

UNIVERSIDADE FEDERAL DE ALFENAS

RAFAELA DA SILVA GOMES

**OBESIDADE E ROTULAGEM DE ALIMENTOS COM A UTILIZAÇÃO DE
REGRESSÕES QUANTÍLICAS: UM ESTUDO DA MACRORREGIÃO SUL DO
ESTADO DE MINAS GERAIS**

ALFENAS/MG

2022

RAFAELA DA SILVA GOMES

**OBESIDADE E ROTULAGEM DE ALIMENTOS COM A UTILIZAÇÃO DE
REGRESSÕES QUANTÍLICAS: UM ESTUDO DA MACRORREGIÃO SUL DO
ESTADO DE MINAS GERAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Estatística Aplicada a Biometria, da Universidade Federal de Alfenas – UNIFAL-MG, como requisito parcial para obtenção do título de Mestre em Estatística Aplicada a Biometria.

Orientador: Prof. Dr. Denismar Alves Nogueira
Coorientador: Prof. Dr. Eric Batista Ferreira

ALFENAS/MG

2022

Sistema de Bibliotecas da Universidade Federal de Alfenas
Biblioteca Central

Gomes, Rafaela da Silva.

Obesidade e rotulagem de alimentos com a utilização de regressões
quantílicas: um estudo da macrorregião sul do estado de Minas Gerais /
Rafaela da Silva Gomes. - Alfenas, MG, 2022.

66 f. : il. -

Orientador(a): Denismas Alves Nogueira.

Dissertação (Mestrado em Estatística Aplicada e Biometria) -
Universidade Federal de Alfenas, Alfenas, MG, 2022.

Bibliografia.

1. IMC.. 2. Software R. 3. Políticas. 4. Regulamentações. I. Nogueira,
Denismas Alves, orient. II. Título.

RAFAELA DA SILVA GOMES

**OBESIDADE E ROTULAGEM DE ALIMENTOS COM A UTILIZAÇÃO DE
REGRESSÕES QUANTÍLICAS: UM ESTUDO DA MACRORREGIÃO SUL DO
ESTADO DE MINAS GERAIS**

A Banca examinadora abaixo-assinada, aprova a
Dissertação apresentada como parte dos requisitos
para obtenção do título de Mestre pela Universidade
Federal de Alfenas. Área de concentração:
Estatística Aplicada e Biometria

Aprovada em: 25 de nov. de 2022

Prof. Dr. Denismar Alves Nogueira

Instituição: Universidade Federal de Alfenas - UNIFAL-MG

Profa. Dra. Adriana Dias

Instituição: Universidade Federal de Alfenas - UNIFAL-MG

Profa. Dra. Daniela Braga Lima

Instituição: Universidade Federal de Alfenas - UNIFAL-MG



Documento assinado eletronicamente por **Daniela Braga Lima, Professor do Magistério Superior**, em 25/11/2022, às 17:42, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do **Decreto nº 8.539, de 8 de outubro de 2015**.



Documento assinado eletronicamente por **Adriana Dias, Professor do Magistério Superior**, em 25/11/2022, às 17:43, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.



Documento assinado eletronicamente por **Denismar Alves Nogueira, Professor do Magistério Superior**, em 25/11/2022, às 17:44, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do **Decreto nº 8.539, de 8 de outubro de 2015**.



A autenticidade deste documento pode ser conferida no site https://sei.unifal-mg.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0874743** e o código CRC **0819DC88**.

AGRADECIMENTOS

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Agradecimento ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), pela bolsa de estudos e auxílio financeiro que possibilitou a dedicação integral ao programa de pós-graduação e a operacionalização do estudo

RESUMO

Atualmente, a obesidade vem sendo um problema grave na saúde pública constituindo até uma epidemia mundial responsável por diversas mortalidades. O aumento do índice de massa corporal, pode levar a problemas graves de saúde, como doenças cardiovasculares, metabólicas, neoplásicas, ortopédicas, entre outras. Por ser causada por fatores como genética, meio social no qual o indivíduo é inserido, fator comportamental e metabólico, pode impactar o bem-estar, o psicológico e a qualidade de vida. Sendo assim, esta pesquisa tem por objetivo analisar o impacto das novas regulamentações, que são as medidas regulatórias RDC 359/2003, RDC 360/2003 e RDC 24/2010, relacionando com a prevalência da obesidade na população adulta e entender o impacto da rotulagem de alimentos por meio do modelo de regressão quantílica. A pesquisa é de natureza quantitativa, utilizou questionário estruturado de uma amostragem da Região Sul de Minas Gerais, na qual se pretendeu entender as relações entre aquisição e utilização de produtos alimentícios. Para se almejar a relação, foi aplicado estudos empíricos utilizando modelos de regressão quantílica, tal análise foi feita usando *software R*. Os resultados indicam que a leitura dos rótulos dos ingredientes tem uma associação negativa com o IMC, principalmente em quartis de IMC mais altos (sobrepeso e obesidade). Indivíduos com maior risco de sobrepeso e obesidade, como pessoas de baixa renda, com baixa escolaridade ou que não praticam esportes, parecem obter o maior benefício da leitura dos ingredientes no rótulo dos alimentos. O estudo feito conclui que é necessário continuar com incentivos governamental na parte de rótulos nutricionais, pois eles ajudam a contribuir contra a obesidade, outro resultado obtido na pesquisa foi que conforme a idade vai avançando a um risco maior na obesidade além da falta de escolaridade portanto é necessário que haja políticas públicas que incentivem os estudo e projetos voltados para pessoas idosas na busca da educação alimentar consciente. No entanto, as descobertas sobre a eficácia do uso de rótulos para diminuir os resultados negativos para a saúde, como sobrepeso e obesidade, são inconclusivas. Como a maioria dos estudos se concentra em estimar os efeitos médios dos rótulos, pouco se sabe se os rótulos têm algum efeito nas caudas da distribuição do peso corporal, que, dado o gradiente social da obesidade, geralmente inclui grupos de risco. Além disso, o uso de rótulos, sobrepeso e obesidade variam entre os subgrupos populacionais

Palavras-chave: IMC, *software R*, Políticas, Regulamentações.

ABSTRACT

Currently, obesity has been a serious public health problem, constituting a worldwide epidemic responsible for several mortalities. The increase in body mass index can lead to serious health problems, such as cardiovascular, metabolic, neoplastic, orthopedic diseases, among others. Because it is caused by factors such as genetics, the social environment in which the individual is inserted, behavioral and metabolic factors, it can impact well-being, psychological and quality of life. Therefore, this research aims to analyze the impact of new regulations, which are regulatory measures RDC 359/2003, RDC 360/2003 and RDC 24/2010, relating to the prevalence of obesity in the adult population and to understand the impact of labeling of food through the quantile regression model. The research is of a quantitative nature, using a structured questionnaire from a sample of the South Region of Minas Gerais, in which it was intended to understand the relationships between acquisition and use of food products. To target the relationship, empirical studies were applied using quantile regression models, such analysis was performed using R software. The results indicate that reading ingredient labels has a negative association with BMI, especially in higher BMI quartiles (overweight and obesity). Individuals with a higher risk of overweight and obesity, such as people with low income, with low education or who do not practice sports, seem to obtain the greatest benefit from reading the ingredients on food labels. The study concluded that it is necessary to continue with government incentives in the part of nutritional labels, as they help to contribute against obesity, another result obtained in the research was that as age advances, there is a greater risk of obesity in addition to the lack of schooling, therefore it is necessary to have public policies that encourage studies and projects aimed at the elderly in the pursuit of conscious food education. However, findings on the effectiveness of using labels to decrease negative health outcomes such as overweight and obesity are mixed. As most studies focus on estimating the average effects of labels, little is known whether labels have any effect on the tails of the body weight distribution, which, given the social gradient of obesity, often includes at-risk groups. Furthermore, the use of labels, overweight and obesity vary across population subgroups.

Keywords: BMI, R software, Policies, Regulations.

LISTA DE FIGURAS

Figura 1- Modelo de rótulo Frontal que será aplicado no Brasil.....	18
Figura 2- Ajustes do modelo de regressão quantílica $Q_{\tau}Y_t X_t = \beta_0\tau + \beta_1\tau X_t$. Assumindo $\beta_0 = 0.1$, $\beta_1 = 0.9$ e $\tau = 0.1, 0.2, \dots, 0.9$	27
Figura 3 - Gráfico de Box Plot dos valores da variável IMC com representação dos possíveis valores outliers segundo a proposta de Tukey.	43
Figura 4 - Gráficos de dispersão e box plot da variável IMC em função das outras variáveis incluídas no estudo.....	45
Figura 5 - Estimativas (linhas contínuas) e intervalo de confiança de 90% (área hachurada) para os coeficientes de regressão considerando um conjunto denso de quantis, $\tau = 0.05, 0.10, \dots, 0.90$	51
Figura 6 - Medida da bondade de ajuste, $R_1(\tau)$, para o modelo de regressão quantílica	53
Figura 7 - Gráfico de envelope simulado para modelos de regressão quantílica considerando os resíduos quantílicos e os quantis condicionais 0.15, 0.25, 0.50, 0.75, 0.85 e 0.90 de IMC.	54

LISTA DE TABELAS

Tabela 1– Estatísticas descritivas essenciais para a regressão quantílica das variáveis quantitativas em estudo.	41
Tabela 2 – Estimativas do estudo e classificação por cores de acordo com o nível de significância dos coeficientes de correlação Spearman das variáveis.....	44
Tabela 3- Estimativas para o modelo de regressão quantílica condicional, do MMQO e respectivas significâncias.....	47
Tabela 4– Medida de ajuste e Valor-p referente ao teste da falta de ajuste.....	53

SUMÁRIO

1	INTRODUÇÃO	10
2	REFERENCIAL TEÓRICO	12
2.1	OBESIDADE E SAÚDE PÚBLICA.....	12
2.2	ROTULAGEM.....	15
2.3	PUBLICIDADE.....	19
2.4	ESTUDOS EMPÍRICOS ENVOLVENDO REVISÕES EPIDEMIOLÓGICAS	21
2.5	REGRESSÃO QUANTÍLICA	22
2.5.1	Teste da Falta de Ajuste em Modelos de Regressão Quantílica	28
2.5.2	Teste de Hipótese Linear Geral	30
3	MATERIAL E MÉTODOS	34
3.1	TIPOLOGIA DA PESQUISA.....	34
3.2	CENÁRIO DO ESTUDO.....	34
3.3	POPULAÇÃO E AMOSTRA	34
3.4	PROCEDIMENTO PARA OBTENÇÃO DOS DADOS	35
3.5	PROCEDIMENTO DE ANÁLISE DE DADOS	35
3.6	ABORDAGEM DA REGRESSÃO QUANTÍLICA	38
4	RESULTADOS	41
4.1	DESCRIÇÃO DOS DADOS.....	41
4.2	ESPECIFICAÇÕES DO MODELO QUANTÍLICO.....	46
5	CONSIDERAÇÕES FINAIS.....	56
	REFERÊNCIAS.....	57
	APÊNDICE I.....	65

1 INTRODUÇÃO

Vem ocorrendo ultimamente mudanças econômicas, sociais e demográficas que estão cada vez mais afetando o cotidiano da vida das pessoas e com isso o seu perfil de consumo como também as práticas alimentares em diferentes contextos geopolíticos (MARTINELLI; CAVALLI, 2019). A população brasileira vem passando por um quadro alarmante de crescimento da obesidade e doenças associadas, por causa dos reflexos no processo saúde-doença e no estado nutricional (IBGE, 2010).

Em apenas doze anos, a prevalência de obesidade passou de 11,8% em 2006 para 19,8% em 2018, atingindo quase um em cada cinco brasileiros de acordo com dados de Brasil (2019). O que é algo bastante preocupante em um cenário nacional, pois isto pode causar gastos ao sistema de saúde tanto público como privado. Esse cenário de crescimento da prevalência da obesidade pode afetar diretamente a saúde pública, que sente o reflexo diário de milhões de pessoas com enfermidades e doenças graves (SANTOS *et al.*, 2019).

