

UNIVERSIDADE FEDERAL DE JUIZ DE FORA

Departamento De Estatística

Curso de Estatística

Jarbas Hagen da Costa

**Fatores Associados ao Ganho na Esperança de Vida Ao Nascer Nos  
Municípios de Minas Gerais nos Períodos 1991-2000 e 2000-2010**

Juiz de Fora

2014

**Jarbas Hagen da Costa**

*Fatores Associados Ao Ganho Na Esperança de Vida Ao Nascer Nos Municípios de  
Minas Gerais Nos Períodos 1991-2000 e 2000-2010*

Trabalho de Conclusão de Curso  
apresentado ao curso de Estatística da  
Universidade Federal de Juiz de Fora, como  
requisito para obtenção do diploma de  
bacharel em Estatística.

**Orientador: Luiz Cláudio Ribeiro**

Juiz de Fora

2014

Ficha catalográfica elaborada através do Programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Hagen da Costa, Jarbas.

Fatores Associados ao Ganho na Esperança de Vida Ao Nascer Nos Municípios de Minas Gerais nos Períodos 1991-2000 e 2000-2010 / Jarbas Hagen da Costa. -- 2014.

78 f.

Orientador: Luiz Cláudio Ribeiro

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Juiz de Fora, ICE/Engenharia, 2014.

1. Regressão Linear. 2. Esperança de Vida ao Nascer. 3. Estatística Espacial. I. Ribeiro, Luiz Cláudio, orient. II. Título.

**Jarbas Hagen da Costa**

Fatores Associados Ao Ganho Na Esperança de Vida Ao Nascer Nos Municípios de  
Minas Gerais Nos Períodos 1991-2000 e 2000-2010

Trabalho de Conclusão de Curso  
apresentado ao curso de Estatística da  
Universidade Federal de Juiz de Fora, como  
requisito para obtenção do diploma de  
bacharel em Estatística.

Aprovado em 14 de julho de 2014.

**BANCA EXAMINADORA**

---

Doutor em Demografia. Luiz Claudio Ribeiro  
Universidade Federal de Juiz de Fora

---

Doutor Em Estatística. Clécio da Silva Ferreira  
Universidade Federal de Juiz de Fora

---

Mestre em Engenharia de Produção. Luiz Fernando Soares de Castro  
Universidade Federal de Juiz de Fora

## **AGRADECIMENTOS**

Agradeço em primeiro lugar, aos meus pais e minha família por todo apoio dispensado a mim nessa longa caminhada que foi o curso de Estatística. À minha namorada Isabela que sempre esteve ao meu lado me apoiando durante o curso, pelo amor e carinho comigo e principalmente pela paciência nos momentos mais difíceis. Aos companheiros de música que sempre foram solidários e me ajudaram com momentos de diversão impagáveis. Ao Professor Alfredo que sempre esteve ao nosso lado quando precisamos durante toda essa jornada. Ao Alex pela ajuda importantíssima com as estatísticas espaciais. Ao Professor Clécio que nos orientou em momentos complicados desse trabalho. Aos amigos Marcos, Cadu, Thales que passaram pelas mesmas dificuldades que passei sempre com bom humor para enfrentá-las diariamente. Ao professor Fernando Tavares pela oportunidade de primeira experiência de trabalho com Estatística concedida. Por fim agradeço ao professor Luiz Cláudio pela dedicação e orientação e também pelas conversas sobre música nas horas vagas.

Agradeço a todos que colaboraram de alguma forma para que esse sonho se tornasse realidade.

## RESUMO

Nas últimas décadas foram observadas melhorias nos âmbitos sociais e econômicos no país. A esperança de vida ao nascer é um importante indicador de qualidade de vida, que pode ser utilizada na medição de índices de desenvolvimento humano de um país ou região. Fatores como queda na mortalidade infantil, renda, desigualdade social, entre outros, estão estreitamente ligados ao crescimento da esperança de vida ao nascer (Wood & Carvalho, 1994). Dito isso, o trabalho desenvolvido tem o objetivo de estudar fatores socioeconômicos e demográficos que estejam associados ao ganho relativo na esperança de vida ao nascer nos municípios do Estado de Minas Gerais, nos períodos 1991-2000 e 2000-2010. As associações foram investigadas a partir da correlação de Pearson e modelos de regressão linear. Também foram aplicadas técnicas de estatística espacial, a fim de se obter comparações regionais entre os municípios. As comparações foram realizadas em diferentes cenários temporais, com dados censitários dos anos 1991, 2000 e 2010. Os dados utilizados no trabalho são provenientes do Atlas do Desenvolvimento Humano 2013, disponibilizado gratuitamente pelo Programa de Desenvolvimento das Nações Unidas.

Palavras Chave: Regressão Linear, Estatística Espacial, Esperança de Vida ao Nascer

## **ABSTRACT**

In recent decades improvements were observed in social and economic spheres in the country. Life expectancy at birth is an important indicator of quality of life, which can be used in measuring human development index of a country or region. Factors such as a decline in infant mortality, income inequality among others are closely linked to the growth in life expectancy at birth (Wood & Carvalho, 1994). Therefore, the work aims to study socioeconomic and demographic factors that are associated with the relative gain in life expectancy at birth in the municipalities of Minas Gerais, in the periods 1991-2000 and 2000-2010. Associations were investigated by the Pearson correlation and linear regression models. Spatial statistical techniques were also applied in order to obtain regional comparisons between municipalities. Comparisons were performed at different temporal scenarios, with census data for the years 1991, 2000 and 2010. Data used in this paper come from the Atlas of Human Development 2013, released by the Development Program of the United Nations for free.

**Keywords:** Linear Regression, Spatial Statistics, Life Expectancy at Birth

## Sumário

1 – Introdução .....	9
2 - Metodologia.....	12
2.1 - Manipulação e Base de Dados .....	12
2.2 - Dependência Espacial .....	12
2.2.1 – Índice de Moran .....	13
2.3 - Coeficiente de Correlação Linear.....	13
2.4 - Regressão Linear.....	15
2.4.1 – Regressão Linear Simples.....	15
2.4.1.1 – Adequação dos Modelos .....	16
2.4.2 – Regressão Linear Múltipla .....	17
2.4.2.1 – Coeficiente de Determinação Múltipla .....	19
2.4.2.2 – Análise Residual .....	20
3- Resultados .....	22
3.1 – Estatísticas Descritivas da Esperança de vida ao nascer .....	22
3.2 - Correlações.....	28
3.3 – Modelos de Regressão .....	31
3.3.1 – Modelos 1991-2000 .....	32
3.3.2- Modelos 2000-2010 .....	43
4 – Conclusões e Considerações Finais .....	61
4.1 – Trabalhos Futuros.....	62
Bibliografia.....	63
Anexo 1: .....	65
Anexo 2: .....	73
Anexo 3.....	76



## 1 – Introdução

Nas últimas duas décadas puderam ser observadas diversas mudanças de âmbito socioeconômico no cenário nacional, muitas delas para melhor. A população brasileira presenciou mudanças como diferentes moedas, o Plano Real, o controle da inflação, um aumento nas oportunidades de empregos, melhor distribuição de renda e melhorias em saneamento básico, saúde e educação. Tais mudanças interferiram positivamente na qualidade de vida do brasileiro o que desencadeou, na maioria dos municípios do país, um acréscimo na Esperança de Vida ao Nascer, que dentre outros indicadores, demonstra a qualidade de vida em uma determinada região.

Segundo Preston, Heuveline e Guilot (2001), o conceito de expectativa de vida ao nascer é a soma das idades vividas por todas as pessoas numa coorte, dividida pelo número de pessoas na coorte. Ou ainda, é a média de idade da pessoa ao morrer, em uma coorte. Essa coorte pode, por exemplo, ser a cidade que se pretende estudar.

A Esperança de Vida ao Nascer é conceituada pelo DATASUS, como:

“Número médio de anos de vida esperados para um recém nascido, mantido o padrão de mortalidade existente na população residente, em determinado espaço geográfico, no ano considerado.”

Além disto, pode ser interpretada como o número médio de anos que se espera que um recém nascido viva.

A Esperança de Vida ao Nascer é um importante indicador socioeconômico, visto que seu aumento indica melhores condições de vida e saúde da população estudada. Sendo assim, seu ganho pode estar estreitamente ligado a fatores de cunho econômicos e sociais, como melhorias em saneamento básico, condições econômicas entre outros (Wood & Carvalho, 1994).

Fatores como a queda da mortalidade, renda familiar, educação, desigualdade, distribuição geográfica da população, residência rural ou urbana estão estreitamente ligadas à esperança de vida ao nascer (Wood & Carvalho, 1994).

Esse indicador também é considerado uma medida resumo que não sofre influência da estrutura etária da população e é obtida a partir de tábuas de vida calculadas no período estudado (REDE Interagencial de Informação para a Saúde, 2008; Preston, Heuveline & Guilot, 2001).

Dentre outras várias aplicações, a Esperança de Vida ao Nascer é utilizada no cálculo de índices, como o Índice de Desenvolvimento Humano (IDH), analisar variações na expectativa de vida ao longo do tempo e nas regiões, bem como avaliar os níveis de saúde e vida de uma população. Também é utilizada para cálculos previdenciários, gestão de gastos em saúde e qualidade de vida, além de cálculos em seguros, empréstimos, entre outros (Ministério da Saúde)

A Esperança de Vida ao Nascer é calculada da seguinte forma:

“a partir de tábuas de vida elaboradas para cada área geográfica, toma-se o número correspondente a uma geração inicial de nascimentos ( $l_0$ ) e determina-se o tempo cumulativo vivido por essa mesma geração ( $T_0$ ) até a idade limite. A esperança de vida ao nascer é o quociente da divisão de  $T_0$  por  $l_0$ .” (REDE Interagencial de Informação para a Saúde, 2008)

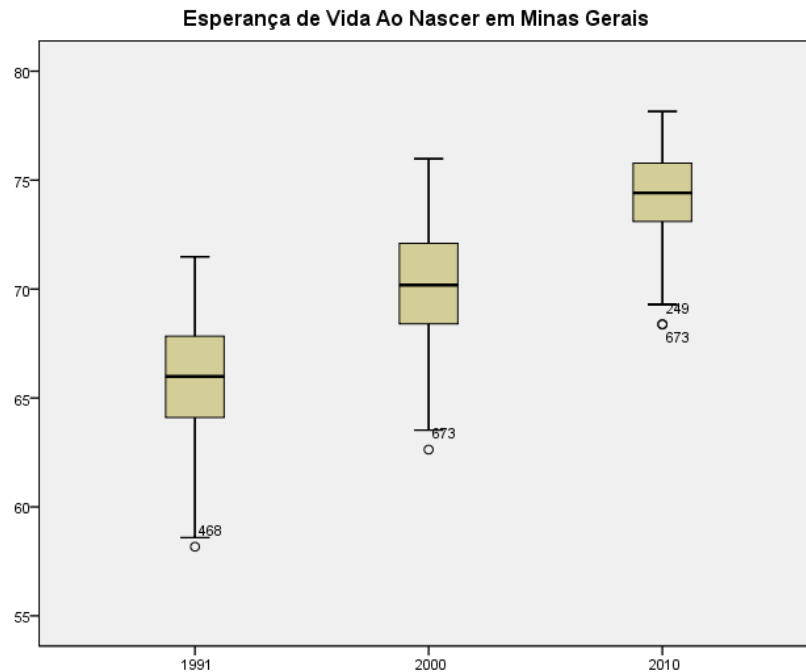
Segundo o IBGE (2010), a Esperança de Vida ao nascer vem crescendo ao longo do tempo, em todas as regiões do Brasil para ambos os sexos. Em 1991 a expectativa de vida dos brasileiros era de 68,97 anos e chegou a 73,29 anos em 2009, segundo dados do DataSUS (2014). A Tabela 1 mostra os resultados pelas regiões e também para o Estado de Minas Gerais, alvo do presente estudo.

Tabela 1 – Esperança de vida ao nascer por região

<b>Regiões</b>	<b>1991</b>	<b>1995</b>	<b>2000</b>	<b>2005</b>	<b>2009</b>
<b>Região Norte</b>	66,92	68,13	69,53	71,02	72,09
<b>Região Nordeste</b>	62,83	64,77	67,15	69,00	70,48
<b>Região Sudeste</b>	68,83	70,29	71,99	73,49	74,69
<b>Região Sul</b>	70,40	71,47	72,74	74,17	75,28
<b>Região Centro-Oeste</b>	68,55	70,03	71,75	73,19	74,25
<b>Brasil</b>	66,93	68,50	70,44	72,05	73,29
<b>Minas Gerais</b>	68,97	70,73	72,73	74,10	75,16

Este crescimento também foi observado nos municípios do Estado de Minas Gerais e está evidenciado observando-se o Gráfico 1.

Gráfico 1 – Box Plot Esperança de Vida ao Nascer Minas Gerais



Em virtude da importância desse indicador, bem como de seu crescimento, este trabalho tem por objetivo encontrar fatores socioeconômicos e demográficos associados a este crescimento, uma vez que os mesmos interferem neste aspecto (Pacievitch, 2010). Estudamos os municípios do Estado de Minas Gerais entre os anos de 1991-2000 e 2000-2010, a partir dos indicadores do Atlas do Desenvolvimento Humano do Programa das Nações Unidas para o Desenvolvimento (PNUD) que utiliza os dados do Censo Demográfico Brasileiro do IBGE em seus cálculos. Foram propostos modelos de regressão linear para estudar o ganho na esperança de vida ao nascer.

A migração interfere na esperança de vida ao nascer, uma vez que a estrutura etária de uma população é utilizada no cálculo deste indicador. A migração de jovens para regiões com melhores oportunidades e o aumento de idosos em algumas regiões de Minas Gerais, como Norte e Nordeste, faz com que a população fique mais jovem ou mais velha, respectivamente (Brito, 2009). Porém, no presente trabalho foram consideradas populações fechadas para os municípios e períodos estudados, desconsiderando a migração.

## 2 - Metodologia

### 2.1 - Manipulação e Base de Dados

Os dados brutos utilizados neste estudo foram obtidos do Atlas do Desenvolvimento Humano no Brasil, 2013 que estão disponíveis no site da PNUD (<http://www.pnud.org.br>). Os dados consistem em índices obtidos a partir dos Censos dos anos 1991, 2000 e 2010 realizados pelo IBGE.

Foram consideradas no estudo apenas variáveis socioeconômicas e demográficas e cada variável selecionada foi transformada em ganhos relativos, nos períodos de 1991-2000 e 2000-2010. Os modelos de regressão foram propostos a partir dessas variáveis de ganho relativo calculadas.

Os dados foram manipulados e analisados com aplicação do *software* SPSS, versão 15.0.

Os ganhos relativos foram calculados pela expressão:

$$\frac{\text{Dados relativos ao período } N}{\text{Dados relativos ao período } (N - 1)} - 1$$

Dessa forma, obtivemos os ganhos em percentual das variáveis estudadas. Houveram situações onde o ganho foi superior a 100%, resultando em ganhos relativos maiores que um.

### 2.2 - Dependência Espacial

Como se tratam de dados referentes aos territórios de Minas Gerais, foram utilizadas estatísticas espaciais, calculadas com o auxílio dos *softwares* livres *QuantumGis* e *TerraView*. Tais programas também auxiliaram na construção dos mapas utilizados nesse estudo.

A dependência espacial é uma característica referente a dados com delimitações territoriais, que pode ser entendida como a tendência de que um valor de uma variável associada a uma determinada região seja semelhante ao valor de vizinhos, seja por fronteiras, distância linear ou outras (Cardoso, 2006).

No presente trabalho estamos interessados nos Ganhos Relativos na Esperança de Vida ao Nascer, nos municípios do Estado de Minas Gerais, e isso pode gerar uma auto correlação espacial dos dados para a variável estudada.

A fim de quantificar essa dependência, podem ser utilizadas várias estatísticas, neste trabalho foi utilizado o Índice de Moran, por ser o mais difundido para tal análise. (Bailey & Gatrell, 1996)

### 2.2.1 – Índice de Moran

O Índice de Moran é uma medida que mostra a auto correlação espacial dos dados, também conhecida como correlação espacial. É a técnica mais utilizada neste tipo de análise (Bailey & Gatrell, 1996).

Seja  $W$  a matriz de proximidade espacial, então o índice de Moran é calculado da seguinte maneira:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{(\sum_{i=1}^n (y_i - \bar{y})^2) (\sum \sum_{i \neq j} w_{ij})}, i, j = 1, 2, \dots, n$$

Onde:

$w_{ij}$  é o elemento da  $i$ -ésima linha e  $j$ -ésima coluna da matriz de proximidade espacial  $W$ ;

$y_i$  são os valores atribuídos a cada local estudado por exemplo o ganho na esperança de vida ao nascer em Juiz de Fora;

$\bar{y}$  é a média desses valores.

Assim como na correlação linear (não espacial), os valores do índice estão restritos ao intervalo  $[-1,1]$  e sua interpretação é feita de forma análoga. Na seção a seguir será apresentado o coeficiente de correlação linear (não espacial).

### 2.3 - Coeficiente de Correlação Linear

Sejam duas amostras das variáveis  $X$  e  $Y$ :

$X_i$	$X_1$	$X_2$	...	$X_n$
$Y_i$	$Y_1$	$Y_2$	...	$Y_n$

O coeficiente de correlação entre as variáveis  $X$  e  $Y$  é calculado por:

$$r_{XY} = \frac{\widehat{COV}(X, Y)}{\sqrt{\widehat{V}(X)\widehat{V}(Y)}}$$

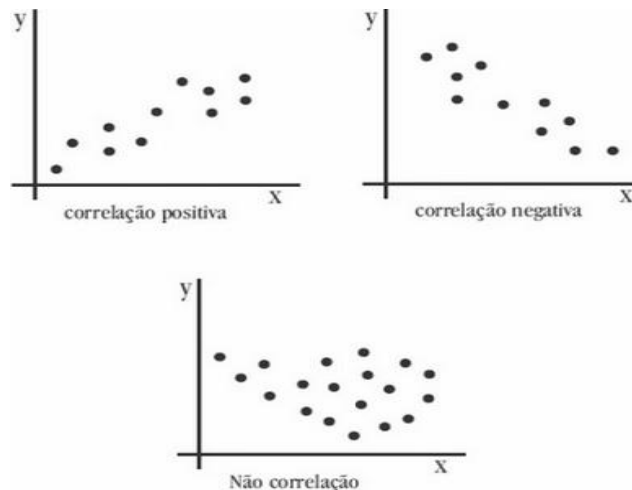
$$r_{XY} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Sendo assim, temos:

$$-1 \leq r_{XY} \leq 1$$

Se uma variável “cresce” e a outra também “cresce”, então temos uma correlação positiva. Analogamente, quando uma variável “cresce” e a outra “decrece”, temos uma correlação negativa. O Gráfico 2 ilustra os tipos de correlações :

Gráfico 2: Tipos de Correlação



A correlação pode ser interpretada da seguinte maneira:

- se  $|r_{XY}| > 0,7$ , temos uma forte correlação;
- se  $0,3 < |r_{XY}| \leq 0,7$ , temos uma correlação moderada;
- se  $0 < |r_{XY}| \leq 0,3$ , temos uma correlação fraca;
- se  $r_{XY} = -1$  ou  $r_{XY} = 1$ , temos uma correlação perfeita negativa e positiva, respectivamente.

No presente estudo, a correlação de Pearson foi utilizada para determinar a seleção das variáveis correlacionadas com a variável resposta “Ganho na Esperança de Vida ao Nascer”, no modelo de regressão linear. Foram excluídas variáveis descorrelacionadas e/ou com fraca correlação. Também foi feita a verificação da

correlação bivariada de cada variável independente candidata a compor o modelo, a fim de evitar multicolinearidade.

## 2.4 - Regressão Linear

A regressão linear é uma análise estatística que tem o objetivo de identificar uma relação funcional entre uma variável dependente e um conjunto de variáveis independentes. Ou seja, busca uma equação que tenta explicar o comportamento da variável dependente em função da variação das variáveis explicativas (Charnet, Freire, Charnet, & Bonvino, 2008).

A análise de regressão pode ser utilizada para a construção de um modelo preditor do comportamento de uma variável resposta, utilizando-se de uma ou mais variáveis auxiliares correlacionadas com a mesma.

Nas próximas seções serão abordados os modelos de Regressões Lineares Simples e Múltipla, sendo utilizados como principais referências (Charnet, Freire, Charnet, & Bonvino, 2008) e (Montgomery & Runger, 2003)

### 2.4.1 – Regressão Linear Simples

A regressão linear simples considera um único regressor ou preditor, ou seja, utiliza-se apenas de uma variável auxiliar para explicar a variável resposta. Supondo que a relação entre  $Y$  e  $x$  seja uma linha reta e que a observação de  $Y$  em cada nível de  $x$  seja uma variável aleatória, obtemos o seguinte modelo:

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

Onde:

$\beta_0$  é a intercepto com o eixo  $y$ ;

$\beta_1$  é a inclinação, coeficientes desconhecidos da regressão;

$\varepsilon$  é o erro;

A estimação de  $\beta_0$  e  $\beta_1$ , coeficientes de regressão, é feita pelo método de mínimos quadrados e deve resultar em uma linha que é o melhor ajuste para os dados e que minimiza a soma dos quadrados dos desvios verticais dos dados observados, em relação à reta estimada pela regressão.

Abaixo seguem as estimativas de mínimos quadrados da interseção e inclinação do modelo de regressão simples:

$$\hat{\beta}_0 = \frac{\sum Y_i}{n} - \hat{\beta}_1 \frac{\sum X_i}{n} \quad \text{ou} \quad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$\hat{\beta}_1 = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

Conseqüentemente, obtemos a seguinte linha estimada, ou ajustada, de regressão:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

#### 2.4.1.1 – Adequação dos Modelos

A significância de um modelo de regressão ajustado pode ser verificada a partir da Análise de Variância (ANOVA) com hipótese nula  $H_0: \beta_1 = 0$ . Ou seja, rejeitando a hipótese nula, temos que  $x$  é importante na explicação de  $Y$ . A ANOVA será abordada nas próximas sessões.

Para ajustar um modelo de regressão devem-se respeitar os seguintes pressupostos: a estimação de parâmetros requer que os erros sejam variáveis aleatórias descorrelacionadas, que tenham valor esperado igual a 0; variância constante, ou seja, os erros devem ser homocedásticos e os erros também devem seguir uma distribuição normal.

Além disso, ao ajustar um modelo de regressão, consideramos que a ordem do modelo esteja correta, ou seja, tomamos como verdadeira a suposição de que o fenômeno estudado realmente se comporta de maneira linear.

Para verificar a adequação do modelo de regressão utilizamos a análise residual, que mostra se o modelo segue os pressupostos citados acima. Os resíduos possuem a seguinte forma:

$$e_i = y_i - \hat{y}_i, \quad i = 1, 2, \dots, n$$

Onde:

$y_i$  é a observação real do fenômeno estudado;

$\hat{y}_i$  é o valor ajustado pelo modelo de regressão.

A análise residual ajuda a verificar as suposições de normalidade, homocedasticidade e independência dos erros. A normalidade dos erros pode ser



verificada a partir de testes de normalidade como Kolmogorov-Smirnov, Shapiro-Wilks e Lilliefors ou graficamente. Já a homocedasticidade e a independência dos erros podem ser averiguadas por meio de verificações gráficas especificadas na análise residual.

