat{\beta}_j}{\sqrt{\hat{\sigma}^2 D_{jj}}} = \frac{\hat{\beta}_j}{Se(\beta_j)}, \quad (18)$$

onde D_{jj} é o j -ésimo elemento diagonal de $(X'X) - I$. Se $t_0 > t(\alpha/2, n-k-1)$ rejeita-se H_0 , o que implica na inclusão (ou permanência) da variável X_j no modelo.

5.5 VERIFICAÇÃO DAS SUPOSIÇÕES DO MODELO DE REGRESSÃO

Tem como objetivo detectar problemas com o ajuste e/ou quebra das hipóteses primárias de um modelo de regressão normal linear: normalidade dos erros, homocedasticidade (variância constante), independência dos erros e linearidade.

5.5.1 ANÁLISE DE RESÍDUOS

O resíduo para uma determinada observação procura avaliar a discrepância entre o valor observado e o ajustado pelo modelo. O resíduo ordinário é dado simplesmente por $r_i = (y_i - \hat{y}_i)$ que tem distribuição normal com média zero e variância $\hat{\sigma}^2(1 - h_{ii})$ onde h_{ii} é o i -ésimo elemento da matriz $H = X(X'X)^{-1}X'$

O resíduo padronizado é dado por:

$$e_i = \frac{r_i}{\{\hat{\sigma}^2(1 - h_{ii})\}^{1/2}}. \quad (19)$$

Os procedimentos gráficos são bastante utilizados na análise de resíduos destacando-se os gráficos:

- i) resíduos versus índices;
- ii) resíduos versus valores previstos;
- iii) resíduos versus quantis da $N(0, 1)$.

Nas situações (i) e (ii) espera-se que a ocorrência de resíduos distribua-se de forma aleatória sem apresentar nenhum padrão. No caso (iii) uma configuração próxima de uma reta $y = x$ indica que o modelo está bem ajustado.

5.5.2 PONTOS INFLUENTES

É a existência ou não de observações que apresentem certo grau de dependência com o modelo ajustado de modo que venham a perturbá-lo, produzindo alterações marcantes nas estimativas obtidas.

Uma medida de influência bastante conhecida é a da distância *D de Cook* (Cook, 1977) que é dada por:

$$D_i = \frac{(\hat{\beta} - \beta_i)(X'X)(\hat{\beta} - \beta_i)}{p\hat{\sigma}^2} \quad (20)$$

Que pode ser escrita como:

$$D_i = \left\{ \frac{e_i}{\hat{\sigma}(1-h_{ii})^{1/2}} \right\}^2 \frac{h_{ii}}{(1-h_{ii})} \left(\frac{1}{p} \right) \quad (21)$$

Geralmente, procura-se identificar medidas com grandes valores para D_i , sugerindo-se retirar medidas com $D_i > F_{p,(n-p)}(\alpha)$.

5.6 MULTICOLINEARIDADE

É o nome dado ao problema geral que ocorre quando duas ou mais variáveis explicativas são muito correlacionadas entre si, o que torna difícil, utilizando-se apenas o modelo de regressão, distinguir suas influências separadamente.

A segunda suposição do modelo de regressão, afirma que nenhuma relação linear exata pode existir entre quaisquer covariáveis ou combinações lineares destas. Quando se viola esta hipótese têm-se o problema de multicolinearidade perfeita. Por outro lado, se as variáveis não estão correlacionadas entre si, denomina-se, este caso, ausência de multicolinearidade, sendo chamada de ortogonal a regressão com estas variáveis. O caso intermediário, muito comum em problemas reais, ocorre quando a correlação entre duas ou mais variáveis é alta, sendo esta situação chamada de alto grau de multicolinearidade. Antes de discutir a Multicolinearidade detalhadamente, é

importante observar a colocação de Kmenta (1978) que afirma: “A multicolinearidade é uma questão de grau e não de natureza. A distinção significativa não está entre a presença ou ausência de multicolinearidade, mas entre seus vários graus.”

A análise do grau de multicolinearidade pode ser feita através da matriz de correlações que mede a dependência linear de primeira ordem entre as variáveis explicativas. O elemento (k,m) da matriz é dado por

$$r_{km} = \frac{\sum_{i=1}^n (X_{ik} - \bar{X}_k)(X_{im} - \bar{X}_m)}{\sqrt{\sum_{i=1}^n (X_{ik} - \bar{X}_k)^2 \sum_{i=1}^n (X_{im} - \bar{X}_m)^2}}, \quad (22)$$

onde \bar{X}_j é a média aritmética da j -ésima covariável. Estes elementos são chamados de coeficiente de correlação entre X_k e X_m .

5.6.1 MULTICOLINEARIDADE PERFEITA

Neste caso é impossível estimar o vetor de parâmetros porque a matriz $X'X$ tem o determinante igual à zero; logo, não possui inversa.

Um bom exemplo é dado por Johnston (1963) para um modelo de regressão com duas variáveis explicativas. Fazendo-se algumas modificações neste exemplo, para torná-lo mais claro, obtêm-se o modelo

$$Y = \beta_1 X_1 + \beta_2 X_2 \quad (23)$$

com a hipótese adicional da relação exata entre as covariáveis dada por

$$X_{1i} = cX_{2i}.$$

A matriz $X'X$ calculada para este modelo é

$$\left(\sum_{i=1}^n X_{i2}^2 \right) \begin{bmatrix} c^2 & c \\ c & 1 \end{bmatrix}, \quad (24)$$

e seu determinante é nulo.

Portanto, neste caso, é impossível estimar os efeitos de X_1 e X_2 .

5.6.2 AUSÊNCIA DE MULTICOLINEARIDADE

Quando as variáveis não estão correlacionadas entre si a matriz $X'X$ é diagonal. Portanto, a estimativa do coeficiente de uma variável independente, no modelo de regressão múltipla, é igual a estimativa obtida pela regressão simples desta variável com a variável dependente. No caso de um modelo com duas variáveis explicativas, se X_1 e X_2 são não correlacionadas, adicionar X_2 ao modelo, contendo X_1 , não altera o coeficiente desta; correspondentemente adicionar X_1 ao modelo, contendo X_2 , não altera o coeficiente desta.

5.6.3 ALTO GRAU DE MULTICOLINEARIDADE

Como uma etapa preliminar de verificação de um modelo ajustado, deve-se observar, na matriz de correlação, se existem pelo menos duas variáveis muito correlacionadas. Se isto ocorrer, o coeficiente de uma variável irá depender da outra, não refletindo, assim, o efeito individual da variável a qual está associado, mas, somente um efeito parcial ou marginal, condicionado a outra variável.

O mesmo ocorre com um conjunto de covariáveis altamente correlacionadas, só que, neste caso, o coeficiente de cada uma delas irá depender das demais.

A consequência mais grave deste problema é que o teste de significância do coeficiente de uma variável independente, sendo esta correlacionada com alguma outra, pode indicar sua exclusão do modelo, mesmo que exista uma forte relação linear desta com a variável resposta.

Se, por exemplo, o modelo analisado possui duas variáveis explicativas, isto ocorre porque os elementos da diagonal principal da matriz de covariância do vetor de estimativas dos parâmetros, dada por

$$Cov(\hat{\beta}) = \sigma^2 (X'X)^{-1}, \quad (25)$$

aumentam devido à diminuição, em termos relativos, do valor do determinante da matriz $X'X$. Além disso, a quantidade de informação acrescentada por X_2 , dado que X_1 está no modelo, que é obtida por

$$\text{SQR}(\beta_2/\beta_1), \quad (26)$$

é pequena porque as duas variáveis estão muito correlacionadas.

6 RESULTADOS E DISCUSSÕES

6.1 COBERTURA DOS ÓBITOS

A Tabela 6.1 mostra a cobertura para as mesorregiões do Nordeste, onde se observou que a cobertura dos óbitos para o sexo masculino nas maioria das mesorregiões foi superior à do sexo feminino. A região da Mata Paraibana e o Extremo Oeste Baiano obtiveram respectivamente a melhor e a pior cobertura para o sexo masculino. Para o sexo feminino a melhor cobertura ficou no Agreste Potiguar e a pior também como para os homens no Extremo Oeste Baiano.

Tabela 6.1: Cobertura de óbitos pelo Método de Brass por sexo para as mesorregiões do Nordeste, 2008.

(Continua)

Mesorregiões	Masculino	Feminino
Agreste Alagoano	79,11	70,54
Agreste Paraibano	92,16	85,93
Agreste Pernambucano	89,58	87,44
Agreste Potiguar	93,01	94,63
Agreste Sergipano	89,10	76,18
Borborema	91,56	80,55
Central Potiguar	81,17	74,17
Centro Maranhense	65,77	58,56
Centro Norte Baiano	83,61	75,45
Centro Norte Piauiense	84,79	70,32
Centro Sul Baiano	81,27	76,47
Centro Sul Cearense	91,15	72,79
Extremo Oeste Baiano	59,77	55,66
Jaguaribe	83,03	64,31
Leste Alagoano	80,25	78,96
Leste Maranhense	67,57	62,92
Leste Potiguar	77,66	80,87
Leste Sergipano	86,18	84,56
Mata Paraibana	97,27	86,86
Mata Pernambucana	95,78	87,19
Metropolitana de Fortaleza	89,19	78,82
Metropolitana de Recife	90,42	74,49
Metropolitana de Salvador	80,34	76,55
Nordeste Baiano	85,40	77,68

Fonte: Ministério da Saúde, (2008).

Tabela 6.1: Cobertura de óbitos pelo Método de Brass por sexo para as mesorregiões do Nordeste, 2008.

Mesorregiões	(Conclusão)	
	Masculino	Feminino
Noroeste Cearense	94,15	89,75
Norte Cearense	89,86	85,43
Norte Maranhense	81,81	78,18
Norte Piauiense	95,26	80,46
Oeste Maranhense	70,13	62,56
Oeste Potiguar	80,90	67,29
São Francisco Pernambucano	77,66	93,89
Sertão Alagoano	83,60	89,31
Sertão Paraibano	86,74	72,31
Sertão Pernambucano	92,97	83,23
Sertão Sergipano	93,42	80,85
Sertões Cearenses	93,67	66,41
Sudeste Piauiense	82,13	73,13
Sudoeste Piauiense	75,85	65,69
Sul Baiano	66,05	63,45
Sul Cearense	92,59	82,36
Sul Maranhense	71,16	77,03
Vale São Francisco da Bahia	79,71	72,25

Fonte: Ministério da Saúde, (2008).

É possível que algumas estimativas estejam necessitando de revisões, particularmente aquelas cujos valores para os homens foram inferiores em relação às mulheres ou que houve uma diferença entre os sexos acima de 10%. Foram elas: Centro Sul Cearense, Jaguaribe, Leste Potiguar, Metropolitana de Recife, Oeste Potiguar, São Francisco Pernambucano, Sertões Cearenses e Sul Maranhense.

Apesar da possibilidade de erros de estimação acredita-se que não comprometam de modo importante o cálculo dos indicadores de mortalidade como as taxas e proporções.

6.2 DISTRIBUIÇÃO DOS ÓBITOS MAL DEFINIDOS

Antes de ser efetuada as análises através das regressões linear os dados dos óbitos por Tuberculose, HIV e Pneumonia foram corrigidos pelo método de Ledermann para a redistribuição dos óbitos mal definidos.

