

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

MAIZZA MICAELLE CARLOS EUCLIDES FERNANDES

ESTUDO DA MORTALIDADE INFANTIL NOS ESTADOS
BRASILEIROS UTILIZANDO APLICAÇÃO DE DADOS EM PAINEL

JOÃO PESSOA, 14 JUNHO DE 2016

MAIZZA MICAELE CARLOS EUCLIDES FERNANDES

ESTUDO DA MORTALIDADE INFANTIL NOS ESTADOS BRASILEIROS
UTILIZANDO APLICAÇÃO DE DADOS EM PAINEL

Projeto apresentado ao Curso de Graduação em Estatística da Universidade Federal da Paraíba – Campus I, como requisito parcial para a obtenção do Título de Bacharel em Estatística.

Orientador: Prof. Dr. João Agnaldo do Nascimento

João Pessoa, 14 de Junho de 2016

Maizza Micaelle Carlos Euclides Fernandes

ESTUDO DA MORTALIDADE INFANTIL NOS ESTADOS BRASILEIROS
UTILIZANDO APLICAÇÃO DE DADOS EM PAINEL

Este Projeto de autoria de Maizza Micaelle Carlos Euclides Fernandes apresentado como requisito parcial para a obtenção do Título de Bacharel em Estatística foi julgado pela Banca Examinadora designada pela Comissão de Coordenação de Monografia do Curso de Estatística da Universidade Federal da Paraíba – Campus I, abaixo assinada:

João Pessoa, 14 de Junho de 2016

BANCA EXAMINADORA

Prof^o. Dr. João Agnaldo do Nascimento
Orientador (UFPB)

Prof^a. Dr. Agnes Liliane Lima Soares de Santana
(UFPB)

Prof^o. Dr. Marcelo Rodrigo Portela Ferreira
(UFPB)

Prof^a. Dr^a Tatiene Correia de Souza
(UFPB)

*Dedico este trabalho primeiramente a Deus,
pois sem ele eu não teria forças
para enfrentar essa longa jornada.
Ao meu orientador João Agnaldo e ao meu amigo Alisson,
por seus ensinamentos, paciência e confiança
ao longo desses anos de convivência diária.*

Agradecimentos

Primeiramente, agradeço a Deus por ter me dado forças e coragem para não desistir dos meus sonhos. A minha família que é a base de tudo. Aos meus pais, Demuntier e Iêda pelo amor e carinho. A minha irmã Marília, pela sua paciência, dedicação, conselhos, palavras de carinho e forças em momentos que pensei que não fosse capaz de conseguir meus objetivos. A minha mãe de criação Elizângela, que se dedicou boa parte da sua vida para cuidar de mim, concedendo amor e força passada mesmo por telefone e por torcer e estar sempre do meu lado em todos os momentos da minha vida desde do meu nascimento. A minha vizinha Valdetrudez (in memoriam), que sempre foi o meu apoio em tudo, deixando muitas saudades e encorajamento para lutar pelos meus objetivos.

Ao meu primo João Euclides, sua esposa Analyane, Ana Emília, Joan, tia Neide e tio Braga que me deu a oportunidade pra me conquistar esse grande sonho de cursar Estatística, me dando apoio, teto para morar, amor, carinho e principalmente conselhos para nunca desistir dos meus ideais.

À minha tia Graça, tio Dedo, primas Isabelle e Priscilla, Anderson e Lucas, aos meus sobrinhos de coração Ana Laura, Ana Beatriz e Arthur por me apoiarem em cada decisão tomada, pela paciência nos momentos que mais precisei da compreensão, do amor, atenção e por fazerem meus dias mais felizes ao lado de vocês.

Ao meu amigo irmão de coração Alisson e a minha irmã de coração Ianne por todos os conhecimentos trocados, pela paciência, alegrias, dificuldades, desesperos, choros, tristezas, companheirismo, forças para prosseguir nesta caminhada árdua ao longo do curso e por serem meus verdadeiros irmãos que eu sempre pude contar de verdade e que a estatística me deu esse maior presente que foi conhecê-los.

Ao meu grande amigo Chico pelo seu apoio dado, pelos conselhos, puxões de orelhas nos momentos que precisei, pelas suas palavras de carinho e por acreditar no meu esforço e que eu era capaz de conseguir tudo que almejava. O meu muito obrigada! Aos meus

amigos de graduação por passarem comigo a maior parte do meu tempo e uma importante fase da minha vida, deixando marcas de lembranças que levarei para o resto da vida de forma especial.

Ao meu orientador e pai de coração, professor João Agnaldo, que devo a minha formação acadêmica ao senhor, pela experiência passada, pelos choros trocados nos dias de desespero e por me dar coragem para não desistir do curso, pela força que me impulsionou a não desistir do curso por mais que os obstáculos fossem grandes, pela ajuda no desenvolvimento deste trabalho, pela paciência, amor, carinho, dedicação, pelos conselhos, palavras de carinho nos momentos mais difíceis dessa jornada, enfim, a minha gratidão pelo senhor é eterna. Não existe palavras concretas para expressar o quanto é importante na minha vida.

À professora Agnes e mãe de coração, pela pessoa e profissional extraordinária que você é. Agradeço por todos os conselhos, ensinamentos e paciência nos momentos que precisei de apoio e uma palavra amiga.

À professora Tatiene, quero aproveitar essa oportunidade para lhe dizer que sou grata pela sua dedicação e por todos conselhos e ensinamentos passado ao longo do curso. É verdade que às vezes nossas opiniões podem divergir, mas até neste momento, estou desfrutando do seu conhecimento e tendo a chance de olhar a vida sobre um novo horizonte. Admiro o seu jeito de ser e lhe desejo que não lhe falte forças para prosseguir nessa sua jornada.

Em especial ao professor Marcelo por toda ajuda dada na realização desse trabalho, por um ser humano admirável, sempre disposto a apoiar e orientar. Palavras de gratidão serão sempre escassas para quem dá tanto de si mesmo e de forma tão constante. Você é acessível e sempre atencioso, uma raridade nestes dias de hoje. Continue nesse caminho professor, nunca deixe que ninguém mude esta sua maneira de ser para com seus alunos. Obrigada por tudo de coração!

À todos os professores do departamento de estatística pelo valioso conhecimento transmitido, pelas disponibilidades em ajudar e por contribuírem para o meu crescimento profissional.

*Nada te perturbe,
Nada te espante,
Tudo passa,
Deus não muda,
A paciência tudo alcança;
Quem a Deus tem,
Nada lhe falta,
Só Deus basta.
Teresa D'Ávila*

Resumo

Este trabalho objetiva determinar uma relação linear entre a Taxa de Mortalidade Infantil (TMI) e um conjunto de variáveis socioeconômicas observadas por unidade federativa no período de 2005 à 2010 utilizando o modelo de dados em painel de efeitos fixo e aleatório . Os resultados mostram que o modelo mais adequado é o de efeito fixo com transformação logarítmica nas variáveis dependente e independentes que foram as seguintes: TMI, taxa de analfabetismo, PIB per capita, proporção pessoas com baixa renda, percentual da população servida por rede de abastecimento de água e a proporção da população servida por coleta de lixo. A conclusão obtida é que as variáveis independentes que causam impacto significativo na TMI são taxa de analfabetismo, PIB per capita e proporção de pessoas com baixa renda. Utilizou-se o *software* R para realizar esta análise de dados com a função `plm`.

Palavras Chave: *Taxa de mortalidade infantil; Dados em painel; Modelo de efeito fixo e aleatório*

Abstract

The aim of this work is to find a linear relation for child mortality rate (CMR) and a observable set of social and economic variables in the states of Brazil for period 2005 to 2010 using the panel data model ling with fixed and random effects. These results show that the more adequate model is a fixed effect model the with logarithmic transformation the on both, the dependent and independents variables such as: CMR, illiteracy rate, per capita income, proportion of people with low income, proportion of the population served by water supply and proportion of population served by waste collection. We conclude that the independent variables causes impact in significant impact in CMR are: CMR, illiteracy rate, per capita income, proportion of people with low income. We use R *software* to perform the data analysis with the plm function.

Keywords:*mortality rarte (CMR); panel data; model of fixed and random effects.*

Lista de Figuras

4.1	Gráfico dos Resíduos para o Modelo OLS pooled	23
4.2	Gráfico de Normalidade do Efeito Fixo	24
4.3	Gráfico de Normalidade do Efeito Aleatório	25

Lista de Tabelas

3.1	Descrição das Bases de Dados	15
4.1	Resultado da Regressão OLS Pooled(Teste de Chow)	22
4.2	Resultado da Regressão do modelo de Efeito Fixo	24
4.3	Resultado da Regressão do modelo de Efeito Aleatório	24
4.4	Resultado dos Testes de Breusch-Pagan e Hausman	26
4.5	Resultado do Teste de Efeito Individual dos Estados	26

Lista de Siglas

CMI – Coeficiente de Mortalidade Infantil

DO – Declaração de Óbito

FUNASA – Fundação Nacional de Saúde

GLS – (Generalized Least Square) Mínimos Quadrados Generalizados

IBGE – Instituto Brasileiro de Geografia e Estatística

IDB – Indicadores e Dados Básicos de Saúde

MI – Mortalidade Infantil

OLS – (Ordinary Least Square) Mínimos Quadrados Ordiniais

PD - Dados em painéis

PNAD – Pesquisa Nacional por Amostra de Domicílios

RC – Registro Civil

SIM – Sistema de Informações sobre Mortalidade

SINASC – Sistema de Informação sobre Nascidos Vivos

SIS – Sistema de Informações em Saúde

SUS – Sistema Único de Saúde

TMI – Taxa de Mortalidade Infantil

UF – Unidades de Federação

Sumário

1	Introdução	1
1.1	Caracterização do Problema	1
1.2	Objetivos	2
1.2.1	Geral:	2
1.2.2	Específicos:	2
1.3	Delimitação da Pesquisa	3
2	Fundamentação Teórica	4
2.1	Mortalidade Infantil	4
2.2	Modelo para Dados em Painel	8
2.2.1	Modelo de Efeitos Fixos	10
2.2.2	Modelo de Efeitos Aleatórios	10
2.3	Contribuição das Variáveis em Estudo para a Mortalidade Infantil	11
3	Metodologia	13
3.1	Fontes de Dados	14
3.2	Construção e organização das bases de dados	14
3.2.1	Variáveis de Estudo	15
3.3	Modelo Dados em Painel	15
3.3.1	Modelos de Efeito Fixo	16
3.3.2	Modelos de Efeito Aleatório	18
3.3.3	Modelo Pooled	20
4	Resultados e Discussões	22
5	Conclusão	28

Referências Bibliográficas	29
A Programas em R para Dados em Painel com Regressão Linear	33

Capítulo 1

Introdução

1.1 Caracterização do Problema

Considerando-se a essência e o propósito deste estudo, assim como o objetivo sobre a mortalidade infantil para as Unidades de Federação (UF), realiza-se na existente elaboração deste trabalho revisão bibliográfica, as origens das informações e publicações que retratem os principais conceitos relacionados ao tema em questão.