Para o combate a obesidade foram criadas normas para a rotulagem nutricional que foram publicadas em 2003 e 2010 que são as medidas regulatórias RDC 359/2003, RDC 360/2003 e RDC 24/2010; e as discussões de saúde e políticas públicas avançaram muito nestes dez anos, especialmente frente ao crescimento global da obesidade e de doenças crônicas não transmissíveis (DCNTs).

No Brasil, segundo a pesquisa “A Mesa dos Brasileiros: Transformações, Confirmações e Contradições”, de abrangência nacional, realizada pela FIESP/CIESP em 2017 com 3.000 consumidores, constatou-se que 53% dos brasileiros lê os rótulos dos alimentos e bebidas, e apenas 23% lê com regularidade.

A ANVISA em 2005 juntamente com o Governo Federal, desencadeou um processo para elaborar uma proposta de regulamentação da publicidade de alimentos não saudáveis e de bebidas de baixo valor nutricional. A proposta foi publicada em 2006 por meio de uma Consulta Pública (CP no 71/06) para que a população pudesse dar sugestões e críticas (ANVISA, 2006). O documento foi amplamente discutido e modificado, sendo aprovado somente quatro anos após a sua publicação, pela RDC 24/2010.

A regressão Quantílica é usada neste estudo, pois modela a relação entre um conjunto de variáveis preditoras (explicativas) e percentis específicos (ou "quantis") de uma variável alvo (dependente), na maioria das vezes a mediana. O método proposto consiste em executar uma

regressão da função de influência (recentrada) do quantil sobre as variáveis explicativas. A função de influência, uma ferramenta amplamente utilizada em estimação robusta, é facilmente calculada para quantis, bem como para outras estatísticas de distribuição. Nessa abordagem, portanto, pode ser prontamente generalizada para outras estatísticas de distribuição.

A principal vantagem desta metodologia de regressão quantílica é que o método permite o entendimento de relações entre variáveis fora da média dos dados, tornando-o útil para a compreensão de resultados que não são normalmente distribuídos e que têm relacionamentos não lineares com variáveis preditoras.

De acordo com que foi apresentado, o objetivo do trabalho foi analisar o impacto das regulamentações, que são as medidas regulatórias RDC 359/2003, RDC 360/2003 e RDC 24/2010, relacionando com a prevalência da obesidade na população adulta e entender o impacto da rotulagem de alimentos por meio do modelo de regressão quantílica. A região trabalhada foi o sul de Minas Gerais, caracterizando um total de 155 cidades, participando desta pesquisa, vários perfis de pessoas e classes econômicas.

Os objetivos secundários foram utilizar análise descritiva para extrair informações a respeito das variáveis abordadas neste estudo, descrever importantes questões relacionadas nas medidas regulatórias (RDCs), descrever também o modelo de regressão Quantílica, identificar quais covariáveis estão relacionadas com o aumento do IMC a partir da modelagem dos dados.

2 REFERENCIAL TEÓRICO

2.1 OBESIDADE E SAÚDE PÚBLICA

A obesidade pode ser considerada de uma forma mais simples, como uma doença que é caracterizada pelo excesso de gordura corporal, sendo acarretada por causa do balanço energético positivo o que por consequência pode trazer agravos a saúde, além da falta de qualidade de vida como também a diminuição do tempo de vida (AGHA; AGHA, 2017).

A obesidade tornou-se estigmatizada e a discriminação apresenta desafios significativos na educação, saúde e emprego. Há evidências fracas de que o QI de crianças com sobrepeso é inferior ao das crianças com peso normal (KEERY *et al.*, 2015). Isso resulta em estresse, sofrimento psicológico, depressão, isolamento social, baixa autoestima e má imagem corporal. Portanto, devido ao aumento da prevalência de obesidade, o desemprego pode aumentar e indivíduos obesos devidamente qualificados podem permanecer desempregados por discriminação ou medo dela no ambiente de trabalho (NORTHSTONE *et al.*, 2012)

A obesidade pode ser um fator importante para o desenvolvimento de doenças graves como hipertensão, diabetes do tipo 2, doenças cardiovasculares, doenças reumatológicas, entre outras (VILLAREAL *et al.*, 2015).

São muitas as causas da obesidade. O excesso de peso pode estar ligado ao patrimônio genético da pessoa, a maus hábitos alimentares ou, por exemplo, às disfunções endócrinas, que advém de questões familiares, fatores sociais onde a comida se torna um fator importante em certas situações de convivência, além da oferta excessiva de comida, além dos fatores comunitários associados à vida sedentária, portanto todos esses fatores contribuem para um ambiente que favorece a obesidade (BRANDÃO; SOARES, 2018).

Segundo Brasil (2019), em apenas treze anos, a prevalência de obesidade passou de 11,8% em 2006 para 19,8% em 2018, atingindo quase um em cada cinco brasileiros. De acordo com a pesquisa da Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (Vigitel), de 2018, esse aumento de 67,8% nos últimos treze anos é um sinal negativo em relação à saúde dos brasileiros, os dados apontaram que o crescimento da obesidade em porcentagem foi maior entre as faixas etárias 25 a 34 anos com 84,2% de aumento de obesos, e dos 35 a 44 anos com 81,1%. A Vigitel (2018) constatou que 55,7% da população brasileira tem excesso de peso, um crescimento de 30,8% bastante preocupante comparando a 2006 que apresentava 42,6% de brasileiros com excesso de peso.

A obesidade pode ser algo preocupante para saúde pública, por levar a alguns tipos de doenças que podem ser letais em alguns casos e também ocasionar cerca de 13 tipos de câncer, o que leva a ser uma das principais causas de morte no Brasil. Uma possível melhora desse quadro seria a aplicação de uma alimentação saudável, exercícios físicos regulares de no mínimo 150 minutos por semana e sempre observar o peso corporal, o que poderia diminuir os casos mais comuns de câncer em aproximadamente 33% no Brasil (INCA, 2017). Conforme Organização Mundial da Saúde (2018), aproximadamente 13% dos casos de câncer no Brasil, são causados por sobrepeso ou obesidade.

A partir da análise desse quadro de obesidade na qual acarreta as doenças crônicas não transmissíveis (DNCT), grande parte da população mundial vivencia uma situação em que se morre mais por excesso de peso do que por problemas ligados à subnutrição e até questões relacionadas a violências urbanas, recorrentes em grande parte de países subdesenvolvidos (BRANDÃO; SOARES, 2018). Uma pesquisa realizada em 20 países da América Latina pela Organização das Nações Unidas para Alimentação e Agricultura (ONU/FAO) mostrou que nestes países o excesso de peso e a obesidade são responsáveis por 300 mil mortes por ano nessas nações, em vista que por assassinato são 116 mil mortes (BERDEGUÉ; AGUIRRE, 2018).

Uma das explicações diante do crescimento acelerado da obesidade nas populações vem ocorrendo por causa da modernização das sociedades, o que mais vem influenciando são a oferta de alimentos juntamente com a melhoria das formas de emprego, concomitante à mecanização e à automação das atividades como o *home office*. Esse modo de viver das pessoas está cada vez mais sendo alterado, o que ocasiona a falta de gastos calóricos devido à economia de gastos energéticos associados às atividades diárias, agregados a maior oferta de alimentos. Por essas justificativas, a obesidade vem sendo denominada “doença da civilização” ou “síndrome do Novo Mundo” (MARTINS, 2018).

Esse cenário de prevalência da obesidade que vem crescendo ao longo dos anos está afetando diretamente a saúde pública o que ocasiona milhões de pessoas apresentando enfermidades e doenças graves devido a isto e falta de qualidade em sua alimentação, ainda não existem políticas direcionadas diretamente a este cenário. Embora haja plano de ações e programas que incentivam o emagrecimento saudável, hábitos alimentares saudáveis e o controle da obesidade, ainda é notório o aumento de pessoas obesas (MARTINS, 2018).

Existem cidades no Brasil que se preocupam com este cenário alarmante, por exemplo a cidade do Rio de Janeiro que tem um Estatuto dos Portadores de Obesidade, criado em 2014, possui algumas diretrizes para auxiliar no controle e combater o avanço da obesidade, além de

manter a liberdade, o respeito e a dignidade, preservando sua integridade psíquica e moral da pessoa obesa (RIO DE JANEIRO (ESTADO), 2014). Ainda está em análise pelo Congresso Nacional, um projeto de lei federal (PL 4.328/2016) para estabelecer uma política nacional para o combate a obesidade (BRASIL, 2016).

A falta de políticas públicas pode ocasionar diversos problemas como o fato ocorrido no Acre em que pessoas obesas tiveram atendimento negado em hospitais e postos de saúde por falta de estrutura mínima e equipamentos apropriados para receber pessoas acima do peso (MOTA, 2012). Além de casos de dentistas não possuírem cadeiras para suportar o peso elevado de pessoas obesas, também existem hospitais que não possuem equipamentos de tomografia ou radiografia que suportam além de certo limite de peso. Logicamente que isso permite que muitos pacientes se sintam discriminados.

Os gastos gerados para o Sistema Único de Saúde (SUS) em relação à obesidade podem ser muito altos devido à problemática da estrutura que se gasta em hospitais, sem contar os gastos com medicamentos para tratar doenças ocasionadas pela obesidade. Quando se faz necessário buscar por um tratamento como a cirurgia bariátrica no SUS a espera na fila chega a ser de quatro a seis anos, podendo variar por região. O Paraná é um dos estados com menor tempo de espera, responsável por 47% das cirurgias realizadas pelo SUS (ROSSI; REIS; FREITAS, 2019). Esse cenário é bastante preocupante, pois os custos gerados e a falta de recursos fazem com que pacientes com obesidade mórbida acabem morrendo antes de conseguirem tratamento eficaz.

Por isso, se faz necessário uma análise mais aprofundada que possa auxiliar políticas públicas para o controle da prevalência da obesidade, pois é fundamental e necessário a prevenção e controle, o que levaria a uma economia em recursos financeiros destinados ao tratamento propriamente da doença. A economia de recursos e melhoria da alimentação dos brasileiros fornecerá à saúde pública uma melhora significativa e assim trazer ações efetivas no âmbito da saúde como um todo. Portanto, incentivar as pessoas a manterem hábitos saudáveis como exercícios físicos e uma dieta balanceada podem ajudar (FLORIDO *et al.*, 2019).

A ação concomitante entre governo e sociedade, pode mudar as estatísticas e criar um cenário mais saudável e otimista. É necessário que as pessoas tenham consciência para querer se ajudar e que haja uma saída desse cenário, possibilitando que todos possam trabalhar conjuntamente para combater a obesidade (FLORIDO *et al.*, 2019).

Uma das ações que o governo pode auxiliar no combate a obesidade seria continuar ajudando na melhoria da rotulagem nutricional que é uma medida fundamental para favorecer

escolhas mais saudáveis e conseqüentemente para o enfrentamento do sobrepeso e obesidade (FLORIDO *et al.*, 2019).

2.2 ROTULAGEM

O rótulo nutricional é um dispositivo de comunicação nutricional em nível populacional que fornece aos consumidores no ponto de compra toda a descrição das propriedades nutricionais de um alimento destinada a informar na qual o objetivo dos rótulos é compreender a declaração de valor energético e os principais nutrientes. Assim trazendo informações necessárias para cumprir as regulamentações da rotulagem nutricional (SILVA *et al.*, 2020)

De acordo com a Agência Nacional de Vigilância Sanitária – Anvisa (2018), a rotulagem nutricional é a descrição que informa ao consumidor sobre as propriedades nutricionais de um alimento e que especifica o valor energético e a quantidade de carboidratos, proteínas, gorduras totais, gorduras trans, gorduras saturadas, fibra alimentar e sódio. Outro ponto de atenção é a presença de lactose ou outros ingredientes que podem causar alergia.