As causas citadas fazem parte do grupo de Doenças Infecciosas (Tuberculose e AIDS) e Doenças do Aparelho Respiratório (Pneumonia) a causa Outros é a soma das demais causas.

O método de Ledermann, como descrito, consiste em calcular através de regressão linear simples os coeficientes de declividade (β) de uma determinada causa definida em relação a causa mal definida.

Na Tabela 6.2 podem ser observados os valores para os coeficientes de redistribuição para os Estados do Nordeste.

Tabela 6.2: Coeficientes de distribuição (β) da regressão linear de Ledermann para ambos os sexos.

Estados	Sexo	Doenças Infecciosas e Parasitárias	Doenças do Aparelho Respiratório	Tuberculose e AIDS	Pneumonia	*Outros
Alagoas	M	-0,1490	-0,0850	-0,0213	-0,0170	-0,7650
	F	-0,1450	-0,0860	-0,0207	-0,0172	-0,7700
Bahia	M	-0,1180	-0,0880	-0,0169	-0,0176	-0,7930
	F	-0,1010	-0,0900	-0,0144	-0,0180	-0,8090
Ceará	M	-0,1250	-0,0880	-0,0179	-0,0176	-0,7880
	F	-0,0480	-0,0950	-0,0069	-0,0190	-0,8560
Maranhão	M	-0,0650	-0,0940	-0,0093	-0,0188	-0,8420
	F	-0,1720	-0,0830	-0,0246	-0,0166	-0,7440
Paraíba	M	-0,0380	-0,1920	-0,0054	-0,0384	-0,7690
	F	-0,0350	-0,1930	-0,0050	-0,0386	-0,7710
Pernambuco	M	-0,2820	-0,0720	-0,0403	-0,0144	-0,6450
	F	-0,1750	-0,0820	-0,0250	-0,0164	-0,7420
Piauí	M	-0,1520	-0,0850	-0,0217	-0,0170	-0,7630
	F	-0,6510	-0,0350	-0,0930	-0,0070	-0,3130
Rio Grande do Norte	M	-0,2120	-0,0790	-0,0303	-0,0158	-0,7090
	F	-0,0860	-0,1830	-0,0123	-0,0366	-0,7310
Sergipe	M	-0,0720	-0,1860	-0,0103	-0,0372	-0,7420
	F	-0,5630	-0,0440	-0,0804	-0,0088	-0,3920

Fonte: Dados Brutos, DataSus (MS, 2008).

*Soma das demais causas.

6.3 INDICADORES DE MORTALIDADE

6.3.1 PROPORÇÕES

As Tabelas 6.3 e 6.4 mostram os resultados das proporções para o sexo masculino e feminino respectivamente. As Doenças Infecciosas aparecem entre as principais causas de morte tanto para homens quanto para mulheres no Nordeste, o que indica que esse grupo de causa apresenta relevante importância em eventuais estudos, como este, para a análise do perfil da mortalidade no Nordeste.

Pode ser observado que as causas têm uma hierarquia de importância entre os sexos e que guardam coerência com o esperado, ou seja, em primeiro lugar as Doenças do Aparelho Circulatório e na sequência: Aparelho Digestivo, Neoplasias, do Aparelho Respiratório e Doenças Infecciosas.

Os indicadores proporcionais foram usados nas redistribuições dos óbitos mal definidos para se estimar os coeficientes que indicaram quanto da causa Mal Definida foi alocada para as causas definidas usadas no trabalho, conforme explicado na metodologia.

Tabela 6.3: Proporções para as principais causas de morte para o sexo masculino para mesorregiões do Nordeste, 2008.

Mesorregiões	DOENÇAS DO APARELHO CIRCULATORIO	DOENÇAS DO APARELHO DIGESTIVO	NEOPLASIAS	DOENÇAS DO APARELHO RESPIRATORIO	ALGUMAS DOENÇAS INFECCIOSAS E PARASITARIAS
Agreste Alagoano	33,02	19,44	11,94	8,20	7,96
Agreste Paraibano	32,22	14,24	17,78	6,16	7,78
Agreste Pernambucano	33,46	16,38	14,09	7,26	6,56
Agreste Potiguar	36,16	9,23	22,14	7,38	7,75
Agreste Sergipano	26,92	15,64	15,38	7,44	6,92
Borborema	37,10	14,48	18,55	8,60	4,52
Central Potiguar	37,76	10,88	22,11	3,74	6,12
Centro Maranhense	40,48	14,41	12,52	7,72	8,40
Centro NorteBaiano	31,64	17,00	17,89	6,14	10,51
Centro Sul Cearense	33,88	7,89	19,41	6,25	9,54
CentroSul Baiano	28,53	15,17	18,01	7,15	9,41
Centro Norte Piauiense	36,45	13,27	17,28	6,80	7,53
Extremo Oeste Baiano	36,43	14,13	14,87	9,67	12,27
Jaguaribe	38,54	11,34	19,65	8,06	8,06
Leste Alagoano	37,10	15,11	11,78	9,28	10,42
Leste Maranhense	40,83	11,30	13,42	6,60	8,17
Leste Potiguar	34,77	13,57	19,32	4,49	9,16
Leste Sergipano	28,65	14,58	20,14	5,90	9,98
Mata Paraibana	34,12	15,65	18,39	7,37	6,91
Mata Pernambucana	33,69	17,55	12,95	7,43	9,13
Metropolitana de Recife	35,55	13,80	17,07	8,75	11,18
Metropolitanade Fortaleza	29,07	12,60	21,92	7,77	11,42
Metropolitanade Salvador	28,89	11,87	20,26	9,15	13,77
Nordeste Baiano	28,85	17,50	16,69	7,30	9,15
Noroeste Cearense	29,13	10,44	17,43	8,03	9,98
Norte Maranhense	34,20	13,90	15,05	6,28	12,15
Norte Cearense	35,32	14,13	19,78	7,38	6,59
Norte Piauiense	34,52	14,76	14,05	7,38	6,43
Oeste Maranhense	38,91	12,55	13,05	7,22	10,27
Oeste Potiguar	28,89	13,85	20,17	5,81	8,03
São Francisco Pernambucano	38,50	10,96	17,65	6,95	8,56
Sertão Alagoano	42,41	13,23	14,01	7,00	6,23
Sertão Paraibano	36,84	12,57	21,20	7,89	5,41
Sertão Pernambucano	36,63	12,35	16,13	6,54	6,83
Sertão Sergipano	36,00	12,80	15,20	7,20	8,80
Sertões Cearenses	33,83	10,53	19,74	7,52	7,89
Sudeste Piauiense	42,75	8,25	15,50	5,00	9,00
Sudoeste Piauiense	37,13	12,28	12,87	5,85	8,48
Sul Baiano	35,38	12,65	14,86	7,99	9,03
Sul Cearense	31,73	14,62	15,09	10,58	8,55
Sul Maranhense	53,23	6,99	10,75	5,91	6,99
Vale São Francisco da Bahia	35,84	12,94	15,03	8,39	10,49

Fonte: Dados Brutos, DataSus (MS, 2008).

Tabela 6.4: Proporções para as principais causas de morte para o sexo feminino para mesorregiões do Nordeste, 2008.

Mesorregiões	DOENÇAS DO APARELHO CIRCULATÓRIO	NEOPLASIAS	D ENDÓCRINAS, NUTRICIONAIS E METABÓLICAS	DOENÇAS DO APARELHO RESPIRATÓRIO	ALGUMAS DOENÇAS INFECCIOSAS E PARASITÁRIAS
Agreste Alagoano	33,23	20,38	13,48	7,84	8,78
Agreste Paraibano	32,98	25,80	13,56	7,85	5,59
Agreste Pernambucano	34,15	25,79	14,05	6,27	5,48
Agreste Potiguar	36,31	24,58	8,94	5,03	5,59
Agreste Sergipano	36,20	27,24	11,11	6,09	3,58
Borborema	37,65	27,16	13,58	3,09	4,94
Central Potiguar	34,40	33,49	7,80	8,26	5,05
Centro Maranhense	35,77	21,65	11,68	7,54	6,57
Centro Norte Baiano	36,06	24,89	8,71	7,74	6,95
Centro Norte Piauiense	32,88	26,71	8,68	5,82	5,82
Centro Sul Baiano	35,26	24,70	9,80	5,88	7,24
Centro Sul Cearense	32,20	31,22	8,78	8,29	4,88
Extremo Oeste Baiano	31,01	20,25	10,13	6,96	15,19
Jaguaribe	29,82	34,18	7,27	8,00	5,45
Leste Alagoano	36,10	23,22	10,49	8,30	5,77
Leste Maranhense	34,25	20,17	10,89	5,66	5,95
Leste Potiguar	27,61	33,84	11,45	7,38	5,60
Leste Sergipano	33,83	29,13	8,03	6,77	5,50
Mata Paraibana	34,24	28,82	9,78	6,24	4,79
Mata Pernambucana	35,16	23,73	13,54	6,53	7,40
Metropolitana de Fortaleza	28,07	36,24	5,85	6,65	8,27
Metropolitana de Recife	35,39	28,46	6,26	8,14	8,27
Metropolitana de Salvador	31,44	28,32	7,99	7,59	8,77
Nordeste Baiano	33,90	23,95	9,80	7,15	6,84
Noroeste Cearense	29,01	30,08	6,72	10,08	5,95
Norte Cearense	29,57	28,26	6,74	10,87	6,74
Norte Maranhense	29,91	27,84	9,47	5,65	8,51
Norte Piauiense	25,08	30,51	11,19	5,08	5,08
Oeste Maranhense	33,02	22,75	11,22	6,00	7,42
Oeste Potiguar	33,57	31,00	10,26	7,93	5,59
São Francisco Pernambucano	32,39	29,87	8,18	4,72	6,60
Sertão Alagoano	38,43	25,46	9,72	7,87	5,09
Sertão Paraibano	37,89	26,50	10,35	6,42	3,73
Sertão Pernambucano	35,15	27,14	9,11	6,92	5,10
Sertão Sergipano	35,71	30,61	6,12	4,08	5,10
Sertões Cearenses	33,93	32,14	6,92	8,04	5,13
Sudeste Piauiense	33,22	24,83	8,05	5,37	8,39
Sudoeste Piauiense	35,53	25,27	8,06	2,93	5,86
Sul Baiano	37,07	21,79	11,52	7,21	7,52
Sul Cearense	33,12	28,98	7,41	8,71	5,88
Sul Maranhense	50,35	17,48	11,19	2,10	6,99
Vale São Francisco da Bahia	37,71	22,20	9,79	5,73	6,92

Fonte: Dados Brutos, DataSus (MS, 2008).

6.3.2 TAXAS DE MORTALIDADE

Conforme já discutido na introdução, a Tuberculose tem fortes indícios de relacionamento com a AIDS e também com a situação socioeconômica em que vive uma população. Em virtude disso será feita algumas observações a respeito das taxas de mortalidade por AIDS e Pneumonia e os índices do FIRJAN.

No Anexo A e B encontram-se as taxas de mortalidade para as mesorregiões e no Anexo C para os Estados.