A mortalidade infantil (MI) pode ser vista como um acontecimento indispensável por infraestrutura e sistema de saúde eficiente para obtenção da qualidade de assistência médica e condições para o bem-estar do indivíduo de uma sociedade. No entanto, é um tema complexo de ser avaliado em sua completude, podendo ser qualificado através de indicadores em definidas regiões conhecidas de acordo com elementos primordiais para o perfil do cenário de satisfação social de uma determinada população.

Dessa forma, dentre os indicadores, o mais clássico é a taxa de mortalidade infantil (TMI), que, por sua vez, caracteriza não apenas as condições de saúde de uma sociedade, como também o nível sócio econômico da população que tem sido empregado como padrão sintetizador de desenvolvimento constante para caracterizar as condições de vida de uma população.

No entanto, a inconsistência das estatísticas de Registro Civil (RC) para os nascimentos quanto para os óbitos, tem colocado grandes restrições nos valores da TMI no Brasil. Assim, fazendo uso das fontes de dados dos censos demográficos tem-se o problema do sub registro dos óbitos que motiva uma correção na TMI que pode ser realizada por métodos de estimação indiretas que proporciona valores mais precisos para a TMI. Por

esta razão há fragilidades nestas informações que podem se distanciar da nossa realidade. Neste trabalho, não houve correções para a TMI e utilizou-se a base de dados do Datasus.

Ao decorrer dos anos, a diminuição dos óbitos inferiores a um ano de idade, caracteriza uma das importantes metas na área de saúde em inúmeros países, que tem como consequência a redução da mortalidade infantil. As variações na TMI serão avaliados por Estado e ao longo do tempo por modelos para Dados em Painel (DP). É importante utilizar modelos econométricos que mostrem a influência de variações que representem sua contribuição e influência na TMI uma vez que a análise destas influências poderão decidir estratégias de gestão de saúde para melhorar as condições de vida de uma população.

Modelos econométricos podem descrever uma variável dependente como função de variáveis independentes representadas por informações associadas às condições sócio-econômicas e de saúde pública. Podem medir de forma indireta o avanço obtido com investimentos públicos realizados ou oferecer um direcionamento para estabelecer quais metas estariam mais ligadas ao plano de diminuir o TMI.

Os países mais desenvolvidos diminuíram a TMI a um ponto onde o mesmo permanece estável em um valor pequeno quando comparando com os valores da TMI em países do terceiro mundo. O limite da redução esbarra na impotência do conhecimento de como tratar a mortalidade infantil decorrente as doenças de efeitos transgênicos (*ALMEIDA, 1955*). *Então por muitos anos a TMI refletirá o desenvolvimento de um povo, o produto de sua luta e o combate e fica, à manutenção das condições de vida.*

1.2 Objetivos

1.2.1 Geral:

Descrever uma relação linear entre a variável dependente a TMI e as variáveis independentes (taxa de analfabetismo, PIB per capita, proporção de pessoas com baixa renda, proporção da população servida por abastecimento de água e a proporção da população servida por coleta de lixo) por grupo, cujas observações ocorre ao longo de alguns anos caracterizando um estudo longitudinal.

1.2.2 Específicos:

- *Descrever o modelo de DP com efeito fixo e aleatório;*

- *Apresentar aplicação do modelo com os principais testes, sendo necessário à escolha de modelo Pooled, efeito fixo ou efeito aleatório;*
- *Utilizar o modelo Pooled, efeito fixo e efeito aleatório.*

1.3 Delimitação da Pesquisa

Este trabalho avalia a relação do TMI nos estados brasileiros no período de 2005 à 2010 através de um modelo econométrico. Uma limitação desta pesquisa é a não correção do CMI através dos métodos indiretos, ou seja, não ocorreu uma correção devido a sub-registros de óbitos.

Capítulo 2

Fundamentação Teórica

2.1 Mortalidade Infantil

De acordo com Silva et al. (2008), ressaltando-se que a mortalidade é motivada por fatores incompatíveis, ressaltando-se dentre eles: as parcialidades sociais e diversidades nas condições de vida; o desenvolvimento da urbanização; transformações na estrutura de idade da população, na formação de trabalhadores e no mercado de trabalho, no nível de educação da população e modificações na ordenação dos serviços de saúde.

De acordo com Paes (2007), as estatísticas de mortalidade são componentes do sistema de estatísticas vitais, subutilizadas em vários países da América Latina em virtude de seus limites por serem incompletos, desatualizados e dispersos. A subutilização se deve da mesma forma à crença injustificada de que esses limites comprometeriam permanentemente qualquer resultado que deles poderiam ser derivados.

O crescimento da mortalidade estabelece um considerável componente para verificação e percepção do crescimento populacional. Os relevantes estudos feitos neste setor para a população brasileira foram executados, especialmente, por Giorgio Mortara, que estuda por meio dos censos demográficos a mortalidade global e infantil para o Brasil. Para analisar o crescimento da mortalidade global e infantil para o Brasil, é necessário conhecer a população em estudo, o número total de óbitos que ocorrem, englobando os menores de um ano de idade e o número de nascidos vivos. Ainda que se conheça normalmente a estimativa da população, não possui os outros elementos, em decorrência das imprecisões de registros de óbitos e especialmente de nascimentos.

A mortalidade infantil elevada é conhecida desde os anos 70, por infraestrutura de

saúde ineficiente. Portanto, pode ser evitável estabelecendo-se uma melhor condição na qualidade da atenção médica (Rutstein, 1976). Nos países onde há um número alto em relação ao risco de morrer com menos de um ano de idade, tem-se mostrado mais difícil de conseguir estimadores com capacidade de mostrar com clareza os fatores que influenciam a mortalidade infantil. É como se o atraso cultural e da qualidade de vida impedem de ver com clareza este grande problema que fica escondido por trás de um péssimo sistema de informações de saúde.

O assunto da mortalidade infantil é essencial não só como sugestão preciosa das situações econômico-sociais e médico sanitárias de uma sociedade, como também para analisar a capacidade dos cuidados feitos para diminuí-la e ainda encaminhar futuros programas. Segundo DE ARAÚJO (1973), Salvador capital da Bahia, ocorreu melhorias nos últimos anos nos dados do registro de óbitos. Insiste ainda assim, que existe um sério problema quanto à evasão do registro civil de nascimento.

Diante disso, o Serviço de Bioestatística da Secretaria da Saúde Pública adotou o padrão de calcular a mortalidade infantil considerando como denominador o número estimado de nascidos vivos em cada ano. Esta estimativa é feita a partir da população total e do índice de natalidade, calculados com base nos dados dos últimos censos. A estimativa de mortalidade infantil tem o proveito de assegurar uma maior comparação de ano a ano, sem correr os riscos de erros resultantes das mudanças no registro civil.

A origem das estatísticas de mortalidade é a Declaração de Óbito (DO) que, além de dar origem à certidão de óbito (documento de caráter jurídico), determina diversas utilidades: benefício para programas de saúde pública, ponto de casos para investigação clínico epidemiológica, documento inicial para estudos retrospectivos, desfecho para estudos prospectivos, uso em seguros de vida e estudo demográficas de tendências da mortalidade (CMCCD, 1958).

O tema mortalidade é fundamental para identificar o perfil epidemiológico de uma população, e que é disponibilizado por meio da Declaração de Óbito, dessa forma o estudo de seus registros auxilia os investigadores na preparação do quadro epidemiológico. A morte mesmo sendo um fato marcante no ponto jurídico e médico, passou por modificações na sua descrição no transcorrer do tempo.

A morte pode ser juridicamente entendida como o "desaparecimento permanente de todo sinal de vida, em um momento qualquer depois do nascimento", ou, conforme definição da Organização Mundial de Saúde, "o desaparecimento de todos os sinais de vida ou a cessação das funções vitais, sem a possibilidade de ressuscitar"(Laurenti; Mello Jorge, 1992).

A DO é o documento encarregado pelo registro da morte e "o documento oficial hábil para que a pessoa seja considerada morta, permitindo, portanto, que seja iniciada toda uma sequência de atos jurídicos, principalmente dentro do Direito de Família e do Direito das Sucessões" (MELLO JORGE, 2006). Em 1975 o Ministério da Saúde aprovou a DO uniformizada para todo o país, visto que em comparação às causas de morte conserva-se o padrão que foi proposto em 1948 pela Organização Mundial de Saúde (MELLO JORGE; LAURENTI; GOTLIEB, 2002).

No decorrer dos anos, o modelo da padronização da DO sofreu mudanças com modificações e integração de variáveis, logo, nos dias de hoje a DO é expedida em três vias numeradas que são impressas pelo Ministério da Saúde que são divididas para as Secretarias Estaduais de Saúde para serem distribuídas às Secretarias Municipais de Saúde. Logo, entrega as associações de saúde, Instituto de Medicina Legal, Serviços de Verificação de Óbito e Cartórios de Registro Civil (MELLO JORGE; LAURENTI; GOTLIEB, 2002). Contudo, em conformidade com a fundação Nacional de Saúde (FUNASA) a sucessão varia conforme o tipo de óbito e onde ocorreu.