Para constatar a adequação do modelo utiliza-se o coeficiente de determinação, que é dado por  $R^2$ . Essa medida nos diz o quanto o modelo proposto explica a variabilidade do fenômeno estudado. Por exemplo, um  $R^2=0,90$  nos diz que o modelo explica 90% da variação da variável resposta.

#### 2.4.2 – Regressão Linear Múltipla

Em várias aplicações de análise de regressão, temos mais de uma variável auxiliar explicativa para a variável resposta, ou seja, temos mais de um regressor. Para tal aplicação, é utilizado o chamado modelo de regressão linear múltipla.

O modelo de regressão linear múltipla, com  $k$  regressores. é dado por:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

Onde:

$Y$  é a variável resposta;

$\beta_j, j = 1, 2, \dots, k$ , são os coeficientes de regressão;

$x_j$ , são as variáveis explicativas ou variáveis regressoras;

$\varepsilon$  representa o erro, com valor esperado igual a zero.

Os parâmetros  $\beta_j$  representam a variação esperada da variável resposta  $Y$ , por unidade de variação unitária das variáveis regressoras  $x_j$ , quando essas são mantidas constantes.

A representação com um sistema de equações lineares do modelo de regressão linear múltipla, com  $k$  regressoras e  $n$  observações é o modelo dado por:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

Este modelo é um sistema com  $n$  equações que pode ser expresso na forma matricial:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

Onde:

$\mathbf{y}$  é a vetor das respostas obtidas a partir do modelo;

$\mathbf{X}$  a matriz de observações das variáveis regressoras, com a primeira coluna preenchida por 1;

$\boldsymbol{\beta}$  a vetor de coeficientes de regressão;

$\boldsymbol{\varepsilon}$  é a vetor dos erros.

Os parâmetros  $\beta_j$  são desconhecidos, porém podemos estima-los a partir do método dos mínimos quadrados. As equações de mínimos quadrados, na forma matricial, são dadas por:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

O teste de Análise de Variância – ANOVA é utilizado para testar a significância da regressão linear múltipla e possui as seguintes hipóteses:

$$\begin{cases} H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0 \\ H_1: \beta_j \neq 0, \text{ para ao menos um } j, j = 1, 2, \dots, k \end{cases}$$

Este teste nos mostra que, quando rejeitarmos  $H_0$ , temos que, ao menos uma das variáveis regressoras ( $x_1, x_2, \dots, x_k$ ), contribui significativamente para o modelo.

A ANOVA é uma generalização do teste usado no modelo simples. Sob  $H_0$ ,  $SQ_R/\sigma^2$  é uma variável aleatória com distribuição Qui-Quadrado, com  $k$  graus de liberdade,  $SQ_E/\sigma^2$  é uma Qui-Quadrado com  $(n-k)$  graus de liberdade e  $SQ_R$  e  $SQ_E$  são independentes.  $SQ_E$  e  $SQ_R$  são definidos como segue.

A estatística de teste para  $H_0$  é dada por:

$$F_0 = \frac{SQ_R/k}{SQ_E/(n-p)} = \frac{MQ_R}{MQ_E}$$

Onde:

$SQ_R$  é a soma de quadrados devido a regressão;

$SQ_E$  é a soma de quadrados devido ao erro;

$MQ_R$  o quadrado médio de regressão;

$MQ_E$  o quadrado médio do erro.

A Tabela 2 traz uma representação do cálculo da estatística de teste da ANOVA:

Tabela 2: Tabela da ANOVA

<b>Fonte de Variação</b>	<b>Graus de Liberdade</b>	<b>Soma de Quadrados</b>	<b>Quadrados Médios</b>	<b>F<sub>0</sub></b>
<b>Regressão</b>	K	SQ <sub>R</sub>	MQ <sub>R</sub>	MQ <sub>R</sub> /MQ <sub>E</sub>
<b>Erro</b>	n-k	SQ <sub>E</sub>	MQ <sub>E</sub>	
<b>Total</b>	n-1	SQ <sub>T</sub>		

Assim, se p-valor de  $F_0$  for maior que o valor tabelado  $f_{\alpha,k,n-k}$ , rejeita-se  $H_0$ . Ou seja, se  $F_0$  for maior que o quantil  $1-\alpha$  de uma distribuição F com k e n-k graus de liberdade, rejeita-se  $H_0$ . Este procedimento é realizado na análise de variância (Tabela 2).

#### 2.4.2.1 – Coeficiente de Determinação Múltipla

O coeficiente de determinação múltipla ou  $R^2$  e a análise residual são utilizados para medir a adequação do modelo.

O coeficiente de determinação múltipla é dado por:

$$R^2 = \frac{SQ_R}{SQ_T} = 1 - \frac{SQ_E}{SQ_T}$$

Esta medida quantifica a redução de variabilidade de y, com o uso das regressoras ( $x_1, x_2, \dots, x_k$ ).

Pela equação temos que  $0 \leq R^2 \leq 1$ . No entanto, temos que um  $R^2$  grande não significa que o modelo seja necessariamente bom, pois a adição de variáveis no modelo implica no aumento da medida, independentemente da variável ser ou não significativa ao modelo.

A raiz quadrada positiva de  $R^2$  é chamada de coeficiente de correlação múltipla entre a variável resposta y e o conjunto de regressoras ( $x_1, x_2, \dots, x_k$ ).

A inclusão de diversas variáveis, mesmo com pouco poder explicativo sobre a variável independente, tende a aumentar o valor de  $R^2$ , incentivando a inclusão indiscriminada de variáveis nos modelos. Quando usamos o  $R^2$  ajustado reduzimos esta tendência, pois a mesma penaliza a inclusão de regressores pouco explicativos.

Por isso neste trabalho utilizamos o  $R^2$  ajustado nas comparações dos modelos. O coeficiente é calculado da seguinte maneira:

$$\overline{R^2} = 1 - \frac{n - 1}{n - (k + 1)} (1 - R^2)$$

#### 2.4.2.2 – Análise Residual

Os resíduos têm um importante papel no julgamento da adequação do modelo, assim como na regressão linear simples e são definidos da seguinte maneira:

$$e_i = y_i - \hat{y}_i$$

Os resíduos auxiliam nas suposições de que os erros foram distribuídos de forma aproximadamente normal, com variância constante, bem como na determinação da utilidade de termos adicionais no modelo.

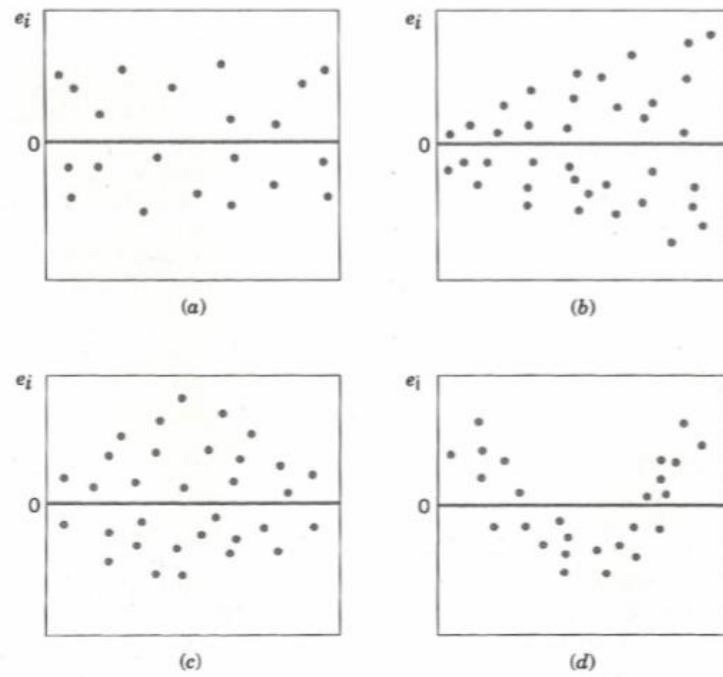
Além disso, podemos analisar o pressuposto de normalidade a partir do gráfico de probabilidade normal dos resíduos ou dos resíduos padronizados. A adoção dos resíduos padronizados é mais adequada, uma vez que as variáveis correspondentes aos pontos do gráfico devem ter a mesma variância. Sendo assim, os resíduos padronizados são dados por:

$$d_i = \frac{e_i}{\sqrt{MQ_E}} = \frac{e_i}{\sqrt{\hat{\sigma}^2}}$$

Quando os erros são normalmente distribuídos, aproximadamente 95% dos resíduos padronizados  $d_i$  se encontram no intervalo (-2,2). Se tivermos observações fora desse intervalo, podem ser um indício de *outliers*.

No Gráfico 2, observamos alguns padrões de comportamento dos resíduos. O padrão apresentado em (a) representa a situação ideal dos resíduos, nas demais ocorrem anomalias. Em (b) temos uma situação conhecida por funil onde a variância é crescente ao longo do tempo. Em (c) temos a situação conhecida por arco duplo onde também temos a heterocedasticidade, por fim em (d) mostra que a relação é não linear.

Gráfico 3: Padrões de comportamento para gráficos de resíduos



Fonte: Montgomery &amp; Runger (2003)

### 3- Resultados

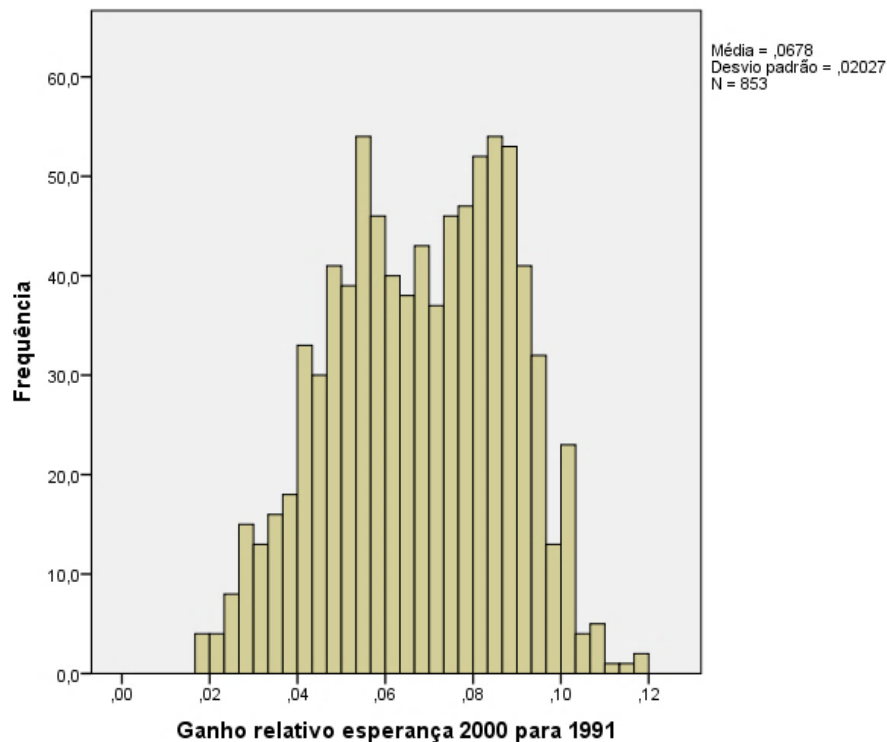
#### 3.1 – Estatísticas Descritivas da Esperança de vida ao nascer

No ano de 2010, a variável de Esperança de Vida ao Nascer apresentou média 74,42 anos e desvio padrão de 1,79 anos, variando de 68,37 a 75,98. No ano 2000, tivemos média 70,16 anos e desvio padrão 2,40 anos, com mínimo de 62,63 e máximo de 75,98. Por fim, o ano 1991 apresentou uma média de 65,73 anos e desvio padrão 2,61 anos, com mínimo de 58,17 e máximo de 71,48 anos esperados de vida.

Quanto à variação relativa da Esperança de Vida ao Nascer, no ano de 2010 comparando-se com 2000, observamos um crescimento médio de 0,061 e desvio padrão 0,023, com mínimo de aproximadamente 0,01 e máximo de 0,1. E comparando o ano 2000 com 1991 tivemos um crescimento médio de 0,068, com desvio padrão de 0,02, variando de aproximadamente 0,02 a 0,12.

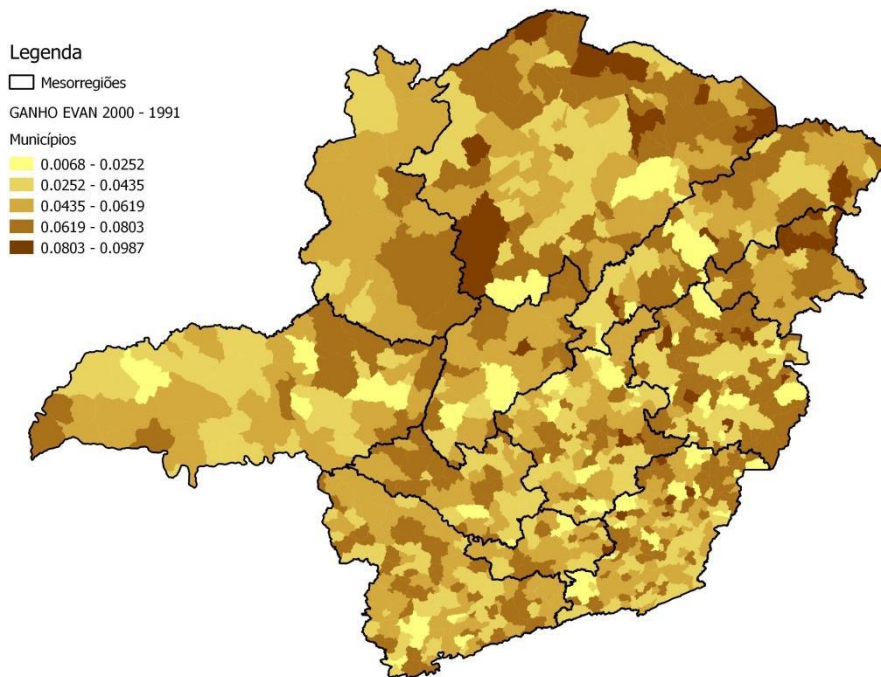
Analisando os dados para o ganho em relação aos anos 2000-1991, observamos o histograma apresentado no Gráfico 4:

Gráfico 4 – Histograma – Ganho Esperança de Vida ao Nascer 2000-1991



O gráfico sugere uma leve assimetria dos dados e bimodalidade. Porém, podemos observar que alguns locais do Estado cresceram mais que outros, como mostra o mapa da distribuição do ganho na esperança de vida ao nascer da Figura 1.

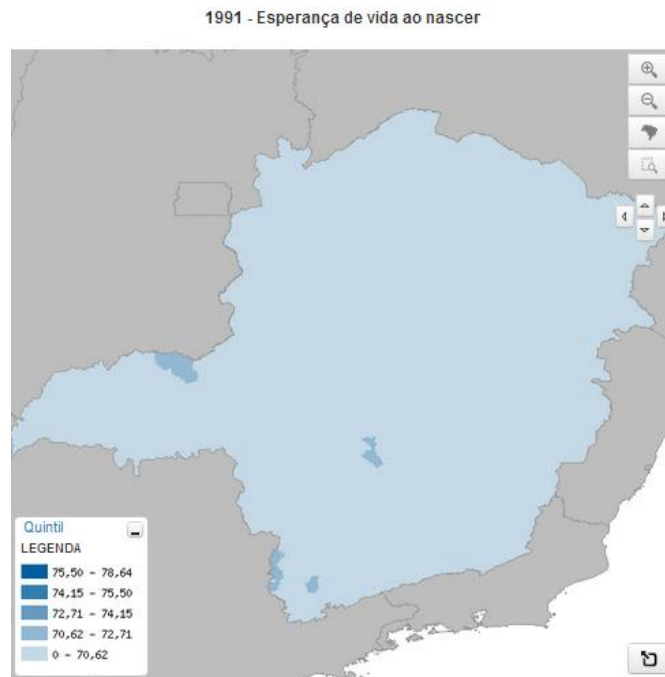
Figura 1 – Ganho Esperança de Vida ao Nascer 2000-1991



Observa-se na Figura 1 que a região Noroeste, Norte, Jequitinhonha, Vale do Mucuri e também a Zona da Mata alcançaram um crescimento maior na esperança de vida ao nascer.

Na Figura 2 temos um mapa com os dados de esperança de vida ao nascer em 1991. Observamos que a esperança de vida ao nascer era menor para todo o Estado de Minas Gerais em relação aos anos 2000 e 2010.

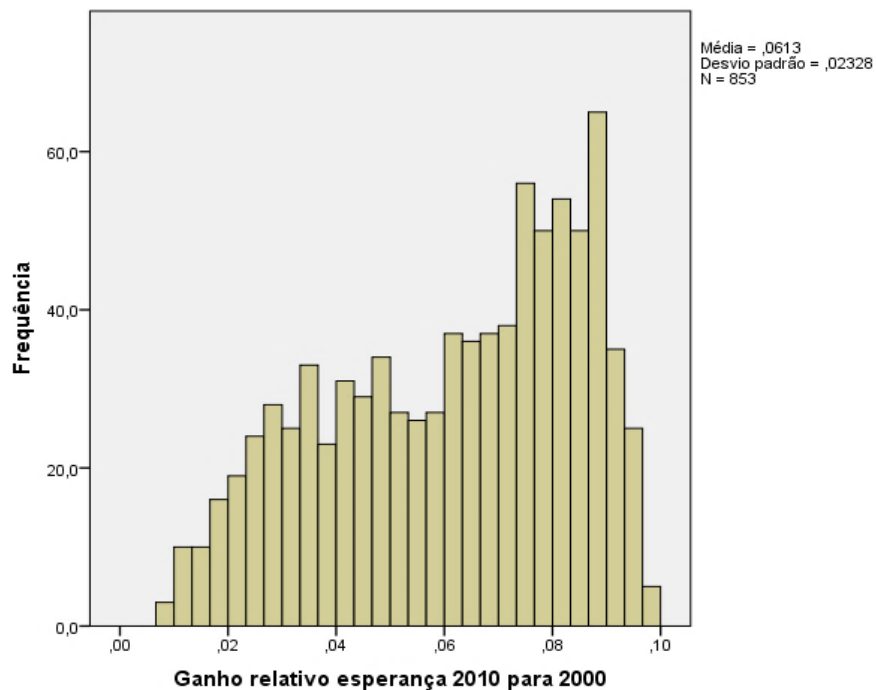
Figura 2 – Esperança de Vida ao Nascer 1991 Minas Gerais



Os mapas de esperança de vida ao nascer utilizados neste estudo possuem a mesma escala do mapa da figura 2, para efeito de comparação.

O Gráfico 5 apresenta o histograma da variável resposta “Ganho na Esperança de Vida ao Nascer 2010-2000”:

Gráfico 5 - Histograma Ganho Esperança de Vida ao Nascer 2010-2000

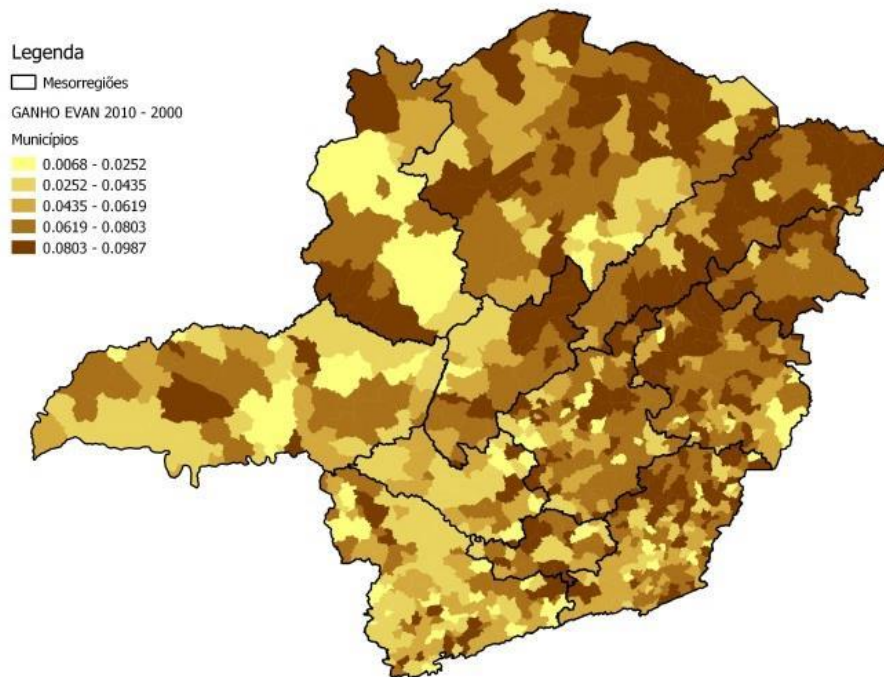




Analisando o Gráfico 5 temos que distribuição desses ganhos no Estado de Minas Gerais indica uma assimetria à direita, além de bimodalidade. Porém, olhando com atenção, podemos notar duas tendências, indicando que uma parte do Estado cresceu mais que a outra, ou seja, temos duas modas.

Na Figura 3 temos a representação da mesma variável de forma espacial, a partir do mapa de Minas Gerais, mostrando as diferenças entre as regiões.

Figura 3 – Ganho Esperança de Vida ao Nascer 2010-2000 MG

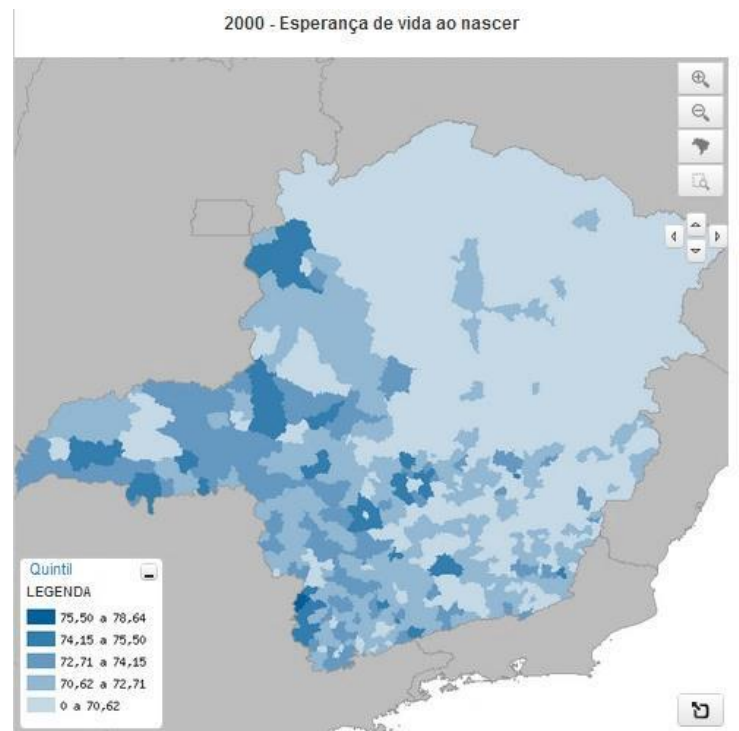


Pelo mapa da Figura 3, observamos claramente uma disparidade no ganho da esperança de vida ao nascer, uma vez que as regiões com a cor mais escura obtiveram um ganho relativo maior. Analogamente, as regiões na cor mais claras obtiveram ganhos relativos menores. Observa-se ainda que as regiões Norte, Noroeste, Jequitinhonha, Vale do Mucuri e Vale do Rio Doce, bem como a região da Zona da Mata apresentaram maior crescimento relativo da esperança de vida ao nascer, o que explica o histograma bimodal do Gráfico 4.