Os mapas da Figura 6.1 ilustram as taxas de mortalidade por Tuberculose, AIDS e Pneumonia para o sexo masculino para as mesorregiões do Nordeste. Para a Tuberculose é observado que na maioria delas para cada 100.000 homens até 4,18 morreram por tuberculose e que em algumas outras como, por exemplo, a região metropolitana de Natal e Salvador a taxa ficou acima de 4,18 até 7,63 mortes para cada 100.000 homens. As taxas mais altas foram verificadas nas regiões metropolitana do Recife e Fortaleza nos Estados de Pernambuco e Ceará respectivamente chegando até 11,09 mortes por 100.000 homens.

Para a mortalidade por AIDS as taxas chegaram a 18,00 mortes para cada 100.000 na região metropolitana de Salvador e do Recife. Observa-se que os níveis mais elevados de morte por AIDS no Nordeste se concentraram nas regiões onde se encontra as capitais dos Estados, fato este esperado, uma vez que nessas localidades se encontram os maiores aglomerados de pessoas.

A mortalidade por Pneumonia obteve maiores concentrações nas regiões do Oeste Maranhense, Centro Norte Piauiense e nas regiões metropolitanas de Fortaleza e Recife, alcançando uma taxa de 18,00 mortes por 100.000 homens. Mas, no geral os óbitos por pneumonia chegaram a uma taxa de 4,18 por 100.000 hab. no sexo masculino.

No tocante ao sexo feminino para as mesorregiões (mapa inferior da Figura 6.1) as taxas de mortalidade foram mais baixas quando comparados com as dos homens. Para a Tuberculose, na maioria das mesorregiões do Nordeste a taxa alcançou níveis de 2,22 mortes para cada 100.000 mulheres. As mais elevadas se concentraram nas regiões metropolitanas do Recife, Salvador, Natal, Fortaleza e nas regiões Centrais e Oeste do Maranhão, os quais obtiveram taxas de até 4,44 mortes.

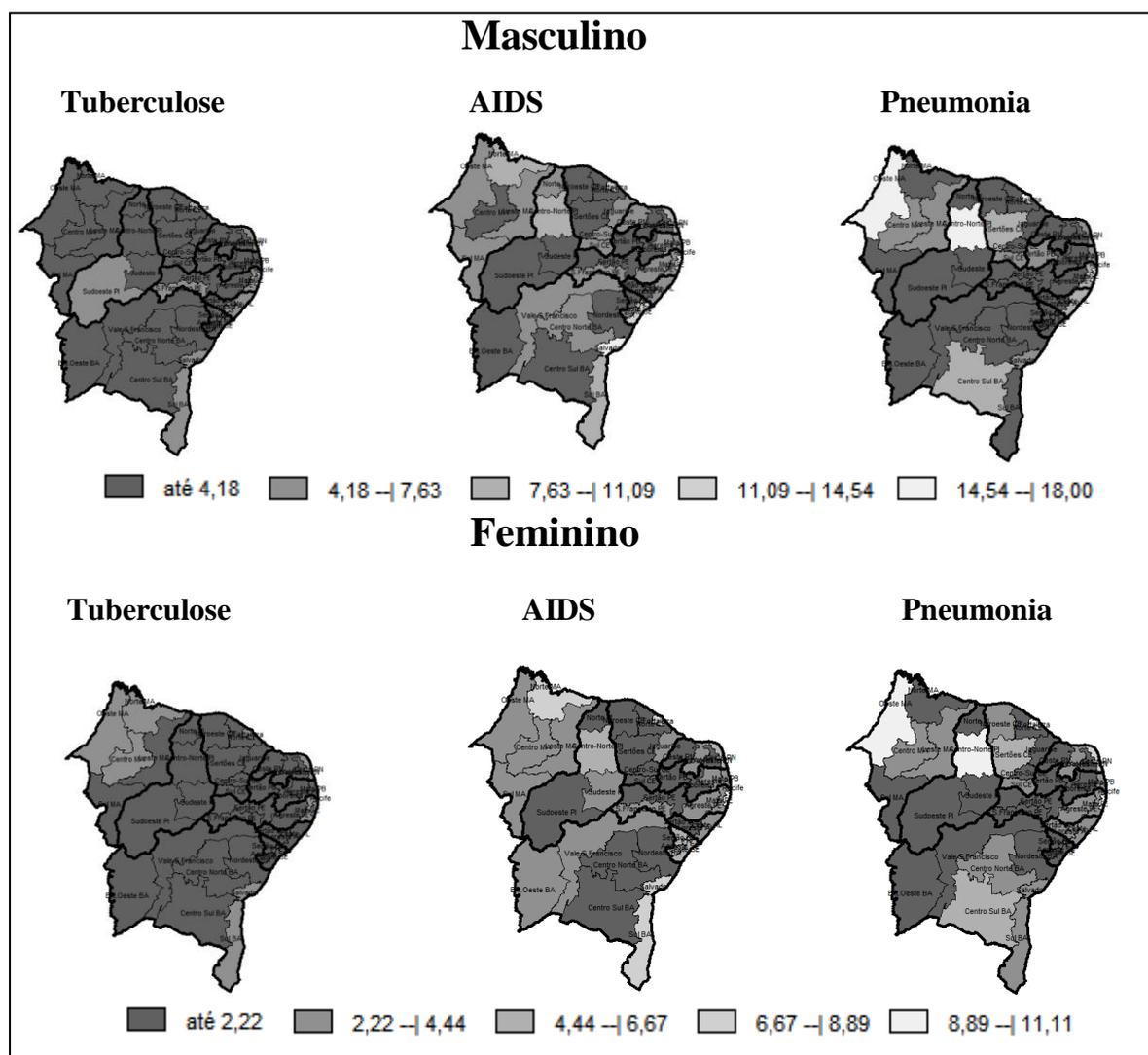
A maior taxa de mortalidade por AIDS foi obtida na região metropolitana do Recife atingindo 11,11 óbitos por 100.000 mulheres. Assim como para os homens, a

maior concentração de mortes por AIDS para o sexo feminino foi mais elevada nas regiões onde se encontram as capitais. O Estado do Maranhão obteve em sua totalidade taxas em torno de 4,44 óbitos por 100.000 mulheres.

A mortalidade por Pneumonia para as mulheres foram mais altas nas regiões metropolitanas de Fortaleza, Recife e Oeste do Maranhão e Centro Norte do Piauí que obtiveram taxas de até 11,11 mortes por 100.000 mulheres.

Como mostram os mapas tanto para os homens quanto para as mulheres, os maiores níveis de mortalidade estiveram concentrados nas regiões com altos contingentes populacionais, principalmente para as regiões metropolitanas de Fortaleza, Recife e Salvador. Esse fato pode ser um indicio da existência da relação entre a Tuberculose e as outras duas causas de morte: a AIDS e a Pneumonia.

Figura 6.1: Taxa de mortalidade por 100.000 hab. para Tuberculose, AIDS e Pneumonia para ambos os sexos - Mesorregiões do Nordeste 2008.



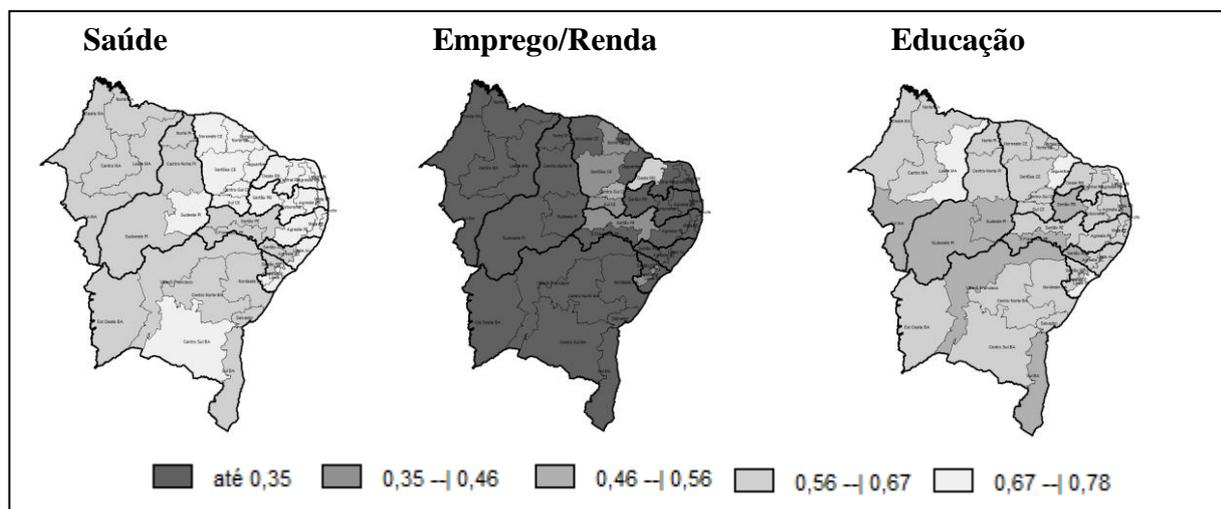
Fonte: Elaboração própria.

6.4 INDICADORES SOCIOECONÔMICOS

Segundo o relatório do Sistema FIRJAN a região Nordeste obteve um nível baixo de desenvolvimento. Isso pode ser comprovado nos mapas da Figura 6.2 que mostra que os índices de Saúde, Emprego/Renda e Educação foram distantes de 1 já que o índice indica que quanto mais próximo 1 mais elevado é o desenvolvimento da região e quanto mais próximo de 0 esse desenvolvimento é baixo. Esse baixo nível de desenvolvimento pode ser notado com mais destaque para o índice de Emprego/Renda que mostra que a maioria das mesorregiões do Nordeste obtiveram um índice de até 0,35 indicando um desenvolvimento próximo de 0 na geração de emprego e aumento da renda no Nordeste para 2007.

No Anexo A e B se encontram os índices de Firjan e de ocupação para cada mesorregião.

Figura 6.2: Índice FIRJAN de Saúde, Emprego/Renda e Educação para ambos os sexos – Mesorregiões do Nordeste 2007.



Fonte: Elaboração própria.

6.5 ANÁLISES DE REGRESSÃO

6.5.1 COLINEARIDADE E MULTICOLINEARIDADE DOS DADOS

Procurou-se investigar a associação entre a tuberculose e as variáveis definidas anteriormente. Para isto, foi utilizado o método de análise de regressão linear. Na regressão linear as variáveis associadas à variável resposta não devem ser relacionadas entre si ou pelo menos não existir uma associação forte, pois isto pode provocar efeitos nas estimativas dos coeficientes de regressão e na aplicabilidade geral do modelo estimado.

Então, para verificar a existência de correlação entre as variáveis independentes e evitar possíveis interpretações equivocadas nas análises foi calculada a matriz de coeficientes de correlação de Pearson.

Na Tabela 6.5 são mostrados os resultados dos coeficientes para o sexo masculino e feminino. No caso masculino, as variáveis de ocupação apresentam alta dependência entre si. Por exemplo, os trabalhadores da construção civil quando foram relacionados com os da administração Pública e do Comércio apresentaram um coeficiente de 0,933 e 0,914 respectivamente a um nível de significância de 0,01, o que indica que essas variáveis não devem estar juntos na construção do modelo. Ao relacionar a taxa de mortalidade por Pneumonia com a taxa de óbito por AIDS foi obtido um coeficiente de Pearson de 0,463 a um nível de significância de 0,01, indicando uma associação entre essas duas variáveis, só que mais baixa que as citadas anteriormente.