O Sistema de Informações sobre Mortalidade (SIM) do Ministério da Saúde (MS), consiste em uma fonte opcional de dados de óbitos no país. Em 1976 foi criada a partir da fundação do modelo padronizado da DO em todo o território nacional, o SIM tem a finalidade indispensável de oferecer auxílio para traçar o perfil de mortalidade no país (MS, 1982). Além da disponibilidade de dois sistemas de informações sobre óbitos em nível nacional, o do Registro Civil e o SIM, uma das grandes complicações que ainda restringem as análises de mortalidade no Brasil é o sub-registro de óbitos, que tem relevância significativa, principalmente, nas regiões Norte e Nordeste, com predomínio entre as crianças com menos de um ano de idade.

O sub-registro de óbitos diz respeito, sobretudo, à ocorrência de funeral sem a determinação da certidão, nos designados cemitérios clandestinos, sendo relacionado à

pobreza e predominante na área rural (Mello-Jorge, 1983, 1987). Os dados sem retificação representam principalmente a ocorrência de mortalidade das regiões mais desenvolvidas. Do mesmo modo é benéfico para os dados do litoral do Nordeste em comparação ao sertão nordestino, das capitais com relação ao interior dos estados, dos centros urbanos em comparação às áreas rurais.

Os dados de registro vão sempre refletir às áreas que melhor informam a mortalidade. Mesmo que retifiquem as informações de registro por razões diferenciadas por região, estar corrigindo apenas aquelas relacionadas às áreas que, de uma certa forma, notificam a mortalidade e que, supostamente, possuem as melhores condições de saúde. Em 1990, no que está relacionado aos dados referentes aos nascimentos, o MS estabeleceu o Sistema de Informação sobre Nascidos Vivos (SINASC).

Através desse sistema tem sustentação na declaração de nascimento, documento da qual a emissão é obrigatória no serviço de saúde onde ocorreu o parto. Diferente do sistema de nascimentos do Registro Civil, dos quais o propósito primordial é a contagem do número de registros de nascimentos, o SINASC tem como finalidade evidenciar as condições de nascimento (Mello-Jorge et al., 1996). Apesar de que seja explícito que a cobertura do SINASC esteja progredindo e que a qualidade da informação vem melhorando desde a sua implantação, percebe-se que a cobertura do sistema ainda não é total.

O confronto entre o número de nascimentos ocorridos nos hospitais do Sistema Único de Saúde (SUS) e os declarados no SINASC, demonstra ainda insuficiências na cobertura do SINASC em alguns estados do país, ainda que de menor relevância que as mostradas pelo SIM (Schramm & Szwarcwald, 2000). Na presença de deficiências das fontes de informações, o IBGE, que é o instituto responsável por propiciar a estimativa da mortalidade infantil no Brasil, tem utilizado a técnicas demográficas, ditas de classificação indireta, para representar as estimativas clássicas (IBGE, 2001).

Assim, as técnicas são fundamentadas em dados de entrevistas domiciliares que são executadas por viabilidade dos Censos Demográficos ou das Pesquisas Nacionais de Amostra por Domicílio (PNAD). Os coeficientes de mortalidade infantil (CMI) é um dos indicadores mais utilizado para avaliar níveis de saúde e de crescimento social de um lugar. A mortalidade infantil verifica o risco de um nascido vivo morrer no seu primeiro ano de vida. Tendo como coeficiente abaixo de 20 por 1000 é considerado baixo; 50 ou mais por 1000 nascidos vivos é considerada alta.

De acordo com alguns autores, a concepção de que o CMI é um indicador ex-

cessivamente sensível, precisa ser atualizada. As vertentes fortemente procedentes da mortalidade infantil, que vem sendo percebidas no país como um todo, estariam, segundo concepções usuais, indicando uma melhoria significativa das condições de saúde e, portanto, da qualidade de vida das populações. Ainda assim, quando se procede a análise de outros indicadores socioeconômico, percebe-se que estes não evidenciam uma evolução coerente com a queda da mortalidade infantil e muitas vezes chegam a comprovar a degradação das condições e da qualidade de vida da população.

A taxa de mortalidade infantil (TMI) é identificada não só como medida das situações de saúde da população inferior a 1 ano, serve para descrever as circunstâncias total da vida, sendo conhecida como do bem-estar, qualidade de vida e do grau de crescimento da população. Dessa maneira, é provável determinar ligações entre a magnitude da TMI e as características do modelo de desenvolvimento de uma população. A TMI é definida por:

$$TMI = \frac{\text{Número de óbitos de residentes com menos de um ano de idade}}{\text{Número de nascidos vivos de mães residentes}} \times 1000$$

2.2 Modelo para Dados em Painel

Os dados em painel relacionam-se a secção diagonal ou transversal de considerações que foi reprisado no decorrer de diversos momentos de tempo. Assim, o método de Dados em Painel baseia-se na análise de fatores ou objetos (por exemplo: empresas, estados, etc.) para dois ou mais intervalos de tempo.

Segundo Hsião (2003), um corte longitudinal, ou painel, é um conjunto de dados, que segue uma dada amostra de indivíduos ao longo do tempo, e, assim, fornece várias observações em cada indivíduo na amostra. Dados em painel tornaram-se amplamente disponíveis em ambos os países desenvolvidos e em desenvolvimento.

De acordo com Hsião (1986), os modelos para dados em painel oferecem uma série de vantagens em relação aos modelos de corte transversal ou aos de séries temporais sendo que delas se refere ao fato de que esses modelos controlam a heterogeneidade presente nos indivíduos.

Conforme Hsião (1986), é que os dados em painel permitem o uso de mais observações, aumentando o número de graus de liberdade e diminuindo a colinearidade entre as variáveis explicativas. Sabe-se que quando existe multicolinearidade torna-se difícil es-

tabelecer se um regressor individual influencia uma variável resposta. Uma vez eliminado esse problema, pode-se obter uma melhora na qualidade da estimação dos parâmetros.

Dados em painel relaciona-se com uma secção transversal de observações, sejam, indivíduos, grupos, países ou regiões repetida ao longo de vários períodos de tempo. Em muitos países, há poucos ou nenhum painel de dados, mas existe uma série de secções transversais independentes.

No modelo com dados em painel, a abordagem no que diz respeito ao tipo ao experimento, considera-se modelo de fixo e modelo de efeito aleatório sendo esta escolha decidida pelo teste de Housman para detectar neste modelo os dados em estudo o geram uma série temporal que pode ser utilizada para inferir relações comportamentais para o coorte como todo.

O importante proveito da aplicação dos modelos de dados em painel relaciona-se com a inclusão da heterogeneidade individual, isto é, à expectativa de mensurar particularmente os impactos gerados por conta de diferenças existentes no meio de toda observação em cada corte transversal, além de ser possível analisar o crescimento, para um dado indivíduo, das variáveis em observações ao longo do tempo (Marques, 2000).

De outro ponto de vista, conforme (Marques, 2000), os dados em painel determinam um melhor número de informação, menor colinearidade entre as variáveis, maior oscilação dos dados, maior número de grau de liberdade e melhor eficácia na estimação. A união da grandeza em corte transversal, com um estudo temporal, verifica-se uma maior utilização dos dados, de acordo com a representação do modelo com a realidade dos dados que incorporam no modelo as necessidades observadas para a modelagem estatística exigida.

Esse acréscimo na modelagem dos dados em relação aos modelos de regressão linear múltipla clássico colabora em prol da diminuição de uma eventual colinearidade existente entre variáveis. Os dados em painel verificam os efeitos fixos ou aleatórios que não se percebem com clareza no modelo tradicional e caso não seja utilizado tem-se os fatores, que proporcionam estimativas mais explícita dos viés nas estimativas que podem sofrer influências diversas da multicolinearidade.

2.2.1 Modelo de Efeitos Fixos

O modelo de efeito fixo também denominado variável dummy de mínimos quadrados ou LSDV (Least Square Dummy Variable), é uma generalização de um modelo constante intercepto inclinação para dados em painel, incluindo uma variável dummy para os efeitos das variáveis omissas, que encontram-se constante no tempo. Os efeitos individuais podem ser espontaneamente correlacionados com os outros regressores. O objetivo do efeito fixo é verificar os impactos das variáveis independentes variando entre indivíduos e também ao longo do tempo. Para este fim, imagina que o intercepto varia de um indivíduo para outro, no entanto, é constante ao longo tempo, a fase que os parâmetros são constantes para as entidades e em todos os períodos de tempo.

A expressão “efeitos fixos” dá uma percepção confusa da modelagem, pois então, para os dois casos, os efeitos no nível do indivíduo são aleatórios. Dessa forma, de acordo com Cameron & Trivedi (2009), os modelos de efeitos fixos mostram quais os regressores estão correlacionados com os efeitos devido a entidade (fator) e, por isso, uma estimativa consistente dos parâmetros do modelo requer uma eliminação ou controle dos efeitos fixos. O termo erro da regressão depende dos fatores e do tempo, enquanto possuindo variância constante.

2.2.2 Modelo de Efeitos Aleatórios

O intercepto não muda de um fator para outro, nem ao longo do tempo; os parâmetros de inclinação do modelo são constantes para todos os fatores e em todos os períodos de tempo, sendo a diferença dos dois modelos o procedimento do intercepto.

Este modelo estima as componentes de variância para os grupos (ou tempos) e o erro supondo mesmos interceptos e inclinações. O termo erro tem duas componentes: uma sistemática e outra não sistemática, sendo ambas não correlacionada com os regressores. As suposições para aplicação dos Mínimos Quadrados Ordiniais (OLS) são violadas (homogeneidade). A diferença entre grupos ou períodos está na variância do termo erro aleatório e não nos interceptos.