A Figura 4 abaixo mostra a distribuição da esperança de vida ao nascer no Estado de Minas Gerais para o ano 2000. Comparando os mapas, notamos que nas regiões mais escuras (com maior ganho) da Figura 3 são as que obtêm os menores índices da esperança de vida ao nascer (regiões mais claras) da Figura 4. Ou seja,

nas regiões em que a esperança de vida ao nascer eram mais baixas o ganho foi relativamente maior, quando comparadas com aquelas em que a esperança de vida ao nascer era maior.

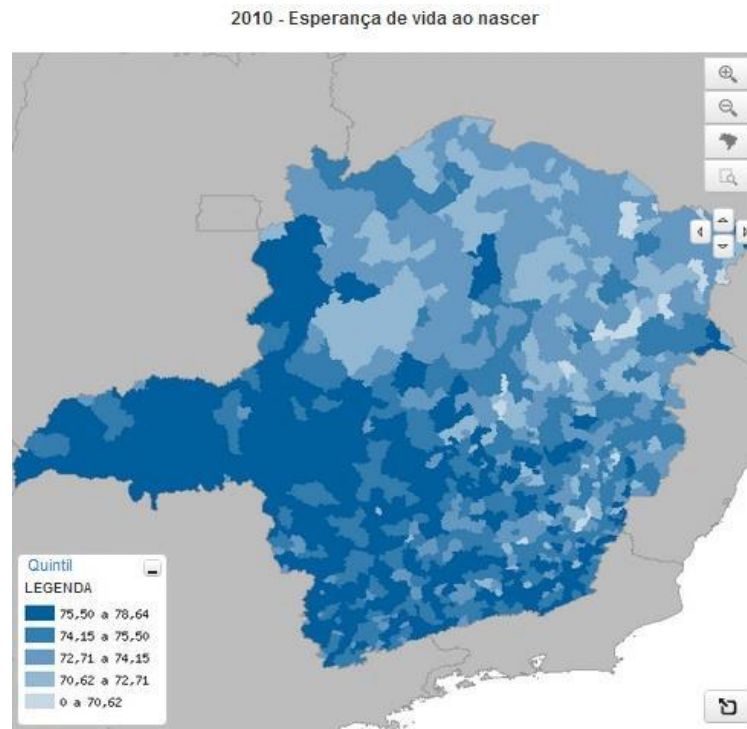
Figura 4 – Esperança de Vida ao Nascer 2000 – Minas Gerais



Ainda pela Figura 4, observamos no ano de 2000 uma disparidade na esperança de vida ao nascer das regiões Triângulo Mineiro, Central, Metropolitana, Oeste, Campo das Vertentes e Sul, em relação às regiões ao norte do Estado, citadas na Figura 3.

Na Figura 5 temos o mapa com o tema esperança de vida ao nascer nos municípios de Minas Gerais, para o ano 2010.

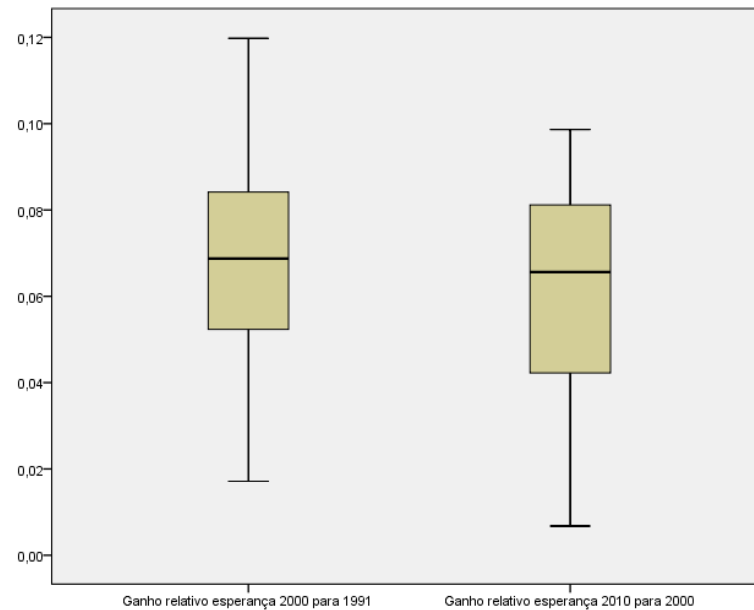
Figura 5 – Esperança de Vida ao Nascer 2010 Minas Gerais



Realizando a mesma análise para os dados do ano 2010, temos novamente uma disparidade na Figura 5, porém com menor intensidade, podendo indicar uma tendência de uniformidade na Esperança de Vida ao nascer no Estado de Minas Gerais, ao longo do tempo. Isto pode explicar o crescimento maior das regiões mais ao norte do Estado, quando comparada às demais.

No Gráfico 6 temos os ganhos na esperança de vida ao nascer dos períodos estudados representados a partir dos diagramas de caixa. Analisando o gráfico, observamos que os dois períodos tiveram mediana semelhantes, porém o ganho da esperança de vida no período 1991-2000 foi maior comparado ao período 2000-2010. Além disso, comparando as assimetrias dos períodos apresentadas nos histogramas acima, temos que o período 1991-2000 foi menos assimétrico.

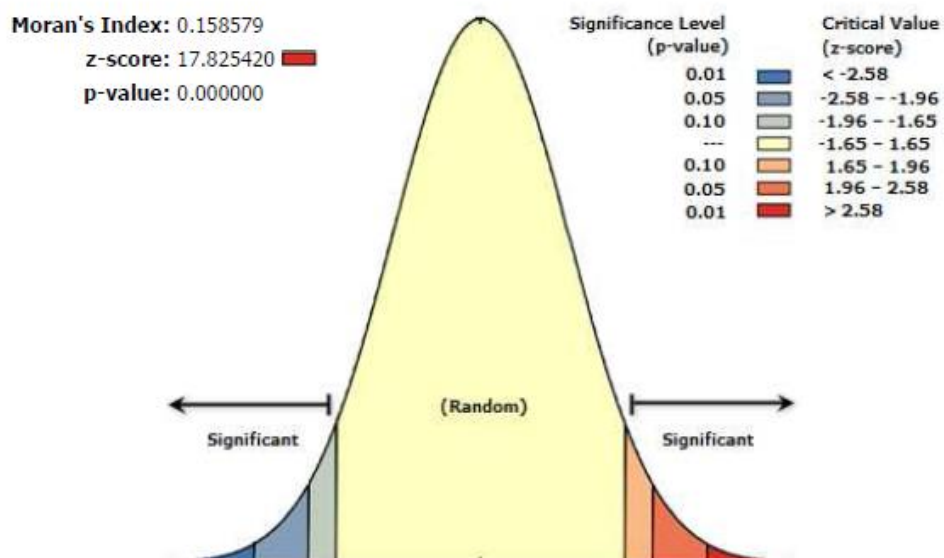
Gráfico 6 – Box-plot Ganho Esperança de Vida Ao Nascer



### 3.2 - Correlações

Uma vez que as observações devem ser independentes entre si, utilizamos o Índice de Moran para verificar a auto correlação espacial entre os ganhos na esperança de vida ao nascer nos dois períodos. Em caso de dependência espacial, buscaremos uma forma de corrigi-la, para ajustar do modelo de regressão linear. A Figura 6 trás a representação do índice para o período de 2010-2000.

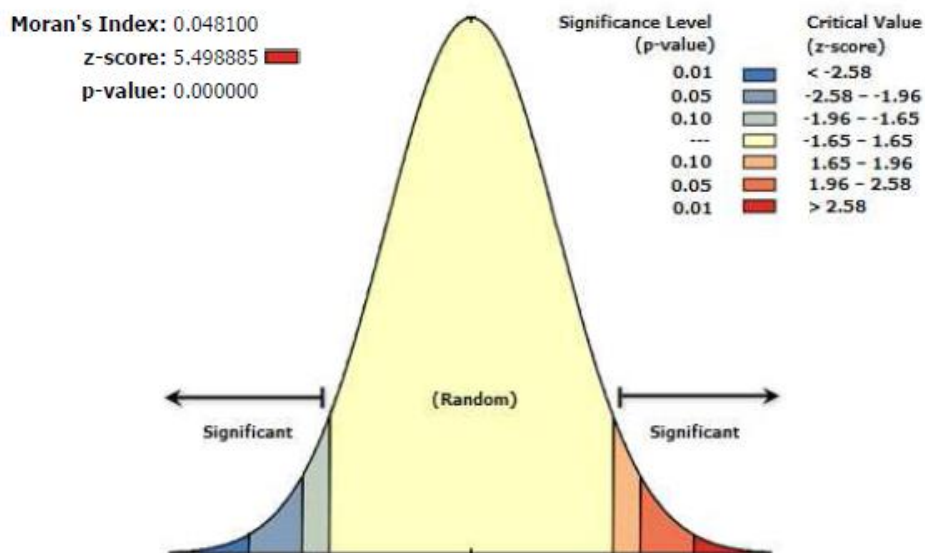
Figura 6 – Índice de Moran 2010 - 2000



Com um índice  $I = 0,159$ , concluímos que existe uma fraca correlação espacial, porém estatisticamente significativa. Pelo p-valor, que é aproximadamente zero e pelo z-score igual a 17,825 maior que o valor crítico 2,58, a um nível de 5%, rejeitamos a hipótese nula de independência espacial dos dados. Isto pode ser devido ao fato de termos muitas observações da variável, uma vez que estamos estudando um total de 853 municípios.

No período de 2000-1991 e o Índice de Moran calculado está representado na Figura 7.

Figura 7 – Índice de Moran Ganho Esperança de Vida ao Nascer 2000-1991



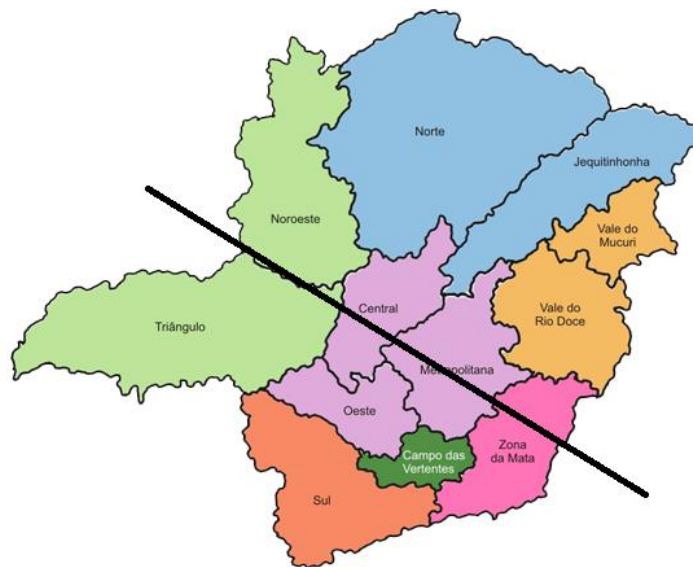
De forma análoga ao que ocorreu no período de 2010-2000, neste período observa-se uma fraca auto correlação espacial a partir do índice calculado  $I = 0,048100$ . Porém, esta auto correlação foi altamente significativa e a hipótese nula de independência espacial foi rejeitada, com um p-valor próximo de 0 e z-score = 5,49888 maior que o valor crítico 2,58. Da mesma maneira que no período de 2010-2000, os ganhos são mais fortes em algumas regiões e mais fracos em outras. O aumento do índice do período de 1991-2000 para 2000-2010, pode indicar maior disparidade entre as regiões do Estado.

Como a correlação espacial foi estatisticamente significativa em ambos os períodos, foi criada uma variável que leva em consideração as regiões onde o ganho foi maior e as regiões onde o ganho foi menor, a partir da análise dos mapas de

ganho da esperança de vida das Figuras 1 e 3, mostrando que a distribuição dos ganhos de esperança de vida ao nascer nos dois períodos foram semelhantes. Tal variável foi criada com o intuito de melhorias no ajuste dos modelos de regressão linear, visto que a mesma detém efeitos da correlação espacial.

Com a análise dos mapas chegamos a uma subdivisão do Estado em duas regiões conforme o mapa da Figura 8.

Figura 8 – Mapa das mesorregiões de Minas Gerais



As regiões situadas acima da linha traçada são as que obtiveram maior ganho e na variável dicotômica receberam o valor 1, e as regiões abaixo da linha são as que obtiveram menor ganho e receberam valor 0. Uma divisão semelhante foi observada em Índices de Desenvolvimento Humano no trabalho de Romero (2006).

Utilizamos a correlação de Pearson para determinar quais variáveis seriam pré-selecionadas em um modelo de regressão linear múltipla, para determinar os fatores socioeconômicos/demográficos que estariam associados ao ganho na esperança de vida ao nascer.

No Anexo 1 estão as matrizes de correlação dos dados de ganhos das variáveis de esperança de vida ao nascer e as variáveis socioeconômicas/demográficas em relação aos períodos 2010-2000 e 2000-1991.

Ao fazer a análise das correlações lineares do período 2010-2000, observamos que as variáveis de ganho de renda per capita, porcentagem da

população residente em domicílio com água encanada, energia elétrica, coleta de lixo, porcentagem de pessoas com mais de 18 anos com ensino médio completo, expectativa de anos de estudo, taxa de fecundidade total, obtiveram correlações significativas. Porém, observamos que o ganho na taxa de mortalidade infantil possui a correlação mais forte com a variável de esperança de vida ao nascer, com  $r = -0,98$  o que demonstra praticamente uma perfeita correlação negativa entre as mesmas, sugerindo que quanto mais decresce a taxa de mortalidade infantil, mais ocorre um incremento da esperança de vida ao nascer neste período.

De maneira geral, durante o período 2000-1991, observa-se, dentre as variáveis disponíveis para o estudo, que os ganhos na taxa de fecundidade total, expectativa de anos de estudo, população total, porcentagem da população maior de 18 anos com ensino médio completo e porcentagem da população que reside em domicílio com água encanada e banheiro, obtiveram taxas de correlação de Pearson significativas, entretanto, essas correlações são fracas, com valor em módulo inferior ou igual a 0,139. No entanto, o ganho no índice de mortalidade infantil obteve correlação significativa de  $r = -0,824$ , representando uma forte correlação com a variável de esperança de vida ao nascer. Da mesma forma que no período 2010-2000, ela é negativa e podemos interpretá-la que quanto maior o decréscimo da mortalidade infantil, maior o incremento da esperança de vida ao nascer no período.

Variáveis de desigualdade foram observadas também quanto à correlação com o ganho na esperança de vida ao nascer, em ambos os períodos estudados. Observa-se que o ganho na desigualdade relativa de renda do quinto mais rico para o quinto mais pobre e ganho no Índice de Gini, não obtiveram correlação significativa com o ganho na esperança de vida ao nascer nos municípios do Estado de Minas Gerais.

### **3.3 – Modelos de Regressão**

A partir das correlações calculadas, foi feita uma pré-seleção de variáveis para compor um modelo de regressão linear proposto para cada período considerado e as variáveis com correlação significativa ao ganho na esperança de vida ao nascer no período foram candidatas a regressoras nos modelos. As tabelas com os modelos considerados, bem como, a significância das variáveis se encontram no Anexo 2. Foram consideradas nos modelos variáveis com  $p$ -valor  $\leq 0,10$ , além de variáveis com termo quadrático ou com interações.

A esperança de vida ao nascer depende da taxa de mortalidade infantil, porém de forma indireta e no presente trabalho estamos utilizando seus ganhos.

### 3.3.1 – Modelos 1991-2000

Para o período entre os anos 1991 e 2000, as variáveis com maior correlação com o ganho de esperança de vida ao nascer foram selecionadas como candidatas para compor um modelo de regressão linear proposto. As variáveis são os seguintes ganhos: taxa de fecundidade total, expectativa de anos de estudo, população total, porcentagem de maiores de 18 anos com ensino médio, porcentagem da população residente em domicílio com água encanada e banheiro e mortalidade infantil. Foi incluído também, para efeito de verificação de auto-correlação espacial, a variável região.

Após estudar a correlação das variáveis candidatas a entrarem no modelo de regressão a ser proposto para o período, excluímos os fatores que estavam correlacionados entre si, para evitar problemas de multicolinearidade no modelo. Permaneceram então, as variáveis de ganho na mortalidade infantil e ganho na porcentagem da população em domicílio com água encanada e banheiro, e foi incluída no modelo a variável de região.

Pelo método *backward* (Charnet et al, 2008), obtivemos o seguinte modelo, onde todas as variáveis propostas foram significativas:

Modelo 1

$$Y = 0,23 - 0,864x_1 + 0,178x_2 + 0,142x_3$$

Onde:

$Y$  é o ganho na esperança de vida ao nascer;

$x_1$  é o ganho na mortalidade infantil;

$x_2$  é o ganho na porcentagem da população em domicílio com água encanada e banheiro;

$x_3$  é a variável de região.



Tabela 3 – ANOVA Modelo 1

**ANOVA<sup>a</sup>**

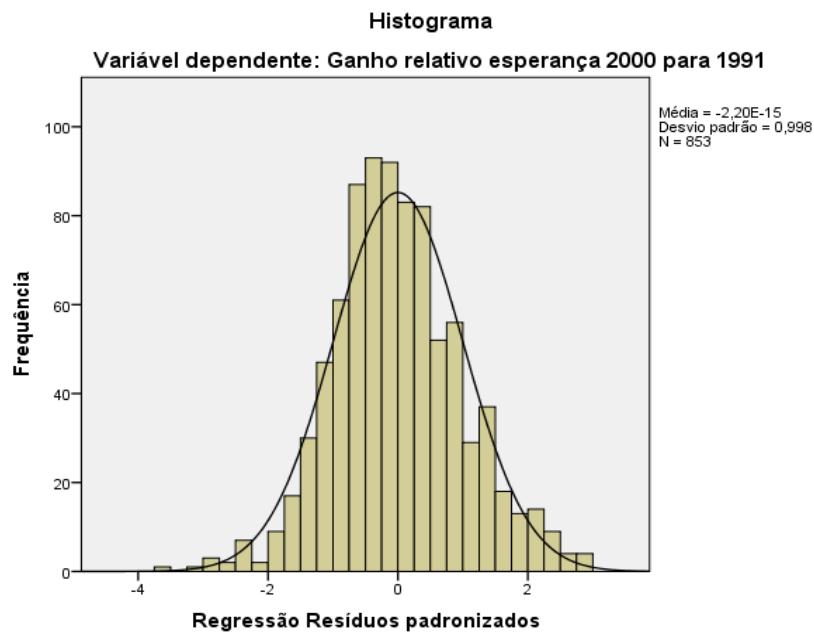
Modelo		Soma dos Quadrados	df	Quadrado Médio	F	Sig.
1	Regressão	,260	3	,087	823,870	,000 <sup>b</sup>
	Resíduos	,089	849	,000		
	Total	,350	852			

a. Variável dependente: Ganho relativo esperança 2000 para 1991

b. Preditores: (Constante), Região, Ganho relativo Mortalidade infantil 2000 para 1991, Ganho relativo % pop em domicílio com água enc. e banheiro 2000 para 1991

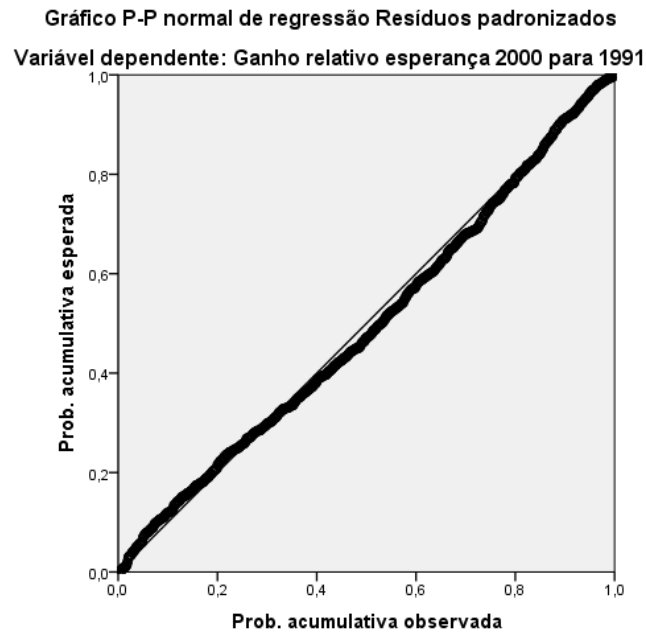
Pela ANOVA do modelo 1 (Tabela 3), observa-se que este é significativo a um nível de 5%, com p-valor menor que 0,001. Observou-se ainda  $R^2$  ajustado foi de 0,743, o que pode nos indicar um bom ajuste do modelo.

Gráfico 7 – Histograma Resíduos Modelo 1



Pelo Gráfico 7, percebemos que os resíduos seguem uma distribuição com média aproximadamente 0 e desvio padrão aproximadamente 1.

Gráfico 8 – Probabilidade Normal Modelo 1



Pelo Gráfico 8 temos que o modelo segue a linha de probabilidade acumulada da normal. Porém a Tabela 4 nos mostra que a um nível de 5% de significância, rejeita-se a hipótese de normalidade dos resíduos.

Tabela 4 – Normalidade Residual Modelo 1

Teste de Normalidade						
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estatística	df	Sig.	Estatística	df	Sig.
Standardized Residual	,035	853	,016	,993	853	,001

a. Lilliefors Significance Correction

Gráfico 9 – Dispersão Resíduos Modelo 1

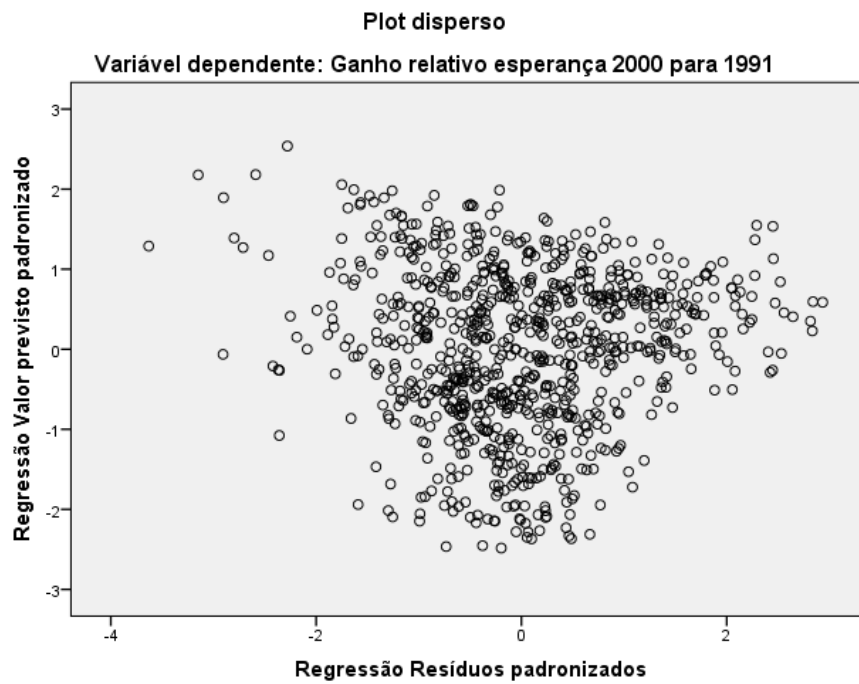
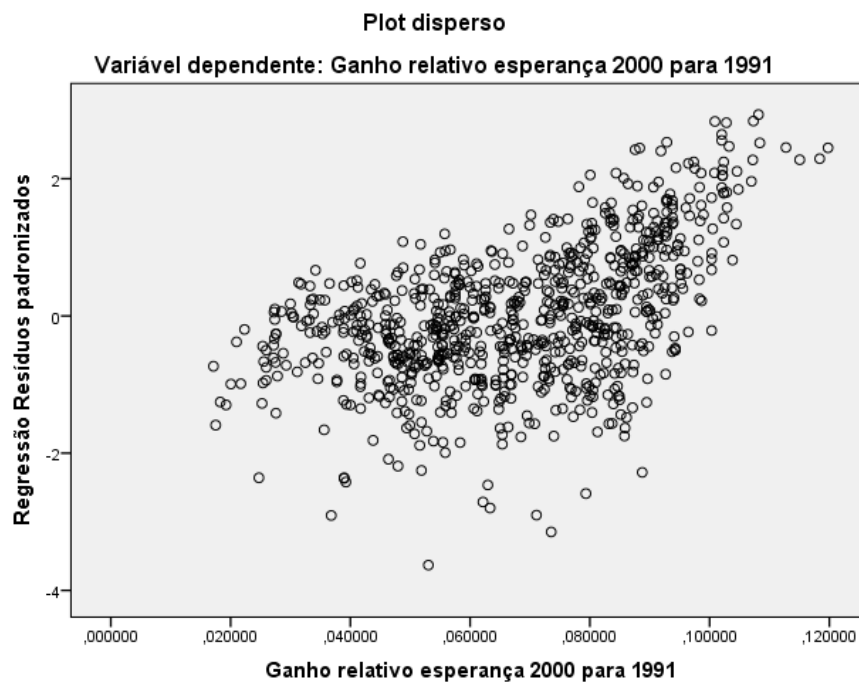


Gráfico – 10 Dispersão Observados x Resíduos Modelo 1



Pela dispersão dos resíduos apresentada nos Gráficos 9 e 10, temos que os mesmos estão aleatoriamente distribuídos, porém surge o problema de

heterocedasticidade, dado que o esperado era que os resíduos tivessem a mesma variabilidade, distribuídos no gráfico de forma aleatória.