Para o desenvolvimento do corpo são necessários nutrientes, além de atividades físicas, em relação a isso uma dieta rica e nutritiva e o uso eficiente dos rótulos nutricionais estará associada positivamente a qualidade de vida de uma pessoa, o que levará ela longe do caminho da obesidade (BLISTEIN; EVANS, 2016).

Com a regulamentação da rotulagem de alimentos, a Anvisa (2018) garante que o consumidor tenha acesso a informações que são essenciais para sua proteção e a promoção de sua saúde. A evolução da legislação é um fator que faz a população cada vez mais buscar uma alimentação mais saudável.

Nos últimos 20 anos as medidas regulatórias relacionadas à transmissão de informações nutricionais na rotulagem dos alimentos vêm sofrendo significativas mudanças (ANVISA, 2018).

As normas atuais de rotulagem nutricional da Anvisa foram publicadas em 2003 e 2010 que são as medidas regulatórias RDC 359/2003, RDC 360/2003 e RDC 24/2010, e as discussões de saúde e políticas públicas avançaram muito nestes dez anos, especialmente frente ao crescimento global da obesidade e de doenças crônicas não transmissíveis (DCNTs). Segundo a Organização Mundial da Saúde (OMS) em 2018, um em cada oito adultos, em todo o planeta, é obeso. A projeção para 2025 é de que cerca de 2,3 bilhões de indivíduos estejam com excesso de peso, sendo mais de 700 milhões com obesidade (OMS, 2018).

A relação que existe entre rotulagem nutricional com dieta adequada e saudável: papel informativo que dá ao consumidor visibilidade do que e quanto está consumindo. O rótulo por si só não endereça a questão da obesidade e DCNTs, que tem causas multifatoriais, e pode ter relação com fatores genéticos, estilo de vida, prática de atividade física, alimentação, entre outros.

No Brasil, segundo a pesquisa “A Mesa dos Brasileiros: Transformações, Confirmações e Contradições”, de abrangência nacional, realizada pela FIESP/CIESP em 2017 com 3.000 consumidores, foi constatado que 53% dos brasileiros lê os rótulos dos alimentos e bebidas, e 23% lê com regularidade. Também, o estudo mostra que as informações mais lidas são a data de validade (70%), seguida de sódio (19%), gorduras (14%), tabela nutricional (12%) e calorias (11%). O que foi destacado na pesquisa é que em torno de 64% dos entrevistados afirmaram que as informações dos rótulos são difíceis de entender. O que se pode concluir com esses números é que para a rotulagem nutricional exercer sua função de ferramenta para uma alimentação saudável e balanceada existem duas frentes de ação em políticas públicas: seria a questão da educação e uma revisão da norma de rotulagem nutricional.

Em relação à educação, a Educação Alimentar e Nutricional (EAN) é primordial para que o consumidor adquira conhecimento sobre alimentação e saúde, compreenda a rotulagem nutricional e esteja apto para uso da informação em suas escolhas alimentares com liberdade e responsabilidade (FNDE, 2020). A população precisa não só de conscientização sobre a importância do hábito de leitura dos rótulos dos alimentos, como também conteúdo que a capacite para interpretar e utilizar a informação nutricional. No Brasil, a Lei nº 13.666/2018 incluiu a EAN nos currículos escolares do ensino fundamental e médio de escolas públicas e privadas (BRASIL, 2018).

Em relação à revisão regulatória, esta iniciativa tem como papel no complemento do processo de educação, uma vez que os consumidores brasileiros relatam dificuldade no entendimento das informações dispostas nos rótulos no padrão atual.

Existem causas multifatoriais, e não somente a alimentação em relação a sobrepeso, obesidade e DCNTs, por exemplo, a prática de atividade é reconhecida pela OMS (2018) como estratégia importante para prevenção e tratamento, o combate efetivo deve abranger todas as possíveis causas do problema.

Enfim, a epidemia mundial de obesidade e DCNTs é reconhecidamente uma pauta urgente em saúde pública, a qual exige diversos esforços para sua solução. A discussão de rotulagem nutricional tem sim sua relevância nas discussões de políticas públicas, contudo, tão

somente como instrumento para que o consumidor faça suas escolhas alimentares de modo simples e informativo.

A Anvisa (2020) aprovou a nova norma sobre rotulagem nutricional de alimentos embalados. Essa nova medida estabelece mudanças na tabela de informação nutricional e nas alegações nutricionais, bem como inova ao adotar a rotulagem nutricional frontal e visa auxiliar o consumidor a realizar escolhas alimentares mais conscientes. A rotulagem nutricional frontal é um símbolo informativo na parte da frente do produto. A ideia é esclarecer o consumidor, de forma clara e simples, sobre o alto conteúdo de nutrientes que têm relevância para a saúde. Para tal, foi desenvolvido um design de lupa para identificar o alto teor de três nutrientes: açúcares adicionados, gorduras saturadas e sódio. O símbolo deverá ser aplicado na frente do produto, na parte superior, por ser uma área facilmente capturada pelo olhar do consumidor (ANVISA, 2020). A nova regra foi publicada no Diário Oficial da União (D.O.U.) no dia 9 de outubro de 2022, portanto entrando em vigor, por meio de uma Resolução da Diretoria Colegiada (RDC) e de uma Instrução Normativa (IN). Os produtos que se encontrarem no mercado na data da entrada da norma em vigor terão, ainda, um prazo de adequação de 12 meses.

O rótulo poderá ajudar concretamente a proteger e a melhorar a saúde, se as escolhas forem apoiadas por algum conhecimento básico de nutrição saudável, com base em uma dieta equilibrada e variada. Em uma perspectiva mais ampla, isso pode se traduzir em uma redução da obesidade e na incidência de doenças crônicas - intimamente relacionadas a hábitos alimentares não saudáveis e que tem uma ligação com altos custos com sistemas de saúde (ANVISA, 2020). Na Figura 1 e mostrado os modelos de como ficaria o rótulo frontal no Brasil.

Figura 1– Modelo de rótulo Frontal que será aplicado no Brasil

a) Modelos com alto teor de um nutriente



b) Modelos com alto teor de dois nutrientes



c) Modelos com alto teor de três nutrientes



Fonte: Anvisa (2020)

Como foi notado na Figura 1 o modelo possui um *design* de lupa para poder identificar melhor teores elevados de nutrientes que prejudiquem a saúde do consumidor. A ideia é esclarecer o consumidor de forma clara e simples, com isso a pessoa poderá comparar os alimentos mais rapidamente e ter uma escolha mais eficiente. Entretanto, a adoção de um modelo de rotulagem nutricional frontal não resolverá as questões de saúde, uma vez que a alimentação é apenas uma dentre as diversas causas do problema (BANDEIRA *et al.*, 2021).

Outra questão a ser trabalhada além do rótulo, seria a forma como estes alimentos estão sendo divulgados, ou seja, a publicidade também tem um papel fundamental e muitas pessoas podem ser influenciadas pelos anúncios e propagandas de alimentos não saudáveis, principalmente nos jovens que ficam mais expostos a este tipo de conteúdo.

2.3 PUBLICIDADE

As grandes mídias veiculam ou produzem notícias, representações e expectativas nos indivíduos com propagandas, informações e noticiário, em que, de um lado estimulam o uso de produtos saudáveis e práticas alimentares para emagrecimento e, de outro, instigam ao consumo de lanches tipo *fast food*. Além das redes sociais, que atualmente estão tendo grande influência, de certa maneira, no consumo dos brasileiros (BAIRD, 2016).

Portanto, trata-se de uma decisão de ação das empresas publicitárias, ou até mesmo *influencers digitais*, que integram um contexto empresarial e um sistema de crenças em que há uma estreita relação entre uma suposta verdade médica e um desejo social e individual. O corpo é um campo de luta que envolve diferentes saberes, práticas e imaginário social (BAIRD, 2016).

No Brasil, as ferramentas que estabelecem normas legais para a propaganda de alimentos começaram em meados da década de 1960, quando o Ministério da Saúde assumiu um papel diante da competência de regular e normatizar os textos e matérias desse tipo (BRASIL, 1969). Com a criação da Agência Nacional de Vigilância Sanitária (ANVISA), a análise de propagandas lhe foi transferida, consolidando o texto constitucional que estabelece a obrigação do Estado de proteger a saúde da população e evitar a exposição a propagandas comerciais de produtos e serviços que possam influenciar de forma inadequada as práticas em saúde (BRASIL, 1999).

Em 2005 a ANVISA juntamente com o Governo Federal, desencadeou um processo para elaborar uma proposta de regulamentação da publicidade de alimentos não saudáveis e de bebidas de baixo valor nutricional. A proposta foi publicada em 2006 através de uma Consulta Pública (CP no 71/06) para que a população pudesse dar sugestões e críticas (ANVISA, 2006). O documento foi amplamente discutido e modificado, sendo aprovado somente quatro anos após a sua publicação, pela RDC no 24/2010 que dispõem “[...]sobre a oferta, propaganda, publicidade, informação e outras práticas correlatas cujo objetivo seja a divulgação e a promoção comercial de alimentos considerados com quantidades elevadas de açúcar, de gordura saturada, de gordura trans, de sódio [...]”. (ANVISA, p. 1, 2010).

Ao propor regular a publicidade de alimentos, a Anvisa impactaria ao mesmo tempo dois setores poderosos na economia brasileira: a indústria de alimentos e a indústria de publicidade. No caso desta última, vale destacar que diversos setores seriam afetados, na medida em que a indústria da publicidade está assentada no tripé anunciantes, agências de publicidade e veículos de comunicação. Ao longo da disputa política, esse setor esteve

representado, principalmente, pelo Conselho Nacional de Autorregulamentação Publicitária (CONAR) (BAIRD, 2016).

O outro setor afetado, os anunciantes dos produtos, é a indústria de alimentos, cujo principal representante é a Associação Brasileira das Indústrias da Alimentação (ABIA). Trata-se de setor de destaque nacional, pois movimenta um décimo da economia, respondendo pelo emprego de mais de 1,5 milhão de pessoas. Tais números já são suficientes para mostrar que qualquer governo seria extremamente cauteloso ao impor medidas que pudessem afetar o setor. Para entender melhor como está regulamentação da RDC 24/2010 teria impacto a indústria de publicidade em 2008 faturava R\$21,4 bilhões o que na época representava 4% do PIB do país e número de pessoas que eles empregavam era de 711 mil pessoas (ABAP, 2008). E a indústria tinha um faturamento em 2010 de R\$ 330 bilhões o que representava naquele período 9 % do PIB do país e empregando um total de aproximadamente 1,5 milhões de pessoas (GRUPO DE MÍDIA SÃO PAULO, 2011).

Pode-se observar nos dias atuais que a principal fonte de lazer são atividades que envolvem a atenção e o raciocínio, ou seja, os jogos eletrônicos, *smartphones*, televisão entre outras tecnologias, vem sendo priorizadas pelas pessoas o que ocasiona o detrimento de atividades que tem seus efetivos gastos calóricos em que de fato haja uma movimentação corporal. Com esses avanços tecnológicos as pessoas estão cada vez mais expostas a conteúdos publicitários, o que pode de certa maneira, influenciar no consumo de certos produtos, portanto o consumo pode aumentar e a propensão a obesidade também. Além disso, a obesidade é uma epidemia global, responsável pela diminuição da expectativa de vida e pelo aumento da mortalidade por doenças cardiovasculares (MENDONÇA, 2016).

Na divulgação de propagandas de produtos alimentares de forma atrativa e que os associa a estilos de vida desejados com vinhetas de fácil memorização, aliada aos componentes motivacionais, às necessidades emocionais/psicológicas e aos desejos individuais fazem com que haja uma influência dos meios de comunicação sobre o estilo de consumo dos brasileiros (LEDO; CHAUD; ABREU, 2019).