2.3 Contribuição das Variáveis em Estudo para a Mortalidade Infantil

A mortalidade infantil era compreendida como dificuldade de atraso socioeconômico até os anos setenta, já que, praticamente, apresentava-se contraditoriamente relacionada com indicadores de Produto Interno Bruto (PIB), a taxa de desemprego o Índice de Gini. Esta ligação era esclarecida pelo papel executado por esses indicadores na composição do perfil de renda, educação, saneamento básico e obtenção aos programas e serviços de saúde ocasionando um avanço nas condições de vida da população.

No entanto, especialmente a partir do final daquela década, a queda na diminuição da mortalidade infantil no países em desenvolvimento que conviveram com a intensa crise econômica definida pela restrição do PIB, do valor real do salário mínimo e do aumento de desemprego, isto auxiliou de incentivo na execução de estudos tendo como objetivo explicar os propósitos para a correção do declínio deste mortalidade.

Os fatores de risco para a mortalidade infantil, didaticamente, são divididos em três grupos: fatores biológicos, sociais e relacionados à assistência médica. Entre os fatores de risco biológicos estão: idade materna, paridade, estatura, peso antes da gravidez, ganho ponderal durante a gestação, doença materna (diabetes, hipertensão), intervalo entre os nascimentos e perdas perinatais prévias, peso ao nascer e duração da gestação. Quanto aos sociais, incluem: instrução da mãe, ocupação dos pais, renda familiar e hábitos, entre outros. Já os relacionados à assistência médica estão condicionados a elementos como disponibilidade de recursos, aceitabilidade e resolubilidade do serviço de saúde. (CAMPOS, Tatiana P.; SA CARVALHO, M.; BARCELLOS, Christovam C.,2000)

O coeficiente de mortalidade infantil tem uma influência nos principais determinantes das condições de vida como: alimentação, moradia, acesso a serviços de saúde, entre outros aspectos que estão relacionados neste contexto e atuam sobre a expectativa de sobrevivência no primeiro ano. De acordo com STOCKWELL (1962), sua constatação teórica, por sua vez, é continuamente procurada seja ao nível da relação inversa que a evolução secular do coeficiente manteve com o avançar da industrialização, seja as con-

dições dos estudos transversais de distribuição de taxas de mortalidade conforme classes sociais.

A ligação de oposição que a mortalidade infantil mantém com o padrão sócio econômico tem permanecido de tal maneira na história dos países hoje denominados desenvolvidos como na história dos não desenvolvidos (VALLIN, 1980). Embora a redução das condições plena da mortalidade e da amplificação do acúmulo da tecnologia médica e sanitária, os diferenciais sócio econômicos da mortalidade infantil continuam a permanecer em ambos os grupos de países.

O esclarecimento que parece mais provável para a conservação destas diferenças, seria a de que a obtenção aos novos fatores que surgem para a redução da mortalidade deve estar se organizando de modo tão rígido quanto o acesso aos demais elementos que definem as condições da qualidade de vida e que tem identificado as classes sociais (ANTONOVSKI, 1977).

Capítulo 3

Metodologia

O conceito de pesquisa pode ser definido como um procedimento formal, como método de pensamento reflexivo que exige um tratamento científico e se constitui no caminho para compreender a realidade ou para encontrar verdades parciais (Lakatos; Marconi, 2001). De acordo com Gil (1991) a pesquisa bibliográfica fortalece com base no material já produzido, baseados especialmente de livros e publicações periódicas. Dessa forma, o presente estudo evidenciou-se como sendo de característica bibliográfica.

Logo, conforme o propósito do estudo, pode-se afirmar que esta é uma pesquisa exploratória, ao procurar compreender o conteúdo com mais profundidade. Segundo Barros e Lehfeld (2007), a pesquisa exploratória pode ser realizada com diversas técnicas permitindo ao pesquisador explicar a problematização do estudo e a formulação de uma hipótese com precisão.

O estudo foi elaborado para avaliar o impacto de algumas variáveis socioeconômicas sobre a TMI nas UF ao longo do período de 2005 à 2010 por meio da aplicação do modelo de DP. As variáveis são: a taxa de analfabetismo; PIB per capita; proporção de pessoas com baixa renda; abastecimento de água e coleta de lixo.

Dados em painel (DP) são medidas repetidas em diferentes períodos de tempo. Nesses dados deve-se considerar o efeito de grupo e do tempo que são analisados como efeito fixo ou aleatório. Para isto, é necessário que se tenha dois tipos de variáveis, uma referente ao objeto, por exemplo, instituição, estado ou grupo, e outra referente ao tempo, podendo ser mês, trimestre, semestre ou ano.

Os modelos de PD são classificados como efeito fixo em que as variáveis dummy são consideradas parte do intercepto ou aleatório em que as variáveis dummy são consideradas

como erros aleatórios. Então para se decidir qual tipo de modelo deve-se utilizar, devemos aplicar o teste de Housman (1978), que decide sobre o efeito a ser aceito.

A hipótese nula desse teste é que os efeitos individuais são não correlacionados com os demais regressores do modelo, ou seja, o modelo de efeito aleatório deve ser acatado, caso contrário ao rejeitamos a hipótese nula utilizamos o modelo com efeito fixo. Portanto, deve-se atingir uma conclusão que consiga representar uma relação linear múltipla entre uma variável dependente e um conjunto de variáveis independentes, permitindo avaliar essa situação utilizando o método de dados em painel.

3.1 Fontes de Dados

As fontes oficiais de referência com relação a mortalidade infantil são o Ministério da Saúde (MS), por intermédio do Sistema de Informação sobre Mortalidade (SIM) – “um sistema de vigilância epidemiológica nacional, cujo objetivo é captar dados sobre os óbitos do país a fim de fornecer informações sobre mortalidade para todas as instâncias do sistema de saúde” (BRASÍLIA, 2001), Sistema de Informações sobre Nascidos Vivos (SINASC) e a Fundação Instituto Brasileiro de Geografia e Estatística (Fundação IBGE), através das Estatísticas do Registro Civil, com disponibilidade anual para todos estados e municípios brasileiros.

Os dados relacionados a mortalidade infantil e os indicadores socioeconômico e de cobertura usado para este trabalho foram extraído do SIM e do IBGE/ Pesquisa Nacional por Amostra de Domicílios (PNAD) para os anos de 2005 e 2010, por Unidade de Federação (UF), acessados através do site do Departamento de Informática do SUS (DATASUS). O SIM foi criado pelo DATASUS para captação dos dados de mortalidade no país e é a única fonte que divulga os resultados por causa básica de morte no Brasil, sendo este um dos principais motivos que justifica seu uso.

3.2 Construção e organização das bases de dados

Os dados obtidos dos sistemas de informações foram por Unidade de Federação, sendo organizados em planilhas eletrônicas do Office Excel 2007. Foram extraídos a mortalidade infantil, Taxa de analfabetismo, percentual da população de 15 ou mais anos não alfabetizada; PIB per capita; Proporção de pessoas com baixa renda; Abastecimento de

água; Coleta de lixo das UFs que é composto por 27 estados brasileiros.

Os dados para os anos 2005 à 2010 foram retirados do Sistema de Informação sobre Mortalidade (SIM) para as UFs. As variáveis do estudo com seus referentes valores foram transferidos para um banco de dados nos software estatísticos R, sendo executadas as análises dos dados por meio de aplicações de dados em painel.

3.2.1 Variáveis de Estudo

Dos indicadores e dados básicos para avaliar a mortalidade infantil nas UFs, selecionou-se 6 (seis) variáveis pertencentes aos indicadores: mortalidade, socioeconômicos e de coberturas, sendo um (1) pertencente ao indicador de mortalidade, sendo ele a Taxa de mortalidade infantil (TMI), 3(três) pertencentes aos Indicadores Socioeconômico, sendo eles: Taxa de analfabetismo; produto interno bruto (PIB) per capita; proporção de pessoas com baixa renda e 2 (dois) pertencentes aos Indicadores de Cobertura: proporção da população servida por rede de abastecimento de água e a proporção da população servida por coleta de lixo, para os anos 2005 à 2010, cuja descrição está no quadro (1) abaixo, expressa as variáveis que serão utilizadas no modelo, definido a seguir:

Tabela 3.1: Descrição das Bases de Dados

Variáveis	Descrição	Fonte de Dados
Y = TMI	Taxa de Mortalidade Infantil até um ano de vida	Datasus
X1 = Analfabetismo	Taxa de analfabetismo para pessoas com 15 anos ou mais de idade	Datasus
X2 = PIB per capita	Produto Interno Buto	Datasus
X3 = Baixa Renda	Proporção de pessoas com baixa renda	Datasus
X4=Abastecimento de água	Proporção da população servida por rede de abastecimento de água	Datasus
X5 = Coleta de lixo	Proporção da população servida por coleta de lixo	Datasus

3.3 Modelo Dados em Painel

O modelo geral para dados em painel é descrito da forma seguinte:

$$y_{it} = \beta_{0it} + \beta_{1it}x_{1it} + \dots + \beta_{nit}x_{kit} + e_{it}$$

Neste caso, o subscrito i expressa os diferentes indivíduos e o subscrito t indica o período de tempo que está sendo estudado. β_0 relaciona-se ao parâmetro de intercepto e β_k ao coeficiente angular equivalente à k -ésima variável explicativa do modelo.

$$y_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{bmatrix} \quad X_i = \begin{bmatrix} x_{1i1} & x_{2i1} & \cdots & x_{ki1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1iT} & x_{2iT} & \cdots & x_{kiT} \end{bmatrix} \quad \beta_i = \begin{bmatrix} \beta_{0i1} & \beta_{1i1} & \beta_{2i1} & \cdots & \beta_{ki1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{0iT} & \beta_{1iT} & \beta_{2iT} & \cdots & \beta_{kiT} \end{bmatrix} \quad e_i = \begin{bmatrix} e_{i1} \\ e_{i2} \\ \vdots \\ e_{iT} \end{bmatrix}$$

Em que y_i e e_i são vetores de dimensão $(T \times 1)$ e contém respectivamente, as T variáveis dependentes e os T erros. Os X_i é uma matriz de dimensão $(K \times T)$ com as variáveis explicativas do modelo. Desta forma, o elemento x_{kiT} refere-se à k -ésima variável explicativa para o indivíduo i no instante de tempo t . Por fim, β_i é a matriz dos parâmetros a serem estimados.