Uma primeira tentativa de contornar este problema foi gerar um modelo que contém interações das variáveis região e as demais regressoras do Modelo 1, como observamos no modelo abaixo:

#### Modelo 2

$$Y = 0,023 + 0,084x_1 - 0,833x_2 + 0,324x_3 - 0,139x_1x_2 - 0,200x_1x_3$$

Onde:

$Y$  é o ganho na esperança de vida ao nascer;

$x_1$  é a variável de região;

$x_2$  é o ganho na mortalidade infantil;

$x_3$  é o ganho na porcentagem da população em domicílio com água encanada e banheiro;

Tabela 5 – ANOVA Modelo 2

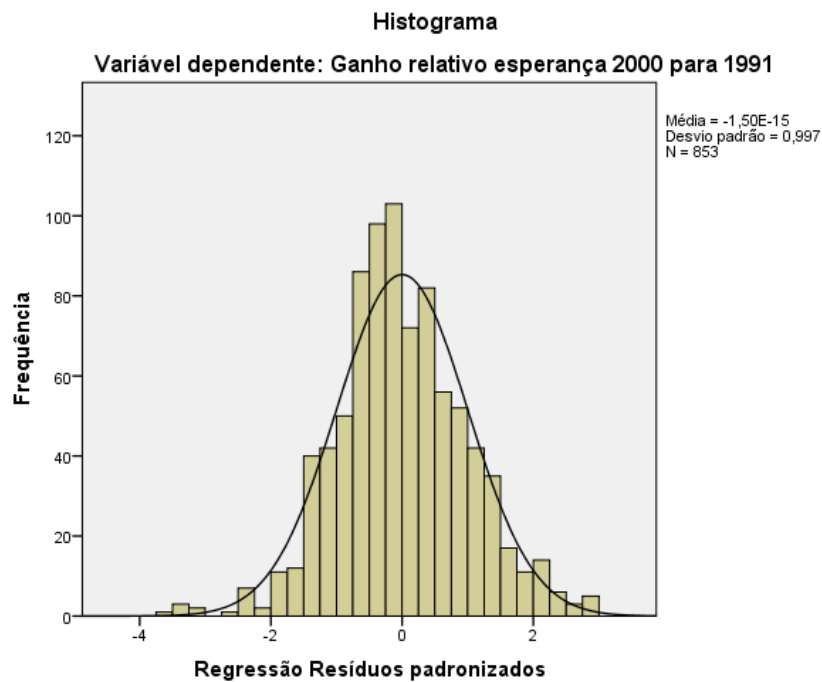
ANOVA <sup>a</sup>						
Modelo		Soma dos Quadrados	df	Quadrado Médio	F	Sig.
1	Regressão	,265	5	,053	527,369	,000 <sup>b</sup>
	Resíduos	,085	847	,000		
	Total	,350	852			

a. Variável dependente: Ganho relativo esperança 2000 para 1991

b. Preditores: (Constante), Reg\_Aguat91, Ganho relativo Mortalidade infantil 2000 para 1991, Reg\_Mort91, Ganho relativo % pop em domicilio com agua enc. e banheiro 2000 para 1991, Região

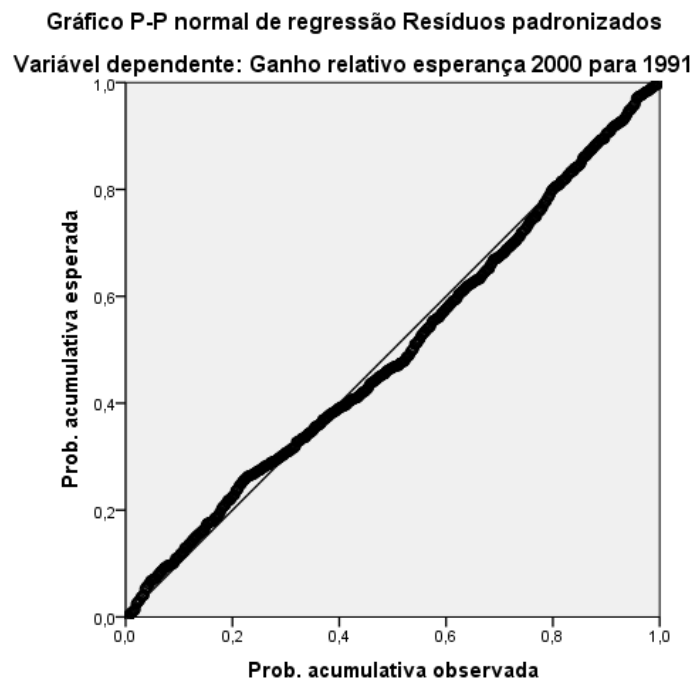
Pela ANOVA da Tabela 5, observamos um p-valor abaixo de 0,001, assim a regressão foi significativa a um nível de 5%. Temos ainda um  $R^2$  ajustado de 0,755, o que pode indicar uma boa explicação da variável resposta pelo modelo proposto.

Gráfico 11 – Histograma Resíduos Modelo 2



Pelo histograma do Gráfico 11 observamos que os resíduos seguem uma distribuição com média e desvio padrão aproximadamente 0 e 1, respectivamente.

Gráfico 12 – Probabilidade Normal Modelo 2



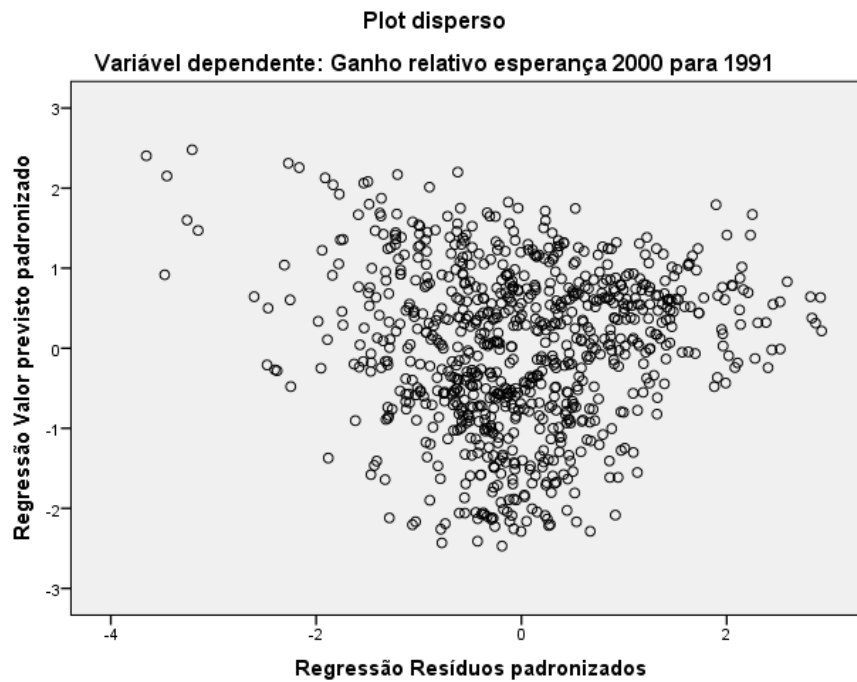
No Gráfico 12, observa-se que os resíduos desviam um pouco da linha normal, podendo nos mostrar uma não normalidade dos mesmos, sendo confirmado pelos testes de normalidade da Tabela 6.

Tabela 6 – Normalidade Residual Modelo 2

Tests of Normality						
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estatística	df	Sig.	Estatística	df	Sig.
Standardized Residual	,044	853	,000	,992	853	,000

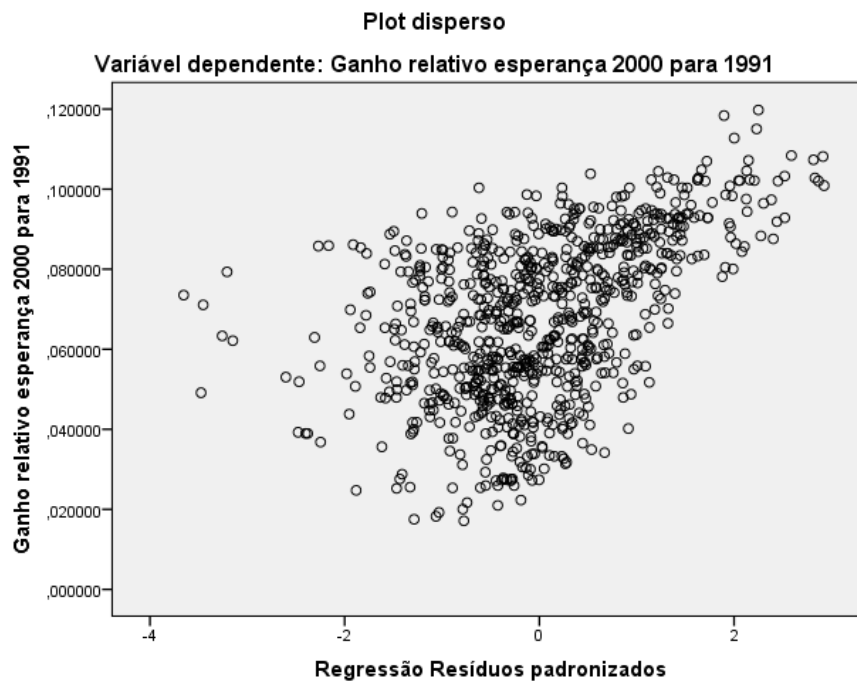
a. Lilliefors Significance Correction

Gráfico 13 – Dispersão Resíduos Modelo 2



O Gráfico 13 mostra que neste caso a heterocedasticidade dos erros ainda persiste. O Gráfico 14 nos mostra independência dos resíduos.

Gráfico 14 - Dispersão Observados x Resíduos Modelo 2



Outra tentativa para contornar o problema foi realizada com a adição do termo quadrático da variável de ganho na mortalidade infantil, gerando o Modelo 3, dado por:

Modelo 3

$$Y = 0,014 + 0,146x_1 - 1,307x_2 + 0,325x_3 - 0,072x_1x_2 - 0,213x_1x_3 - 0,474x_1^2$$

Onde:

$Y$  é o ganho na esperança de vida ao nascer;

$x_1$  é a variável de região;

$x_2$  é o ganho na mortalidade infantil;

$x_3$  é o ganho na porcentagem da população em domicílio com água encanada e banheiro;

Tabela 7 – ANOVA Modelo 3

**ANOVA<sup>a</sup>**

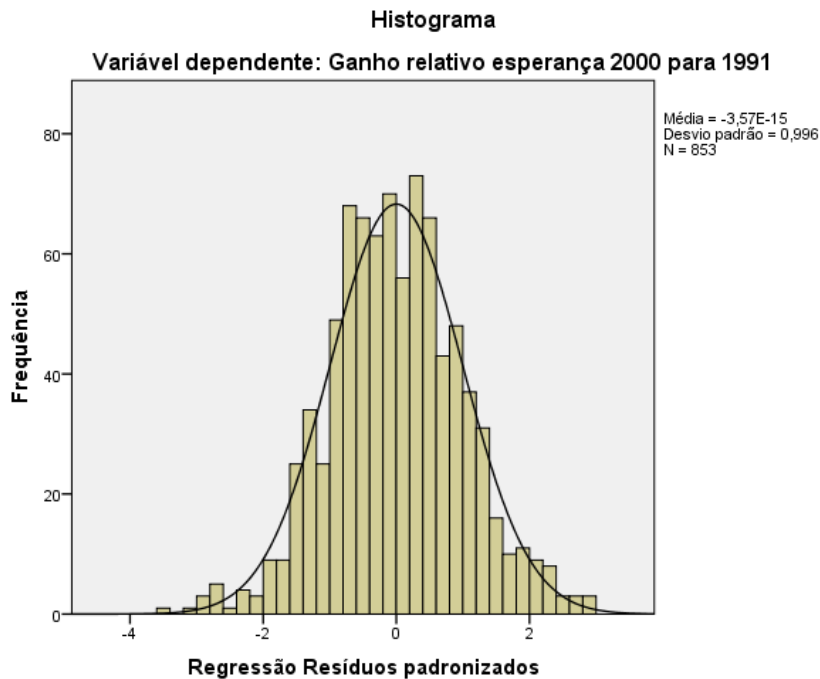
Modelo		Soma dos Quadrados	df	Quadrado Médio	F	Sig.
1	Regressão	,270	6	,045	477,927	,000 <sup>b</sup>
	Resíduos	,080	846	,000		
	Total	,350	852			

a. Variável dependente: Ganho relativo esperança 2000 para 1991

b. Preditores: (Constante), Mort\_Quad\_91, Região, Ganho relativo % pop em domicílio com água enc. e banheiro 2000 para 1991, Reg\_Aguat91, Reg\_Mort91, Ganho relativo Mortalidade infantil 2000 para 1991

Com a Tabela 7, podemos observar que o p-valor é menor que 0,001, portanto a regressão é estatisticamente significativa. Observou-se também  $R^2$  ajustado de 0,771 o que pode nos mostrar que o modelo explica bem a variável resposta.

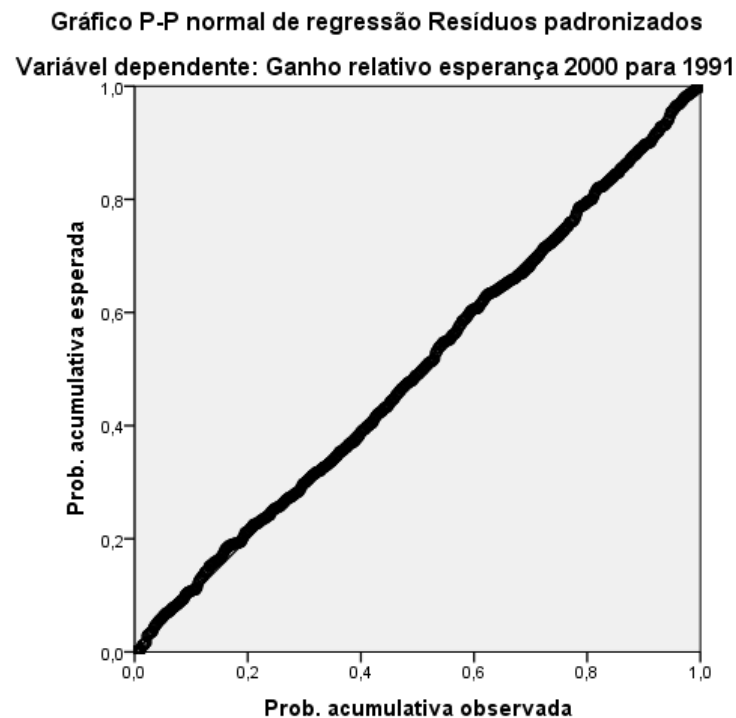
Gráfico 15 – Histograma Resíduos Modelo 3



Pelo histograma apresentado no Gráfico 15, percebe-se que os resíduos seguem uma distribuição com média e desvio padrão aproximadamente 0 e 1, respectivamente.



Gráfico 16 – Probabilidade Normal Modelo 3



Pelo Gráfico 16 da distribuição acumulada, observa-se que os resíduos não desviam da linha da normal, o que indica uma aproximação de normalidade dos dados. O que é confirmado pelos testes de normalidade com a não rejeição da hipótese da mesma na Tabela 8.

Tabela 8 – Normalidade Residual

Tests of Normality						
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estatística	df	Sig.	Estatística	df	Sig.
Standardized Residual	,021	853	,200*	,997	853	,057

\*. Este é um limite inferior da significância verdadeira.

a. Lilliefors Significance Correction

Gráfico 17 – Dispersão Resíduos Modelo 3

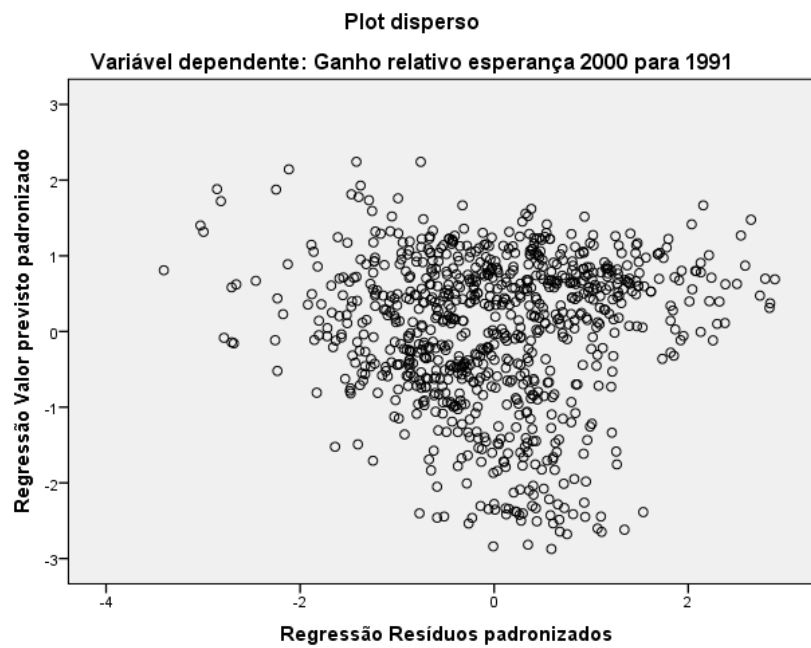
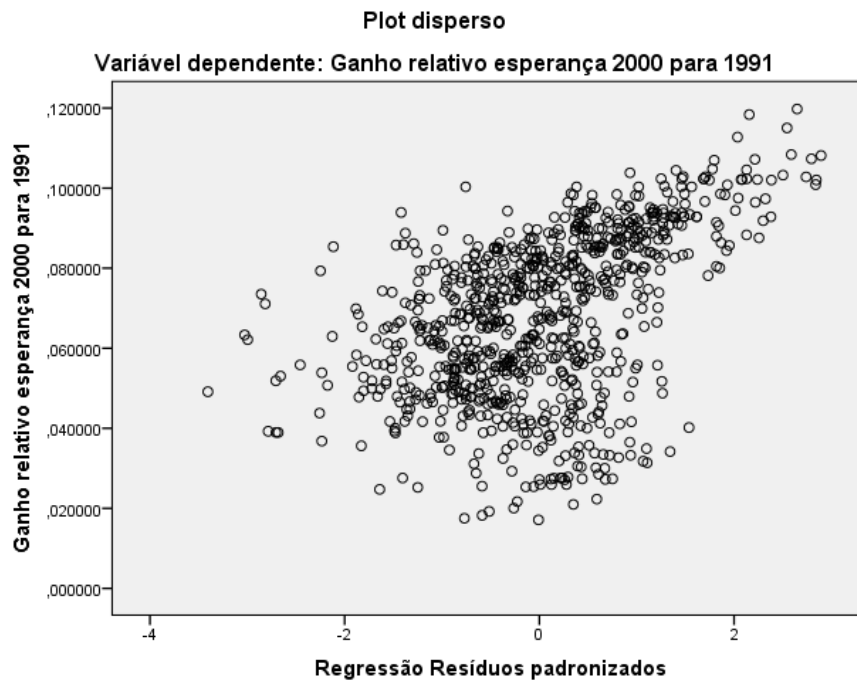


Gráfico 18 - Dispersão Observados x Resíduos Modelo 3



Os gráficos 17 e 18 nos mostram que os resíduos são independentes, porém a heterocedasticidade dos mesmos persiste.

Foram também testadas alternativas como transformação exponencial, logarítmica, raiz, quadrática. Porém, as mesmas não surtiram efeito necessário para resolver as dificuldades encontradas e por isso, foram ignoradas na nossa análise.

O modelo que obteve o melhor ajuste com maior  $R^2$  ajustado foi o Modelo 3, temos neste interações de região com as variáveis estudadas e a forma quadrática da mortalidade infantil, o que nos mostra que a não linearidade pode ser observada neste caso.

### 3.3.2- Modelos 2000-2010

Para o período dos anos 2000-2010, as variáveis candidatas a regressoras para o modelo do ganho de esperança de vida ao nascer foram: ganho de mortalidade infantil, ganho de renda per capita, ganho de porcentagem da população em domicílio com água encanada, ganho de porcentagem da população em domicílio com energia elétrica, ganho de porcentagem da população com coleta de lixo, ganho de porcentagem da população maior de 18 anos com ensino médio completo, ganho de expectativa de anos de estudo e ganho de taxa de fecundidade total. Foi inclusa também, para efeito de verificação da correlação espacial, a variável região descrita no item anterior.

No Anexo 3 se encontra o modelo proposto a partir do método *backward* (Charnet et al, 2008), em função das variáveis disponíveis descritas acima, bem como suas respectivas ANOVA's. O modelo final é dado por:

Modelo 4

$$Y = -0,27 - 0,976x_1 + 0,031x_2 - 0,014x_3$$

Onde:

$Y$  é o ganho na esperança de vida ao nascer;

$x_1$  é o ganho na mortalidade infantil;

$x_2$  é o ganho na porcentagem de pessoas em domicílio com água encanada;

$x_3$  é o ganho na expectativa de anos de estudo.