As variáveis do FIRJAN, Saúde e Emprego/Renda não mostraram ter relação com nenhuma outra variável selecionada para o estudo. A única que mostrou ter associação com a AIDS foi o índice de Educação.

Como é notado pelos resultados das nove variáveis usadas no estudo todos os índices do FIRJAN e a Extrativa Mineral não evidenciaram uma relação com a taxa de mortalidade por Tuberculose. Isso pode ter ocorrido pelo fato de que esses índices são distribuídos homogeneamente nas mesorregiões do Nordeste.

Também na Tabela 6.5 são apresentados os resultados dos coeficientes de Pearson para as variáveis do sexo feminino. Nota-se que a tuberculose está relacionada à AIDS, Pneumonia, Extrativa Mineral, Construção Civil, Comércio e Administração

Pública. Dessas variáveis citadas existe uma relação significativa entre a AIDS, Construção Civil e Comércio, o que pode comprometer as análises de regressão.

As variáveis do FIRJAN para as mulheres, assim como para os homens, não sugerem relação com a tuberculose. Como o índice FIRJAN é um indicador global, ou seja, inclui o desenvolvimento das mesorregiões para homens e mulheres, a mesma explicação dada para a falta de relação dessas variáveis com a tuberculose para o sexo masculino é válida para as mulheres, já que se esperaria que pelo menos o índice de Educação e Emprego/Renda estivesse associado à tuberculose.

Tabela 6.5: Matriz de Coeficientes de Correlação de Pearson e significância de alguns indicadores de mortalidade e socioeconômicas – Nordeste 2008.

		Tuberculose	AIDS	Pneumonia	Saúde	Emprego/ Renda	Educação	Extrativa Mineral	Construção Civil	Comércio	Administração Pública
		Masculino									
Tuberculose	Correlação	1	,690**	,549**	,077	-,286	,131	,115	,665**	,650**	,568**
	Significância		,000	,000	,627	,066	,410	,469	,000	,000	,000
AIDS	Correlação	,715**	1	,463**	,075	-,083	,306*	,325*	,781**	,690**	,653**
	Significância	,000		,002	,637	,600	,048	,036	,000	,000	,000
Pneumonia	Correlação	,398**	,476**	1	,144	-,247	,248	,056	,753**	,862**	,829**
	Significância	,009	,001		,363	,114	,114	,723	,000	,000	,000
Saúde	Correlação	-,190	,053	,173	1	,262	,289	,268	,165	,119	,219
	Significância	,229	,739	,272		,093	,063	,086	,297	,454	,164
Emprego/ Renda	Correlação	-,056	-,234	-,229	,262	1	,079	,220	-,212	-,226	-,242
	Significância	,725	,135	,144	,093		,618	,162	,177	,149	,122
Educação	Correlação	,215	,238	,272	,289	,079	1	,250	,256	,294	,329*
	Significância	,171	,128	,081	,063	,618		,110	,102	,059	,033
Extrativa Mineral	Correlação	,383*	,498**	,364*	,261	,002	,297	1	,303	,280	,261
	Significância	,012	,001	,018	,095	,992	,056		,051	,072	,095
Construção Civil	Correlação	,505**	,719**	,703**	,205	-,207	,239	,628**	1	,914**	,933**
	Significância	,001	,000	,000	,192	,189	,128	,000		,000	,000
Comércio	Correlação	,535**	,674**	,854**	,134	-,225	,299	,582**	,919**	1	,916**
	Significância	,000	,000	,000	,398	,152	,054	,000	,000		,000
Administração Pública	Correlação	,387*	,580**	,890**	,199	-,216	,322*	,491**	,847**	,914**	1
	Significância	,011	,000	,000	,205	,169	,038	,001	,000	,000	

Fonte: Resultados dos coeficientes de Pearson.

** . Correlação é significativa para o nível 0.01 (bilateral).

* . Correlação é significativa para o nível 0.05 (bilateral).

6.5.2 TRANSFORMAÇÕES DOS ÍNDICES DO FIRJAN

Os índices do FIRJAN medem o nível de desenvolvimento de uma determinada região indicando à qualidade de vida da população que habita essa região, no caso do trabalho, as mesorregiões do Nordeste. Em virtude disso era de se esperar que elas estivessem associadas com a incidência de Tuberculose. Mas, como foi observado pela Tabela 6.5 essas variáveis não mostraram ter relação linear com a Tuberculose. Então, transformações foram feitas nessas variáveis com a tentativa de contornar esse problema.

As seguintes transformações foram usadas: a logarítmica, raiz quadrada, arco seno e com quociente $1/x$. Após aplicação dessas transformações apenas a variável Emprego/Renda para o sexo masculino mostrou significância com a Tuberculose.

Com a transformação logarítmica a variável Emprego/Renda passou a ter um R de Pearson igual -0,325 e um p-valor de 0,036 significativo a um nível $\alpha = 0,05$.

6.5.3 TESTE DE NORMALIDADE NOS RESÍDUOS DAS VARIÁVEIS CORRELACIONADAS

Como foi observado pela matriz de correlação de Pearson algumas variáveis apresentaram correlação entre si tanto para homens como para mulheres. Então, estas variáveis não poderiam entrar no modelo juntas, pois, causariam interpretações erradas nas análises dos resultados.

Para confirmar se entre estas variáveis existem relações lineares foram aplicados os testes de normalidade de Kolmogorov-Smirnov e de Shapiro-Wilk nos resíduos gerados pelos modelos entre elas. Na Tabela 6.6 são encontrados os resultados dos testes da variável AIDS em relação às outras variáveis que também obtiveram relação significativa com a tuberculose.

Como é notado, dos resíduos testados os que obtiveram p-valores superiores ao nível de significância de 5% tanto para o teste de Kolmogorov-Smirnov como para o de Shapiro-Wilk (indicando normalidade) foram os gerados pelos modelos da variável AIDS versus Construção Civil e Administração Pública, confirmando que estas variáveis apresentaram uma relação linear com a AIDS. Na mesma tabela também se

encontra os testes dos resíduos gerados pelo modelo entre a Pneumonia e a variável Comércio que mostrou não ter normalidade.

A partir dessas análises foi concluído que para a construção do modelo de regressão múltipla para o sexo masculino seriam usadas as variáveis: AIDS, Pneumonia, Emprego/Renda e Comércio.

A variável AIDS foi escolhida como indispensável na construção do modelo pelo fato dela aparecer na literatura como sendo fator muito associado à Tuberculose. Então, por este motivo a AIDS foi usado como referencia nos testes de normalidade.

Tabela 6.6: Resultado do teste de normalidade para o sexo masculino.

Variáveis		Kolmogorov-Smirnov			Shapiro-Wilk		
		Estatística	df	Sig.	Estatística	df	Sig.
AIDS	Pneumonia	0,149	42	0,021	0,911	42	0,003
	Construção Civil	0,119	42	0,150	0,967	42	0,251
	Comércio	0,162	42	0,007	0,934	42	0,017
	Administração Pública	0,133	42	0,059	0,951	42	0,071
Pneumonia	Comércio	0,138	42	0,044	0,945	42	0,043

Fonte: Teste de normalidade de Kolmogorov-Smirnov e Shapiro-Wilk.

Na Tabela 6.7 se encontram os resultados dos testes de normalidade de Kolmogorov-Smirnov e Shapiro-Wilk para o sexo feminino. O teste de Kolmogorov apresentou p-valores acima da significância de 5% para a variável AIDS relacionada com as variáveis Extrativa Mineral, Construção Civil e Administração Pública. Como o teste de Kolmogorov-Smirnov foi escolhido como referencia na escolha da normalidade dos resíduos confirmou-se que estas variáveis apresentaram uma relação linear com a AIDS.

A Pneumonia versus a variável Comércio pelo teste de Shapiro-Wilk obteve um p-valor superior ao nível de significância de 5% mas no teste de Kolmogorov este valor foi abaixo deste nível, daí concluiu-se que não houve normalidade nos resíduos.

A partir do que foi observado pelos testes de normalidade para o sexo feminino as variáveis que foram usadas na construção do modelo de regressão linear múltipla foram: AIDS, Pneumonia e Comércio.

A AIDS para o sexo feminino, assim como para o masculino foi escolhido como variável referencia por aparecer como fator de grande influencia na incidência da Tuberculose.

Tabela 6.7: Resultado do teste de normalidade para o sexo feminino.

Variáveis		Kolmogorov-Smirnov			Shapiro-Wilk		
		Estatística	df	Sig.	Estatística	df	Sig.
AIDS	Pneumonia	0,145	42	0,026	0,890	42	0,001
	Estrativa Mineral	0,095	42	0,200	0,938	42	0,024
	Construção Civil	0,126	42	0,091	0,931	42	0,014
	Comércio	0,138	42	0,042	0,913	42	0,004
	Administração Pública	0,124	42	0,100	0,922	42	0,007
Pneumonia	Comércio	0,148	42	0,021	0,950	42	0,062

Fonte: Teste de normalidade de Kolmogorov-Smirnov e Shapiro-Wilk.

Para avaliar a aleatoriedade dos resíduos nos modelos que atenderam o suposto de normalidade foram feitos os gráficos de resíduos versus valores preditos. Eles mostraram-se aleatórios tendo uma concentração de resíduos mais para os valores pequenos. Esses gráficos são encontrados no apêndice A.

6.5.4 ANÁLISE DAS VARIÁVEIS OCUPACIONAIS ATRAVÉS DE REGRESSÃO SIMPLES

As variáveis ocupacionais, exceto Comércio, não entraram na análise de regressão múltipla por terem apresentado relações entre si e também com as demais variáveis usadas no estudo. Mas, como foi observado pela Tabela 6.5 da matriz de coeficientes de correlação de Pearson elas apresentaram p-valores que demonstram significância com a Tuberculose indicando que possivelmente, através de uma regressão linear simples, poderia ser encontrado um bom ajuste do modelo e, com isso, indicar que elas têm relação com a Tuberculose.

Como o objetivo do trabalho é verificar se existe relação da taxa de mortalidade de Tuberculose com as variáveis selecionadas para o estudo foi feita também uma análise destas que não foram usadas na análise múltipla para se comprovar os resultados obtidos na matriz de correlação de Pearson. A Tabela 6.8 mostra os resultados do R de Pearson, R^2 e p-valor para as variáveis: Construção Civil, Comércio e Administração Pública para homens e mulheres e Extrativa Mineral para o sexo feminino.

Tabela 6.8: R de Pearson, R^2 e p-valor para as variáveis de Construção Civil, Comércio e Administração Pública por sexo e Extrativa Mineral para as mulheres.

Variáveis ocupacionais	R		R^2		p-valor	
	Masculino	Feminino	Masculino	Feminino	Masculino	Feminino
Extrativa Mineral	-	0,383	-	0,126	-	0,012
Construção Civil	0,665	0,505	0,443	0,255	0,000	0,001
Comércio	0,650	0,535	0,423	0,286	0,000	0,000
Administração Pública	0,568	0,387	0,323	0,150	0,000	0,011

Fonte: Matriz de Correlação de Pearson (Tabela 6.5).