Neste modelo geral, o intercepto e os parâmetros resposta são distinto para cada indivíduo e para cada período de tempo. Encontrando assim, mais parâmetros desconhecido do que observações, não sendo possível, neste caso, estimar os seus parâmetros. Assim, é necessário identificar suposições em relação ao modelo geral, afim de transformá-lo em operacional. Dentre os modelos que condiz com dados de séries temporais e dados em corte transversal, dois deles serão utilizados, Modelo de Efeitos Fixos e Modelos de Efeitos Aleatórios.

3.3.1 Modelos de Efeito Fixo

Em conformidade com Hill, Griffiths e Judge (1999), as suposições do modelo são:

$$\beta_{0it} = \beta_{0i} \quad \beta_{1it} = \beta_1, \dots, \beta_{kit} = \beta_k$$

O modelo de efeitos fixos, no entanto, é dado da seguinte forma:

$$y_{it} = \alpha_i + \beta_1 x_{1it} + \dots + \beta_k x_{kit} + e_{it}$$

A forma matricial para o i -ésimo indivíduo, como sugerido por Griffiths, Hill e Judge (1993), será:

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \alpha_i + \begin{bmatrix} x_{1i1} & x_{2i1} & \cdots & x_{ki1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1iT} & x_{2iT} & \cdots & x_{kiT} \end{bmatrix} \times \begin{bmatrix} \beta_{0i1} & \beta_{1i1} & \beta_{2i1} & \cdots & \beta_{ki1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{0iT} & \beta_{1iT} & \beta_{2iT} & \cdots & \beta_{kiT} \end{bmatrix} + \begin{bmatrix} e_{i1} \\ e_{i2} \\ \vdots \\ e_{iT} \end{bmatrix}$$

Assim, neste modelo o α_1 refere-se aos interceptos a serem estimados, um para cada indivíduo. Assim como os parâmetros respostas não muda entre os indivíduos e nem ao longo do tempo, todas as diferenças de procedimentos entre os indivíduos devem ser percebidas pelo intercepto. Dessa maneira, α_1 pode ser explicado como o efeito das variáveis omissas no modelo.

Outro pressuposto relevante do modelo efeito fixo é que o parâmetro é fixo e desconhecido que compreende as particularidades dos indivíduos da amostra. Logo, as inferências feitas com relação ao modelo são apenas a respeito dos indivíduos dos quais se utiliza de dados. Pode-se realizar uma classificação do modelo de efeitos fixos empregando variáveis binárias para mostrar os interceptos próprios para cada indivíduo. Assim sendo, a equação geral será expressa por:

$$y_{it} = \beta_0 + \beta_1 x_{1it} + \dots + \beta_k x_{kit} + \gamma_1 D_{1i} + \gamma_2 D_{2i} + \gamma_3 D_{3i} + \dots + \gamma_n D_{ni} + e_{it}$$

Neste sentido, $\beta_0, \beta_1, \dots, \beta_k, \gamma_2, \dots, \gamma_n$ são os parâmetros que serão estimados. Quando $i = 1$, o intercepto é dado por $\beta_0 = \gamma_1$. Para $i \geq 2$, o intercepto é dado por $\beta_0 + \gamma_i$. Através do OLS é estimado o modelo de efeitos fixos. As estimativas alcançadas são não tendenciosas e consistentes, por sua vez que o modelo supõe que os erros tem distribuição normal, variância constante e não são correlacionados.

Segundo Wooldridge (2002), a estimativa do modelo com variáveis binárias produz os mesmos resultados da estimação de efeitos fixos. No entanto, o emprego de variáveis dummy não é prático. Por isso, mesmo assim com total baixo de indivíduos no modelo, cresce muito o número de parâmetros a serem estimados. Em alguns pacotes econométricos quando se tem uma quantidade alta de indivíduos, essa regressão poder ser improvável de ser executado. Para verificar se as suposições sobre o modelo estão adequadas é satisfatório realizar um teste para investigar se os interceptos são diferentes entre os indivíduos.

De acordo com Griffiths, Hill e Judge (1993), as hipóteses nula e alternativa são as seguintes:

$$H_0 : \beta_{01} = \beta_{02} = \dots = \beta_{0k}$$

H_1 : os interceptos β_{0i} não são todos iguais

As hipóteses são testadas através da estatística F . Quando rejeita-se H_0 , conclui-se que os interceptos não são todos iguais, atendendo a suposição do modelo de n interceptos diferentes. De acordo com Wooldridge (2002) que o estimador do intercepto do modelo, $(\hat{\alpha}_i)$, além de não viesado, é inconsistente quando a quantidade de indivíduos é muito grande ($N \rightarrow \infty$). Isso ocorre porque cada observação em corte seccional adicionada aumenta um novo coeficiente α_i . Dessa forma, os estimadores de α_i serão melhores quando o número de períodos de tempo observado for grande.

Quando o intercepto α_i é correlacionado com as variáveis explicativas em qualquer período de tempo, o modelo de efeitos fixos é mais adequado para modelar dados em painel. Além disso, como o intercepto do modelo é abordado como um parâmetro fixo, além disso, utiliza efeitos fixos quando as observações são adquiridas de toda população e o que se pretende realizar são inferências para os indivíduos dos quais propõe os dados.

3.3.2 Modelos de Efeito Aleatório

O modelo de efeito aleatório, as variáveis têm as mesmas suposições do modelo de efeito fixo, ou seja, o intercepto muda de um indivíduo para outro, porém não ao decorrer do tempo, e os parâmetros resposta são constantes para todos os indivíduos e perante todas as etapas de tempo. O que muda entre o modelo de efeito fixo e o modelo de efeito aleatório é o tratamento do intercepto. O modelo de efeitos aleatórios analisa os interceptos como variáveis aleatórias, assim, o modelo entende-se que os indivíduos possuem dados que são amostras aleatórias de indivíduos com uma população maior. Assim como proposto por Hill, Griffiths e Judge (1993), serão modelados os n interceptos da seguinte forma:

$$\beta_{0i} = \bar{\beta}_0 + \alpha_i \quad i = 1, \dots, n$$

Consegue-se observar que este intercepto é formado pelo intercepto do modelo de efeitos fixos α_i que verifica as diferenças do comportamento dos indivíduos e por um outro

componente, $\overline{\beta_0}$, refere-se ao intercepto populacional. O modelo geral de efeitos aleatórios é dado da seguinte forma:

$$y_{it} = \overline{\beta_0} + \beta_1 x_{1it} + \cdots + \beta_k x_{kit} + v_{it}$$

Logo, $v_{it} = e_{it} + \alpha_i$ caracteriza o erro. A forma matricial desse modelo, para i -ésimo indivíduo, será dado por:

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \overline{\beta_0} + \begin{bmatrix} x_{1i1} & x_{2i1} & \cdots & x_{ki1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1iT} & x_{2iT} & \cdots & x_{kiT} \end{bmatrix} \times \begin{bmatrix} \beta_{0i1} & \beta_{1i1} & \beta_{2i1} & \cdots & \beta_{ki1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{0iT} & \beta_{1iT} & \beta_{2iT} & \cdots & \beta_{kiT} \end{bmatrix} + \begin{bmatrix} e_{i1} \\ e_{i2} \\ \vdots \\ e_{iT} \end{bmatrix}$$

Hil, Griffiths, Judge (1999) demonstra as quatro propriedades do termo estocástico v_{it} :

1. $E(v_{it}) = 0$
2. $Var(v_{it}) = \sigma_e^2 + \sigma_\alpha^2$
3. $Cov(v_{it}, v_{is}) = \sigma_\alpha^2 \quad \forall t \neq s$
4. $Cov(v_{it}, v_{jt}) = 0 \quad \forall i \neq j$

Desse modo, as duas primeiras propriedades significam que v_{it} tem média zero e variância constante, ou seja, o erro é homocedástico. Na terceira propriedade, tem-se que os erros do mesmo indivíduo em diferentes períodos de tempo são correlacionados, determinando uma autocorrelação. Por fim, a quarta propriedade expressa que os erros de diferentes indivíduos no mesmo instante de tempo não são correlacionados, assim, não ocorre correlação contemporânea.

Assim como ocorre correlação entre os indivíduos dos mesmos indivíduos em períodos de tempo diferentes, o método OLS não é o mais adequado para estimar os coeficientes do modelo de efeitos aleatórios. Assim sendo, o método que fornece os melhores estimadores é o mínimos quadrados generalizados (GLS).

Para testar se o modelo de efeitos aleatórios é apropriado, Breusch e Pagan (1980) apud Greene (1997) desenvolveram um teste baseado no multiplicador de Lagrange, determinando-se as seguintes hipóteses:

$$H_0 : \sigma_\alpha^2 = 0$$

$$H_1 : \sigma_\alpha^2 \neq 0$$

A estatística do teste é dada por:

$$LM = \frac{nT}{2(T-1)} \left[\frac{\sum_{i=1}^n [\sum_{t=1}^T e_{iT}]^2}{\sum_{i=1}^n \sum_{t=1}^T e_{iT}^2} - 1 \right]^2,$$

Conforme a hipótese nula (H_0), LM tem distribuição Qui-Quadrado com um grau de liberdade. Caso a hipótese (H_0) não rejeite, o modelo de efeitos fixos é preferível ao modelo de efeitos aleatórios. Quando não, é necessário atribuir que o modelo de efeitos aleatórios é mais adequado para os dados que sendo abordados.

Segundo Wooldridge (2002), argumenta que a principal razão para escolher entre o modelo de efeito fixos e o modelo de efeitos aleatórios é o efeito não observado α_i . O modelo de efeitos aleatórios é o mais apropriado em circunstâncias que α_i não é correlacionado com todas as variáveis explicativas.