Tabela 9 – ANOVA Modelo 4

**ANOVA<sup>a</sup>**

Modelo		Soma dos Quadrados	df	Quadrado Médio	F	Sig.
1	Regressão	,445	3	,148	7559,426	,000 <sup>b</sup>
	Resíduos	,017	849	,000		
	Total	,462	852			

a. Variável dependente: Ganho relativo esperança 2010 para 2000

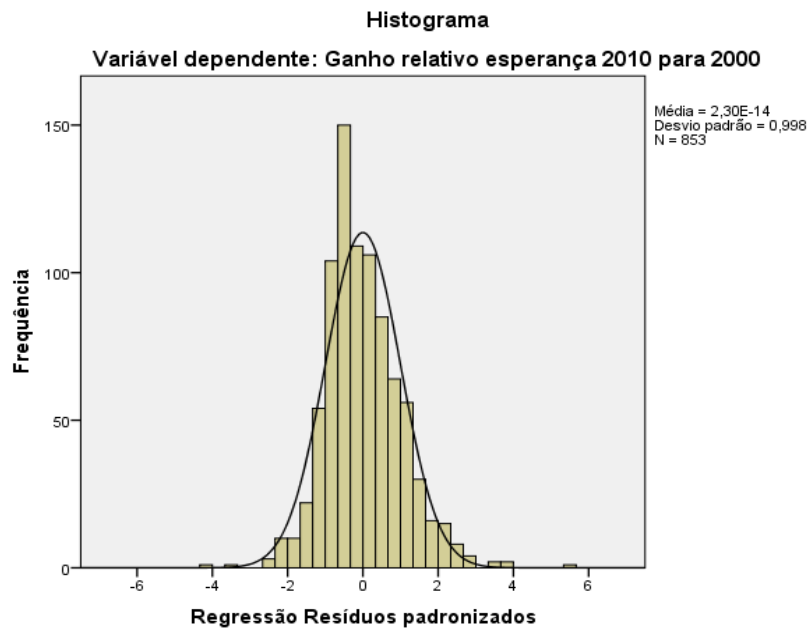
b. Preditores: (Constante), Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo expectativa de anos de estudo 2010 para 2000

Pela Tabela 9, observamos que o p-valor da ANOVA foi menor que 0,001. Portanto, a regressão foi significativa a um nível de 5%. O modelo apresentou um R<sup>2</sup> ajustado 0,964, indicando uma boa adequação do mesmo.

Observamos que as variáveis de ganho de renda per capita, porcentagem da população em domicílio com energia elétrica, coleta de lixo, porcentagem da população maior de 18 anos com ensino médio completo e taxa de fecundidade total, foram removidas do modelo por não serem significantes na relação com o ganho na esperança de vida ao nascer. Também foi excluída a variável região, o que mostra que neste período, temos a possibilidade de independência espacial dos dados.

Analisando o Gráfico 19 logo abaixo, notamos que os resíduos seguem distribuição com média 0 e desvio padrão 1. Porém, não podemos garantir a normalidade dos mesmos.

Gráfico 19 – Histograma dos resíduos Modelo 4



Pelo Gráfico 20 de probabilidade normal observamos que os resíduos estão fora da linha da probabilidade normal, primeiro problema encontrado no modelo. A Tabela 9 nos mostra a um nível de 5% a rejeição de hipótese de normalidade

Gráfico 20 – Probabilidade Normal Modelo 4

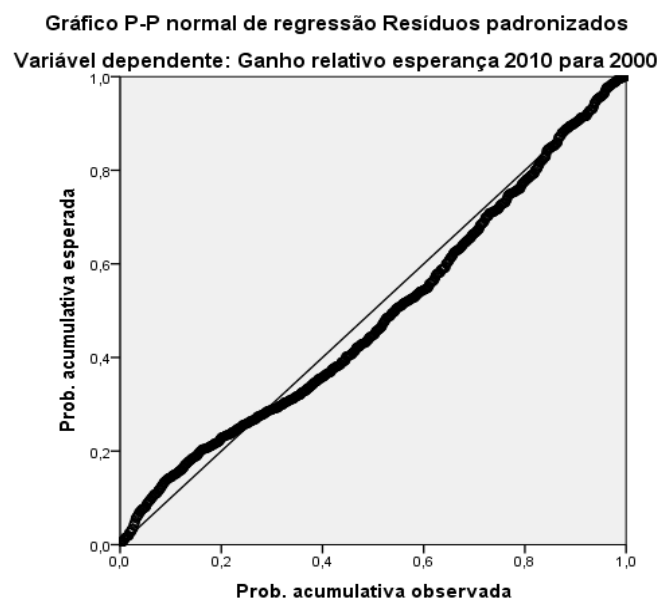


Tabela 9 – Normalidade Residual Modelo 4

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estatística	df	Sig.	Estatística	df	Sig.
Standardized Residual	,060	853	,000	,973	853	,000

a. Lilliefors Significance Correction

Os dados de resíduos não estão aleatoriamente distribuídos em torno do 0 no gráfico de resíduos por observações (Gráfico 22), indicando independência dos resíduos do modelo. Porém, encontramos o problema da heterocedasticidade e não linearidade dos resíduos, devido à distribuição dos pontos no Gráfico 21. O esperado seria uma distribuição aleatória, como uma nuvem de pontos, com variabilidade semelhante. Além disso, o gráfico sugere que temos uma relação não linear entre o ganho na esperança de vida ao nascer e os fatores estudados. A heterocedasticidade dos resíduos é visto como um problema, uma vez que não respeita um dos pressupostos da regressão linear, como vimos na metodologia. Também nos mostra uma relação não linear entre as variáveis.

Gráfico 21 – Dispersão Resíduos Modelo 4

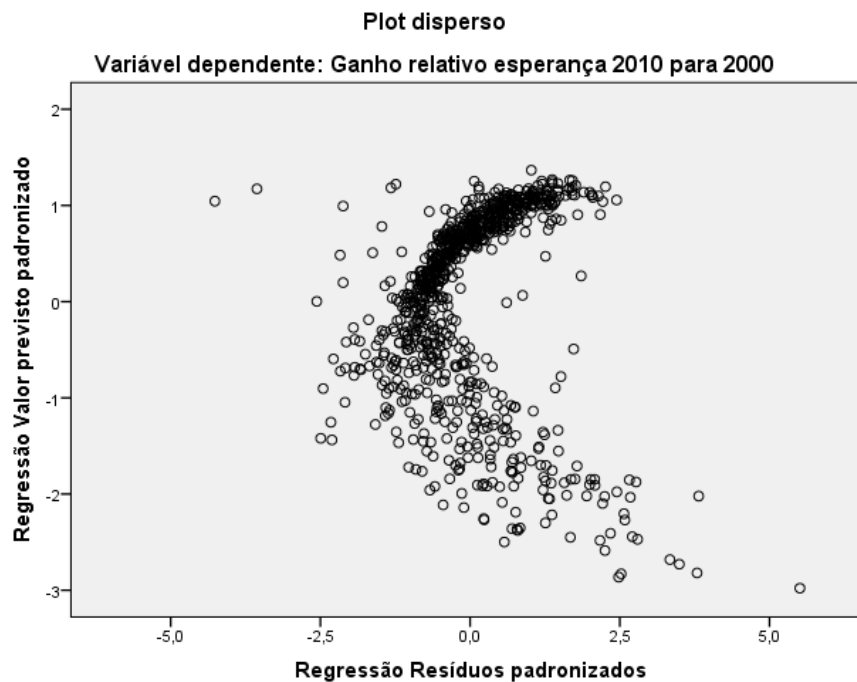
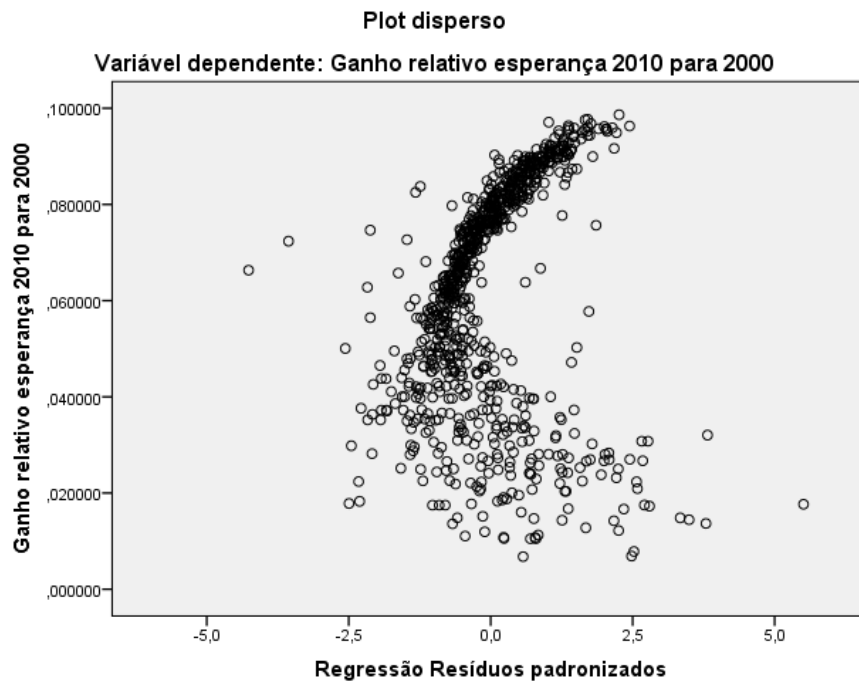


Gráfico 22 - Dispersão Observados x Resíduos Modelo 4



Se excluirmos a variável de ganho na porcentagem de pessoas em domicílio com água encanada, temos outro modelo com a variável de ganho na porcentagem de pessoas em domicílios com energia elétrica que foi significativa, fato que pode ser explicado pela correlação entre essas as duas. O modelo de regressão foi dado por:

Modelo 5

$$Y = -0,27 - 0,979x_1 + 0,021x_2 - 0,011x_3$$

Onde:

$Y$  é o ganho na esperança de vida ao nascer;

$x_1$  o ganho na mortalidade infantil;

$x_2$  o ganho na porcentagem de pessoas em domicílio com energia elétrica;

$x_3$  o ganho na expectativa de anos de estudo.

Tabela 10 – ANOVA Modelo 5

**ANOVA<sup>a</sup>**

Modelo		Soma dos Quadrados	df	Quadrado Médio	F	Sig.
1	Regressão	,445	3	,148	7454,877	,000 <sup>b</sup>
	Resíduos	,017	849	,000		
	Total	,462	852			

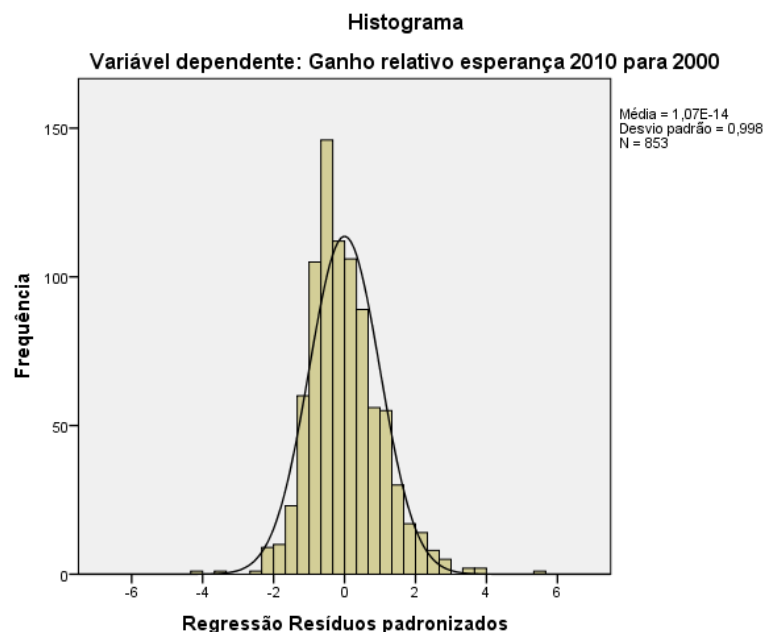
a. Variável dependente: Ganho relativo esperança 2010 para 2000

b. Preditores: (Constante), Ganho relativo % da pop com energia elétrica 2010 para 2000, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo expectativa de anos de estudo 2010 para 2000

A um nível de 5%, o modelo foi significativo, com p-valor da ANOVA menor que 0,001, conforme a Tabela 10. Observa-se ainda um  $R^2$  ajustado de 0,963 neste modelo, indicando uma boa adequação do mesmo.

Pelo Gráfico 23 podemos observar que assim como no primeiro modelo os resíduos seguem aproximadamente uma distribuição com média 0 e desvio padrão 1, porém não podemos concluir sobre sua normalidade:

Gráfico 23 – Histograma Resíduos Modelo 2



No gráfico de probabilidade normal (Gráfico 24) vemos que os resíduos estão desviando da linha da probabilidade normal, o que pode indicar a ausência da normalidade.



Gráfico 24 – Probabilidade Normal Modelo 5

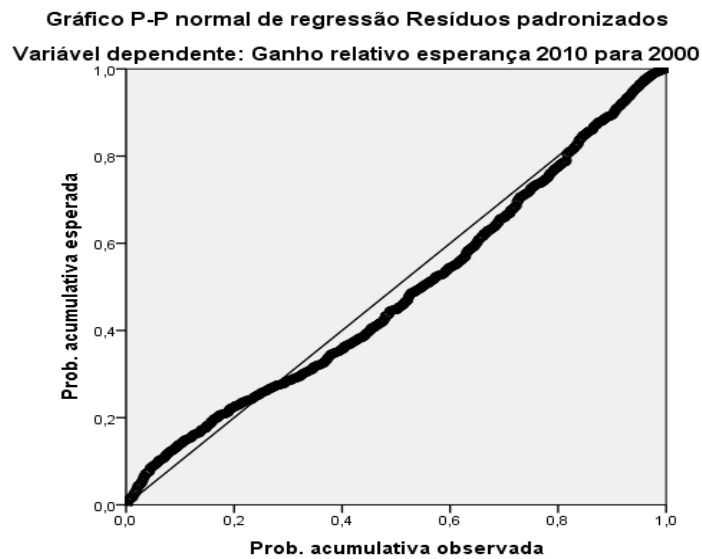


Tabela 11 – Normalidade Residual Modelo 5

**Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estatística	df	Sig.	Estatística	df	Sig.
Standardized Residual	,061	853	,000	,971	853	,000

a. Lilliefors Significance Correction

A Tabela 11, nos mostra a rejeição da hipótese normalidade com os testes de Lilliefors e Shapiro-Wilk com p-valores próximos de 0.

O gráfico 25 nos mostra que os resíduos sugerem ser heterocedásticos, com a tendência próxima do 0, também nos mostra uma relação não linear.

Gráfico 25 – Dispersão Resíduos Modelo 5

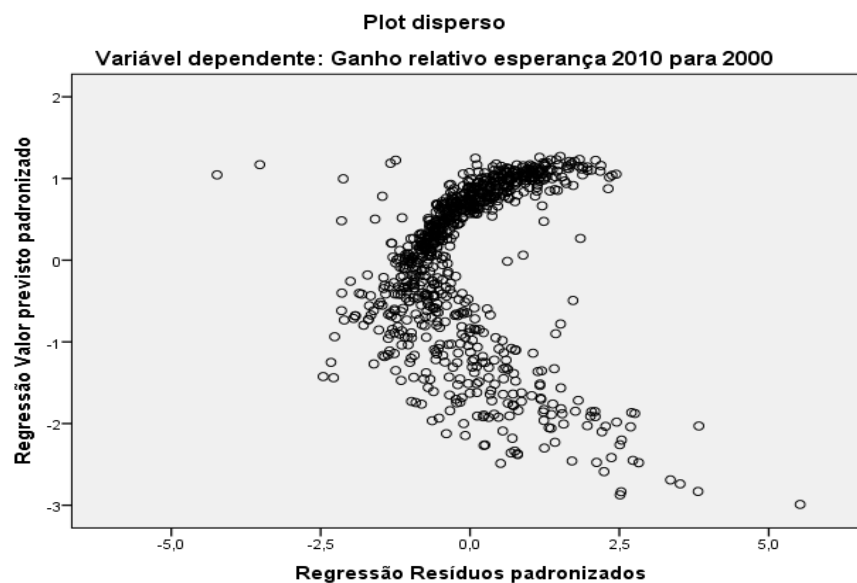
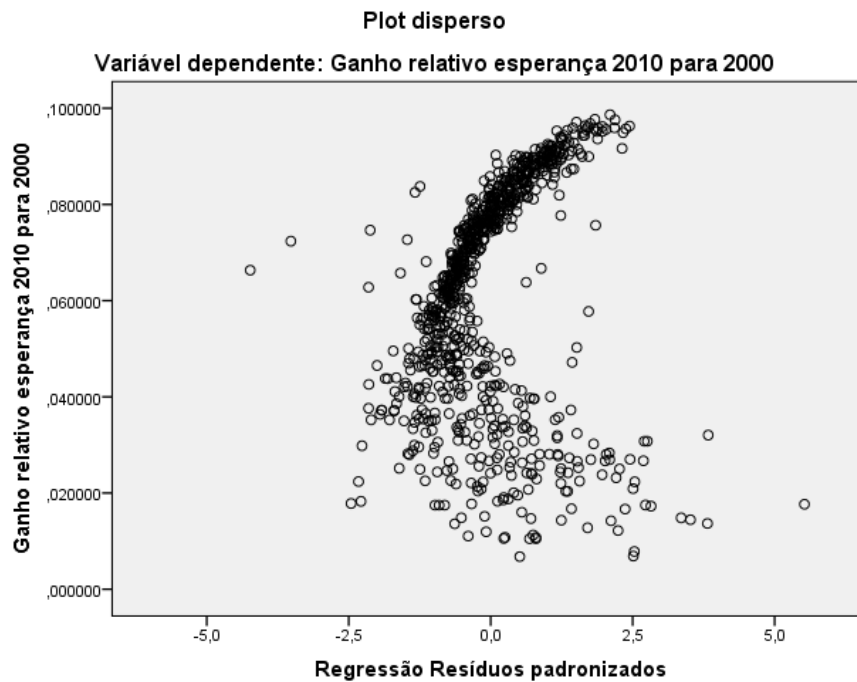


Gráfico 26 - Dispersão Observados x Resíduos Modelo 5



O Gráfico 26 nos indica a falta de independência dos resíduos, com uma tendência próxima do 1.

As variáveis de ganho na porcentagem de pessoas em domicílios com água encanada, ganho na porcentagem de pessoas em domicílios com energia elétrica e ganho na porcentagem de pessoas em domicílios com coleta de lixo possuem correlação entre si. Então, retiramos as duas primeiras variáveis para realizar o estudo com a coleta de lixo. Porém, esta também não apresentou significância no modelo de regressão linear, gerando um terceiro modelo somente com a variável de ganho na mortalidade infantil como regressora, que foi ignorado.

Uma primeira tentativa de solução para o problema da heterocedasticidade nos modelos considerados acima, foi a adição da forma quadrática da variável de mortalidade infantil ao modelo. O modelo foi dado por:

Modelo 6

$$Y = 0,01 + 0,27x_1 + 0,018x_2 + 1,014x_1^2$$

Onde:

$Y$  é o ganho na esperança de vida ao nascer;

$x_1$  o ganho na mortalidade infantil;

$x_2$  o ganho na porcentagem da população em domicílio com água encanada.

Neste modelo, a variável de ganho na expectativa de anos de estudos não foi significativa.

Tabela 12 – ANOVA Modelo 6

ANOVA <sup>a</sup>						
Modelo		Soma dos Quadrados	df	Quadrado Médio	F	Sig.
1	Regressão	,454	3	,151	17389,298	,000 <sup>b</sup>
	Resíduos	,007	849	,000		
	Total	,462	852			

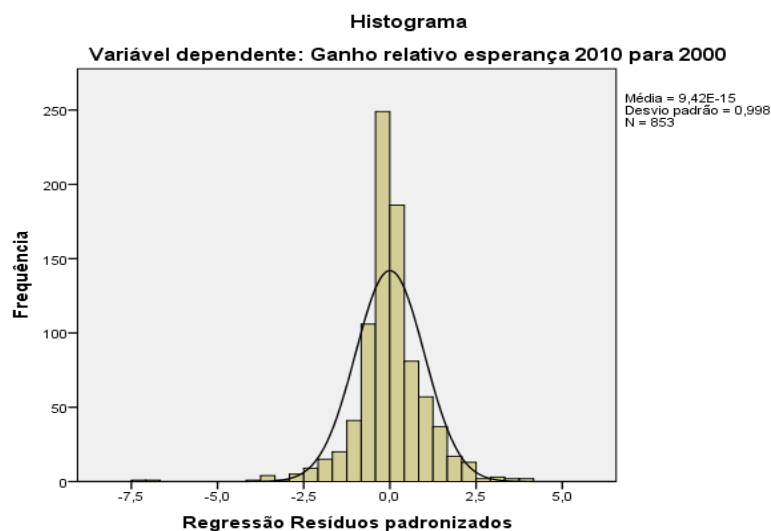
a. Variável dependente: Ganho relativo esperança 2010 para 2000

b. Preditores: (Constante), mort2, Ganho relativo na porcentagem de domicilio com água encanada 2010 para 2000, Ganho relativo Mortalidade Infantil 2010 para 2000

Como o p-valor da ANOVA do modelo foi menor que 0,001, temos que ele é significativo, a um nível de 5%, conforme a Tabela 12. Observamos ainda um  $R^2$  ajustado 0,984, próximo de 1, o que indica bom ajuste do mesmo.

No entanto, o problema da normalidade dos erros não foi resolvido nesta tentativa como podemos observar nos gráficos abaixo.

Gráfico 27 – Histograma dos resíduos Modelo 6



Notamos no Gráfico 27 que os resíduos seguem uma distribuição com média aproximadamente 0 e desvio padrão próximo do 1.

Pelo gráfico de probabilidade normal (Gráfico 28), observamos que os resíduos desviam da linha de probabilidade normal, mostrando que há indícios de que não seguem uma distribuição normal, e os testes da tabela 13 nos mostram rejeição da hipótese de normalidade.

Gráfico 28 – Probabilidade normal dos resíduos Modelo 6

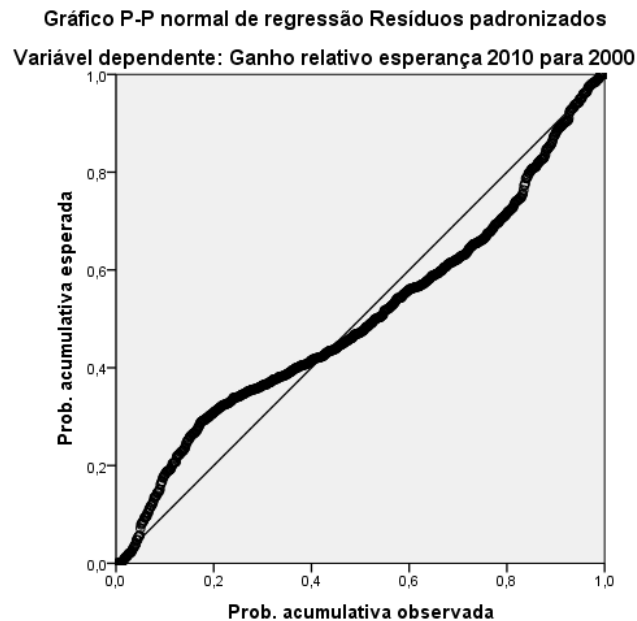


Tabela 13 – Normalidade Residual

**Tests of Normality**

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estatística	df	Sig.	Estatística	df	Sig.
Standardized Residual	,116	853	,000	,910	853	,000

a. Lilliefors Significance Correction

No entanto, a heterocedasticidade e a não linearidade persiste neste caso, como podemos observar no Gráfico 29. E no Gráfico 30, observamos a dependência residual.