6.5.5 ANÁLISE DE RESÍDUOS DA TUBERCULOSE EM RELAÇÃO ÀS VARIÁVEIS OCUPACIONAIS

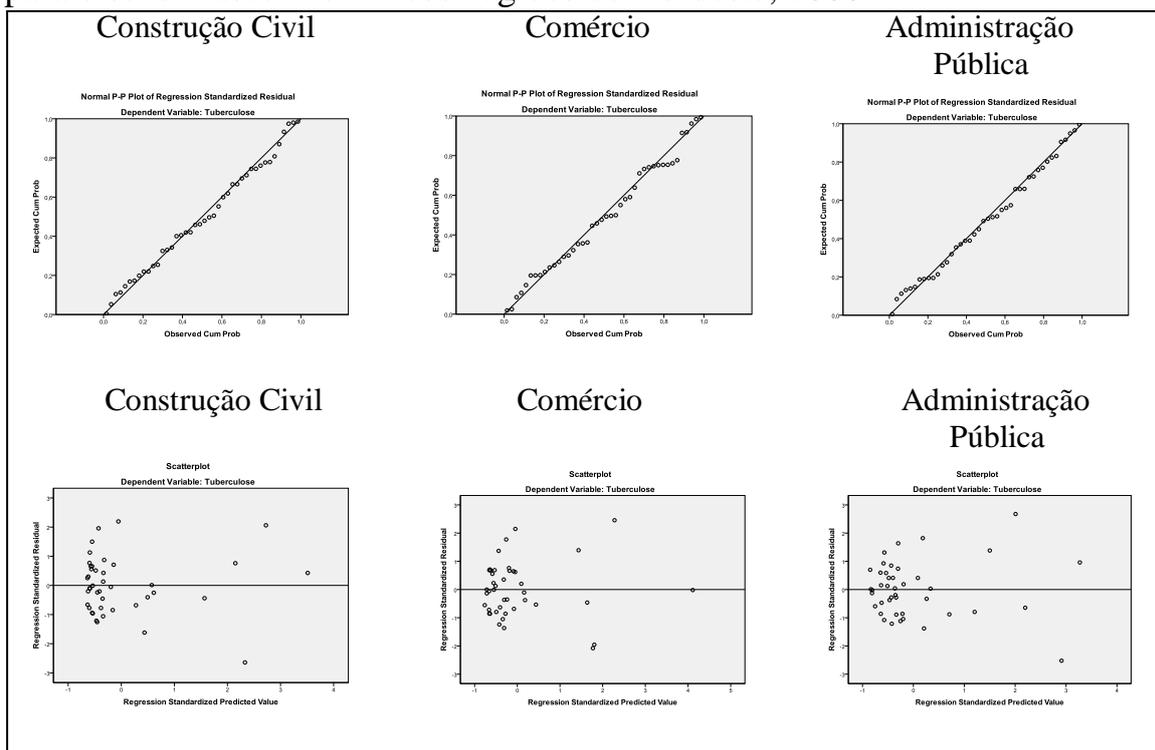
Para validar os modelos gerados pelas regressões simples foram feitas as análises de resíduos.

As Figuras 6.3 e 6.4 mostram os resultados dos PP-plot's que não fornece qualquer indicação que contrarie o pressuposto de normalidade dos resíduos tanto para os indivíduos do sexo masculino das mesorregiões do Nordeste como os do sexo feminino.

Os gráficos de dispersão dos resíduos em função dos valores preditos mostraram-se aleatórios tendo uma concentração de resíduos mais para os valores pequenos tanto para homens como para mulheres.

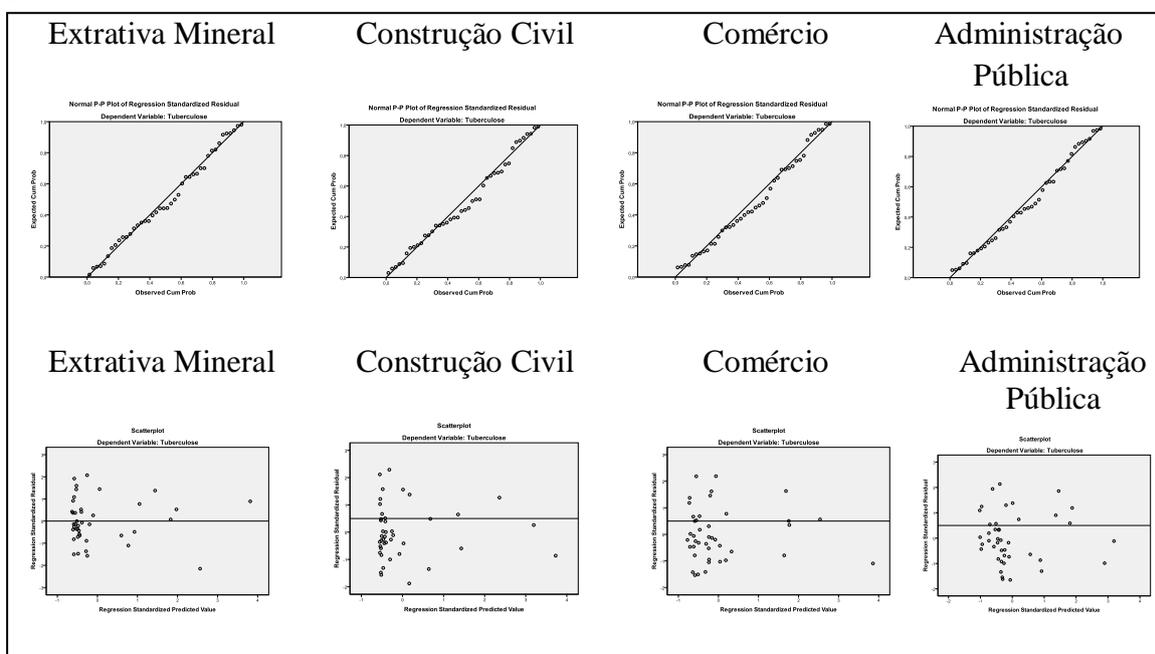
Tendo sido feita essas observações conclui-se que os resultados do ajustamento dos modelos de regressão linear simples são confiáveis e indica que a Tuberculose tem associação com essas variáveis.

Figura 6.3: Gráficos para as análises de resíduos das variáveis de ocupação para o sexo masculino – mesorregiões do Nordeste, 2008.



Fonte: Elaboração própria.

Figura 6.4: Gráficos para as análises de resíduos das variáveis de ocupação para o sexo feminino - mesorregiões do Nordeste, 2008.



Fonte: Elaboração própria.

6.5.6 MODELO DE REGRESSÃO LINEAR MÚLTIPLA

Após a eliminação das variáveis que causariam resultados distorcidos por terem um nível de associação elevado entre si foram efetuadas as análises através da regressão múltipla. Em um primeiro momento foi verificado que a variável Comércio e Pneumonia teriam de ser retiradas do modelo, pois, elas obtiveram uma significância de 0,804 e 0,414 respectivamente que são valores acima do nível de 5% usado no trabalho. Como a variável Comércio obteve um valor mais alto, ela foi retirada do modelo.

Na segunda verificação, o p-valor da variável Pneumonia baixou para 0,065 e a variável Emprego/Renda obteve um p-valor de 0,068, valores estes ainda acima do nível de significância adotado no trabalho. Foi observado ainda que nesta segunda análise a constante obteve um p-valor de 0,245.

Ao se retirar a constante do modelo na terceira análise, chegou-se aos resultados da Tabela 6.9 para o sexo masculino das regiões do Nordeste:

Na Tabela 6.9 pode-se observar que o modelo constituído pela AIDS, Pneumonia e Emprego/Renda juntas fornecem uma relação com a Tuberculose de 0,940 como é constatado pelo valor do R de Pearson. O R^2 mostra que o conjunto destas variáveis fornece uma associação de 88,4% com a Tuberculose.

Na mesma tabela verifica-se o modelo gerado através da análise de regressão. Os coeficientes de declividade forneceram a relação esperada entre as variáveis que ficaram no estudo com a Tuberculose. A variável Emprego/Renda, por exemplo, obteve um valor de -2,184 indicando que existe uma relação inversa com a Tuberculose, ou seja, quanto mais baixo for o índice deste indicador nas regiões estudadas mais elevado será o nível da incidência dos indivíduos com a Tuberculose. Esse valor do coeficiente revela que com a diminuição de 1% no índice de Emprego/Renda a taxa de mortalidade por Tuberculose para os homens aumentaria em 2,184 habitantes para cada 100.000.

As variáveis AIDS e Pneumonia mostraram uma relação positiva com a Tuberculose, o que também era esperado, pois, se se aumenta os níveis de pessoas com a AIDS ou Pneumonia a incidência de pessoas com Tuberculose também aumenta. O coeficiente de declividade gerado pela relação da Tuberculose com a AIDS foi um pouco mais alto que o da Pneumonia. Com o aumento de 1% na taxa de mortalidade

por AIDS e Pneumonia a incidência da Tuberculose aumentaria respectivamente em torno de 0,285 e 0,118 por 100.000 hab. do sexo masculino.

Tabela 6.9: Modelo ajustado para o sexo masculino.

	R	R²	R² ajustado	F	sig.
	0,940	0,884	0,875	99,26	0,000
Modelo	Coefficientes não padronizados	Coefficientes padronizados	Signif.	Intervalo de confiança de B com 95%	
	B	beta		Limite inferior	Limite superior
AIDS	0,285	0,512	0,000	0,159	0,410
Pneumonia	0,118	0,201	0,040	0,006	0,231
Emprego/Renda	-2,184	-0,287	0,010	-3,81	-0,558

Fonte: Resultados da regressão.

A Tabela 6.10 é referente aos resultados da análise de regressão para o sexo feminino. As etapas que antecederam ao modelo final contaram com as variáveis Comércio e Pneumonia, pois, assim como nas análises realizadas para os homens elas foram as únicas que não apresentaram correlações entre si. Fato este confirmado pelos testes de normalidade de Kolmogorov-Smirnov e Shapiro-Wilk nos seus resíduos.

Na primeira tentativa de gerar o modelo foi observado que as variáveis Comércio e Pneumonia obtiveram p-valores iguais 0,781 e 0,921 respectivamente. Valores muito acima do nível de significância adotado no trabalho que foi de 5%. Tendo sido feita essa observação a variável Pneumonia foi retirada do modelo.

O segundo modelo gerado mostrou que a variável Comércio ainda obteve um p-valor acima do nível de significância de 5% sendo necessária a sua retirada do modelo.

Sendo assim, a AIDS foi a única variável a ficar no modelo para o sexo feminino. Na Tabela 6.10 pode-se observar o valor do R de Pearson que indica uma relação da Tuberculose com a AIDS de 0,715 e o R² de 0,511 determinando 51,1% de associação entre elas.

Na mesma tabela é apresentado coeficiente de declividade do modelo, indicando uma relação positiva entre as variáveis. Esse resultado indica que ao aumentar em 1% a

taxa de mortalidade por AIDS para o sexo feminino a taxa de mortalidade por Tuberculose aumentaria em 0,282 para cada 100.000 mulheres.

Como foi notado, tanto para homens como para mulheres, a AIDS foi a variável que apareceu nos dois modelos mostrando ser um fator que merece uma atenção mais detalhada dos órgãos de saúde e da administração pública.

Tabela 6.10: Modelo ajustado para o sexo feminino.