Em alguns casos as pessoas com baixa escolaridade pode ser induzidas pelas ideologias capitalistas ao ampliarem o seu poder de escolha e decisão, pois o consumo de alimentos não nutritivos é fruto de hábitos e costumes influenciados e reforçados por informações que subsidiam a tomada de decisão, a influência da mídia sobre a consolidação de uma orientação alimentar e nutricional da população mostra-se como temática relevante para investigação (CAIVANO *et al.*, 2017).

Mas o que é necessário ressaltar é que muitos adolescentes e crianças dos anos de 2000 a 2010 sofreram influências publicitárias, e que neste período a alimentação saudável não era tão utilizada, pois o que chamava a atenção eram comidas açucaradas e gordurosas, o que poderiam ter levados ao sobrepeso na vida adulta nos dias atuais.

2.4 ESTUDOS EMPÍRICOS ENVOLVENDO REVISÕES EPIDEMIOLÓGICAS

Para entender melhor todos os assuntos trazidos até esta seção, será importante um estudo empírico em que traga um pouco da realidade e acompanhe melhor todos os assuntos, portanto será apresentado alguns estudos e metodologias de regressões envolvendo revisões epidemiológicas.

Na literatura existe uma ampla quantidade de métodos estatísticos que podem ser utilizados para modelar dados. Contudo, durante a revisão bibliográfica foi visto que os modelos de Regressão Quantílica têm bastante destaque nos estudos epidemiológicos em relação aos outros métodos.

Antiporta *et al.* (2015) buscando determinar a associação entre obesidade e tempo de residência em áreas urbanas, no Peru, verificaram por meio da análise multivariada que para cada 10 anos de residência nas áreas urbanas os imigrantes rurais-urbanos tinham em média 12% a mais de prevalência da obesidade.

Embora vários estudos tenham investigado a relação entre rótulos e qualidade da dieta como, por exemplo, na revisão sistemática de Campos *et al.* (2011) e Drichoutis *et al.* (2011), estudos com foco na influência dos rótulos nutricionais na obesidade são escassos principalmente que usam regressão, alguns dos poucos estudos encontrados foram Variyam e Cawley feito em 2006, Drichoutis *et al.* em 2009, Loureiro *et al.* em 2012 e por fim, Banterle e Cavaliere no ano 2014. Esses estudos se concentraram em medir o efeito médio da leitura do rótulo nutricional no Índice de Massa Corporal (IMC) dos consumidores e em mitigar o viés de auto seleção para produzir estimativas imparciais do efeito da leitura dos rótulos no IMC. Medindo os efeitos médios, eles encontram pouco ou nenhuma melhoria do peso corporal dos consumidores.

Os estudos existentes não abordaram a possibilidade de que a relação entre a leitura dos rótulos dos alimentos e o IMC possa ser não linear e que a leitura dos rótulos possa afetar os consumidores com IMC maior, (IMC entre 25,0 e 29,9 Kg/m²: sobrepeso; IMC entre 30,0 e 34,9 Kg/m²: obesidade grau I; IMC entre 35,0 e 39,9 Kg/m²: obesidade grau II; IMC maior do que 40,0 Kg/m²: obesidade grau III) de forma diferente daqueles com menor IMC (ou seja, IMC

abaixo de 25Kg/m^2) (OSWALDO CRUZ, 2020). Além disso, alguns subgrupos da população podem ser mais afetados pelo uso de rótulos do que outros. A relação entre a leitura dos rótulos dos alimentos e os resultados de peso varia com o sexo (LOUREIRO *et al.*, 2012), mas nenhum outro efeito diferencial da leitura do rótulo entre os subgrupos populacionais foi investigado.

A revisão de literatura deste tema apresentou uma questão sociodemográfica tanto no uso de rótulos quanto na obesidade, levando a hipótese de que a relação entre leitura de rótulos de alimentos e IMC pode mudar dependendo das características sociodemográficas. Como os estados membros europeus estão atualmente debatendo se a união europeia deve revisar sua política de rotulagem nutricional (MICHAIL, 2016, 2017a, 2017b), uma compreensão completa da relação entre o uso da rotulagem nutricional e o IMC ajudaria a projetar um esquema de rotulagem mais eficaz.

Bonanno *et al.* (2018) utilizou a relação entre a leitura dos ingredientes nos rótulos dos alimentos e os níveis de IMC entre adultos italianos. Utilizando um estimador de regressão quantílica incondicional (RQI), consideraram diferentes subamostras de dados, destacando a existência de uma relação heterogênea entre leitura de rótulos de alimentos e IMC de adultos. Como descoberto pelos autores, os coeficientes que capturam a relação entre leitura de rótulos e IMC são maiores em níveis mais altos de IMC, os resultados apoiaram políticas para promover escolhas alimentares informadas, pois a leitura de rótulos pode ter um efeito mais acentuado de redução do IMC em indivíduos que possuem sobrepeso e obesidade. Em suas análises os autores tiveram uma série de limitações que pretendem superar em pesquisas futuras. Uma de suas limitações foi devido à falta de dados, então tiveram que usar um banco de dados transversal para medir a relação entre a leitura do rótulo nutricional e o IMC adulto, o que acabou limitando a capacidade de identificar adequadamente o efeito dos rótulos dos alimentos no IMC.

2.5 REGRESSÃO QUANTÍLICA

A regressão de quantis ou quantílica, introduzida por Koenker e Bassett (1978), permite calcular estimativas para quaisquer quantis da variável resposta, tais como a mediana podendo ser condicional e incondicional. A princípio o modelo de distribuição condicional da variável resposta, apresenta quantis que são expressos como funções de covariáveis observadas e já na abordagem do modelo incondicional é necessário integrar as variáveis de condicionamento para obter o efeito marginalizado.

A regressão quantílica não faz suposições sobre a distribuição dos resíduos e permite explorar diferentes aspectos da relação entre a variável dependente e as variáveis explicativas ou covariáveis (SOUZA, 2017).

A teoria clássica dos modelos começou com fundamento em teoria das esperanças condicionais. Um texto de Mosteller e Tukey (1977), que foi citado por Koenker e Basset (1982), informa o que a curva de regressão faz é dar um grande resumo das médias das distribuições correspondentes ao conjunto dos x 's de covariáveis. O que pode ser feito é calcular diversas curvas de regressões diferentes correspondendo aos vários pontos percentuais da distribuição e assim obter uma visão mais completa desse conjunto. Assim, percebe-se que a média dá uma visão incompleta de uma única distribuição, a curva de regressão também dá uma visão incompleta correspondente a um conjunto de distribuições, portanto, se a esperança condicional ou outra medida de tendência central condicional é capaz de caracterizar adequadamente uma relação estatística entre a variável resposta e as demais covariáveis em estudo.

Quando se utiliza a estimação de um modelo de Regressão linear clássica utilizando os métodos de mínimos quadrados ordinários (MMQO), é possível se gerar estimativas errôneas, pois quando um banco de dados apresenta *outliers* nos dados ou quando a variável apresenta comportamento de assimetria o MMQO pode não conseguir obter estimativas corretas para os parâmetros envolvidos. Este método atribui pesos iguais a cada observação durante o processo de estimação dos parâmetros (DRAPER; SMITH, 1998). O pesquisador acaba tendo uma interpretação errônea, pois a relação entre a variável resposta e as demais covariáveis não seja estabelecida de forma correta.

Algumas características importantes da regressão quantílica foram sintetizadas por Buchinsky (1998). O modelo pode ser utilizado para caracterizar toda a distribuição condicional da variável resposta dado um vetor de covariáveis. Os modelos de regressão quantílica têm uma representação de programação linear que torna as estimativas mais fáceis, pois é uma vantagem a estimar uma medida de tendência central (mediana) mais robusta em relação aos valores atípicos, os *outliers*. O modelo de regressão quantílica fornece a possibilidade de obter um ajuste mais global, explorando o efeito das covariáveis nos quantis mais extremos, apresentando assim uma ideia geral a respeito da distribuição condicional da variável resposta. Desde então, os modelos de regressão quantílica vem se tornando uma técnica cada vez mais atrativa, ganhando cada vez mais importância, contando com inúmeros trabalhos publicados; ver Koenker e Hallock (2001), Yu e Moyeed (2001), Koenker (2004) e Bonanno *et al.* (2018). O estimador soma dos desvios absolutos ponderados é responsável por dar um caráter mais

robusto à regressão quantílica, de modo que a estimativa do vetor de coeficientes não seja sensível a *outliers* na variável resposta. Quando os erros não são normais o estimador dos desvios absolutos ponderados pode ser mais eficiente do que o estimador MMQO. As soluções potencialmente diferentes nos distintos quantis podem ser obtidas por modificar as covariáveis nos diferentes pontos da distribuição condicional da variável resposta (SOUZA, 2017).

Então, seja Y uma variável aleatória com uma função de distribuição acumulada $F_Y(y) = P(Y \leq y)$, então o quantil de ordem τ da variável Y é dado por:

$$Q_\tau(Y) = F^{-1}(\tau) = \inf\{y | F(y) \geq \tau\}, \quad (03)$$

sendo $\tau \in [0,1]$.

Segundo Maciel *et al.* (2001), uma importante propriedade de função quantil concerne ao fato que para $-\infty \leq y \leq +\infty$ e $0 \leq \tau \leq 1$, $F(y) \geq \tau$ se e somente se $Q_\tau \leq y$. Assim, tem-se Y identicamente distribuída a Q_τ .

O modelo de regressão quantílica pode ser dado da seguinte forma:

$$Q_\tau(Y_\tau | X_\tau) = \beta(\tau) = \beta_{0(\tau)} + \beta_{1(\tau)}X_{\tau 1} + \dots + \beta_{p(\tau)}X_{\tau p} \quad (04)$$

onde, $\beta(\tau)$ é o efeito marginal das variáveis explicativas X no τ -ésimo quantil da variável Y , efeito este que pode ser variante a depender do quantil escolhido. Essa abordagem se mostra bastante pertinente para variáveis dependentes cuja distribuição pode apresentar assimetria, caudas pesadas ou heterocedasticidade.

Seja Y_j , ($j = 1, 2, \dots, n$) uma amostra aleatória da variável aleatória Y , com distribuição simétrica em torno de $\beta(\tau)$. Para $0 \leq \tau \leq 1$, o τ -ésimo quantil amostral pode ser definido como a solução do problema de minimização dado por:

$$\min_{\hat{\beta}(\tau) \in \mathbb{R}} \left[\sum_{j \in \{j: Y_j \geq \hat{\beta}(\tau)\}} \tau |Y_j - \hat{\beta}(\tau)| + \sum_{j \in \{j: Y_j \leq \hat{\beta}(\tau)\}} (1 - \tau) |Y_j - \hat{\beta}(\tau)| \right], \tau \in (0,1) \quad (05)$$

Seja $Y = (Y_1, Y_2, \dots, Y_n)^T$ o vetor de variável resposta do modelo de regressão quantílica, em que Y_j é dado por: $Y_j = \beta_0 + X_1\beta_1 + X_2\beta_2 + \dots + X_k\beta_k + \epsilon_j$, $j = 1, \dots, n$. O estimador de $\beta(\tau)$ no modelo de Regressão Quantílica de ordem τ é qualquer solução em $\hat{\beta}(\tau)$ do problema de minimização:

$$\min_{\hat{\beta}(\tau) \in \mathbb{R}^n} \left[\sum_{j \in \{j: Y_j \geq X_j \hat{\beta}(\tau)\}} \tau |Y_j - X_j \hat{\beta}(\tau)| + \sum_{j \in \{j: Y_j \leq X_j \hat{\beta}(\tau)\}} (1 - \tau) |Y_j - X_j \hat{\beta}(\tau)| \right], \tau \in (0,1) \quad (06)$$

em que, X_j é a j -ésima linha da matriz X na qual é um conjunto de dados que será representado em forma matricial e para minimização será usado a Eq. (06)

Conforme Rodrigues (2019), o parâmetro τ pondera a regressão quantílica, ou seja, é tal que pelo menos $100\tau\%$ dos valores dos Y_j encontram-se acima do hiperplano de regressão quantílica e pelo menos $100(1 - \tau)\%$ abaixo.