Gráfico 29 – Dispersão Resíduos Modelo 6

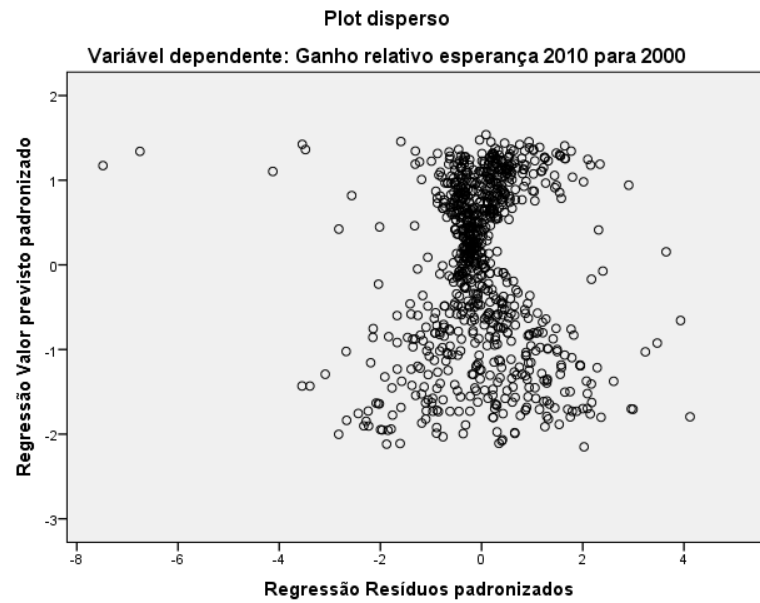
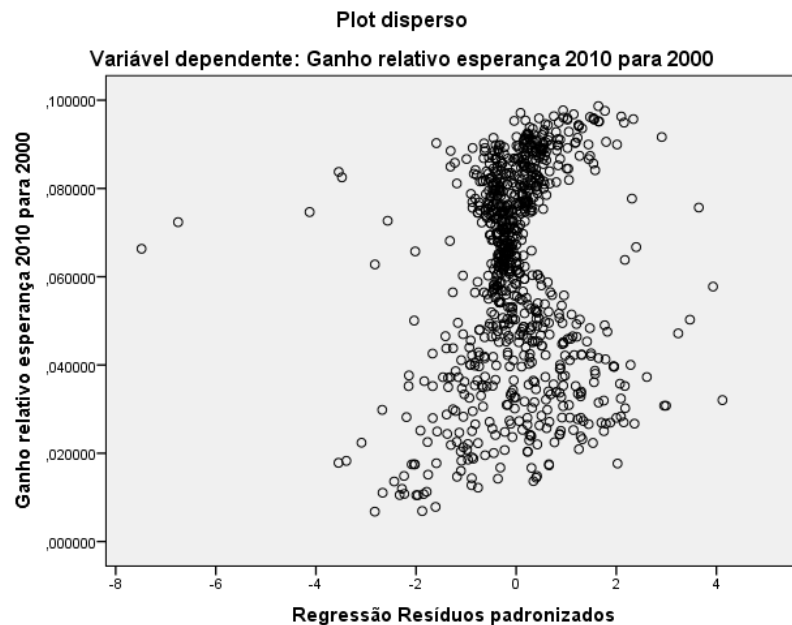


Gráfico 30 - Dispersão Observados x Resíduos Modelo 6



Uma segunda tentativa para a melhoria deste modelo foi a adição de interações com a variável de região criada como podemos observar no modelo 7, dado por:

Modelo 7

$$Y = -0,025 - 0,948x_1 - 0,034x_2 - 0,159x_3 - 0,017x_2x_3 - 0,186x_1x_3$$

Onde:

$Y$  é o ganho na esperança de vida ao nascer;

$x_1$  é o ganho na mortalidade infantil;

$x_2$  é ganho na porcentagem de pessoas em domicílio com água encanada;

$x_3$  é a região;

Tabela 14 – ANOVA Modelo 7

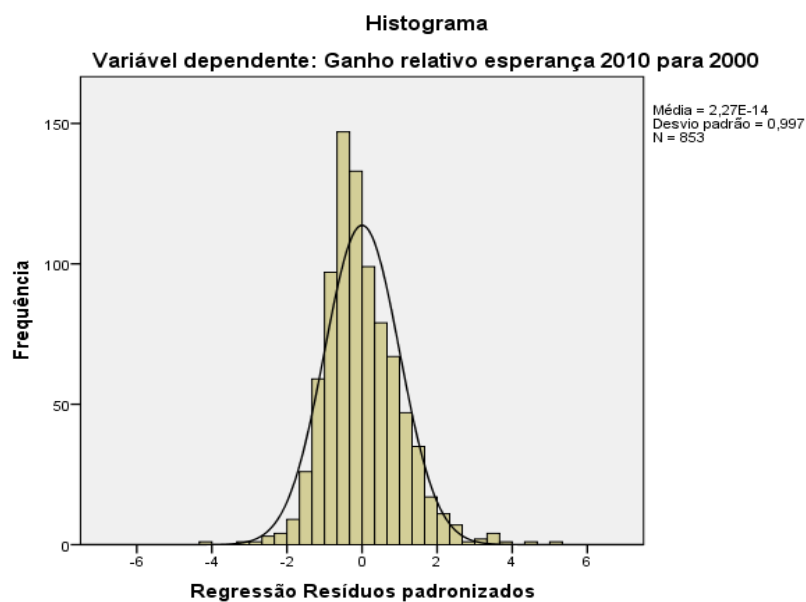
ANOVA <sup>a</sup>						
Modelo		Soma dos Quadrados	df	Quadrado Médio	F	Sig.
1	Regressão	,446	5	,089	4724,633	,000 <sup>b</sup>
	Resíduos	,016	847	,000		
	Total	,462	852			

a. Variável dependente: Ganho relativo esperança 2010 para 2000

b. Preditores: (Constante), Reg\_Mort2010, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000, Reg\_Agua2010, Região

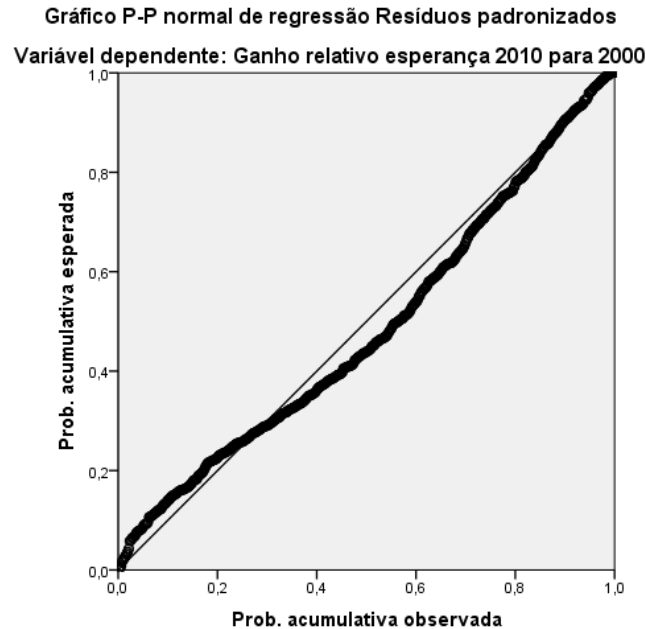
Como podemos observar na Tabela 14, a regressão foi significativa a um nível de 5%, com p-valor menor que 0,001. Observamos também  $R^2$  ajustado de 0,965, próximo de 1, o que pode indicar um bom ajuste do modelo. Neste modelo a interação região e água encanada não foi significativa. Porém, continuamos sem resolver o problema da normalidade dos resíduos, como podemos observar nos gráficos 31 e 32.

Gráfico 31 – Histograma Resíduos Modelo 7



Pelo gráfico 31, observamos que os resíduos seguem uma distribuição com média aproximadamente 0 e desvio aproximadamente 1.

Gráfico 32 - Probabilidade Normal Modelo 7



Analisado o gráfico 32, percebemos que os resíduos desviam da linha de probabilidade acumulada da normal, o que nos indicam ausência de normalidade dos mesmos. Pela Tabela 15, os testes nos mostram rejeição da hipótese de normalidade dos resíduos com p-valores menores que 0,01.

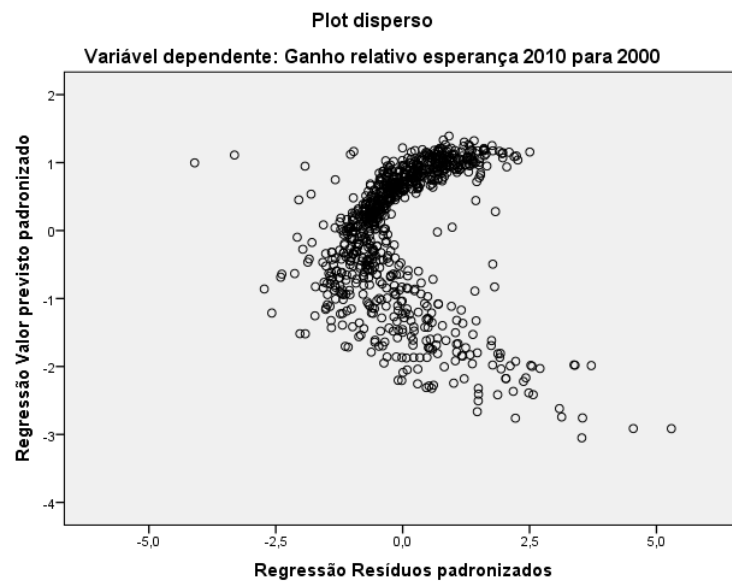
Tabela 15 - Normalidade Residual Modelo 7

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estatística	df	Sig.	Estatística	df	Sig.
Standardized Residual	,071	853	,000	,966	853	,000

a. Lilliefors Significance Correction

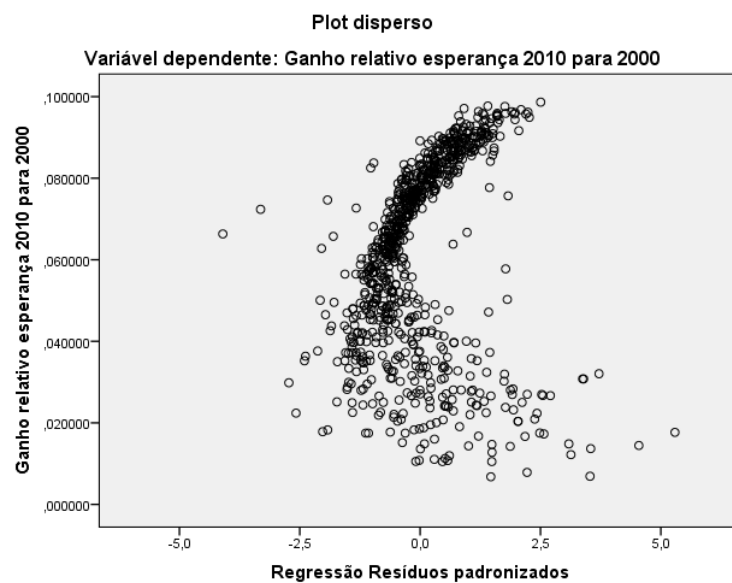
Pela dispersão dos resíduos apresentada no Gráfico 33, observa-se que a heterocedasticidade persiste e também a não linearidade.

Gráfico 33 – Dispersão Resíduos Modelo 7



O Gráfico 34 nos mostra dependência dos dados de resíduo do modelo.

Gráfico 34 - Dispersão Observados x Resíduos Modelo 7



Por fim, fizemos a tentativa de incluir a variável quadrática de ganho na mortalidade infantil, junto das variáveis de interação do modelo 7, gerando o seguinte modelo:

Modelo 8

$$Y = 0,011 + 0,036x_1 + 0,014x_2 - 0,138x_3 - 0,150x_1x_3 + 1,003x_1^2$$



Onde:

$Y$  é o ganho na esperança de vida ao nascer;

$x_1$  é o ganho na mortalidade infantil;

$x_2$  é o ganho na porcentagem de pessoas em domicílios com água encanada;

$x_3$  é a variável de região;

Tabela 16 – ANOVA Modelo 8

ANOVA <sup>a</sup>						
Modelo		Soma dos Quadrados	df	Quadrado Médio	F	Sig.
1	Regressão	,455	5	,091	11151,340	,000 <sup>b</sup>
	Resíduos	,007	847	,000		
	Total	,462	852			

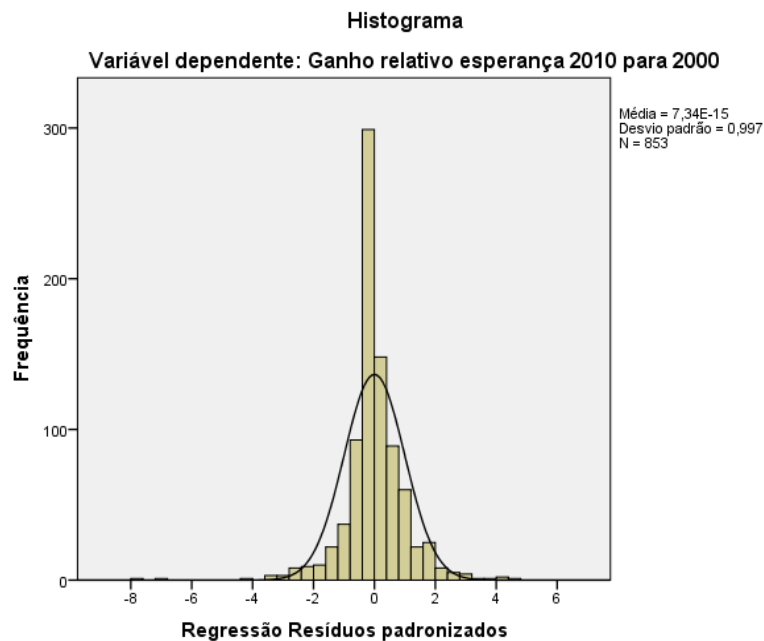
a. Variável dependente: Ganho relativo esperança 2010 para 2000

b. Preditores: (Constante), Mort\_Quad2010, Região, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000, Reg\_Mort2010, Ganho relativo Mortalidade Infantil 2010 para 2000

Pelo p-valor menor que 0,001 observado na Tabela 16, temos que a regressão foi significativa a um nível de 5%. Temos ainda que o  $R^2$  ajustado foi de 0,985, próximo de 1, o que pode indicar boa explicação da variação do ganho da esperança de vida ao nascer com este modelo.

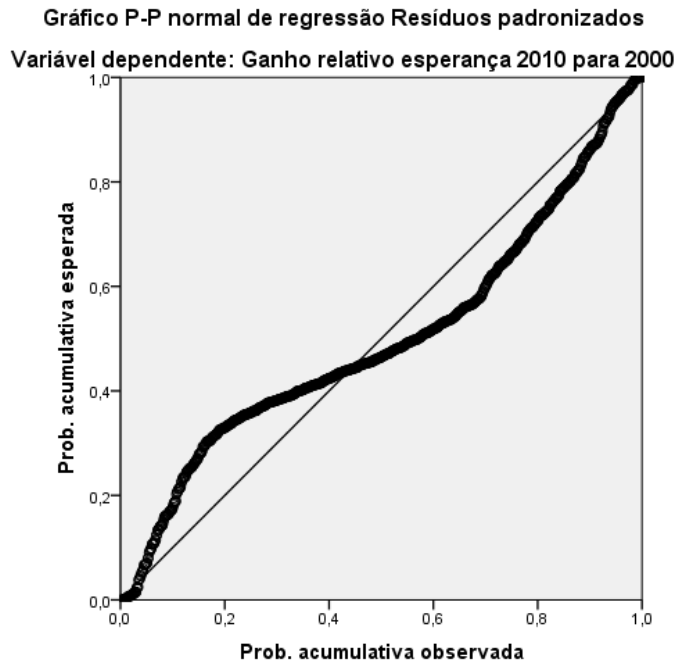
O histograma dos resíduos está representados no Gráfico 35.

Gráfico 35 – Histograma Resíduos Modelo 8



Pelo gráfico 35, podemos observar que os resíduos seguem uma distribuição com média e desvio padrão aproximadamente 0 e 1, respectivamente.

Gráfico 36 – Probabilidade Normal Modelo 8



Pelo Gráfico 36, podemos observar um grande desvio da linha de probabilidade normal acumulada, o que nos indica ausência da normalidade dos resíduos. A Tabela 17 nos mostra a rejeição da hipótese de normalidade pelos testes com p-valor menor que 0,001.

Tabela 17 – Normalidade Residual Modelo 8

	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Estatística	df	Sig.	Estatística	df	Sig.
Standardized Residual	,115	853	,000	,899	853	,000

a. Lilliefors Significance Correction

Temos que a heterocedasticidade e não linearidade dos resíduos ainda persiste, como observado no Gráfico 37 e no gráfico 38 nos mostra dependência residual.

Gráfico 37 – Dispersão Resíduos Modelo 8

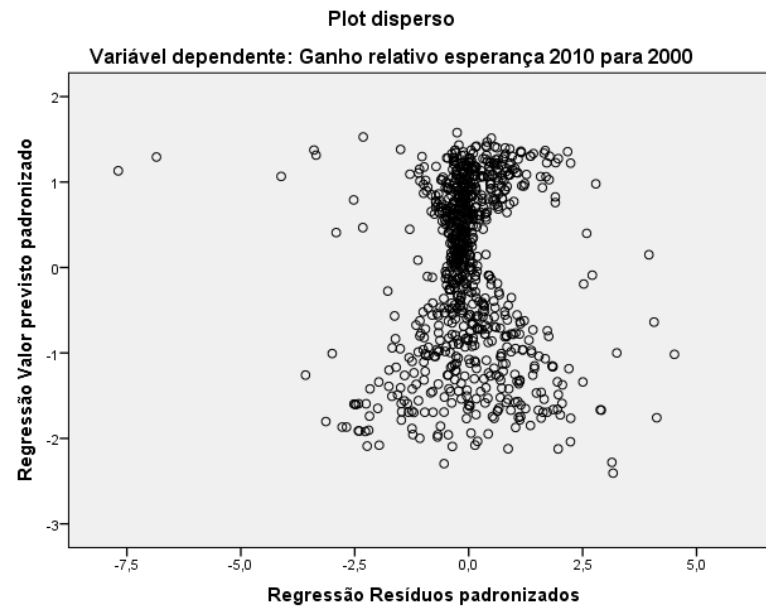
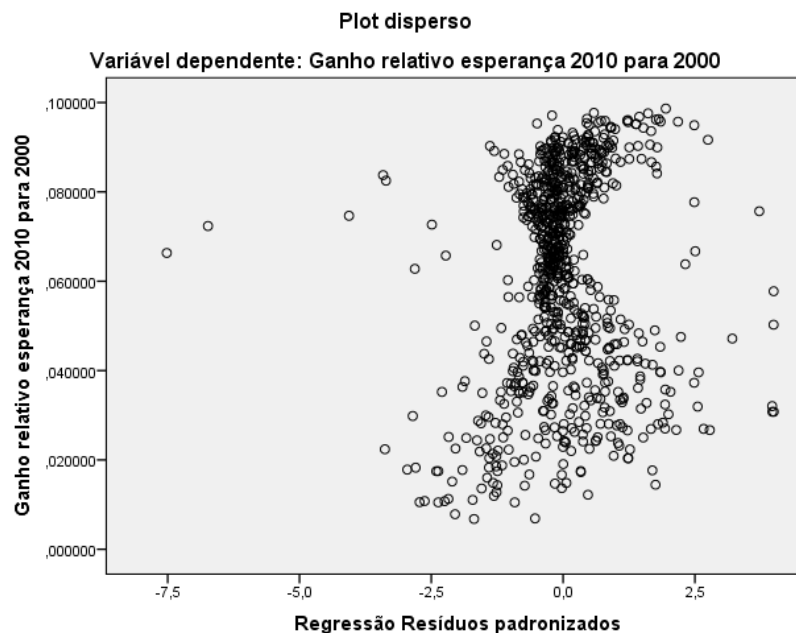


Gráfico 38 - Dispersão Observados x Resíduos Modelo



Numa tentativa de contornar as dificuldades encontradas nos modelos, foram testadas alternativas como transformações logarítmicas, exponenciais, quadráticas, raízes, mas ainda assim, nenhuma delas surtiu o efeito necessário para amenizar tais dificuldades, e por isso, elas foram ignoradas na análise.

Comparando os modelos dos períodos estudados, observamos uma tendência de crescimento no  $R^2$ , nos modelos com adição de interação e forma

quadrática, indicando uma melhora no ajuste dos mesmos. Além disso, houve algumas melhorias quanto a independência e homocedasticidade dos erros porém, os modelos ainda não respeitaram os pressupostos necessários para aplicação da regressão linear múltipla.

#### **4 – Conclusões e Considerações Finais**

Concluimos com o presente trabalho, dadas as devidas limitações por conta dos problemas encontrados nos modelos propostos, que fatores socioeconômicos estão associados de alguma maneira ao crescimento na esperança de vida ao nascer. No período 2010-2000 vimos que os fatores associados estão ligados ao crescimento das residências que possuem água encanada, ou também energia elétrica, o que nos indica que melhorias em saneamento básico ou melhorias como a chegada da energia elétrica nos municípios, melhoram a qualidade de vida das pessoas, acarretando no crescimento da variável estudada de ganho de esperança de vida ao nascer no período. Pudemos observar ainda que, nesse período, a variável de região não foi significativa nos modelos sem interação, ou seja, temos o indício de que o estado possa estar se homogeneizando nesse aspecto com o passar do tempo. Com o crescimento maior nos locais onde a esperança de vida ao nascer é menor, trazendo assim um equilíbrio de maneira geral a essa variável.

Um fator preponderante no período foi a queda na mortalidade infantil. Observamos no modelo que quando essa taxa decresce, temos um crescimento do ganho na esperança de vida ao nascer. A variável de mortalidade infantil obteve estimativas de parâmetros padronizados mais altos nos modelos estudados, como no Modelo 3 em que obteve 1,307 enquanto as outras não chegaram nem a 0,4, mostrando o peso da mesma. Analogamente temos isto ocorrendo em todos os outros modelos.

No período 2000-1991 os modelos propostos contém uma variável de cunho socioeconômico e outra demográfica. Foi observado ainda que a variável de ganho na mortalidade infantil, principalmente, está associada negativamente ao crescimento na esperança de vida ao nascer, ou seja, quanto maior a queda dessa variável, maior o crescimento da esperança de vida ao nascer. Variáveis de cunho social, como ganho na expectativa de anos de estudo, ganho na porcentagem de pessoas maiores de 18 anos com ensino médio completo, estão associadas positivamente ao ganho na esperança de vida ao nascer, mas não estão nos modelos pelo fato de serem multicolineares com as variáveis de cunho econômico.

O fator de ganho na porcentagem de pessoas em domicílio com água encanada e banheiro também foi associado, positivamente, ao ganho na esperança de vida ao nascer.

No período 2000-1991, observamos o Estado de Minas Gerais mais heterogêneo, sendo a variável de região também associada ao ganho na esperança de vida ao nascer. Os municípios em mesorregiões acima da linha traçada possuem crescimento diferente dos municípios situados em mesorregiões abaixo da linha traçada no mapa.

A conclusão que podemos tomar, dada as devidas limitações do estudo por conta dos problemas de ajustes dos modelos propostos, é que no período estudado as melhorias na qualidade de vida da população são associadas ao crescimento na esperança de vida ao nascer.