	R	R²	R² ajustado	F	sig.
	0,715	0,511	0,499	41,773	0,000
Modelo	Coefficientes não padronizados	Coefficientes padronizados	Signif.	Intervalo de confiança de B com 95%	
	B	beta		Limite inferior	Limite superior
constante	0,522		0,005	0,163	0,881
AIDS	0,282	0,715	0,000	0,194	0,370

Fonte: Resultados da regressão.

6.5.7 ANÁLISE DE RESÍDUOS

Tanto na Regressão Linear Simples quanto na Regressão Múltipla, as suposições do modelo ajustado precisam ser validadas para que os resultados sejam confiáveis. As análises dos resíduos, como já mencionado, são um conjunto de técnicas utilizadas para investigar a adequabilidade de um modelo de regressão com base nos resíduos.

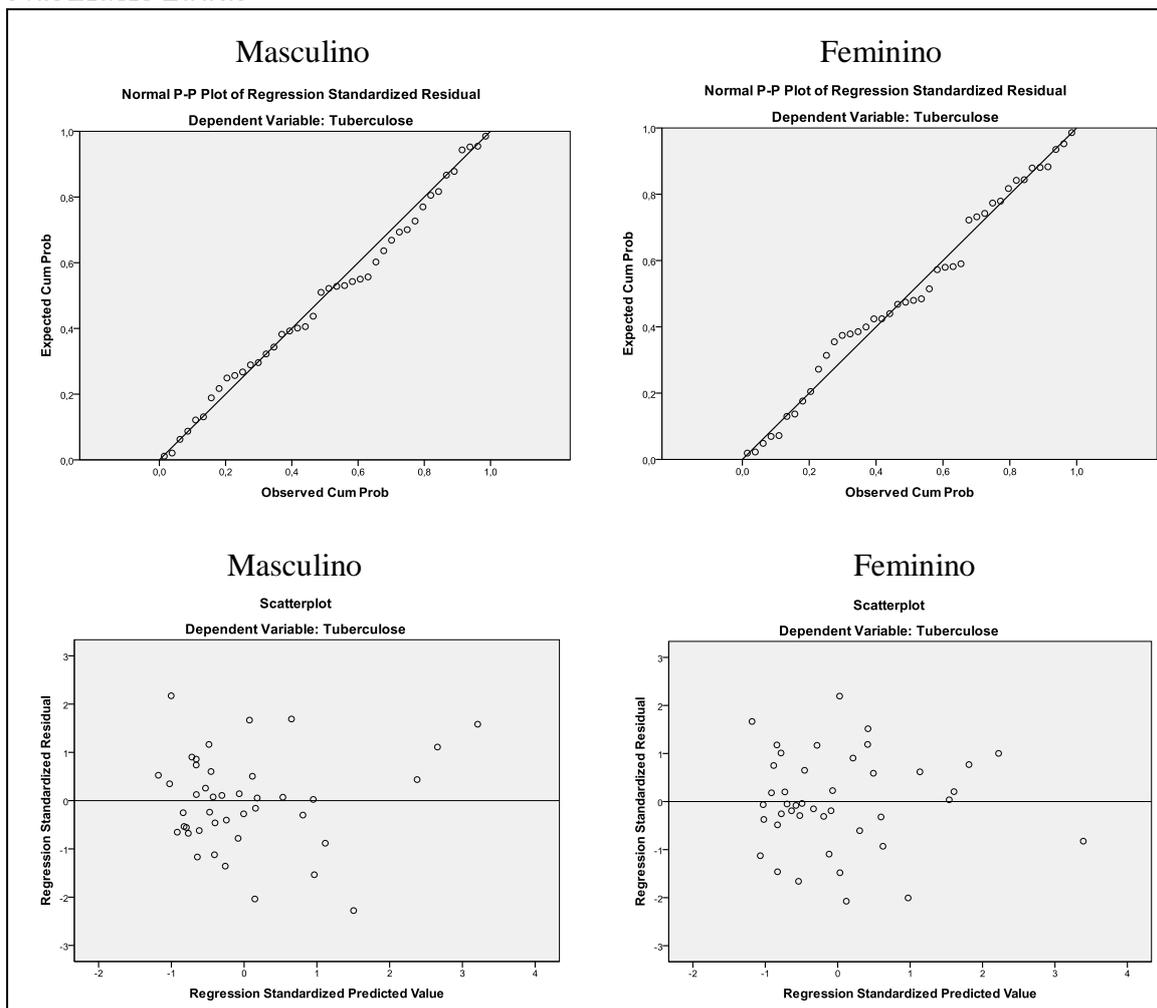
Então, para validar se os resíduos são independentes e identicamente distribuídos (são aleatórios e com variância constante) foram construídos os gráficos de resíduos versus valores preditos. A Figura 6.5 mostra os resultados das análises de resíduos feitas para os modelos mencionados anteriormente.

Os PP-plot's não fornece qualquer indicação que contrarie o pressuposto de normalidade dos resíduos tanto para os indivíduos do sexo masculino das mesorregiões do Nordeste como para os do sexo feminino.

Os gráficos de dispersão dos resíduos em função dos valores preditos estandardizados mostram-se bastante aleatórios, o que satisfaz às condições.

A partir desses resultados, pode-se concluir que houve evidências satisfatórias para os resultados dos ajustamentos.

Figura 6.5: Gráficos para as análises de resíduos dos modelos gerados através da regressão múltipla para ambos os sexos – mesorregiões do Nordeste. 2008.



Fonte: Elaboração própria.

7 CONSIDERAÇÕES FINAIS

As análises feitas no estudo, através da análise de regressão, evidenciaram uma relação da Tuberculose com algumas das variáveis usadas no estudo. Para o modelo ajustado pela regressão múltipla para o sexo masculino a Tuberculose mostrou ter 47,61% de relação com a AIDS. Em relação à Pneumonia, obteve-se um valor de 30,14% de associação. Quando essa comparação foi feita com o índice de Emprego/Renda a diferença aumenta ainda mais. A variável Emprego/Renda adquiriu um valor de 10,56% de associação com a morte por Tuberculose mostrando uma diferença de mais de 100% de contribuição na incidência da Tuberculose.

Para o sexo feminino só a AIDS continuou no modelo ajustado por regressão múltipla sendo responsável por 51,12% dos casos de Tuberculose. A Pneumonia através de regressão simples mostrou associação com a mortalidade por Tuberculose para as mulheres com um valor de 15,84% de associação.

Esses resultados para o Nordeste confirmam o que a literatura vem mostrando a respeito da forte relação que existe dos indivíduos portadores com o vírus da AIDS e da incidência da Tuberculose.

Com os resultados obtidos pelas regressões simples para as variáveis dos trabalhadores da Construção Civil, Comércio e Administração Pública chegaram-se aos valores de 44,22%, 42,25% e 32,26% de relação com a incidência da Tuberculose para o sexo masculino. Os percentuais para as mulheres foram de 14,67% da Extrativa Mineral, 25,50% na Construção Civil, 28,62 no Comércio e 14,98% da Administração Pública associadas com a Tuberculose.

Os índices do FIRJAN não mostraram ter relação com a incidência da taxa de mortalidade por Tuberculose no Nordeste para ambos os sexos exceto Emprego/Renda para os homens. Um dos motivos para que a Saúde e a Educação não tenham mostrado associação com a Tuberculose nesse estudo pode ser pelo fato de que os níveis da educação e da saúde das mesorregiões foram tão uniformes que provavelmente não houve variabilidade para discriminar diferenças entre elas, de modo que inviabilizaram uma relação linear de dependência com a Tuberculose.

Certamente que outros modelos poderiam ser testados como o MLG ou o Beta. No entanto eles não foram alvo desse trabalho, ficando esta sugestão para trabalhos posteriores.

Embora, se conheçam as limitações de estudos ecológicos, ainda não houve registros na literatura de trabalhos dessa natureza para o Nordeste. Espera-se que os resultados do censo 2010 possibilitem novas explorações, aportes e atualizações.

De forma geral, a regressão linear clássica mostrou-se ser uma boa ferramenta estatística na análise dos dados propostos neste trabalho. Mostraram-se as possíveis relações, seja na análise simples ou múltipla, das variáveis estudadas com a Tuberculose.

Os modelos propostos neste trabalho possibilitaram investigar fatores importantes para entender melhor os condicionantes da mortalidade por Tuberculose. Estas análises podem auxiliar nas políticas de atenção primária à saúde do nordestino e nortear políticas de prevenção no que se diz respeito à mortalidade por Tuberculose.

REFERÊNCIAS

- BARTLETT, J.G. **Tuberculosis and HIV Infection: Partners in Human Tragedy.** In: Friedland G, Churchyard GJ, Nardell E, editors. Tuberculosis and HIV Coinfection: Current State of Knowledge and Research Priorities. *Infect Dis J* 2007; 196 (S124).
- BRASIL. Ministério da Saúde. **Sistema de Informação de Agravos de Notificação.** Área Técnica de Pneumonia Sanitária. Brasília: 2004.
- BRASS, W. (1975), **Methods for Estimating Fertility and Mortality from Limited and Defective Data, Occasional Publication.** International Program of Laboratories for Population Statistics, Chapel Hill.
- COCK, K.M. **HIV infection, tuberculosis and World AIDS Day, 2006.** *Int J Tuberc Lung Dis* 2006; 10:1305.
- COOK, R.D. **Detection of influential observations in Linear Regression** *Technometrics*, 19. 15-18, 1977.
- DALEY, C.L. et al. **An outbreak of tuberculosis with accelerated progression among persons infected with the human immunodeficiency virus.** *N. Engl. J. Med.*, 326:231-5, 1992.
- DIPERRI, G. et al. **Nosocomial epidemic active tuberculosis among HIV- infected patients.** *Lancet*, 2:1502-4, 1992.
- DOENÇAS INFECCIOSAS E PARASITÁRIAS: **GUIA DE BOLSO. Secretaria de Vigilância em Saúde,** Ministério da Saúde. Brasília: Ministério da Saúde; 2005. 320 pp. ISBN: 85-334-1048-4.
- FEDERAÇÃO DAS INDÚSTRIAS DO ESTADO DO RIO DE JANEIRO. Rio de Janeiro: 2010. Disponível em: <<http://www.firjan.org.br>>. Acesso entre: 1 e 30/09/2011.
- HIJJAR, M.A. **Controle das doenças endêmicas no Brasil - tuberculose.** *Rev. Soc. Bras. Med. Trop.*, 27(Supl. 3): 23-36, 1994.
- HIJJAR, M.A. **Tuberculose: problema atual.** *JBM*, 66: 123-45, 1994.
- KMENTA, J. **Elementos de econometria.** São Paulo: Atlas, 1978.
- LINDGREN, B.W. (1976). **Statistical Theory. 3rd ed., Mac Millan Publ. Inc. Co.,** New York.
- MINISTÉRIO DA SAÚDE. **Programa Nacional de Doenças Sexualmente Transmissíveis.** *Bol. Epidemiol. AIDS*, 8 (3), 1995.

MINISTÉRIO DA SAÚDE – **Sistema de Informação de Mortalidade** – SIM , 2008. Disponível em: <<http://www2.datasus.gov.br/DATASUS/index.php>>. Acesso entre: 1 e 30/09/2011.