Koenker e Machado (1999) propuseram uma medida similar à estatística R^2 , bastante utilizada na regressão normal linear, que permite ao investigador avaliar a qualidade do ajuste nos modelos de regressão quantílica.

A função de perda é definida como $\rho_\tau(Y) = Y(\tau - \Pi_{(Y_0)})$, sendo Π uma função de indicador. Quando usado, por exemplo, $\tau = 0,5$, a função de perda $\rho_\tau(Y)$ é proporcional à função de valor absoluto e, portanto, a regressão mediana é o mesmo que a regressão linear por desvios absolutos mínimos.

Para tanto, considere o modelo linear para o quantil condicional de $y|x$

$$Q_\tau(Y_t|X_t) = \beta(\tau)X_t^T = \beta_1(\tau)X_{t1}^T + \beta_2(\tau)X_{t2}^T \quad (07)$$

em que β é responsável por minimizar a soma dos desvios absolutos ponderados para o modelo completo

$$\hat{V}(\tau) = \sum_{t=1}^n \rho_\tau(Y_t - X_t^T \hat{\beta}(\tau)) \quad (08)$$

e β , sob a restrição linear que $H_0 : \beta_2 = 0$, é responsável por minimizar a soma dos desvios absolutos ponderados para o modelo reduzido

$$\tilde{V}(\tau) = \sum_{t=1}^n \rho_\tau(Y_t - X_{t1}^T \hat{\beta}_1(\tau)) \quad (09)$$

Dessa forma, podemos definir a medida de bondade de ajuste para modelos de regressão quantílica da seguinte maneira

$$R^1(\tau) = 1 - \frac{\hat{V}(\tau)}{\tilde{V}(\tau)} \quad (10)$$

o valor gerado de $R^1(\tau)$ apresenta são restritos ao intervalo $(0,1)$, constituindo assim uma medida da qualidade do ajuste para um quantil da regressão quantílica.

O modelo de regressão quantílica permite ao investigador ter uma visão mais completa

da relação existente entre a variável resposta e as covariáveis observadas, uma vez que é possível construir um modelo para cada quantil de interesse, possibilitando identificar diferenças existentes entre os coeficientes estimados (NASCIMENTO *et al.*, 2012). Quando as superfícies das linhas de regressão não são paralelas, o efeito das covariáveis não é uniforme ao longo dos quantis, tem-se que os erros podem apresentar alguma forma de heteroscedasticidade (KOENKER, 2005).

Os modelos lineares de heteroscedasticidade apresentam-se como um importante caso na regressão quantílica. Neste cenário, por exemplo, pode-se assumir que a média e a variância são funções lineares das covariáveis. Segundo Koenker e Bassett (1982) partindo do pressuposto de que os erros são independentes e identicamente distribuídos (i.i.d.), pode-se definir um modelo geral de heteroscedasticidade sistemática dado por:

$$Y = \mu(x) + \sigma(x)\varepsilon, \quad (11)$$

em que $\mu(x)$ pode ser pensado como a média condicional do processo de regressão, $\sigma(x)$ como a escala condicional e ε como o termo de erros independente de x , provenientes de uma distribuição com função quantílica $Q\tau(\varepsilon)$.

Para explicar melhor, apresenta-se um exemplo, utilizando dados retirados do livro *Quantile Regression* de Koenker (2005), o cenário apresenta variância dos erros não constantes. Ou seja, ela é definida em função da covariável, x , permitindo assim que a variância aumente em conjunto com x , violando assim a suposição de homoscedasticidade dos modelos de regressão. Para tanto, ajustou-se o seguinte modelo da regressão:

$$Q_{\tau}(Y_t|X_t) = \beta_0(\tau) + \beta_1(\tau)X_{t1}, \quad (12)$$

em que x é uma sequência de 1 até 100 com espaçamento 1 e parâmetros β_0 e β_1 iguais a 10 e 0,09, respectivamente. A Figura 1 apresenta os resultados obtidos após o ajuste do modelo para τ variando de 0,1 até 0,9 com espaçamento 0,1. A partir da análise gráfica pode-se observar que as curvas de regressão não são paralelas, ou seja, os coeficientes estão variando em função dos quantis τ . Portanto, o efeito da covariável x não é uniforme através dos quantis.

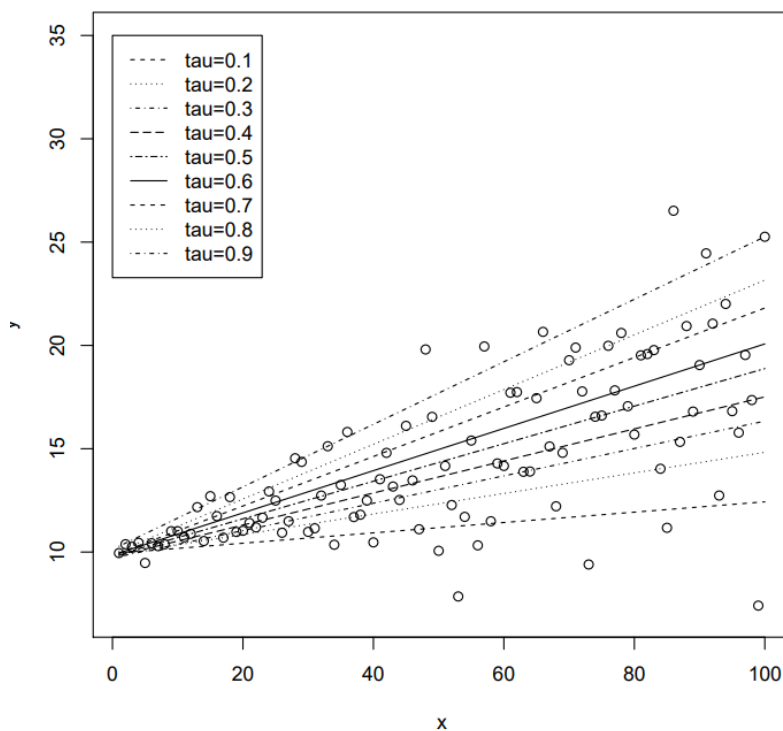
Conforme verificado na literatura os modelos clássicos de regressão são bastante úteis quando os dados são normais e não ocorre violação do pressuposto de homoscedasticidade. Isto ocorre por que os testes de hipóteses utilizados para realizar inferência a respeito dos parâmetros desconhecidos estão fortemente relacionados à suposição de normalidade dos dados. Contudo, em muitas situações envolvendo conjunto de dados reais não é possível validar tais pressupostos. Dessa forma, muitos pesquisadores buscam aplicar transformações na variável

resposta de modo a adequar os dados ao modelo proposto e obter uma distribuição mais próxima da desejada.

A princípio, uma transformação na variável resposta pode ser uma alternativa prática para se conseguir um bom ajuste nos modelos de regressão linear normal. Entretanto, tal classe de modelos não permite obter o valor esperado da variável original por aplicar uma transformação inversa no preditor linear. Por exemplo, considere o caso em que aplica-se a transformação logarítmica na variável resposta, que não segue distribuição normal, obtendo assim a variável $z = \log(Y)$ com distribuição de probabilidade mais próxima da desejada. Portanto, é possível obter a esperança condicional de $z|X$ da seguinte maneira

$$E(z|X) = \beta_0 + \beta_1 X, \quad (13)$$

Figura 2- Ajustes do modelo de regressão quantílica $Q_\tau(Y_t|X_t) = \beta_0(\tau) + \beta_1(\tau)X_{t1}$. Assumindo $\beta_0 = 0.1$, $\beta_1 = 0.9$ e $\tau = 0.1, 0.2, \dots, 0.9$.



Fonte: Koenker (2005)

com n observações. Neste cenário, tem-se que

$$E(z|X) = E(\log(Y) | X) \neq \log E(Y|X) \quad (14)$$

Ou seja, não é possível obter o valor esperado da variável original, y , simplesmente aplicando a exponencial em $E(z|X)$. Por outro lado, a regressão quantílica apresenta uma característica distinta denominada de equivariância a transformações monótonas, como apresentado em Powell (1986), Hao e Naiman (2007), Mua e Hea (2007), Bottai, Cai e McKeown (2010) e Santos (2012). Caso torne-se necessário aplicar uma função $h(\cdot)$, a exemplo da exponencial ou da logarítmica, não decrescente no conjunto R da variável resposta y , tem-se que os quantis serão obtidos por aplicar a mesma transformação na função quantílica, $Q_\tau(Y_t|X_t)$. Por exemplo, se q é o τ -ésimo quantil de Y , então $h(q)$ é o τ -ésimo quantil de $h(Y)$. Portanto, considere Y uma variável resposta restrita ao intervalo $(0, 1)$, com distribuição de probabilidade não normal, e X um vetor de covariáveis.

Dessa forma, é possível estabelecer uma relação funcional entre a variável Y e o vetor de covariáveis, X , da seguinte maneira

$$hQ_\tau(Y_t|X_t) = \beta_0(\tau) + \beta_1(\tau)X_{\tau 1} + \beta_2(\tau)X_{\tau 2} + \dots + \beta_{p(\tau)}X_{\tau p} \quad (15)$$

em que $Q_\tau(Y_t|X_t)$ é o quantil condicional de $x|Q_\tau$. Aqui, $h(\cdot)$ é uma função estritamente monótona e não decrescente no conjunto R . Existem algumas escolhas possíveis para as funções de ligação $h(\cdot)$, a exemplo das ligações probit, log-log e logit (MCCULLAGH; NELDER, 1989). Portanto estimando os coeficientes $\hat{\beta}_j(\tau)$ é possível obter os valores de $Q_\tau(Y_t|X_t)$ por aplicar uma transformação inversa $h^{-1}(\cdot)$ na Eq. (10). Por exemplo, pode-se utilizar a função $h(Y) = \text{logit}(Y)$ para transformar a variável resposta y , restrita ao intervalo $(0, 1)$, da seguinte maneira

$$h(Y) = \text{logit}(Y) = \log\left(\frac{Y}{1-Y}\right) \quad (16)$$

Dessa forma, a inferência a respeito do τ -ésimo quantil da variável resposta original pode ser obtida por meio da seguinte transformação inversa

$$Q_\tau(Y_t|X_t) = \frac{\exp(\beta_0(\tau) + \beta_1(\tau)X_{\tau 1} + \beta_2(\tau)X_{\tau 2} + \dots + \beta_{p(\tau)}X_{\tau p})}{1 + \exp(\beta_0(\tau) + \beta_1(\tau)X_{\tau 1} + \beta_2(\tau)X_{\tau 2} + \dots + \beta_{p(\tau)}X_{\tau p})} \quad (17)$$

com n observações. Assim é possível retornar aos valores da variável original sem perda de interpretação do modelo original.

2.5.1 Teste da Falta de Ajuste em Modelos de Regressão Quantílica

Para entender melhor os erros, tanto homocedásticos como também os heterocedásticos, usa-se como fonte os autores He e Zhu (2003) que propuseram duas abordagens de testes. Primeiro, para os erros homocedásticos, supondo o seguinte modelo:

$$Y_j = X_j' \beta(\tau) + \epsilon_j \quad (18)$$

em que os erros são iid com quantil de ordem τ igual a zero. Seja $\psi_\tau(u)$ a derivada da função de perda definida como $\psi_\tau(r) = \tau - I(r \leq 0)$. O parâmetro $\beta(\tau)$ do modelo da Eq. (18) e estimado pela $\hat{\beta}(\tau)$ para o quantil condicional de ordem τ e definida os resíduos como:

$$r_j = Y_j - X_j' \hat{\beta}(\tau) \quad (19)$$

Nessas condições, He e Zhu (2003) definem inicialmente o bloco principal para o teste de falta de ajuste em modelos de regressão quantílica como

$$R_n(t) = n^{-\frac{1}{2}} \sum_{j=1}^n \psi(r_j) X_j I(x_j \leq t) \quad (20)$$

em que I representa a função indicadora, sendo que $I(X_j \leq t) = 1$ somente quando todos os termos de x_j são menores ou iguais aos correspondentes componentes de t .