Nos dois períodos além dos fatores socioeconômicos/demográficos associados, percebemos em comum a preponderância do ganho na mortalidade infantil na explicação do ganho na esperança de vida ao nascer, que obteve a associação mais forte com a variável estudada. Também se observou em ambos os períodos que a interação entre região e as variáveis regressoras foram estatisticamente significantes, o que pode nos mostrar que existe diferenças entre as regiões que foram divididas arbitrariamente a partir do estudo descritivo dos dados no mapa de Minas Gerais.

#### **4.1 – Trabalhos Futuros**

Para contornar os problemas encontrados nesta monografia, sugerimos em trabalhos futuros a realização de outros métodos de modelagem para a explicação por meio de variáveis socioeconômicas e demográficas do ganho na esperança de vida ao nascer.

Uma destas sugestões é realizar análises de modelos com estimação por mínimos quadrados ponderados e/ou modelos não lineares, dado os problemas que indicam uma suposta não linearidade das relações entre as variáveis.

## Bibliografia

- Achen, C. H. (1982). *Interpreting And Using Regression*. Sage University Paper Series on Quantitative Applications in the Social Sciences. Newbury Park: SAGE Ltd.
- Bailey, T. C., & Gatrell, A. C. (1996). *Interactive Spatial Data Analysis*. Longman Group Limited.
- Brito, F. (2009, Setembro). As Migrações Internas No Brasil: Um Ensaio Sobre Desafios Teóricos Recentes. Belo Horizonte, Minas Gerais, Brasil: UFMG.
- Carvalho, J. M., Sawyer, D. O., & Rodrigues, R. d. (1998). *Introdução a Alguns Conceitos Básicos e Medidas em Demografia*. Associação Brasileira de Estudos Populacionais.
- Charnet, R., Freire, C. A., Charnet, E. M., & Bonvino, H. (2008). *Análise de Modelos de Regressão Linear Com Aplicações*. Campinas: Editora Unicamp.
- DATASUS. (n.d.). *Indicadores básicos para a saúde no Brasil: conceitos e aplicações*. Acesso em outubro de 2013, de DATASUS: <http://tabnet.datasus.gov.br/tabdata/livroidb/2ed/indicadores.pdf>
- Filho, D. B., & Júnior, J. d. (2009). Desvendando os Mistérios do Coeficiente de Correlação de Pearson (r). *Revista Política Hoje*, 18, 115-146.
- IBGE. (2010). *Síntese de Indicadores Sociais Uma Análise das Condições de Vida da população Brasileira*. Rio de Janeiro.
- Lewis-Beck, M. S. (1980). *Applied Regression: An Introduction*. Sage University Paper Series on Quantitative Applications in the Social Sciences. Newbury Park: SAGE Ltd.
- Ministério da Saúde. (n.d.). *Datasus*. Acesso em Março 2014, de Esperança de vida ao nascer: <http://tabnet.datasus.gov.br/cgi/idb1998/fqa13.htm>
- Montgomery, D. C., & Runger, G. C. (2003). *Estatística Aplicada e Probabilidade Para Engenheiros*. Rio de Janeiro: LTC - Livros Técnicos e Científicos Editora S.A.
- Montgomery, D. C., Runger, G. C., & Hubele, N. F. (2004). *Estatística Aplicada à Engenharia*. Rio de Janeiro: LTC - Livros Técnicos e Científicos Editora S.A.

- Pacievitch, T. (2010, Fevereiro 5). *Info Escola - Expectativa de Vida*. Acesso em Junho 2014, de Info Escola:  
<http://www.infoescola.com/demografia/expectativa-de-vida/>
- Peternelli, L. A. (2004). *Material Didático disponível na Web Disciplina Estatística I*. Retrieved Fevereiro 2014, from Departamento de Estatística, UFV:  
<http://www.dpi.ufv.br/~peternelli/inf162.www.16032004/materiais/CAPITULO9.pdf>
- PNUD. (n.d.). *Atlas do Desenvolvimento Humano no Brasil 2013*. Acesso em Novembro 2013, de Programa das Nações Unidas Para o Desenvolvimento:  
<http://www.atlasbrasil.org.br/2013/>
- PNUD. (n.d.). *Release Longevidade*. Acesso em Dezembro 2013, de PNUD  
[http://www.pnud.org.br/publicacoes/atlas\\_bh/release\\_longevidade.pdf](http://www.pnud.org.br/publicacoes/atlas_bh/release_longevidade.pdf)
- Preston, S. H., Heuveline, P., & Guillot, M. (2001). *Demography Measuring and Modeling Population Processes*. Oxford: Blackweel Publishing.
- REDE Interagencial de Informação para a Saúde. (2008). *Indicadores básicos para a saúde no Brasil: conceitos e aplicações (2ª ed.)*. Brasília: Organização Pan-Americana da Saúde.
- Romero, J. A. (2006, Setembro). *Análise Espacial da Pobreza Municipal no Estado de Minas Gerais 1991-2000*. Acesso em Julho 2014, de ABEP:  
[http://www.abep.nepo.unicamp.br/encontro2006/docspdf/abep2006\\_745.pdf](http://www.abep.nepo.unicamp.br/encontro2006/docspdf/abep2006_745.pdf)
- Schroeder, L. D., Sjoquist, D. L., & Stephan, P. E. (1986). *Understanding Regression Analysis: An Introduction Guide*. Sage University Paper Series on Quantitative Applications in the Social Sciences. Newbury Park: SAGE Publications Ltd.
- Wood, C. H., & Carvalho, J. M. (1994). *A Demografia da Desigualdade no Brasil*. Rio de Janeiro: PNPE IPEA.



## Anexo 1:

Quadro de correlação das variáveis de ganho relativo do período 2000-2010

Variáveis		Região	Ganho relativo esperança	Ganho relativo renda per capita	Ganho relativo Mortalidade Infantil	Ganho relativo taxa de fecundidade total	Ganho relativo taxa de envelhecimento	Ganho relativo expectativa de anos de estudo	Ganho relativo na população total	Ganho relativo na população rural	Ganho relativo na população urbana
Região	Correlação de Pearson	1	,209**	,227**	-,194**	-,137**	,158**	,118**	-,057	,065	,021
	P-valor	-	,000	,000	,000	,000	,000	,001	,096	,059	,547
Ganho relativo esperança	Correlação de Pearson	,209**	1	,241**	-,981**	-,182**	,063	,126**	-,089**	,037	,034
	P-valor	,000	-	,000	0,000	,000	,067	,000	,009	,275	,326
Ganho relativo renda per capita	Correlação de Pearson	,227**	,241**	1	-,236**	-,221**	,178**	,125**	-,091**	-,023	,101**
	P-valor	,000	,000	-	,000	,000	,000	,000	,008	,506	,003
Ganho relativo Mortalidade Infantil	Correlação de Pearson	-,194**	-,981**	-,236**	1	,183**	-,058	-,135**	,100**	-,026	-,024
	P-valor	,000	0,000	,000	-	,000	,091	,000	,003	,449	,493
Ganho relativo taxa de fecundidade total	Correlação de Pearson	-,137**	-,182**	-,221**	,183**	1	-,087*	-,118**	,147**	,026	-,050
	P-valor	,000	,000	,000	,000	-	,011	,001	,000	,453	,147
Ganho relativo taxa de envelhecimento	Correlação de Pearson	,158**	,063	,178**	-,058	-,087*	1	,073*	-,337**	-,048	-,191**
	P-valor	,000	,067	,000	,091	,011	-	,034	,000	,161	,000
Ganho relativo expectativa de anos de estudo	Correlação de Pearson	,118**	,126**	,125**	-,135**	-,118**	,073*	1	-,159**	,011	,005
	P-valor	,001	,000	,000	,000	,001	,034	-	,000	,743	,885
Ganho relativo na população total	Correlação de Pearson	-,057	-,089**	-,091**	,100**	,147**	-,337**	-,159**	1	,356**	,530**
	P-valor	,096	,009	,008	,003	,000	,000	,000	-	,000	,000
Ganho relativo na população rural	Correlação de Pearson	,065	,037	-,023	-,026	,026	-,048	,011	,356**	1	-,109**
	P-valor	,059	,275	,506	,449	,453	,161	,743	,000	-	,002
Ganho relativo na população urbana	Correlação de Pearson	,021	,034	,101**	-,024	-,050	-,191**	,005	,530**	-,109**	1
	P-valor	,547	,326	,003	,493	,147	,000	,885	,000	,002	-
Ganho relativo % de 18 anos ou mais com ensino médio completo	Correlação de Pearson	,182**	,188**	,379**	-,185**	-,327**	,171**	,266**	-,028	,038	,225**
	P-valor	,000	,000	,000	,000	,000	,000	,000	,413	,268	,000

Quadro de correlação das variáveis de ganho relativo do período 2000-2010 (Continuação)

Variáveis		Região	Ganho relativo esperança	Ganho relativo renda per capita	Ganho relativo Mortalidade Infantil	Ganho relativo taxa de fecundidade total	Ganho relativo taxa de envelhecimento	Ganho relativo expectativa de anos de estudo	Ganho relativo na população total	Ganho relativo na população rural	Ganho relativo na população urbana
Ganho RelativoTaxa de Atividade	Correlação de Pearson	,021	-,051	,284**	,044	-,038	-,040	,062	,057	,024	,103**
	P-valor	,549	,140	,000	,196	,270	,248	,070	,095	,488	,003
Ganho relativo na porcentagem de domicílio com água encanada	Correlação de Pearson	,374**	,265**	,350**	-,243**	-,273**	,196**	,253**	-,086*	,040	,129**
	P-valor	,000	,000	,000	,000	,000	,000	,000	,012	,243	,000
Ganho relativo % pop em domicílio com água enc. e banheiro	Correlação de Pearson	,374**	,258**	,367**	-,242**	-,327**	,219**	,279**	-,137**	,039	,136**
	P-valor	,000	,000	,000	,000	,000	,000	,000	,000	,254	,000
Ganho relativo % da pop com coleta de lixo	Correlação de Pearson	,202**	,177**	,216**	-,176**	-,216**	,048	,081*	-,126**	,003	,029
	P-valor	,000	,000	,000	,000	,000	,161	,018	,000	,933	,405
Ganho relativo % da pop com energia elétrica	Correlação de Pearson	,310**	,220**	,292**	-,206**	-,296**	,149**	,229**	-,133**	,039	,139**
	P-valor	,000	,000	,000	,000	,000	,000	,000	,000	,258	,000
Ganho relativo % pop em dom. com densidade>2	Correlação de Pearson	,035	-,037	-,118**	,045	,147**	-,170**	-,054	,325**	,136**	,130**
	P-valor	,301	,283	,001	,194	,000	,000	,114	,000	,000	,000
Ganho relativo % pessoas em dom. sem energia elétrica	Correlação de Pearson	-,048	-,052	-,102**	,061	,034	-,081*	-,059	,149**	,070*	,089**
	P-valor	,168	,130	,003	,078	,328	,019	,086	,000	,044	,009
Ganho relativo % pessoas em dom. com abast. e esgotamento inadequados	Correlação de Pearson	,014	-,030	-,033	,038	-,038	-,009	-,055	-,010	,040	-,006
	P-valor	,679	,381	,340	,272	,272	,796	,113	,773	,246	,874
Ganho relativo Índice de Gini	Correlação de Pearson	-,001	,030	,466**	-,028	,017	,008	-,048	,009	,013	-,039
	P-valor	,972	,388	,000	,409	,614	,812	,161	,785	,694	,257
Ganho relativo desigualdade	Correlação de Pearson	,143**	,166**	,898**	-,162**	-,131**	,122**	,059	-,045	-,012	,063
	P-valor	,000	,000	,000	,000	,000	,000	,085	,190	,721	,065

Quadro de correlação das variáveis de ganho relativo do período 2000-2010 (Continuação)

Variáveis		Ganho Relativo Taxa de Atividade	Ganho relativo % de domicilio com água encanada	Ganho relativo % pop em domicilio com agua enc. e banheiro	Ganho relativo % da pop com coleta de lixo	Ganho relativo % da pop com energia elétrica	Ganho relativo % pop em dom. com densidade >2	Ganho relativo % pessoas em dom. sem energia elétrica	Ganho relativo % pessoas em dom. com abast. e esgotamento inadequados	Ganho relativo Índice de Gini	Ganho relativo desigualdade	Ganho relativo % de 18 anos ou mais com ensino médio completo
Região	Correlação de Pearson	,021	,374**	,374**	,202**	,310**	,035	-,048	,014	-,001	,143**	,182**
	P-valor	,549	,000	,000	,000	,000	,301	,168	,679	,972	,000	,000
Ganho relativo esperança	Correlação de Pearson	-,051	,265**	,258**	,177**	,220**	-,037	-,052	-,030	,030	,166**	,188**
	P-valor	,140	,000	,000	,000	,000	,283	,130	,381	,388	,000	,000
Ganho relativo renda per capta	Correlação de Pearson	,284**	,350**	,367**	,216**	,292**	-,118**	-,102**	-,033	,466**	,898**	,379**
	P-valor	,000	,000	,000	,000	,000	,001	,003	,340	,000	,000	,000
Ganho relativo Mortalidade Infantil	Correlação de Pearson	,044	-,243**	-,242**	-,176**	-,206**	,045	,061	,038	-,028	-,162**	-,185**
	P-valor	,196	,000	,000	,000	,000	,194	,078	,272	,409	,000	,000
Ganho relativo taxa de fecundidade total	Correlação de Pearson	-,038	-,273**	-,327**	-,216**	-,296**	,147**	,034	-,038	,017	-,131**	-,327**
	P-valor	,270	,000	,000	,000	,000	,000	,328	,272	,614	,000	,000
Ganho relativo taxa de envelhecimento	Correlação de Pearson	-,040	,196**	,219**	,048	,149**	-,170**	-,081*	-,009	,008	,122**	,171**
	P-valor	,248	,000	,000	,161	,000	,000	,019	,796	,812	,000	,000
Ganho relativo expectativa de anos de estudo	Correlação de Pearson	,062	,253**	,279**	,081*	,229**	-,054	-,059	-,055	-,048	,059	,266**
	P-valor	,070	,000	,000	,018	,000	,114	,086	,113	,161	,085	,000
Ganho relativo na população total	Correlação de Pearson	,057	-,086*	-,137**	-,126**	-,133**	,325**	,149**	-,010	,009	-,045	-,028
	P-valor	,095	,012	,000	,000	,000	,000	,000	,773	,785	,190	,413
Ganho relativo na população rural	Correlação de Pearson	,024	,040	,039	,003	,039	,136**	,070*	,040	,013	-,012	,038
	P-valor	,488	,243	,254	,933	,258	,000	,044	,246	,694	,721	,268
Ganho relativo na população urbana	Correlação de Pearson	,103**	,129**	,136**	,029	,139**	,130**	,089**	-,006	-,039	,063	,225**
	P-valor	,003	,000	,000	,405	,000	,000	,009	,874	,257	,065	,000
Ganho relativo % de 18 anos ou mais com ensino médio completo	Correlação de Pearson	,184**	,495**	,557**	,266**	,392**	-,056	-,056	,011	,009	,263**	1
	P-valor	,000	,000	,000	,000	,000	,104	,106	,745	,797	,000	-

Quadro de correlação das variáveis de ganho relativo do período 2000-2010 (Continuação)

Variáveis		Ganho Relativo Taxa de Atividade	Ganho relativo na porcentagem de domicílio com água encanada	Ganho relativo % pop em domicílio com água enc. e banheiro	Ganho relativo % da pop com coleta de lixo	Ganho relativo % da pop com energia elétrica	Ganho relativo % pop em dom. com densidade>2	Ganho relativo % pessoas em dom. sem energia elétrica	Ganho relativo % pessoas em dom. com abast. e esgotamento inadequados	Ganho relativo Índice de Gini	Ganho relativo desigualdade	Ganho relativo % de 18 anos ou mais com ensino médio completo
Ganho Relativo Taxa de Atividade	Correlação de Pearson	1	,100**	,124**	,047	,050	,044	-,003	,000	-,078*	,158**	,184**
	P-valor	-	,004	,000	,168	,146	,201	,929	,999	,022	,000	,000
Ganho relativo na porcentagem de domicílio com água encanada	Correlação de Pearson	,100**	1	,903**	,354**	,634**	,045	-,064	-,010	-,066	,199**	,495**
	P-valor	,004	-	0,000	,000	,000	,193	,062	,771	,055	,000	,000
Ganho relativo % pop em domicílio com água enc. e banheiro	Correlação de Pearson	,124**	,903**	1	,371**	,702**	-,007	-,080*	-,029	-,048	,217**	,557**
	P-valor	,000	0,000	-	,000	,000	,833	,021	,401	,164	,000	,000
Ganho relativo % da pop com coleta de lixo	Correlação de Pearson	,047	,354**	,371**	1	,313**	,002	-,048	,022	-,020	,128**	,266**
	P-valor	,168	,000	,000	-	,000	,958	,167	,523	,560	,000	,000
Ganho relativo % da pop com energia elétrica	Correlação de Pearson	,050	,634**	,702**	,313**	1	,024	-,125**	,020	-,038	,170**	,392**
	P-valor	,146	,000	,000	,000	-	,485	,000	,569	,265	,000	,000
Ganho relativo % pop em dom. com densidade>2	Correlação de Pearson	,044	,045	-,007	,002	,024	1	,088*	-,054	-,050	-,105**	-,056
	P-valor	,201	,193	,833	,958	,485	-	,011	,116	,145	,002	,104
Ganho relativo % pessoas em dom. sem energia elétrica	Correlação de Pearson	-,003	-,064	-,080*	-,048	-,125**	,088*	1	,097**	-,053	-,088*	-,056
	P-valor	,929	,062	,021	,167	,000	,011	-	,005	,125	,011	,106
Ganho relativo % pessoas em dom. com abast. e esgotamento inadequados	Correlação de Pearson	,000	-,010	-,029	,022	,020	-,054	,097**	1	-,031	-,034	,011
	P-valor	,999	,771	,401	,523	,569	,116	,005	-	,365	,329	,745
Ganho relativo Índice de Gini	Correlação de Pearson	-,078*	-,066	-,048	-,020	-,038	-,050	-,053	-,031	1	,782**	,009
	P-valor	,022	,055	,164	,560	,265	,145	,125	,365	-	,000	,797
Ganho relativo desigualdade	Correlação de Pearson	,158**	,199**	,217**	,128**	,170**	-,105**	-,088*	-,034	,782**	1	,263**
	P-valor	,000	,000	,000	,000	,000	,002	,011	,329	,000	-	,000

Quadro de correlação das variáveis de ganho relativo do período 1991-2000

Variáveis		Região	Ganho relativo esperança	Ganho relativo renda per capita	Ganho relativo Mortalidade infantil	Ganho relativo taxa de fecundidade total	Ganho relativo taxa de envelhecimento	Ganho relativo expectativa de anos de estudo	Ganho relativo na população total	Ganho relativo na população rural	Ganho relativo na população urbana
<b>Região</b>	Correlação de Pearson	1	,101**	-,096**	,110**	-,234**	,169**	,224**	-,118**	-,018	,093**
	P-valor	-	,003	,005	,001	,000	,000	,000	,001	,595	,007
<b>Ganho relativo esperança</b>	Correlação de Pearson	,101**	1	-,033	-,824**	-,113**	,013	,089**	-,139**	-,054	-,034
	P-valor	,003	-	,332	,000	,001	,699	,010	,000	,116	,321
<b>Ganho relativo renda per capita</b>	Correlação de Pearson	-,096**	-,033	1	,038	-,038	-,100**	,022	-,030	-,015	,038
	P-valor	,005	,332	-	,271	,261	,004	,521	,374	,662	,266
<b>Ganho relativo Mortalidade infantil</b>	Correlação de Pearson	,110**	-,824**	,038	1	-,103**	,044	,090**	,014	,039	,093**
	P-valor	,001	,000	,271	-	,003	,203	,009	,688	,259	,007
<b>Ganho relativo taxa de fecundidade total</b>	Correlação de Pearson	-,234**	-,113**	-,038	-,103**	1	-,060	-,178**	,123**	,008	-,033
	P-valor	,000	,001	,261	,003	-	,081	,000	,000	,813	,346
<b>Ganho relativo taxa de envelhecimento</b>	Correlação de Pearson	,169**	,013	-,100**	,044	-,060	1	,077*	-,423**	-,216**	-,068*
	P-valor	,000	,699	,004	,203	,081	-	,024	,000	,000	,050
<b>Ganho relativo expectativa de anos de estudo</b>	Correlação de Pearson	,224**	,089**	,022	,090**	-,178**	,077*	1	-,068*	-,014	,080*
	P-valor	,000	,010	,521	,009	,000	,024	-	,047	,673	,021
<b>Ganho relativo na população total</b>	Correlação de Pearson	-,118**	-,139**	-,030	,014	,123**	-,423**	-,068*	1	,432**	,359**
	P-valor	,001	,000	,374	,688	,000	,000	,047	-	,000	,000
<b>Ganho relativo na população rural</b>	Correlação de Pearson	-,018	-,054	-,015	,039	,008	-,216**	-,014	,432**	1	-,024
	P-valor	,595	,116	,662	,259	,813	,000	,673	,000	-	,485
<b>Ganho relativo na população urbana</b>	Correlação de Pearson	,093**	-,034	,038	,093**	-,033	-,068*	,080*	,359**	-,024	1
	P-valor	,007	,321	,266	,007	,346	,050	,021	,000	,485	-
<b>Ganho relativo % de 18 anos ou mais com ensino médio completo</b>	Correlação de Pearson	,216**	,076*	,043	,091**	-,159**	,068*	,253**	,011	,027	,317**
	P-valor	,000	,027	,214	,008	,000	,047	,000	,744	,429	,000

Quadro de correlação das variáveis de ganho relativo do período 1991-2000 (Continuação)

Variáveis		Região	Ganho relativo esperança	Ganho relativo renda per capita	Ganho relativo Mortalidade infantil	Ganho relativo taxa de fecundidade total	Ganho relativo taxa de envelhecimento	Ganho relativo expectativa de anos de estudo	Ganho relativo na população total	Ganho relativo na população rural	Ganho relativo na população urbana
<b>Ganho relativo na porcentagem de domicílio com água encanada</b>	Correlação de Pearson	,294**	,097**	-,041	,131**	-,283**	,166**	,275**	-,130**	-,022	,242**
	P-valor	,000	,004	,237	,000	,000	,000	,000	,000	,517	,000
<b>Ganho relativo % pop em domicílio com água enc. e banheiro</b>	Correlação de Pearson	,308**	,101**	-,022	,140**	-,257**	,160**	,340**	-,148**	-,020	,287**
	P-valor	,000	,003	,520	,000	,000	,000	,000	,000	,567	,000
<b>Ganho relativo % da pop com coleta de lixo</b>	Correlação de Pearson	,075*	-,083*	-,021	,115**	-,048	-,009	,044	-,008	,020	,194**
	P-valor	,038	,021	,562	,001	,179	,813	,223	,832	,581	,000
<b>Ganho relativo % da pop com energia elétrica</b>	Correlação de Pearson	,316**	,066	,054	,205**	-,301**	,141**	,310**	-,207**	-,007	,280**
	P-valor	,000	,053	,118	,000	,000	,000	,000	,000	,848	,000
<b>Ganho relativo % pop em dom. com densidade&gt;2</b>	Correlação de Pearson	,104**	,030	-,162**	,103**	-,122**	-,058	,038	,106**	,041	,209**
	P-valor	,002	,385	,000	,003	,000	,090	,269	,002	,236	,000
<b>Ganho relativo % pessoas em dom. sem energia elétrica</b>	Correlação de Pearson	,247**	,064	-,195**	,090**	-,125**	,123**	,188**	-,142**	-,057	,049
	P-valor	,000	,062	,000	,009	,000	,000	,000	,000	,097	,158
<b>Ganho relativo % pessoas em dom. com abast. e esgotamento inadequado</b>	Correlação de Pearson	-,006	-,056	,024	,035	,025	-,020	-,003	-,012	,013	,016
	P-valor	,865	,115	,511	,328	,489	,570	,938	,735	,709	,653
<b>Ganho relativo Índice de Gini</b>	Correlação de Pearson	,146**	,008	,389**	,096**	-,092**	,122**	,135**	-,014	-,017	,197**
	P-valor	,000	,815	,000	,005	,007	,000	,000	,686	,610	,000
<b>Ganho relativo desigualdade</b>	Correlação de Pearson	-,009	-,014	,886**	,064	-,071*	-,021	,058	-,018	-,018	,129**
	P-valor	,801	,693	,000	,060	,039	,548	,093	,596	,600	,000