MINISTÉRIO DA SAÚDE – **Sistema de Informação de Mortalidade** – SIM , 2008. Disponível em: <http://portal.saude.gov.br/portal/arquivos/pdf/taxa_mortalidade_tuberculose.pdf >. Acesso entre: 07/12/2011.

MINISTÉRIO DO TRABALHO E EMPREGO, 2008. Disponível em: <<http://portal.mte.gov.br/portal-mte/>>. Acesso entre: 1 e 30/09/2011.

ORGANIZAÇÃO MUNDIAL DE SAÚDE. Stop TB e HIV/SIDA. Quadro geral estratégico para reduzir o peso da TB/HIV. Geneva; 2002.

PABLOS-MENDES A. **Anti-tuberculosis drug resistance surveillance** in 35 countries. Int J Tberc Lung Dis 1997; S18.

PAES, N.A. e MARINHO, P.R.D. **Cenários regionais dos eventos vitais do Nordeste: Brechas redutíveis da mortalidade**. Relatório de Pesquisa PIBIC/CNPq/UFPB, 2009.

SEARLE, S. R. **Linear Models for Unbalanced Data** : Wiley Series in Probability and Statistics, 1987.

SEBER G, A.F. e LEE, A.J. **Linear Regression Analysis, (2nd Ed.)**: Wiley Series in Probability and Statistics, 1977.

SISTEMA DE INFORMAÇÃO DE AGRAVOS DE NOTIFICAÇÃO (SINAN), 2010. Disponível em: < http://portal.saude.gov.br/portal/arquivos/pdf/incidencia_tabela2.pdf>. Acesso entre: 1 e 30/11/2011.

SUDRE, P. et al. **Tuberculosis in the present time: a global overview of the tuberculosis situation**. Geneva, World Health Organization, 1991. (Documento WHO/TUB/ 91. 158).

VALLIN, J.(1987), **Seminário sobre causas de muerte: Aplicación AL caso de Francia, INED** – Instituto Nacional Del Estudios Demográficos e CELADE, Série E, Nº 31, Santiago, Chile, Nov.

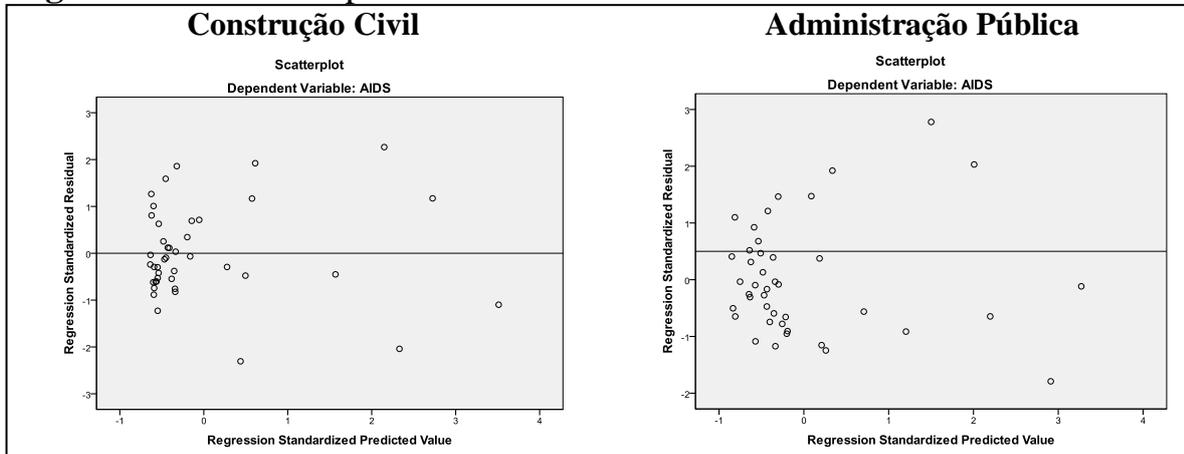
WORLD HEALTH ORGANIZATION. **Global Tuberculosis Control**, WHO Report 2002; Geneva: WHO: 1-167.

WORLD HEALTH ORGANIZATION. **Global Programme on AIDS. The HIV/AIDS pandemic: 1993 overview**. Geneva,1993. (Documento WHO/GPA/CNP/EUA/93.1).

APÊNDICE A

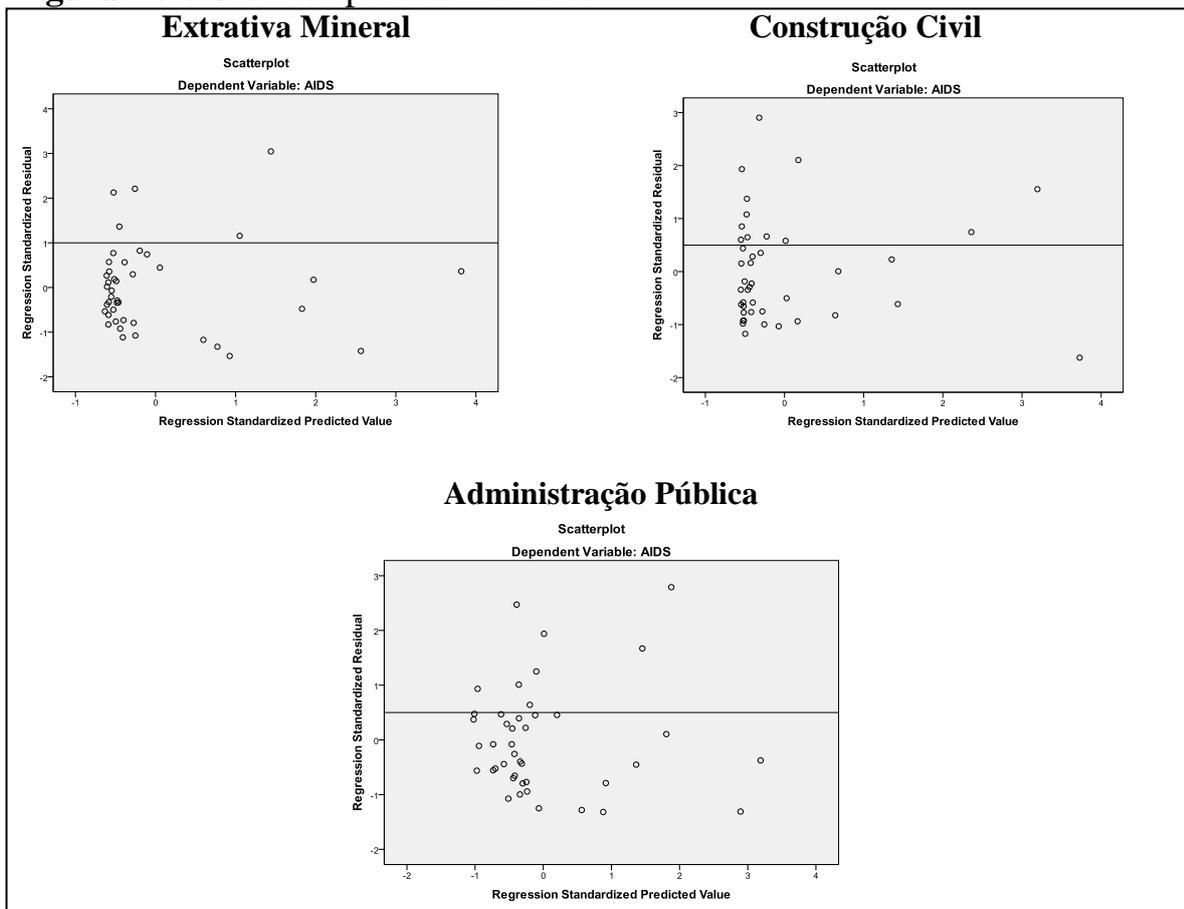
GRÁFICOS DE RESÍDUOS PADRONIZADOS VERSOS VALORES PREDITOS PARA AMBOS OS SEXOS SEGUNDO A AIDS EM RELAÇÃO AS VARIÁVEIS OCUPACIONAIS.

Figura A.1: Gráficos para os homens.



Fonte: Elaboração própria.

Figura A.2: Gráficos para as mulheres.



Fonte: Elaboração própria.

ANEXO A

TAXAS DE ÓBITO PARA ADULTOS, ÍNDICE FIRJAN E TAXAS DE OCUPAÇÃO PARA O SEXO MASCULINO, MESORREGIÕES DO NORDESTE – 2007/2008.

Tabela A: Base de Dados para o sexo masculino.

Mesorregiões	Taxa de Óbito/100000 hab. (Ministério da Saúde/2008)			(Firjan/2007)			Taxa de Ocupação/100000 hab. (Ministério do Trabalho/2008)			
	Tuberculose	AIDS	Pneumonia	Saúde	Emprego/Renda	Educação	Extrativa Mineral	Construção Civil	Comércio	Administração Pública
Agreste Alagoano	4,32	3,47	3,93	0,65	0,27	0,57	32,71	128,55	2329,04	2072,91
Agreste Paraibano	2,49	5,01	3,20	0,71	0,27	0,58	66,77	495,64	2160,65	2469,77
Agreste Pernambucano	3,51	5,14	4,16	0,69	0,24	0,59	19,26	353,23	1875,86	1883,76
Agreste Potiguar	3,47	2,80	3,74	0,73	0,28	0,57	8,88	186,00	875,79	3066,19
Agreste Sergipano	0,79	8,49	2,08	0,70	0,36	0,58	53,20	407,71	1903,04	2121,23
Borborema	1,34	2,64	3,33	0,74	0,30	0,61	240,73	106,70	584,16	3028,85
Central Potiguar	1,64	3,48	2,62	0,76	0,28	0,60	843,26	552,55	2023,92	2974,22
Centro Maranhense	2,89	3,46	4,30	0,64	0,26	0,60	11,08	35,29	1154,46	1265,74
Centro-Norte Baiano	3,02	6,22	3,79	0,60	0,29	0,58	484,92	930,15	3737,15	2373,52
Centro-Sul Baiano	3,11	2,93	7,99	0,73	0,29	0,67	352,08	638,69	4690,78	4905,49
Centro-Norte Piauiense	2,45	8,01	17,12	0,64	0,25	0,65	128,97	6062,87	9677,91	15698,93
Centro-Sul Cearense	2,99	6,07	2,99	0,74	0,56	0,66	32,92	67,34	980,10	1775,03
Extremo-Oeste Baiano	2,18	3,98	3,37	0,62	0,26	0,67	35,06	613,96	2767,69	1943,01
Metropolitana Fortaleza	8,86	13,88	17,86	0,73	0,27	0,66	294,12	8462,99	18419,19	17163,89
Jaguaribe	1,14	5,89	2,37	0,77	0,32	0,68	29,08	240,33	606,66	1243,97
Leste Alagoano	6,72	7,55	8,56	0,71	0,28	0,65	86,39	1219,39	2887,24	4603,07
Leste Maranhense	3,75	6,63	4,19	0,63	0,26	0,68	19,87	113,78	1074,47	1666,50
Leste Potiguar	4,84	9,60	6,97	0,74	0,27	0,70	279,47	4515,04	9178,29	12793,66
Leste Sergipano	3,82	12,52	3,10	0,72	0,31	0,70	563,97	2574,05	3410,13	5220,63
Zona da Mata Paraibana	3,42	6,30	5,99	0,73	0,30	0,51	56,10	2337,87	3654,34	8750,40
Mata Pernambucana	5,84	4,97	7,79	0,74	0,29	0,64	26,52	460,30	2086,37	2633,47
Nordeste Baiano	3,64	3,58	1,87	0,59	0,28	0,60	89,59	212,37	663,13	1216,65
Noroeste Cearense	2,62	3,32	3,62	0,75	0,32	0,59	46,42	236,78	1006,65	2411,19
Norte Cearense	1,47	3,97	1,28	0,72	0,36	0,65	9,69	41,24	210,40	788,62
Norte Maranhense	4,18	10,56	3,45	0,60	0,29	0,64	24,73	2496,32	2371,84	4208,19
Norte Piauiense	0,84	4,24	1,58	0,66	0,32	0,60	10,09	378,09	1483,36	1514,90
Oeste Maranhense	2,69	6,09	17,25	0,61	0,33	0,59	72,31	1892,73	9799,32	6727,70
Oeste Potiguar	1,83	5,32	1,54	0,75	0,60	0,61	735,71	1001,39	1551,63	1306,11
Metropolitana do Recife	10,33	17,06	16,51	0,78	0,25	0,63	109,29	6865,14	11573,73	12023,86
São Francisco Pernambucano	4,28	7,23	2,30	0,63	0,27	0,52	13,48	1037,91	2723,49	1461,69
Metropolitana de Salvador	7,51	18,00	6,32	0,67	0,34	0,61	636,14	5695,38	8420,89	9959,91
Sertão Alagoano	3,61	2,75	3,67	0,63	0,30	0,53	14,68	164,21	713,24	2214,04
Sertão Paraibano	1,45	1,65	4,43	0,70	0,28	0,47	15,86	2219,14	1219,54	2483,85
Sertão Pernambucano	2,41	2,37	0,72	0,65	0,38	0,64	33,69	129,92	383,95	449,88
Sertão Sergipano	1,15	3,02	1,64	0,67	0,35	0,65	22,10	215,22	666,13	2811,83
Sertões Cearense	1,25	3,08	8,58	0,73	0,38	0,60	18,22	633,52	1790,15	4701,01
Sudeste Piauiense	2,39	2,01	1,29	0,68	0,29	0,52	9,02	121,17	585,53	557,31
Sudoeste Piauiense	4,96	1,31	3,79	0,64	0,31	0,50	52,37	214,79	1425,00	1526,33
Sul Baiano	4,29	9,55	3,57	0,62	0,33	0,53	40,01	676,44	2869,46	2618,89
Sul cearense	2,38	4,38	4,99	0,75	0,40	0,58	92,57	415,45	1913,67	2078,40
Sul Maranhense	3,58	5,04	1,61	0,67	0,31	0,51	19,61	648,26	607,42	392,41
Vale São Francisco da Bahia	2,20	7,18	0,90	0,59	0,30	0,50	11,44	59,87	392,17	540,23