A estatística de teste é definida como

$$T_n = \max_{\|a\|=1} a' n^{-1} \sum_{j=1}^n (a' R_n(X_j))^2 \quad (21)$$

Como T_n pode ser escrita na forma

$$T_n = \max_{\|a\|=1} a' \left[n^{-1} \sum_{j=1}^n R_n(x_j) R_n'(X_j) \right] \quad (22)$$

então, o seu valor coincide com o maior autovalor da matriz $n^{-1} \sum_{j=1}^n R_n(X_j) R_n'(X_j)$.

O cálculo do nível descritivo do teste envolve simulações para decidir se a estatística T_n excede o quantil superior de ordem α de T . Para melhor entendimento da teoria que envolve esse teste, é indicado Stute (1997), He e Zhu (2003) na qual mostram que, sob a hipótese de que o modelo é da forma Eq. (18) e para n grande, T_n converge em distribuição para a variável aleatória T , em que T é o maior autovalor de $\int R(t) R'(t) w(t) dt$ e w é a função de distribuição de X .

Segundo Stute (1997) e He e Zhu (2003), como a distribuição limite do teste é independente da distribuição dos erros, então para calcular o nível descritivo do teste, várias reamostras de (Y_1^*, \dots, Y_n^*) são geradas, em que os Y_j^* possuem qualquer distribuição contínua

com o quantil de ordem τ igual a zero. Em cada reamostra gerada, a estatística de teste é calculada a partir dos resíduos r_j^* , $r_j^* = Y_j^* - X_j' \hat{\beta}(\tau)$, da seguinte forma,

$$R_n^*(t) = n^{-\frac{1}{2}} \sum_{j=1}^n \psi(r_j^*) X_j I(X_j \leq t). \quad (23)$$

Portanto para se calcular a reamostra basta seguir estas equações em que será possível chegar a resultados mais bem-sucedido sobre seus erros.

2.5.2 Teste de Hipótese Linear Geral

Para entender melhor os dados existem algumas alternativas na literatura para testar hipóteses lineares gerais da forma

$$H_0 : C\beta(\tau) = c, \quad (24)$$

em que C é uma matriz de constantes conhecidas, de posto completo e c é um vetor de constantes conhecidas, quando os parâmetros são estimados pelo método da regressão quantílica. Por exemplo, se um modelo fosse ajustado, então uma hipótese de interesse seria verificar se todos os $\beta_j(\tau)$, $j = 1, \dots, p$ são iguais a zero, contra a hipótese alternativa de que pelo menos um deles seja diferente de zero, como é feito no teste da hipótese da análise de variância para modelos de regressão clássica.

Koenker (2005) formula testes do tipo Wald, que podem ser utilizados para verificar a hipótese da Eq. (24), como também hipóteses envolvendo diversos quantis e diversos parâmetros de forma simultânea. Considerando um problema em que são estimados m diferentes modelos, então a hipótese linear geral sobre o vetor $\zeta = (\beta(\tau_1)', \dots, \beta(\tau_m)')'$, em que $\beta(\tau_j)$ é o vetor com p parâmetros para $\tau = \tau_j$ ($j = 1, \dots, m$), pode ser escrita da seguinte forma

$$H_0 : C\zeta = c, \quad (25)$$

em que C é uma matriz de constantes conhecidas, $q \times mp$, de posto completo q e c é um vetor de constantes conhecidas, $q \times 1$.

Nestas condições, a estatística de teste é

$$T_n = n(C\hat{\zeta} - c)'[CV_n^{-1}C'] - 1(C\hat{\zeta} - c), \quad (26)$$

em que V_n é a matriz $mp \times mp$,

$$V_n(\tau_1, \dots, \tau_m) = \begin{bmatrix} V_n(\tau_1, \tau_1) & V_n(\tau_1, \tau_2) & \cdots & V_n(\tau_1, \tau_m) \\ V_n(\tau_2, \tau_1) & V_n(\tau_2, \tau_2) & \cdots & V_n(\tau_2, \tau_m) \\ \vdots & \vdots & \ddots & \vdots \\ V_n(\tau_m, \tau_1) & V_n(\tau_m, \tau_2) & \cdots & V_n(\tau_m, \tau_m) \end{bmatrix} \quad (27)$$

e cada matriz $V_n(\tau_i, \tau_j)$, $p \times p$, é dada por

$$V_n(\tau_i, \tau_j) = [\tau_i \wedge \tau_j - \tau_i \tau_j] H_n(\tau_i)^{-1} J_n H_n(\tau_j)^{-1}, \quad (28)$$

com $\tau_i \wedge \tau_j$ representando o mínimo entre τ_i e τ_j , $i \neq j$, $j = 1, 2, \dots, m$ e J_n e $H_n(\tau)$ são definidos como

$$J_n = \frac{\sum_{i=1}^n X_i X_i'}{n} \quad \text{e} \\ H_n(\tau) = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n X_i X_i' f_i(\xi_i(\tau))}{n} \quad (29)$$

O termo $f_i(\xi_i(\tau))$ denota a densidade condicional da variável resposta, Y_i , avaliada no quantil de ordem τ , $\xi_i(\tau)$.

Uma hipótese de bastante interesse é a de que todos os parâmetros do modelo com $p - 1$ variáveis explicativas são iguais a zero, ou seja

$$\beta_2(\tau) = \beta_3(\tau) = \cdots = \beta_p(\tau) = 0, \quad (30)$$

para τ definido no modelo $Y = X\beta(\tau) + \varepsilon$ sendo que β_1 se refere ao intercepto do modelo. A suposição dos erros independentes e identicamente distribuídos simplifica a notação da matriz de covariâncias.

Após alguns cálculos, verifica-se que a estatística de teste para a hipótese da Eq. (30) é dada por

$$T_n = n \sum_{i=2}^p \frac{\hat{\beta}_i^2(\tau)}{\text{Var}(\hat{\beta}_i(\tau))} \quad (31)$$

Essa estatística pode ser reescrita da seguinte forma

$$T_n = n \frac{f^2(0)}{\tau(1-\tau)} \sum_{i=2}^p \frac{\hat{\beta}_i^2(\tau)}{v_{ii}} \quad (32)$$

em que v_{ii} é i -ésimo elemento da diagonal da matriz $(X'X)^{-1}$ e $f(0)$ deve ser substituído por uma estimativa, para que o valor acima possa ser considerado uma estatística.

A estatística T_n tem assintoticamente distribuição χ_q^2 sob H_0 , em que q é o posto da matriz C . Para a hipótese da Eq. (30), a estatística $T_n \xrightarrow{D} \chi_{p-1}^2$.

Além dessa possibilidade, Chen *et al.* (2008) desenvolveram um método que, segundo definição dos próprios autores, pode ser visto como uma análise de variância para modelos de regressão L_1 . A estatística de teste que os autores propuseram, inicialmente, é a seguinte

$$M_n = \sum_{i=1}^n |Y_i - X_i' \hat{\beta}_r| - \sum_{i=1}^n |Y_i - X_i' \hat{\beta}_c| \quad (33)$$

em que $\hat{\beta}_r$ é o estimador de β no modelo reduzido sob H_0 e $\hat{\beta}_c$ é o estimador de β no modelo completo. Verifica-se que essa estatística de teste coincide com a estatística de teste da razão de verossimilhança, para a hipótese na Eq. (30), quando os erros têm distribuição de Laplace que é uma probabilidade contínua com o nome de Pierre-Simon Laplace. Às vezes também é chamada de distribuição exponencial dupla, porque pode ser pensada como duas distribuições exponenciais (com um parâmetro de localização adicional) unidas ao longo da abcissa, embora o termo também seja usado às vezes para se referir à distribuição Gumbel. A diferença entre duas variáveis aleatórias exponenciais identicamente distribuídas independentes é governada por uma distribuição de Laplace, avaliado em um tempo aleatório exponencialmente distribuído. Incrementos de movimento de Laplace ou um processo de variação de gama avaliados ao longo da escala de tempo também têm uma distribuição de Laplace. Diante do que foi exposto os autores ainda mostram que

$$M_n \xrightarrow{D} \frac{X_q^2}{4f(0)} \quad (34)$$

em que q é o número de linhas da matriz C e $f(\cdot)$ é a função densidade dos erros. Porém, para evitar a estimação desse valor da função densidade, os autores propuseram a seguinte transformação da estatística M_n ,

$$M_n^* = \min_{\beta \in \Omega_0} \sum_{i=1}^n w_i |y_i - x_i' \beta| - \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n w_i |y_i - x_i' \beta| - \left(\sum_{i=1}^n w_i |y_i - x_i' \hat{\beta}_r| - \sum_{i=1}^n w_i |y_i - x_i' \beta| \right) \quad (35)$$

em que w_1, \dots, w_n é uma sequência de variáveis aleatórias não negativas independentes e identicamente distribuídas com média 1 e variância 1 e Ω_0 é o espaço paramétrico gerado pela hipótese nula. Com relação aos pesos utilizados na definição da nova estatística de teste, w_i , a distribuição exponencial de parâmetro 1 pode ser utilizada para gerar os valores, pois tem média e variância 1. Chen *et al.* (2008) provam que, sob H_0

$$M_n^* \xrightarrow{D} \frac{X_q^2}{4f(0)} \quad (36)$$

Por esse motivo, os autores defendem que, ao invés de estimar a densidade em $f(0)$, a região crítica para a estatística de teste M_n possa ser construída a partir da distribuição empírica de M_n^* . Tendo em vista esses resultados para regressão L_1 , Chen *et al.* (2008) argumentam que os mesmos também podem ser utilizados na regressão quantílica, com a simples troca do desvio absoluto pela função de perda definida. Nesse caso, a estatística de teste seria igual a

$$M_n = \min_{\beta \in \Omega_0} \sum_{i=1}^n \rho_\tau(y_i - x_i' \beta) - \min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_\tau(y_i - x_i' \beta) \quad (37)$$

De forma análoga, deve-se reescrever a estatística M_n^* utilizando a função de perda $\rho_\tau(u)$ para a construção da região crítica da estatística de teste M_n nos modelos de regressão quantílica.

Finalizando, Gutenbrunner *et al.* (1993) propõem um teste para hipóteses lineares do tipo da Eq. (29) que é baseado em Gutenbrunner e Jureckova (1992) e os escores ordinais de regressão quantílica. Para mais detalhes sobre o teste, como estatística de teste, resultados assintóticos, ver Gutenbrunner *et al.* (1993) e Koenker (2005). Para utilização do teste, dentro da função *anova.rq*, deve-se usar o comando *test = "rank"*.

3 MATERIAL E MÉTODOS

3.1 TIPOLOGIA DA PESQUISA

O estudo foi descritivo com abordagem quantitativa com foco na análise de regressão quantílica. Foi realizado com parte dos dados do projeto de pesquisa mais amplo intitulado “Avaliação do custo da obesidade para o Sistema Único de Saúde, com desenvolvimento de ferramenta de monitoramento” aprovado no edital MS-SCTIE-Decit/CNPq N° 26/2019 – Pesquisa em Alimentação e Nutrição e aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de Alfenas - UNIFAL/MG (CAAE: 33571520.1.0000.5142/ protocolo nº: 4.209.111/2020).