Do contrário, se α_i for correlacionado com algumas variáveis explicativas, logo deve ser aplicado o modelo de efeitos fixos. Portanto, nesse sentido, o modelo de efeitos aleatórios gera estimadores inconsistentes. Greene (1997) propõe o teste de Hausman e formula as hipóteses nula e alternativa para verificar se existe correlação entre α_i e as variáveis explicativas:

$$H_0 : \alpha_i \text{ não é correlacionado com as variáveis explicativas}$$

$$H_1 : \alpha_i \text{ é correlacionado com as variáveis explicativas}$$

Em Greene (1997) acha-se a estatística do teste. Quando a hipótese nula não rejeita, não há evidências que α_i seja correlacionado com as variáveis explicativas. Dessa forma, deve-se aplicar o modelo de efeitos aleatórios. No entanto, se rejeita a hipótese nula, logo utiliza o modelo de efeitos fixos.

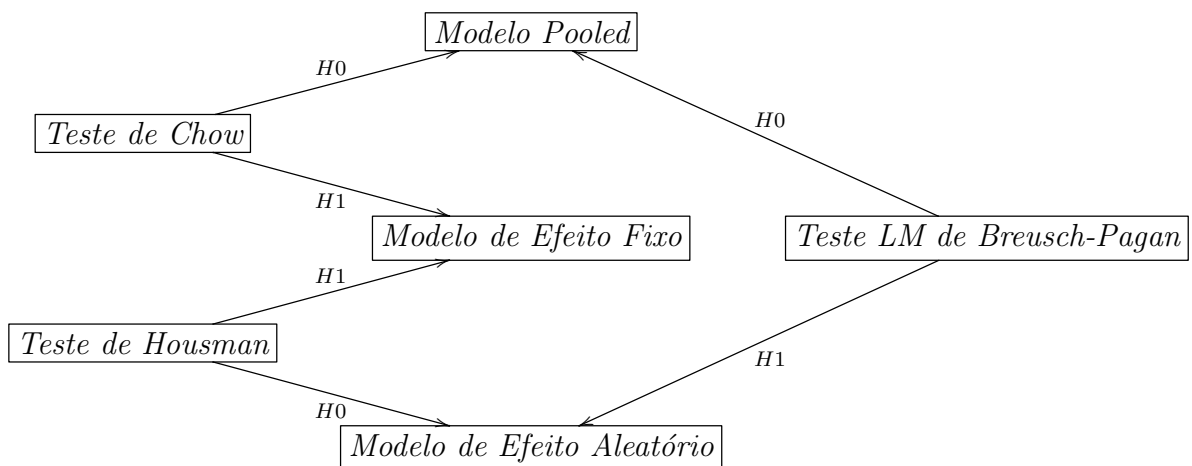
3.3.3 Modelo Pooled

A especificação do modelo pooled (empilhado) cross-section básica é:

$$Y_{it} = a + X_{it}b_1 + e_{it}$$

Confirme Y (taxa de mortalidade infantil) a variável dependente, X é o vetor de variáveis que estabelece o Y , o subscrito t indica o tempo (2005-2010) e o i representa as UFs. Se e_{it} for um ruído branco, pode-se estimar por POLS, ou seja, o estimador OLS ampliado para o caso de painel. Logo, neste modelo a estimação é feita atribuindo que os parâmetros a e b são comuns para todas as UFs.

Para decidir qual o modelo é o mais apropriado para indicar a redução da taxa de mortalidade infantil neste trabalho, serão aplicados dois testes: o teste de Breusch-Pagan, que é utilizado para determinar qual dos modelos é o mais adequado: Pooled ou o de Painel. Pelo teste rejeita-se a hipótese nula de que a variância do coeficiente individual estimado é igual a zero; e o teste de Hausman com a finalidade de decidir entre os modelos de efeito aleatório e efeito fixo, que tem como hipótese nula a não existência de correlação com as variáveis explicativas. O organograma abaixo mostra a estratégia decisória utilizada na análise de Dados em Painel, onde são essenciais o teste de Chow, Bruesch-Pagan e Housman.



Capítulo 4

Resultados e Discussões

Este modelo de dados em painel é formado por uma série temporal 2005 a 2011 e pelas 27 UFs, sendo equilibrado no sentido de todas as unidades estão presente em todos os anos. A tabela 4.1 abaixo apresenta os resultados das variáveis explicativas do modelo estimado pelo método Pooled, que não leva em conta a especificação do efeito individual omitido, tendo um total de 162 observações.

Tabela 4.1: Resultado da Regressão OLS Pooled (Teste de Chow)

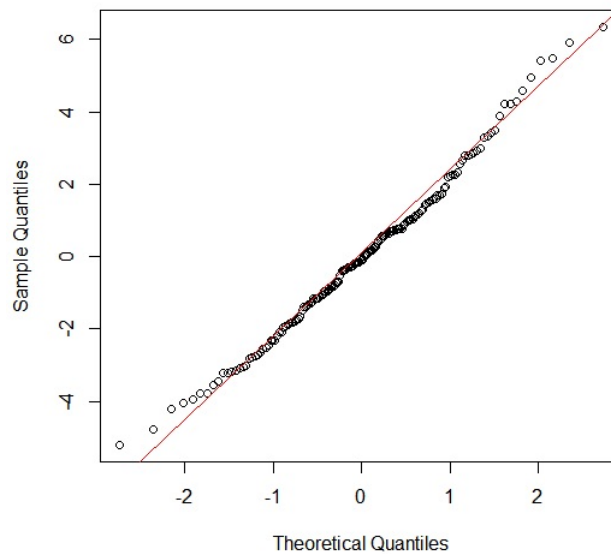
Variáveis	Estimativa	Erro Padrão	t-valor	P-valor
Intercepto	0,1532	3,0391	5,0419	$0,1265 \times 10^{-5}$ ***
X_1	-0,3126	0,0589	-5,2997	$0,3898 \times 10^{-6}$ ***
X_2	$8,1289 \times 10^{-5}$	$3,3951 \times 10^{-5}$	-2,3943	0,0178 *
X_3	0,3132	0,0264	11,8650	$0,0022 \times 10^{-13}$ ***
X_4	-0,0651	0,1899	-3,4275	$0,0779 \times 10^{-2}$ ***
X_5	-0,0016	0,0344	-0,0470	0,9625
R^2	0,7994		R^2 Ajustado	0,7698
Estatística F	124,323		p-valor	$0,0222 \times 10^{-14}$

Os resultados estimados com o modelo OLS Pooled apresentados na Tabela 4.1 mostra que as variáveis X_1 =Taxa de analfabetismo, X_2 =PIB per-capita, X_3 =Percentual de pessoas com baixa renda e X_4 =Percentual da população servida por rede de Abastecimento de água foram significativas para este modelo ao nível de 5%. Somente a variável X_5 = Proporção da população servida por coleta de lixo não apresentou contribuição significativa na TMI.

A taxa de analfabetismo e proporção de pessoas de baixa renda são as que mais contribuíram para diminuir a taxa de mortalidade infantil nas UFs, pelos OLS pooled, uma vez que essas variáveis apresentaram coeficientes mais elevados. Pelo modelo ao ser

excluído o analfabetismo (e aumentando os serviços de educação básica) e baixa renda (tendo salários melhores), a esperança de diminuir a taxa de mortalidade será mais satisfatória. Neste modelo o R^2 foi de 0.7994 e o R^2 ajustado foi de 0.7698, mostrando bom ajuste.

Figura 4.1: Gráfico dos Resíduos para o Modelo OLS pooled



A Figura 4.1, mostra que há desvios de normalidade nas extremidades (caudas pesadas). Para confirmar se há normalidade dos resíduos aplicou-se o teste Kolmogorov-Smirnov, com p -valor = 0.5797, logo, não temos evidências estatísticas para rejeitar a hipótese de normalidade dos dados, considera-se portanto distribuição normal para os resíduos do modelo OLS Pooled. A Tabela 4.2 apresenta os resultados da estimação do modelo de Efeito Fixo, tendo um total de 162 observações.

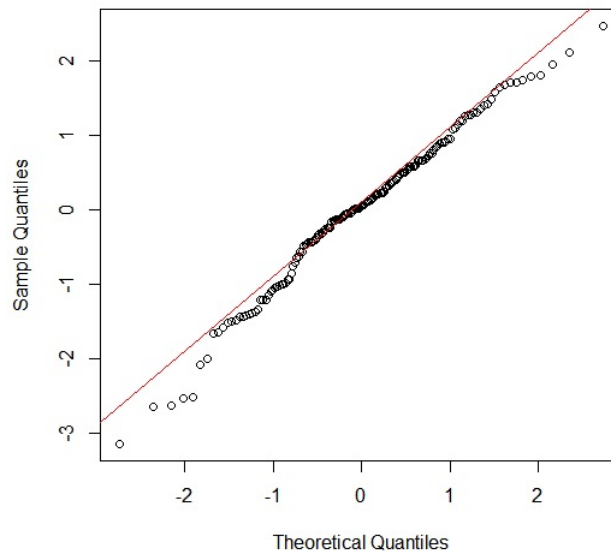
Para o modelo de efeito fixo, as variáveis PIB e percentual da população servida por coleta de lixo não apresentaram evidências estatísticas significante. Já a taxa de analfabetismo, percentual de pessoas com baixa renda e percentual da população servida por rede de abastecimento de água indicam evidências estatísticas para a redução da TMI com 5% de significância. Dessa forma quando a taxa de analfabetismo diminui com 5%, então: a TMI se reduz com 25,552%, a proporção de pessoas com baixa renda se reduz em 34,242% e o abastecimento de água em 12,635%. Neste modelo, R^2 foi igual a 0,73796.