Fonte: Ministério da Saúde, Ministério do Trabalho e Emprego e Firjan.

ANEXO B

TAXAS DE ÓBITO PARA ADULTOS, ÍNDICE FIRJAN E TAXAS DE OCUPAÇÃO
PARA O SEXO FEMININO, MESORREGIÕES DO NORDESTE – 2007/2008.

Tabela B: Base de Dados para o sexo feminino.

Mesorregiões	Taxa de Óbito/100000 hab. (Ministério da Saúde/2008)			(Firjan/2007)			Taxa de Ocupação/100000 hab. (Ministério do Trabalho/2008)			
	Tuberculose	AIDS	Pneumonia	Saúde	Emprego/Renda	Educação	Extrativa Mineral	Construção Civil	Comércio	Administração Pública
Agreste Alagoano	0,64	1,51	2,52	0,65	0,27	0,57	7,57	2,21	1331,93	4008,41
Agreste Paraibano	0,24	3,67	2,08	0,71	0,27	0,58	7,35	15,82	1128,96	4450,64
Agreste Pernambucano	1,60	2,36	2,49	0,69	0,24	0,59	1,66	24,60	1316,89	4098,81
Agreste Potiguar	0,02	1,51	1,90	0,73	0,28	0,57	0,95	7,62	544,72	4741,04
Agreste Sergipano	1,29	4,82	0,92	0,70	0,36	0,58	9,17	14,67	1324,58	4439,87
Borborema	0,08	0,97	0,00	0,74	0,30	0,61	8,02	6,01	421,02	6000,48
Central Potiguar	0,56	3,47	3,60	0,76	0,28	0,60	67,39	44,24	1321,05	4963,19
Centro Maranhense	2,89	3,46	2,28	0,64	0,26	0,60	1,14	1,82	616,38	3060,01
Centro-Norte Baiano	0,82	1,63	3,48	0,60	0,29	0,58	29,62	83,07	2267,31	5027,75
Centro-Sul Baiano	1,14	1,74	4,59	0,73	0,29	0,67	25,92	47,86	2587,83	9334,89
Centro-Norte Piauiense	1,66	4,76	10,72	0,64	0,25	0,65	11,10	339,54	5531,72	21699,67
Centro-Sul Cearense	1,62	1,62	1,11	0,74	0,56	0,66	0,00	6,68	620,55	4126,11
Extremo-Oeste Baiano	1,30	3,19	0,71	0,62	0,26	0,67	2,48	26,60	1490,67	3931,15
Metropolitana Fortaleza	2,49	6,90	11,01	0,73	0,27	0,66	35,49	731,08	10488,04	23260,51
Jaguaribe	1,14	2,28	1,62	0,77	0,32	0,68	3,24	9,72	475,03	2447,39
Leste Alagoano	1,29	4,09	3,77	0,71	0,28	0,65	5,18	57,09	1676,41	5713,70
Leste Maranhense	1,58	3,24	3,56	0,63	0,26	0,68	0,46	6,03	535,53	3503,16
Leste Potiguar	2,60	5,98	7,05	0,74	0,27	0,70	54,93	326,39	5816,23	15929,75
Leste Sergipano	2,17	4,52	1,91	0,72	0,31	0,70	51,89	211,38	2298,50	7413,40
Zona da Mata Paraibana	0,71	3,13	3,28	0,73	0,30	0,51	3,00	204,83	1942,64	11210,80
Mata Pernambucana	2,50	4,34	3,62	0,74	0,29	0,64	2,21	13,88	1395,07	5297,10
Nordeste Baiano	1,69	1,49	0,95	0,59	0,28	0,60	4,90	23,85	449,35	2610,12
Noroeste Cearense	1,01	1,32	4,08	0,75	0,32	0,59	2,87	7,71	575,79	5086,45
Norte Cearense	0,95	1,95	0,94	0,72	0,36	0,65	0,55	1,60	126,96	1336,32
Norte Maranhense	2,64	7,05	2,21	0,60	0,29	0,64	2,29	125,32	1465,69	6399,74
Norte Piauiense	1,00	1,81	1,59	0,66	0,32	0,60	2,22	27,31	749,89	3278,42
Oeste Maranhense	2,71	4,35	9,65	0,61	0,33	0,59	14,48	97,91	5785,01	13579,26
Oeste Potiguar	1,39	1,39	1,43	0,75	0,60	0,61	32,85	52,07	812,09	2444,19
Metropolitana do Recife	3,13	11,11	9,02	0,78	0,25	0,63	43,69	640,14	7535,27	16311,08
São Francisco Pernambucano	1,17	2,64	0,71	0,63	0,27	0,52	1,79	100,07	1933,44	3890,11
Metropolitana de Salvador	3,54	8,45	3,81	0,67	0,34	0,61	93,81	497,85	5627,61	14056,18
Sertão Alagoano	0,59	1,08	2,27	0,63	0,30	0,53	0,91	8,17	532,89	4526,86
Sertão Paraibano	0,08	2,17	2,03	0,70	0,28	0,47	3,17	123,67	717,62	4672,42
Sertão Pernambucano	0,78	1,06	0,54	0,65	0,38	0,64	4,02	6,28	266,68	1171,13
Sertão Sergipano	0,83	5,61	0,00	0,67	0,35	0,65	3,80	3,80	474,48	5801,02
Sertões Cearense	1,06	2,09	6,21	0,73	0,38	0,60	1,04	16,57	1169,94	11010,34
Sudeste Piauiense	2,19	3,88	0,80	0,68	0,29	0,52	1,07	3,61	247,90	1235,87
Sudoeste Piauiense	1,78	0,71	1,96	0,64	0,31	0,50	4,72	11,00	800,46	3624,66
Sul Baiano	3,13	7,52	2,64	0,62	0,33	0,53	7,91	41,11	1715,34	4268,96
Sul cearense	0,96	2,21	3,81	0,75	0,40	0,58	3,58	20,15	1311,82	4543,25
Sul Maranhense	1,16	2,96	0,90	0,67	0,31	0,51	0,90	22,67	289,09	982,86
Vale São Francisco da Bahia	2,04	2,75	0,47	0,59	0,30	0,50	0,56	2,50	229,01	919,12

Fonte: Ministério da Saúde, Ministério do Trabalho e Emprego e Firjan.

ANEXO C

SÉRIE HISTÓRICA DA TAXA DE MORTALIDADE DE TUBERCULOSE.
REGIÕES E UNIDADES FEDERADAS DE RESIDÊNCIA POR ANO DO ÓBITO (1990 A 2006).

Tabela C.

(por 100.000 hab.)

Região e UF	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Maranhão	1,94	1,85	1,65	1,89	1,78	2,06	1,84	1,30	2,52	2,20	2,05	2,11	2,15	1,97	2,68	2,97	2,36
Piauí	2,60	1,94	1,77	1,81	2,08	1,21	1,20	1,71	2,10	2,60	2,22	1,95	2,73	2,43	2,17	2,43	2,27
Ceará	1,96	2,17	3,55	2,79	2,22	2,43	2,31	2,89	4,09	3,63	2,80	3,39	3,03	2,46	2,72	2,87	3,05
Rio Grande do Norte	1,99	2,19	1,88	2,00	1,97	3,87	1,84	3,28	2,86	2,94	2,48	2,38	1,68	1,59	1,61	1,73	1,25
Paraíba	1,71	1,47	1,65	1,92	1,90	1,35	1,45	2,16	1,58	1,60	1,74	1,53	2,46	3,21	2,23	3,95	2,93
Pernambuco	4,81	5,32	4,82	5,26	6,32	5,55	5,51	6,00	6,82	5,79	5,43	5,27	4,96	5,23	5,29	4,73	4,39
Alagoas	2,48	2,82	1,81	2,07	2,91	2,35	2,58	2,14	3,61	2,29	2,37	2,77	3,08	3,05	2,37	2,52	2,62
Sergipe	2,33	1,68	2,11	1,61	2,28	2,06	1,79	2,47	2,73	2,16	1,85	1,87	1,41	1,60	2,05	2,08	2,10
Bahia	3,78	3,51	3,65	3,74	3,67	3,33	3,70	3,19	3,32	3,61	3,33	3,25	3,53	3,11	3,04	2,71	2,79
Região Nordeste	3,01	2,99	3,07	3,11	3,25	3,07	3,01	3,14	3,69	3,43	3,1	3,14	3,19	3,04	3,05	3,08	2,89

Fonte: SIM/SVS/MS atualizado em 18/08/2008.

Dados preliminares sujeitos a revisão.