O presente trabalho utilizou como técnica para coletar dados, o emprego do questionário estruturado, com questões fechadas e abertas, envolvendo cinco seções. O emprego dessa técnica foi necessário para o levantamento de uma grande quantidade de dados, como também o seu uso de cunho qualitativo. Como os pontos fortes deste tipo de estudo é a garantia do anonimato, além de trazer questões objetivas, pois assim não deixam em aberto o tempo para as pessoas pensarem sobre as respostas, questões fechadas garantem uniformidade e assim uma rapidez para a conversão de dados para arquivos de *software*. (CHAER; DINIZ; RIBEIRO, 2012).

Na pesquisa em questão foi abordado o uso da metodologia de regressão. Para o tipo de variável em estudo considerou-se a regressão quantílica a metodologia mais apropriada, que permite ajustar diferentes modelos de regressão para cada quantil de interesse, contudo é necessário que a variável resposta esteja definida no conjunto dos reais.

3.2 CENÁRIO DO ESTUDO

O estudo realizado utilizou dados obtidos após a aplicação de um questionário no período de maio a setembro de 2020, nos municípios de Minas Gerais na Macrorregião Sul que envolve: doze microrregiões com quatro superintendências a de Alfenas, Pouso Alegre, Passos e Varginha. Este questionário auxiliou a entender os impactos causados em relação às medidas regulatórias RDC 359/2003, RDC 360/2003 e RDC 24/2010, obtendo assim dados satisfatório para a pesquisa.

3.3 POPULAÇÃO E AMOSTRA

A população em estudo é a macrorregião do sul de Minas Gerais e é composta por 154 municípios, que totaliza uma população estimada de 2.834.846 em todo território que compreende as 04 Superintendências Regionais de Saúde (SRS), a saber, Alfenas, Passos, Pouso Alegre e Varginha.

A técnica utilizada é a amostragem aleatória simples que é o processo mais elementar e frequentemente utilizado. Corresponde a uma amostra de elementos retirados ao acaso da população, isto é, cada indivíduo é escolhido completamente ao acaso e cada membro da população tem a mesma probabilidade de ser incluído na amostra. Para obtermos essa amostragem desta população da macrorregião do sul de Minas Gerais foi retirado uma amostragem com idades em fase ativa entre 18 e 60 anos, assim foram enviados questionários eletronicamente por meio da rede mundial de computadores além de escolher alguns pontos como os PSFs e escolher ao acaso as pessoas que aceitaram fazer nos locais.

Dessa forma, no final da pesquisa obteve-se um total de 580 participantes de diversas cidades que estão na macrorregião do sul de Minas Gerais assim os dados obtidos foram bastante diversificados e apresentou diversos valores extremos.

3.4 PROCEDIMENTO PARA OBTENÇÃO DOS DADOS

O recrutamento dos participantes foi por meio das redes sociais. Os participantes que aceitaram fazer a pesquisa, após recebimento de um convite, foram direcionados para o ambiente virtual do projeto, onde, ao ler o Termo de Consentimento Livre Esclarecido e confirmar o interesse em seguir como voluntário da pesquisa, somente após esse é que o voluntário teve acesso ao questionário (Apêndice I). Foram elegíveis para responder o questionário os indivíduos com 18 anos e menos de 60 anos, com escolaridade mínima de ensino fundamental e com autonomia não discutível.

3.5 PROCEDIMENTO DE ANÁLISE DE DADOS

Os dados foram tabulados em uma planilha eletrônica, e após checagem foram transferidos para o software *R* (R CORE TEAM, 2022). Inicialmente, foi realizada uma análise descritiva dos dados a fim de extrair informações importantes a respeito das variáveis abordadas no estudo conhecendo suas distribuições. Em seguida, são apresentados os procedimentos inferenciais, as medidas da qualidade do ajuste e análise de diagnóstico relacionado ao modelo

de regressão quantílica. Vale ressaltar que para tais modelos utilizou-se o pacote *quantreg* disponíveis no CRAN (repositório) do software R (KOENKER, 2005).

Obteve-se 32 variáveis, destas variáveis 10 variáveis foram usadas para descrição da amostra em estudo para entender melhor como são seus hábitos alimentares e desta maneira perceber se existiria alguma limitação. As variáveis em estudo foram divididas em: variáveis categóricas que contêm um número finito de categorias ou grupos distintos; variáveis discretas são variáveis numéricas que têm um número contável de valores entre quaisquer dois valores; variáveis contínuas são variáveis numéricas que têm um número infinito de valores entre dois valores quaisquer; variável dicotômica é aquela que pode assumir apenas dois valores. Esses valores geralmente são zero, como ausência, ou um, como presença.

Portanto foram escolhidas as variáveis de um estudo aplicado em uma amostra de 580 voluntários, em que foi coletado informações sobre características em nível familiar e individual (por exemplo, idade, sexo, nível de escolaridade, hábitos de compra de seu próprio alimento, prática de atividade física, tempo gasto em mídias sociais, etc.), bem como dados autorrelatados, deste formulário irá usar 10 variáveis para implementar a regressão e para obter resultados que explorem a relação entre a leitura de rótulos de alimentos e o IMC de adultos, foi usado uma especificação de modelo que inclui variáveis parecidas com estudos anteriores que investigam os fatores socioeconômicos, ambientais e comportamentais da obesidade. Além disso, muitas das variáveis incluídas neste modelo são representadas por fatores que influenciam o uso de rótulos nutricionais, que são idade, escolaridade, renda familiar, tamanho da família, atividade física, IMC, responsável pela compra de alimentos na sua família (*com*), leitura de rótulos nutricionais, horas de acompanhamento das mídias e a influência das propagandas (*pro*) apresentadas no Quadro 1, para melhor identificar o que influencia o IMC. A justificativa para se usar essas variáveis foi devido a literatura que deixa evidente a ligação do IMC com essas variáveis. Os autores que fazem essa ligação são Kim, Nayga e Capps Jr. (2000), Loureiro *et al* (2012), Bonnano *et al.* (2018) e Souza *et al.* (2019), durante a revisão de literatura, foram poucos os artigos que tratavam deste assunto, pois a regressão quantílica ainda é um tema complexo, principalmente quando se trata de temas ligados a saúde.

O formulário aplicado coletou informações das seguintes variáveis apresentadas no Quadro 1. Neste estão apresentadas uma descrição a respeito do que está sendo abordado em cada variável.

Quadro 1 – Variáveis selecionadas e sua descrição das siglas

Variáveis dicotômicas	Descrição das Variáveis
<i>propaganda</i>	Você já se sentiu atraído(a) por algum alimento durante a propaganda dele e consequentemente fez a compra? (0: caso não tenha se sentido atraído e, 1: para caso tenha se sentido atraído)
<i>Compra seu próprio alimento</i>	Você é o(a) responsável pela compra de alimentos na sua casa? (0: caso não seja responsável e, 1: para caso seja responsável pela compra)
<i>Leitura de rótulos nutricionais</i>	Você costuma ler os rótulos dos alimentos? (0: caso não leia e talvez não leia rótulos e, 1: para caso leia rótulos)
<i>Atividades físicas</i>	Você pratica alguma atividade física? Segundo a Organização Mundial da Saúde, é preciso fazer 150 minutos semanais de atividade física leve ou moderada (cerca de 20 minutos por dia) ou, pelo menos, 75 minutos de atividade física de maior intensidade por semana (cerca de 10 minutos por dia). (0: caso não pratique exercício e, 1: para caso pratique exercício)
Variáveis categóricas	
<i>Escolaridade</i>	Qual a sua escolaridade?
<i>Renda Familiar</i>	Renda familiar
Variáveis discretas	
<i>Tamanho da família</i>	Tamanho da Família
Variáveis Contínua	
<i>IMC</i>	IMC (índice de massa corporal)
<i>idade</i>	Idade
<i>Horas nas mídias</i>	Quantas horas costuma acompanhar a mídia por dia (incluindo internet)?

Fonte: Elaboração da autora.

Para o cálculo do IMC foi necessário perguntar no questionário o peso, altura e a circunferência da cintura dos participantes, para obtê-lo foi necessário dividir o peso autorreferido em quilogramas dos entrevistados pelo quadrado da altura em metros (kg/m²).

Foi feito o teste de Shapiro-Wilk que é um teste de normalidade nas estatísticas. Na qual foi publicado em 1965 por Samuel Sanford Shapiro e Martin Wilk (SHAPIRO; WILK, 1965).

Foi usada a correlação de Spearman, sendo que o coeficiente de correlação de posto ou ρ de Spearman, em homenagem a Charles Spearman é muitas vezes denotado como r_s e é uma medida não paramétrica de correlação de postos (dependência estatística entre os rankings de duas variáveis). Ele avalia quão bem a relação entre duas variáveis pode ser descrita usando uma função monotônica.

A correlação de Spearman entre duas variáveis é igual à correlação de Pearson entre os valores de rank dessas duas variáveis; enquanto a correlação de Pearson avalia as relações lineares, a correlação de Spearman avalia as relações monotônicas (lineares ou não). Se não

houver valores de dados repetidos, uma correlação de Spearman perfeita de +1 ou -1 ocorre quando cada uma das variáveis é uma função monótona perfeita da outra (GIBBONS; CHAKRABORTI, 2014)

Intuitivamente, a correlação de Spearman entre duas variáveis será alta quando as observações tiverem uma classificação semelhante (ou idêntica para uma correlação de 1) (ou seja, rótulo de posição relativa das observações dentro da variável: 1º, 2º, 3º, etc.) entre os dois variáveis e baixo quando as observações têm uma classificação diferente (ou totalmente oposta para uma correlação de -1) entre as duas variáveis.

Realizou-se o teste White que irá identificar se os dados apresentam heterocedasticidade. O teste de White permite não-linearidades por utilizar quadrados e produtos cruzados de todas as variáveis explicativas (X) (SOUZA; JUNQUEIRA, 2005).

Para avaliar se os erros são normalmente distribuídos foi utilizado o teste para a normalidade dos resíduos, onde a hipótese nula apresentada é de que o erro possui distribuição normal. Portanto, para a normalidade dos resíduos a um nível de significância de 10% aceita H_0 .

Para fazer o teste de hipótese linear geral, foi usado o pacote estatístico *quantreg*. Para utilizá-lo basta fornecer o argumento *test = "anovar"* dentro da função *anova.r*, com o argumento *test = "Wald"*. Na escolha do número de reamostras que serão utilizadas para o cálculo do nível descritivo do teste, ou simplesmente valor-*p*.