A Figura 4.2, mostra que tem distorções nos resíduos. Logo, para confirmar se há normalidade, aplicou-se o teste Kolmogorov-Smirnov, com p -valor = 0,04082, logo,

Tabela 4.2: Resultado da Regressão do modelo de Efeito Fixo

Variáveis	Estimativa	Erro Padrão	t-valor	P-valor
X_1	0,2555	$1,2000 \times 10^3$	2.1306×10^{-1}	0.0350 *
X_2	$4,7765 \times 10^{-5}$	$5,3000 \times 10^{-1}$	0.9014	0.3691
X_3	0,3424	$4,2500 \times 10^2$	80.521	4.5130×10^{-13} ***
X_4	0,1264	$4,8500 \times 10^2$	26.057	1.0240×10^{-2} *
X_5	-0,0364	$6,0900 \times 10^2$	-0.5978	0.5511
R^2	0,73796		R^2 Ajustado	0,59219
Estatística F	73,2209		p-valor	$0,0222 \times 10^{-14}$

Figura 4.2: Gráfico de Normalidade do Efeito Fixo



temos evidências estatísticas para rejeitar a hipótese de normalidade dos resíduos deste modelo, portanto não há normalidade dos resíduos no modelo de efeito fixo. A Tabela 4.3 apresenta a estimação para o modelo de efeitos aleatórios.

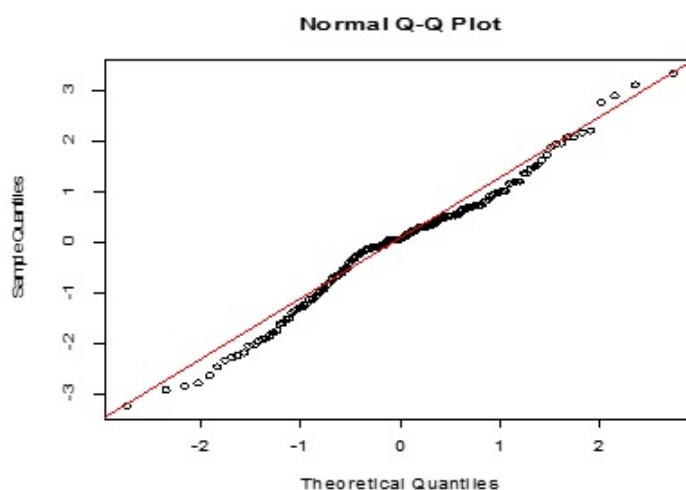
Tabela 4.3: Resultado da Regressão do modelo de Efeito Aleatório

Variáveis	Estimativa	Erro Padrão	t-valor	P-valor
Intercepto	5,9149	4,9104	1,2046	0,2302
X_1	$-1,6521 \times 10^{-1}$	$8,2347 \times 10^{-2}$	-2,0063	0,0466 *
X_2	$2,2580 \times 10^{-5}$	$4,6133 \times 10^{-5}$	0,4895	0,6252
X_3	$3,6210 \times 10^{-1}$	$3,5006 \times 10^{-2}$	1.03437×10^{-1}	$2,0000 \times 10^{-16}$ ***
X_4	$-2,2103 \times 10^{-2}$	$3,2027 \times 10^{-2}$	-0,6901	0,4911
X_5	$7,1262 \times 10^{-3}$	$5,0725 \times 10^{-2}$	0,1405	0,8885
R^2	0,70764		R^2 Ajustado	0,68143
Estatística F	75,518		p-valor	$0,0222 \times 10^{-14}$

Para o modelo com efeito aleatório, as variáveis PIB, abastecimento de água, coleta

de lixo não foram significativos para o modelo. Logo, a variável de percentual de pessoas com baixa renda teve evidência estatística para redução da mortalidade infantil, com maior contribuição, em seguida a taxa de analfabetismo também apresenta contribuição importante para diminuir a mortalidade infantil. Este modelo apresenta um $R^2 = 0,70764$, tendo um modelo razoável para explicar os valores observados.

Figura 4.3: Gráfico de Normalidade do Efeito Aleatório



A Figura 4.3, mostra que há distorções nos resíduos. Logo, para confirmar se há normalidade entre os dados aplicou-se o teste Kolmogorov-Smirnov, com p-valor igual a 0.000413, logo, temos evidências estatísticas para rejeitar a hipótese de normalidade dos dados, portanto não há normalidade dos erros no modelo aleatório. Em todos os modelos estimados as variáveis taxa de analfabetismo e renda apresentaram-se significantes ao nível de 5%.

Em conformidade com os testes executados, o teste de Breusch-Pagan, que serve para escolher entre o modelo de OLS pooled e modelo de dados em painel de efeito aleatório, teve a hipótese nula rejeitada, com 5% de significância, por consequência, o modelo de Dados em Painel de efeito aleatório é o mais apropriado. Em seguida, realizou-se o teste de Hausman que decide sobre o efeito fixo ou aleatório. A sua hipótese nula deste teste é que o efeito é aleatório, ou seja, não há correlação entre as variáveis explicativas e a componente erro do modelo de efeitos aleatório. Após a aplicação deste teste, e o teste de Breusch-Pagan cuja hipótese nula é que o modelo é o OLS pooled versus modelo de efeito aleatório.

A Tabela 4.4 mostra que o teste de Housman decide para o modelo de efeito fixo e o de Breusch-Pagan para o modelo de efeito aleatório. Portanto de acordo com o organo-

grama decide-se pelo modelo de efeito fixo. As soluções dos testes estão na seguinte:

Tabela 4.4: Resultado dos Testes de Breusch-Pagan e Hausman

Testes de Breusch-Pagan	Testes de Hausman
124,323	35,609
$2,200 \times 10^{-16}$	$1,137 \times 10^{-6}$

Com base nos resultados obtidos dos modelos estimados e dos teste executados o modelo de dados em painel mais adequado para explicar a redução da TMI das UFs foi o modelo de efeitos fixos. Assim sendo, a percentual de pessoas com baixa renda tem sido uma variável decisiva para diminuição da TMI, visto que uma proporção pequena de pessoas com baixa renda faz com que beneficie melhor a qualidade e expectativa de vida com mais assistência médica proporciona uma educação de boa qualidade. Considerando que o modelo com as variáveis originais não apresentaram resíduos com provável distribuição normal, admitiu-se realizar uma transformação logarítmica ($Z_i = \ln x$ e $Z = \ln y$) nas variáveis dependentes e independentes do modelo, obtendo-se os resultados da Tabela 4.5 onde se apresenta os efeitos individuais dos estados.

Tabela 4.5: Resultado do Teste de Efeito Individual dos Estados

Variáveis	Estimativa	Erro Padrão	t-valor	P-valor
Z1	-0,1127	0,0487	-2,3161	0,0221 *
Z2	-0,202383	0,051824	-3,9052	0,0002 ***
Z3	0,5331	0,0641	8,3216	$1,0310 \times 10^{-13}$ ***
Z4	-0,1259	0,0641	-1,9648	0,0516
Z5	0,1844	0,1244	1,4823	0,1407
R^2	0,64812		R^2 Ajustado	0,5201
Estatística F	47,889		p-valor	$0,0222 \times 10^{-14}$

O resultado do teste F geral permite decidir que é provável a existência do efeito individual e as variáveis que mais contribuem para isso são: X_1 (taxa de analfabetismo) e X_3 percentual de pessoas com baixa renda são responsáveis. Portanto, fica comprovado a variação significativa da TMI por estados. Com base na transformação feita para os resíduos do modelo de efeito fixo com logaritmo, sendo as variáveis chamada de Z1, Z2, Z3, Z4 e Z5.

A variável Z1 = aumentando uma unidade na taxa de analfabetismo diminui 0.11% na TMI; Z2 = aumentando uma unidade no PIB per capita diminui 0.20% na TMI. Isto pode ser decorrente da elevada concentração que não permite o benefício ser compartilhado por toda população baixa. Este dado mostra de forma irracional do nosso crescimento

econômico onde a justiça social não recebe o benefício do enriquecimento do Estado.

Capítulo 5

Conclusão

Dentro os modelos analisados são: OLS pooled, efeito fixo e aleatório decidiu-se pelo modelo de efeito fixo, onde as variáveis que causam impacto significativo na TMI foram taxa de analfabetismo, percentual de pessoas com baixa renda e o PIB per capita, sendo que esta última variável influencia a TMI de forma inversa, ou seja; quanto maior o PIB per capita menor a TMI parecendo absurdo, porém, isto reflete a previsão do capitalismo brasileiro que produz a doença social da má distribuição da riqueza de cada estado. Isto é confirmado com os imutáveis índices de distribuição de renda do Brasil ao longo dos anos.

O modelo admitido de efeito fixo não possui resíduos normais e tornou-se necessário uma transformação logarítma no modelo nas variáveis dependente e independentes. Este modelo transformado apresentou um melhor ajuste e resíduos com distribuição normal. A contribuição deste trabalho foi mostrar os passos necessários para a análise de DP de forma simples em um programa em linguagem R, tornando acessível realizar esta aplicação, uma vez que, os textos básicos de Análise de Regressão não abordam este tema.

Deixa-se a contribuição de como efetuar uma análise de dados em DP com a utilização do software R e a convicção da sensibilidade de poder diminuir a TMI com investimentos no combate ao analfabetismo e com ações na gestão administrativa que permite numa maior distribuição de renda.