Quadro de correlação das variáveis de ganho relativo do período 1991-2000 (Continuação)

Variáveis		Ganho relativo % de 18 anos ou mais com ensino médio completo	Ganho relativo na porcentagem de domicílio com água encanada	Ganho relativo % pop em domicílio com água enc. e banheiro	Ganho relativo % da pop com coleta de lixo	Ganho relativo % da pop com energia elétrica	Ganho relativo % pop em dom. com densidade>2	Ganho relativo % pessoas em dom. sem energia elétrica	Ganho relativo % pessoas em dom. com abast. e esgotamento inadequado	Ganho relativo Índice de Gini	Ganho relativo desigualdade
Região	Correlação de Pearson	,216**	,294**	,308**	,075*	,316**	,104**	,247**	-,006	,146**	-,009
	P-valor	,000	,000	,000	,038	,000	,002	,000	,865	,000	,801
Ganho relativo esperança	Correlação de Pearson	,076*	,097**	,101**	-,083*	,066	,030	,064	-,056	,008	-,014
	P-valor	,027	,004	,003	,021	,053	,385	,062	,115	,815	,693
Ganho relativo renda per capta	Correlação de Pearson	,043	-,041	-,022	-,021	,054	-,162**	-,195**	,024	,389**	,886**
	P-valor	,214	,237	,520	,562	,118	,000	,000	,511	,000	,000
Ganho relativo Mortalidade infantil	Correlação de Pearson	,091**	,131**	,140**	,115**	,205**	,103**	,090**	,035	,096**	,064
	P-valor	,008	,000	,000	,001	,000	,003	,009	,328	,005	,060
Ganho relativo taxa de fecundidade total	Correlação de Pearson	-,159**	-,283**	-,257**	-,048	-,301**	-,122**	-,125**	,025	-,092**	-,071*
	P-valor	,000	,000	,000	,179	,000	,000	,000	,489	,007	,039
Ganho relativo taxa de envelhecimento	Correlação de Pearson	,068*	,166**	,160**	-,009	,141**	-,058	,123**	-,020	,122**	-,021
	P-valor	,047	,000	,000	,813	,000	,090	,000	,570	,000	,548
Ganho relativo expectativa de anos de estudo	Correlação de Pearson	,253**	,275**	,340**	,044	,310**	,038	,188**	-,003	,135**	,058
	P-valor	,000	,000	,000	,223	,000	,269	,000	,938	,000	,093
Ganho relativo na população total	Correlação de Pearson	,011	-,130**	-,148**	-,008	-,207**	,106**	-,142**	-,012	-,014	-,018
	P-valor	,744	,000	,000	,832	,000	,002	,000	,735	,686	,596
Ganho relativo na população rural	Correlação de Pearson	,027	-,022	-,020	,020	-,007	,041	-,057	,013	-,017	-,018
	P-valor	,429	,517	,567	,581	,848	,236	,097	,709	,610	,600
Ganho relativo na população urbana	Correlação de Pearson	,317**	,242**	,287**	,194**	,280**	,209**	,049	,016	,197**	,129**
	P-valor	,000	,000	,000	,000	,000	,000	,158	,653	,000	,000
Ganho relativo % de 18 anos ou mais com ensino médio completo	Correlação de Pearson	1	,572**	,608**	,113**	,456**	,086*	,123**	-,001	,328**	,177**
	P-valor	-	,000	,000	,002	,000	,012	,000	,975	,000	,000

Quadro de correlação das variáveis de ganho relativo do período 1991-2000 (Continuação)

Variáveis		Ganho relativo % de 18 anos ou mais com ensino médio completo	Ganho relativo na porcentagem de domicílio com água encanada	Ganho relativo % pop em domicílio com água enc. e banheiro	Ganho relativo % da pop com coleta de lixo	Ganho relativo % da pop com energia elétrica	Ganho relativo % pop em dom. com densidade>2	Ganho relativo % pessoas em dom. sem energia elétrica	Ganho relativo % pessoas em dom. com abast. e esgotamento inadequado	Ganho relativo Índice de Gini	Ganho relativo desigualdade
<b>Ganho relativo na porcentagem de domicílio com água encanada</b>	Correlação de Pearson	,572**	1	,906**	,146**	,614**	,153**	,230**	-,032	,220**	,063
	P-valor	,000	-	0,000	,000	,000	,000	,000	,364	,000	,067
<b>Ganho relativo % pop em domicílio com água enc. e banheiro</b>	Correlação de Pearson	,608**	,906**	1	,147**	,642**	,131**	,237**	-,034	,264**	,090**
	P-valor	,000	0,000	-	,000	,000	,000	,000	,344	,000	,008
<b>Ganho relativo % da pop com coleta de lixo</b>	Correlação de Pearson	,113**	,146**	,147**	1	,121**	,047	,066	,017	,029	-,007
	P-valor	,002	,000	,000	-	,001	,194	,068	,661	,421	,854
<b>Ganho relativo % da pop com energia elétrica</b>	Correlação de Pearson	,456**	,614**	,642**	,121**	1	,181**	,237**	,005	,285**	,159**
	P-valor	,000	,000	,000	,001	-	,000	,000	,892	,000	,000
<b>Ganho relativo % pop em dom. com densidade&gt;2</b>	Correlação de Pearson	,086*	,153**	,131**	,047	,181**	1	,174**	,019	,018	-,102**
	P-valor	,012	,000	,000	,194	,000	-	,000	,605	,602	,003
<b>Ganho relativo % pessoas em dom. sem energia elétrica</b>	Correlação de Pearson	,123**	,230**	,237**	,066	,237**	,174**	1	,023	,130**	-,088*
	P-valor	,000	,000	,000	,068	,000	,000	-	,515	,000	,010
<b>Ganho relativo % pessoas em dom. com abast. e esgotamento inadequado</b>	Correlação de Pearson	-,001	-,032	-,034	,017	,005	,019	,023	1	,024	,030
	P-valor	,975	,364	,344	,661	,892	,605	,515	-	,505	,407
<b>Ganho relativo Índice de Gini</b>	Correlação de Pearson	,328**	,220**	,264**	,029	,285**	,018	,130**	,024	1	,747**
	P-valor	,000	,000	,000	,421	,000	,602	,000	,505	-	,000
<b>Ganho relativo desigualdade</b>	Correlação de Pearson	,177**	,063	,090**	-,007	,159**	-,102**	-,088*	,030	,747**	1
	P-valor	,000	,067	,008	,854	,000	,003	,010	,407	,000	-



**Anexo 2:****Modelos de regressão para os períodos 1991-2000 e 2000-2010**

Tabela 1: Modelo de regressão 1, 1991-2000

**Coefficientes<sup>a</sup>**

Modelo	Coefficients não padronizados		Coefficientes padronizados	t	Sig.
	B	Modelo padrão	Beta		
1 (Constante)	,023	,001		22,767	,000
Ganho relativo Mortalidade infantil 2000 para 1991	-,183	,004	-,864	-49,191	,000
Ganho relativo % pop em domicilio com agua enc. e banheiro 2000 para 1991	,005	,000	,178	9,704	,000
Região	,006	,001	,142	7,740	,000

a. Variável dependente: Ganho relativo esperança 2000 para 1991

Tabela 2: Modelo de regressão 2, 1991-2000

**Coefficientes<sup>a</sup>**

Modelo	Coefficients não padronizados		Coefficientes padronizados	t	Sig.
	B	Modelo padrão	Beta		
1 (Constante)	,023	,001		19,141	,000
Região	,004	,002	,084	1,828	,068
Ganho relativo Mortalidade infantil 2000 para 1991	-,177	,004	-,833	-40,017	,000
Ganho relativo % pop em domicilio com agua enc. e banheiro 2000 para 1991	,009	,001	,324	10,031	,000
Reg_Mort91	-,025	,008	-,139	-3,240	,001
Reg_Aguat91	-,006	,001	-,200	-5,366	,000

a. Variável dependente: Ganho relativo esperança 2000 para 1991

Tabela 3: Modelo de regressão 3, 1991-2000

**Coefficientes<sup>a</sup>**

Modelo	Coefficients não padronizados		Coefficientes padronizados	t	Sig.
	B	Modelo padrão	Beta		
1 (Constante)	,014	,002		8,626	,000
Região	,006	,002	,146	3,215	,001
Ganho relativo Mortalidade infantil 2000 para 1991	-,277	,014	-1,307	-19,797	,000
Ganho relativo % pop em domicilio com agua enc. e banheiro 2000 para 1991	,009	,001	,325	10,398	,000
Reg_Mort91	-,013	,008	-,072	-1,693	,091
Reg_Aguat91	-,006	,001	-,213	-5,906	,000
Mort_Quad_91	-,233	,031	-,474	-7,540	,000

a. Variável dependente: Ganho relativo esperança 2000 para 1991

Tabela 4: Modelo de regressão 4, 2000-2010

**Coefficientes<sup>a</sup>**

Modelo	Coefficients não padronizados		Coefficients padronizados	t	Sig.
	B	Modelo padrão	Beta		
1 (Constante)	-,027	,001		-43,587	,000
Ganho relativo Mortalidade Infantil 2010 para 2000	-,215	,001	-,976	-144,709	,000
Ganho relativo na porcentagem de domicilio com água encanada 2010 para 2000	,002	,000	,031	4,562	,000
Ganho relativo expectativa de anos de estudo 2010 para 2000	-,002	,001	-,014	-2,067	,039

a. Variável dependente: Ganho relativo esperança 2010 para 2000

Tabela 5: Modelo de regressão 5, 2000-2010

**Coefficientes<sup>a</sup>**

Modelo	Coefficients não padronizados		Coefficients padronizados	t	Sig.
	B	Modelo padrão	Beta		
1 (Constante)	-,027	,001		-44,079	,000
Ganho relativo Mortalidade Infantil 2010 para 2000	-,215	,001	-,979	-145,271	,000
Ganho relativo % da pop com energia elétrica 2010 para 2000	,003	,001	,021	3,036	,002
Ganho relativo expectativa de anos de estudo 2010 para 2000	-,002	,001	-,011	-1,650	,099

a. Variável dependente: Ganho relativo esperança 2010 para 2000

Tabela 6: Modelo de regressão 6, 2000-2010

**Coefficientes<sup>a</sup>**

Modelo	Coefficients não padronizados		Coefficients padronizados	t	Sig.
	B	Modelo padrão	Beta		
1 (Constante)	,010	,001		8,149	,000
Ganho relativo Mortalidade Infantil 2010 para 2000	,006	,007	,027	,869	,385
Ganho relativo na porcentagem de domicilio com água encanada 2010 para 2000	,001	,000	,018	4,030	,000
Mort_Quad2010	,301	,009	1,014	32,769	,000

a. Variável dependente: Ganho relativo esperança 2010 para 2000

Tabela 7: Modelo de regressão 7, 2000-2010

**Coefficientes<sup>a</sup>**

Modelo	Coefficients não padronizados		Coefficients padronizados	t	Sig.
	B	Modelo padrão	Beta		
1 (Constante)	-,025	,001		-35,256	,000
Ganho relativo Mortalidade Infantil 2010 para 2000	-,209	,002	-,948	-120,659	,000
Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000	,002	,001	,034	2,544	,011
Região	-,008	,001	-,159	-5,601	,000
Reg_Agua2010	-,001	,001	-,017	-1,170	,242
Reg_Mort2010	-,020	,003	-,186	-6,184	,000

a. Variável dependente: Ganho relativo esperança 2010 para 2000

Tabela 8: Modelo de regressão 8, 2000-2010

**Coefficientes<sup>a</sup>**

Modelo	Coefficients não padronizados		Coefficients padronizados	t	Sig.
	B	Modelo padrão	Beta		
1 (Constante)	,011	,001		9,469	,000
Ganho relativo Mortalidade Infantil 2010 para 2000	,008	,007	,036	1,208	,227
Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000	,001	,000	,014	3,111	,002
Região	-,007	,001	-,138	-7,401	,000
Reg_Mort2010	-,016	,002	-,150	-7,704	,000
Mort_Quad2010	,298	,009	1,003	33,398	,000

a. Variável dependente: Ganho relativo esperança 2010 para 2000

## Anexo 3

Tabela 1: Modelo Completo 2000-2010

Modelo		Coeficientes não padronizados		Coeficientes padronizados	t	Sig.
		B	Modelo padrão	Beta		
1	(Constante)	-,026	,001		-37,253	,000
	Região	,001	,000	,010	1,454	,146
	Ganho relativo Mortalidade Infantil 2010 para 2000	-,215	,002	-,975	-140,754	0,000
	Ganho relativo renda per capita 2010 para 2000	,000	,001	,003	,378	,705
	Ganho relativo na porcentagem de domicilio com água encanada 2010 para 2000	,002	,000	,030	3,127	,002
	Ganho relativo % da pop com energia eletrica 2010 para 2000	,001	,001	,004	,464	,643
	Ganho relativo % da pop com coleta de lixo 2010 para 2000	,000	,000	-,005	-,759	,448
	Ganho relativo % de 18 anos ou mais com ensino médio completo 2010 para 2000	,000	,000	-,005	-,675	,500
	Ganho relativo expectativa de anos de estudo 2010 para 2000	-,002	,001	-,014	-1,986	,047
	Ganho relativo taxa de fecundidade total 2010 para 2000	,000	,001	,003	,389	,697
2	(Constante)	-,026	,001		-37,473	,000
	Região	,001	,000	,011	1,495	,135
	Ganho relativo Mortalidade Infantil 2010 para 2000	-,215	,002	-,975	-142,068	0,000
	Ganho relativo na porcentagem de domicilio com água encanada 2010 para 2000	,002	,000	,030	3,172	,002
	Ganho relativo % da pop com energia eletrica 2010 para 2000	,001	,001	,004	,480	,631
	Ganho relativo % da pop com coleta de lixo 2010 para 2000	,000	,000	-,005	-,741	,459
	Ganho relativo % de 18 anos ou mais com ensino médio completo 2010 para 2000	,000	,000	-,005	-,607	,544
	Ganho relativo expectativa de anos de estudo 2010 para 2000	-,002	,001	-,014	-1,993	,047
	Ganho relativo taxa de fecundidade total 2010 para 2000	,000	,001	,003	,368	,713
	3	(Constante)	-,026	,001		-38,093
Região		,001	,000	,011	1,492	,136
Ganho relativo Mortalidade Infantil 2010 para 2000		-,215	,002	-,975	-142,738	0,000
Ganho relativo na porcentagem de domicilio com água encanada 2010 para 2000		,002	,000	,030	3,172	,002
Ganho relativo % da pop com energia eletrica 2010 para 2000		,000	,001	,004	,437	,662
Ganho relativo % da pop com coleta de lixo 2010 para 2000		,000	,000	-,006	-,777	,438
Ganho relativo % de 18 anos ou mais com ensino médio completo 2010 para 2000		,000	,000	-,005	-,694	,488
Ganho relativo expectativa de anos de estudo 2010 para 2000		-,002	,001	-,014	-1,995	,046
4	(Constante)	-,026	,001		-38,178	,000
	Região	,001	,000	,011	1,538	,125
	Ganho relativo Mortalidade Infantil 2010 para 2000	-,215	,002	-,975	-142,919	0,000
	Ganho relativo na porcentagem de domicilio com água encanada 2010 para 2000	,002	,000	,032	3,833	,000
	Ganho relativo % da pop com coleta de lixo 2010 para 2000	,000	,000	-,005	-,736	,462
	Ganho relativo % de 18 anos ou mais com ensino médio completo 2010 para 2000	,000	,000	-,005	-,658	,511
5	(Constante)	-,027	,001		-39,063	,000
	Região	,001	,000	,011	1,554	,121
	Ganho relativo Mortalidade Infantil 2010 para 2000	-,215	,001	-,975	-143,151	0,000
	Ganho relativo na porcentagem de domicilio com água encanada 2010 para 2000	,002	,000	,030	3,882	,000
	Ganho relativo % da pop com coleta de lixo 2010 para 2000	,000	,000	-,006	-,814	,416
	Ganho relativo expectativa de anos de estudo 2010 para 2000	-,002	,001	-,014	-2,108	,035
6	(Constante)	-,027	,001		-43,645	,000
	Região	,001	,000	,011	1,501	,134
	Ganho relativo Mortalidade Infantil 2010 para 2000	-,214	,001	-,974	-143,710	0,000
	Ganho relativo na porcentagem de domicilio com água encanada 2010 para 2000	,001	,000	,028	3,809	,000
	Ganho relativo expectativa de anos de estudo 2010 para 2000	-,002	,001	-,014	-2,093	,037
7	(Constante)	-,027	,001		-43,587	,000
	Ganho relativo Mortalidade Infantil 2010 para 2000	-,215	,001	-,976	-144,709	0,000
	Ganho relativo na porcentagem de domicilio com água encanada 2010 para 2000	,002	,000	,031	4,562	,000
	Ganho relativo expectativa de anos de estudo 2010 para 2000	-,002	,001	-,014	-2,067	,039

a. Variável dependente: Ganho relativo esperança 2010 para 2000

Tabela 2 : ANOVA Backward

**ANOVA<sup>a</sup>**

Modelo		Soma dos Quadrados	df	Quadrado Médio	F	Sig.
1	Regressão	,445	9	,049	2513,713	,000 <sup>b</sup>
	Resíduos	,017	843	,000		
	Total	,462	852			
2	Regressão	,445	8	,056	2830,784	,000 <sup>c</sup>
	Resíduos	,017	844	,000		
	Total	,462	852			
3	Regressão	,445	7	,064	3238,475	,000 <sup>d</sup>
	Resíduos	,017	845	,000		
	Total	,462	852			
4	Regressão	,445	6	,074	3781,806	,000 <sup>e</sup>
	Resíduos	,017	846	,000		
	Total	,462	852			
5	Regressão	,445	5	,089	4541,121	,000 <sup>f</sup>
	Resíduos	,017	847	,000		
	Total	,462	852			
6	Regressão	,445	4	,111	5678,495	,000 <sup>g</sup>
	Resíduos	,017	848	,000		
	Total	,462	852			
7	Regressão	,445	3	,148	7559,426	,000 <sup>h</sup>
	Resíduos	,017	849	,000		
	Total	,462	852			

a. Variável dependente: Ganho relativo esperança 2010 para 2000

b. Preditores: (Constante), Ganho relativo taxa de fecundidade total 2010 para 2000, Ganho relativo expectativa de anos de estudo 2010 para 2000, Região, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo % da pop com coleta de lixo 2010 para 2000, Ganho relativo renda per capita 2010 para 2000, Ganho relativo % da pop com energia elétrica 2010 para 2000, Ganho relativo % de 18 anos ou mais com ensino médio completo 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000

c. Preditores: (Constante), Ganho relativo taxa de fecundidade total 2010 para 2000, Ganho relativo expectativa de anos de estudo 2010 para 2000, Região, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo % da pop com coleta de lixo 2010 para 2000, Ganho relativo % da pop com energia elétrica 2010 para 2000, Ganho relativo % de 18 anos ou mais com ensino médio completo 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000

d. Preditores: (Constante), Ganho relativo expectativa de anos de estudo 2010 para 2000, Região, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo % da pop com coleta de lixo 2010 para 2000, Ganho relativo % da pop com energia elétrica 2010 para 2000, Ganho relativo % de 18 anos ou mais com ensino médio completo 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000

e. Preditores: (Constante), Ganho relativo expectativa de anos de estudo 2010 para 2000, Região, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo % da pop com coleta de lixo 2010 para 2000, Ganho relativo % de 18 anos ou mais com ensino médio completo 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000

f. Preditores: (Constante), Ganho relativo expectativa de anos de estudo 2010 para 2000, Região, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo % da pop com coleta de lixo 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000

g. Preditores: (Constante), Ganho relativo expectativa de anos de estudo 2010 para 2000, Região, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000

h. Preditores: (Constante), Ganho relativo expectativa de anos de estudo 2010 para 2000, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000

Tabela 3: Resumo Modelo Completo 2000-2010

**Resumo do modelo<sup>h</sup>**

Modelo	R	R quadrado	R quadrado ajustado	Erro padrão da estimativa
1	,982 <sup>a</sup>	,964	,964	,004434934
2	,982 <sup>b</sup>	,964	,964	,004432682
3	,982 <sup>c</sup>	,964	,964	,004430414
4	,982 <sup>d</sup>	,964	,964	,004428295
5	,982 <sup>e</sup>	,964	,964	,004426812
6	,982 <sup>f</sup>	,964	,964	,004425932
7	,982 <sup>g</sup>	,964	,964	,004429194

- a. Preditores: (Constante), Ganho relativo taxa de fecundidade total 2010 para 2000, Ganho relativo expectativa de anos de estudo 2010 para 2000, Região, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo % da pop com coleta de lixo 2010 para 2000, Ganho relativo renda per capita 2010 para 2000, Ganho relativo % da pop com energia elétrica 2010 para 2000, Ganho relativo % de 18 anos ou mais com ensino médio completo 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000
- b. Preditores: (Constante), Ganho relativo taxa de fecundidade total 2010 para 2000, Ganho relativo expectativa de anos de estudo 2010 para 2000, Região, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo % da pop com coleta de lixo 2010 para 2000, Ganho relativo % da pop com energia elétrica 2010 para 2000, Ganho relativo % de 18 anos ou mais com ensino médio completo 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000
- c. Preditores: (Constante), Ganho relativo expectativa de anos de estudo 2010 para 2000, Região, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo % da pop com coleta de lixo 2010 para 2000, Ganho relativo % da pop com energia elétrica 2010 para 2000, Ganho relativo % de 18 anos ou mais com ensino médio completo 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000
- d. Preditores: (Constante), Ganho relativo expectativa de anos de estudo 2010 para 2000, Região, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo % da pop com coleta de lixo 2010 para 2000, Ganho relativo % de 18 anos ou mais com ensino médio completo 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000
- e. Preditores: (Constante), Ganho relativo expectativa de anos de estudo 2010 para 2000, Região, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo % da pop com coleta de lixo 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000
- f. Preditores: (Constante), Ganho relativo expectativa de anos de estudo 2010 para 2000, Região, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000
- g. Preditores: (Constante), Ganho relativo expectativa de anos de estudo 2010 para 2000, Ganho relativo Mortalidade Infantil 2010 para 2000, Ganho relativo na porcentagem de domicílio com água encanada 2010 para 2000
- h. Variável dependente: Ganho relativo esperança 2010 para 2000