Referências Bibliográficas

- 1 AIDIL Jesus da Silveira Barros, NEIDE Aparecida de Souza Lehfeld. *Fundamentos de metodologia científica – 3ª Ed. – São Paulo: Pearson Prentice Hall, 2007.*
- 2 ALMEIDA, A. S. – *Estimativa da mortalidade infantil em 50 cidades do Estado da Bahia. Rev. Serv. Saúde públ., Rio de Janeiro, 8:153-68, 1955.*
- 3 ANTONOVSKI, A. & BERNSTEIN, J. *Social class and infant mortality. Soc. Sci. Med., 11:453-70, 1977.*
- 4 BRASS W. *Uses of census or survey data for the estimation of vital rates. 1964. Paper prepared for the African Seminar on Vital Statistics, 14–19 December, 1964. Addis Ababa.*
- 5 BRASS W. *Methods for estimating fertility and mortality from limited and defective data. Chapel Hill, North Carolina: Carolina Population Center, Laboratories for Population Statistics; 1975.*
- 6 CAMERON, A. C.; TRIVEDI, P. K. *Microeconometrics using Stata. College Station: Stata CAMPOS, Tatiana P.; SA CARVALHO, M.; BARCELLOS, Christovam C. Mortalidade infantil no Rio de Janeiro, Brasil: áreas de risco e trajetória dos pacientes até os serviços de saúde. PAN AMERICAN JOURNAL OF PUBLIC HEALTH, v. 8, p. 164-171, 2000.*
- 7 CAMPOS, TATIANA P.; SA CARVALHO, M.; BARCELLOS, Christovam C. *Mortalidade infantil no Rio de Janeiro, Brasil: áreas de risco e trajetória dos pacientes até os serviços de saúde. PAN AMERICAN JOURNAL OF PUBLIC HEALTH, v. 8, p. 164-171, 2000.*
- 8 CHASE, H.C. *Ranking countries by infant mortality rates. Publ. Hlth Rep., 84:19-27, 1969. CMCCD (Committee on Medical Certification of Causes of Death), 1958. Problems*

in the medical certification of causes of death. American Journal of Public Health, 48:71-80.

9 COSTA, Maria da Conceição Nascimento et al. *Mortalidade infantil no Brasil em períodos recentes de crise econômica. Revista de Saúde Pública, v. 37, n. 6, p. 699-706, 2003.*

10 DATASUS - Departamento de Informática do Sistema Único de Saúde: Disponível em: <<http://datasus.saude.gov.br/>>.

11 DE ARAÚJO, José Duarte. *A mortalidade infantil no Estado da Bahia, Brasil. Revista de Saúde Pública, v. 7, n. 1, p. 29-36, 1973.*

12 DEATON, Angus. *Panel data from time series of cross-sections. Journal of econometrics, v. 30, n. 1, p. 109-126, 1985.*

13 DUARTE, P.C.; LAMOUNIER, W.M.; TAKAMATSU, R.T. *Modelos Econométricos para Dados em Painel: Aspectos Teóricos e Exemplos de Aplicação à Pesquisa em Contabilidade e Finanças.*

14 GIL, Antonio Carlos. *Como elaborar projetos de pesquisa. São Paulo: Atlas, 1991.*
HSIAO, C. *Analysis of panel data. Cambridge: UP, 1999. (Econometric Society HSIAO, Cheng. Analysis of panel data. Cambridge: Cambridge University Press, 1986.*

15 IBGE – Instituto Brasileiro de Geografia e Estatística: Disponível em: <<http://www.ibge.gov.br/home/>>.

16 IBGE (Fundação Instituto Brasileiro de Geografia e Estatística), 1998. *Estatísticas do Registro Civil, 1996. v. 23. Rio de Janeiro: IBGE.*

17 IBGE (Fundação Instituto Brasileiro de Geografia e Estatística), 2001. *População. Indicadores Sociais.* <<http://www.ibge.gov.br>>.

18 LAURENTI R. *Análise da informação em saúde: 1893-1993, cem anos da Classificação Internacional de Doenças. Rev Saúde Pública 1991; 25:407-17.*

19 MARCONI, Marina de Andrade; LAKATOS, Eva Maria. – 6ª Ed. – São Paulo: Atlas, 2001. *Metodologia do Trabalho Científico: procedimentos básicos, pesquisa bibliográfica, projeto e relatório, publicações e trabalhos científicos.*

- 20 MARQUES, L. D. *Modelos dinâmicos com dados em painel: revisão da literatura. Série Working Papers do Centro de Estudos Macroeconômicos e Previsão (CEMPRE) da Faculdade de Economia do Porto, Portugal, n. 100, 2000.*
- 21 MELLO JORGE, M. H. P. et al. *O Sistema de Informações sobre Mortalidade: passado, presente e futuro. São Paulo: CBCD; 2006. (Série Divulgação nº 11).*
- 22 MELLO, Jorge Maria Helena Prado de Mello; GOTLIEB, Sabina Léa Davidson and LAURENTI, Ruy. *O sistema de informações sobre mortalidade: problemas e propostas para o seu enfrentamento I - Mortes por causas naturais. Rev. bras. epidemiol. [online]. 2002, vol.5, n.2, pp. 197-211.*
- 23 MELLO-JORGE, M. H. P., 1983. *Sub-registro dos eventos vitais. Revista de Saúde Pública, 17:148-151.*
- 24 MELLO-JORGE, M. H. P.; GOTLIEB, S. L. D. & OLIVEIRA, H., 1996. *O Sistema de Informação sobre Nascidos Vivos: Primeira avaliação dos dados brasileiros. Informe Epidemiológico do SUS, 2:15-48. Monographs, n. 11).*
- 25 MORTARA, G. — *Estimativa do número de óbitos e da taxa de mortalidade geral para o Brasil (1950). In: INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Pesquisa sobre a mortalidade no Brasil: 2.a série. Rio de Janeiro, 1956. p. 56-7. (Estudos de Estatística Teórica e Aplicada. Estatística Demográfica, 20)*
- 26 MS (Ministério da Saúde), 1982. *O Subsistema de Informações sobre Mortalidade. In: Estatísticas de Mortalidade: Brasil, 1979 (Ministério da Saúde, org.), pp. 2-56, Brasília: Centro de Documentação, MS.*
- 27 PAES, Neir Antunes. *Qualidade das estatísticas de óbitos por causas desconhecidas dos Estados brasileiros. Rev. Saúde Pública [online]. 2007, vol.41, n.3, pp. 437.*
- 28 PNUD/FAO/IBAMA/UFPB/GOV. PARAÍBA. *Diagnóstico do setor florestal do Estado da Paraíba. João Pessoa: PNUD / FAO / BRA. 87- 007, 1994. 84 p. Press, 2009.*
- 29 Rajaratnam JK, Tran LN, Lopez AD, Murray CJL. *Measuring Under-Five Mortality: Validation of New Low-Cost Methods. PLoS Med. 2010;7(4):e1000253. doi: 10.1371/journal.pmed.1000253. [PMC free article] [PubMed]*

30 RUTSTEIN, D. D.; BERENBERG, W.; CHALMERS, T. C.; CHILD, C. G.; FISHMAN, A. P.; PERRIN, E. B. *Measuring the quality of medical care: a clinical method. N. Engl. J. Med.*, 294: 582-8, 1976.

31 SCHRAMM, J. M. A. & SZWARCOWALD, C. L., 2000. *Sistema hospitalar como fonte de informações para estimar a mortalidade neonatal e a natimortalidade. Revista de Saúde Pública*, 34:272-279. *Spatial Panel data models in R*

32 STOCKWELL, E.G. *Infant mortality and socioeconomic status: a changing relationship. Milbank mem. Fd Quart.*, 11:101-11, 1962.

33 Vallin J, Pollard JH, Heligman L, editors. (1984) *Methodologies for the collection and analysis of mortality data. Ordina, Liege: International Union for the Scientific Study of Population.*

34 VALLIN. J. & BEHM, H. *Mortality differentials among human groups. [Paper presented to the Seminar on Biological and Social Aspects of Mortality and the Length of Life. International Union for the Scientific Study of Population. Fiuggi Terme, Italy, 1980].*

Apêndice A

Programas em R para Dados em Painel com Regressão Linear

```
#####  
Dados  
#####  
library(plm) ## pacote plm##  
Maizza<-read.csv("C:/Users/MAIZZA MICAELE/Documents/Monografia/Maizza3.txt",sep=";",header = T)  
Maizza      ##arquivo do Banco de dados##  
attach(Maizza)  
View(Maizza)  
names(Maizza)  
length(names(Maizza))  # Tamanho  
dim(Maizza)            # Linhas e Colunas  
summary(Maizza)  
#####  
Aplicação dos Modelos  
#####  
###Teste de Chow##  
gpooling<-plm(Y~X1 + X2 + X3 + X4 + X5, data = Maizza, index= c("Estado","Ano"), model = "pooling")  
summary(gpooling)  
qqnorm(gpooling$residuals)
```

```

abline(0.1,2.3,col="red")

### Teste Lilliefors###

library(nortest)

lillie.test(gpooling$residuals)

###Modelo de Efeito Fixo###

gfixed<-plm(Y~X1 + X2 + X3 + X4 + X5, data = Maizza, index= c("Estado","Ano"), model = "within")

summary(gfixed)

qqnorm(gfixed$residuals)

abline(0.1,1.0,col="red")

### Teste Lilliefors###

library(nortest)

lillie.test(gfixed$residuals)

###Modelo de Efeito Aleatório###

grandom<-plm(Y~X1 + X2 + X3 + X4 + X5, data = Maizza, index= c("Estado","Ano"), model = "random")

summary(grandom)

qqnorm(grandom$residuals)

abline(0.1,1.2,col="red")

## Teste Lilliefors##

library(nortest)

lillie.test(gradon$residuals)

## Teste de Housman##

phtest(gfixed, grandom)

### Teste Breusch Pagan ###

##Random effect and Pooling ##

plmtest(gpooling, type=c("bp"))

##Teste para saber se existe o efeito individual dos estados###

grandom<-plm(Y~X1 + X2 + X3 + X4 + X5, data = Maizza, index= c("Estado","Ano"), model = "random", effect=c("individ

summary(grandom)

#=====#

Transformação Logaritma dos Resíduos do Modelo de Efeito Fixo

#=====#

```



```
Maizza<-read.csv("C:/Users/MAIZZA MICAELE/Documents/Monografia/Maizza3.txt", sep=";", header = T)

Maizza          ##arquivo do Banco de dados##

attach(Maizza)

names(Maizza)

length(names(Maizza))  # Tamanho

dim(Maizza)           # Linhas e Colunas

Estado= Maizza$Estado

Ano = Maizza$Ano

log(X1)

Z1=log(X1)

Z1

Z2=log(X2)

Z2

Z3=log(X3)

Z3

Z4=log(X4)

Z4

Z5=log(X5)

Z5

Z=log(Y)

Z
```