3.6 ABORDAGEM DA REGRESSÃO QUANTÍLICA

Assumiu-se uma relação simples em que o IMC do indivíduo tem como variáveis independentes a leitura de informações sobre os ingredientes no rótulo ($Label_j$); características socioeconômicas (SE_j ; quantas pessoas existem na família, idade, escolaridade, renda familiar); acesso e hábitos alimentares (AC_j ; nível médio de pessoas que fazem suas próprias compras; influências da publicidade); e variáveis comportamentais (Be_j ; prática de atividade física, tempo gasto em mídias no geral). Uma expressão, segundo Bonanno *et al.* (2018), que representa o IMC em função desses grupos de variáveis é

$$IMC_j = f(Label_j, SE_j, AC_j, Be_j, R|\beta) + \epsilon_j, \quad (38)$$

onde R indica efeitos fixos específicos da região estudada, que capturam fatores não observáveis que podem afetar o IMC dos adultos, β é um vetor de coeficientes a serem

estimados e ϵ_j é um termo erro. Ainda segundo Bonanno *et al.* (2018), foi postulada uma forma funcional linear, renderizando a especificação empírica da Eq. (38), representada como:

$$IMC_j = \beta_0 + \beta_{Label}Label_j + \sum_{n=1}^N \beta_{SEn}SE_{nj} + \sum_{m=1}^M \beta_{ACl}AC_{lj} + \sum_{l=1}^L \beta_{BEl}BE_{lj} + \sum_{r=1}^R \beta_r R_{rj} + \epsilon_j \quad (39)$$

Para analisar e avaliar a natureza heterogênea da relação entre a leitura de informações sobre os ingredientes no rótulo (*Label*) e o IMC, foram estimadas diferentes versões da Eq. (39), particionando os dados em diferentes subamostras para identificar o efeito de ler o rótulo nutricional, sendo um indivíduo com um determinado nível socioeconômico referentes as categorias das variáveis independentes ou covariáveis. Como exemplo, se baseando em Bonanno *et al.*, (2018), suponha que a variável usada para particionar a amostra seja um dos indicadores binários socioeconômicos e representados pela SE_1 . A equação estimada que captura a relação entre ler os ingredientes no rótulo e o IMC para a subamostra de respondentes caracterizada por $SE_1 = 1$ é:

$$IMC_{I|SE_1=1} = \beta_0^{SE_1=1} + \beta_{Label}^{SE_1=1}Label_j|_{SE_1=1} + \sum_{n=2}^N \beta_{SEn}^{SE_1=1}SE_{nj}|_{SE_1=1} + \sum_{m=1}^M \beta_{ACl}^{SE_1=1}AC_{lj}|_{SE_1=1} + \sum_{l=1}^L \beta_{BEl}^{SE_1=1}BE_{lj}|_{SE_1=1} + \sum_{r=1}^R \beta_r^{SE_1=1}R_{rj}|_{SE_1=1} + \epsilon_j|_{SE_1=1}. \quad (40)$$

Na abordagem da regressão quantílica existem algumas variações, as mais conhecidas são a regressão quantílica condicional (RQC) que é contingente sobre os valores de todas as variáveis restantes, e se mais de uma variável é incluída no subconjunto então esta regressão condicional é a distribuição conjunta condicional das variáveis; já a outra variação é a regressão quantílica incondicional (RQI), que é obtida marginalizando a distribuição das variáveis proposta no modelo.

Nas estimativas dos parâmetros das Eqs. (39) e (40) foi usado um estimador da regressão quantílica incondicional (FIRPO *et al.*, 2009), porque nos permite entender a relação entre a

leitura dos rótulos nutricionais e o IMC em toda sua distribuição incondicional. Em vez disso, a RQC estima relacionamentos com base na distribuição condicional (nos valores das covariáveis incluídas no modelo) da variável dependente (KOENKER; BASSETT, 1978). Mas conforme Koenker (2005) a RQI é para análises focadas em políticas (BORAH; BASU, 2013) em que a questão central da pesquisa é sobre a relação entre variáveis de interesse político e a distribuição incondicional de uma variável de resultado (PARK, 2015).

Neste caso, busca-se entender o efeito da leitura dos ingredientes dos rótulos alimentares e outras variáveis influenciam no IMC da população adulta do sul de Minas Gerais. Conforme Bonnano *et al.* (2018), usando a RQC, só poderá medir esta relação em valores específicos das variáveis dependentes.

Realizar a análise usando RQI em estratos específicos dos dados permitirá avaliar a relação entre leitura de rótulos e IMC, em diferentes níveis da distribuição incondicional do IMC para um determinado subgrupo da população.

O RQC seria um método de estimação adequado, conforme Souza *et al.* (2019), se o objetivo fosse entender como a leitura dos rótulos dos alimentos se relaciona com o IMC de indivíduos específicos nos dados (por exemplo, IMC de um homem, com 35 anos, com ensino superior entre outras características). Em vez disso, os coeficientes obtidos usando RQI podem ser interpretados como o efeito de uma mudança unitária da covariável na distribuição de resultados, mantendo todas as outras covariáveis constantes, o que leva a uma interpretação mais significativa dos coeficientes para análise de políticas (BORAH; BASU, 2013). Além disso, conforme ilustrado por Borah e Basu (2013), o RQI é uma técnica de estimação mais adequada quando um modelo é caracterizado por múltiplas covariáveis, pois as relações condicionais e incondicionais não são necessariamente equivalentes e podem diferir bastantes.

4 RESULTADOS

4.1 DESCRIÇÃO DOS DADOS

Na Tabela 1 são apresentadas algumas estatísticas descritivas das variáveis estudadas na pesquisa.

Tabela 1– Estatísticas descritivas essenciais para a regressão quantílica das variáveis quantitativas em estudo.

Variáveis	Mínimo	1°Q.	Mediana	Média	3°Q.	Máximo	Desvio Padrão
<i>idade</i>	18,00	28,00	34,00	35,72	43,00	59,00	10,68
<i>Tamanho da família</i>	1,00	2,00	3,00	2,70	4,00	8,00	1,30
<i>IMC</i>	17,26	23,02	25,93	26,74	29,30	64,52	4,78
<i>Horas passadas nas médias</i>	1,00	2,00	2,00	2,48	3,00	4,00	0,96

Fonte: Elaboração da autora.

Nota: 1°Q e 3°Q: Primeiro e terceiro quartis

Para melhor entendimento nas estatísticas descritivas do IMC, peso e altura, foi realizada a estratificação para homens e mulheres, sendo que, das 580 pessoas, 80% são do sexo feminino, com idade média de 35,7 anos ($\pm 10,68$) e, em sua maioria, possui graduação, mestrado ou doutorado (57,01%), 20% são do sexo masculino com idade média de 35,4 anos ($\pm 10,50$) e, possui graduação ou outros estudos superiores (72,17%).

A maioria dos participantes possui companheiro (80,69%), cerca de 25% recebem uma renda média de 1 a 2 salários mínimos ($\pm 1,18$) e as famílias têm em média três pessoas ($\pm 1,29$) e 91,55% moram na zona urbana (91,55%). Quando perguntado sobre a utilização dos serviços oferecidos pelas Unidades Básicas de Saúde (UBS), a grande maioria dos participantes (85,6%) relataram que o utilizam. Esse valor está de acordo com dados levantados pela plataforma e-gestor sobre os dados de cobertura do *e-gestor* (BRASIL, 2021), onde na região Macrossul Minas, 72,5% da população está coberta pela Estratégia da Saúde da Família e 84,1% está coberta pelo *e-gestor* (BRASIL, 2021). Cerca de 57% dos respondentes relataram que praticam atividade física de acordo com as recomendações da OMS, ou seja, 150 minutos semanais de atividade física leve ou moderada (cerca de 20 minutos por dia) ou, pelo menos, 75 minutos de atividade física de maior intensidade por semana (cerca de 10 minutos por dia). O número de praticantes de atividade física é maior do que no país em 2013 e 2019. De acordo com o IBGE, em 2019, 30,1% dos brasileiros praticaram o nível recomendado de atividade física no

lazer, enquanto esse percentual foi de apenas 22,7% em 2013. Entre homens, esse percentual foi de 34,2%, enquanto para as mulheres foi de 26,4%.

Em contrapartida, mostrou o IBGE (2019), os classificados como "insuficientemente ativos" ainda são 40,3% dos adultos brasileiros, percentual considerado alto. Essas pessoas não praticaram atividade física ou o fizeram por menos de 150 minutos por semana, seja no lazer, no trabalho ou em deslocamentos.

O peso médio estimado da amostra foi de 73kg ($\pm 16,6$), possuindo um IMC médio de 26,74kg/m² ($\pm 5,66$), demonstrando que os participantes estão com possível sobrepeso. Estratificando a amostra na variável IMC foi possível verificar a presença de *outliers*, com IMC das mulheres com no mínimo 17,26 kg/m² e máximo 64,52 kg/m², e no caso dos IMC dos homens com no mínimo 18,94 kg/m² e máximo 44,37 kg/m², mostrando que a população se caracteriza por pessoas com baixo peso até obesidade mórbida.

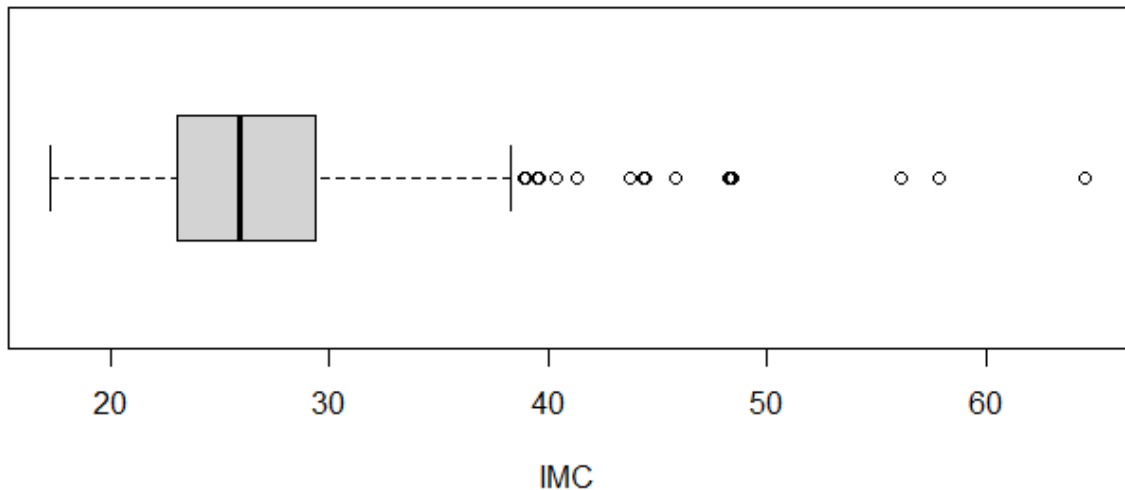
A maioria das pessoas é responsável pela compra mensal dos alimentos da sua casa cerca de 71,6%, porém, apenas 224 acham que os rótulos são importantes, sendo, 38,62% dos participantes que têm o hábito de lê-los sempre. Cerca de 65% dos participantes disseram que utilizam os rótulos para manter uma alimentação saudável, porém, apenas 49% relataram que compreendem os rótulos com facilidade. A grande maioria, cerca de 524 respondentes, ou seja, em torno de 90% das pessoas, utilizam a internet como a sua principal fonte de informação, dedicando cerca de 1 a 2h por dia ($\pm 0,95$), e conferem um grau de confiança médio às informações apresentadas (69%). A maioria dos participantes (81%) já se sentiu atraído e conseqüentemente fizeram a compra do produto devido a sua propaganda.

Os termos que são utilizados nos rótulos como "reduzido valor calórico", "light", "diet", "enriquecido" e "fonte de vitaminas" têm uma influência média na hora da compra dos alimentos (33,62%), e 70% já deixaram de comprar algum alimento após ler no rótulo que ele continha uma quantidade grande de calorias, carboidratos, gordura e/ou sódio.

As variáveis preço, qualidade nutricional, sabor do alimento e praticidade tiveram grande influência na hora da compra, sendo que 88% declararam como mais relevante, o sabor. As variáveis marca, mídias e aparência apresentam baixa influência na hora da compra (12%), sendo que dentro delas as mídias a de menor influência com 6% das declarações.

Na Figura 3 é representada a distribuição dos valores da variável IMC por meio dos seus quantis, sendo possível observar a presença de *outliers*, segundo a metodologia proposta por Tukey, demonstrando a necessidade de uma regressão mais robusta que pudesse melhor modelar estes valores discrepantes e uma distribuição assimétrica.

Figura 3 - Gráfico de Box Plot dos valores da variável IMC com representação dos possíveis valores *outliers* segundo a proposta de Tukey.



Fonte: Elaboração da autora.

Por meio da Figura 3, pode-se observar que há indivíduos com valores de IMC muito acima do valor mediano. Na regressão linear clássica a variável resposta que apresenta tanta discrepância não seria recomendável o seu uso devido a possibilidade de não garantias das suposições e conseqüentemente a baixa confiabilidade na inferência dos resultados para explicar as variações do IMC. Além disso, a regressão linear iria trazer apenas as variações em torno da média e com isso não seria possível entender o que ocorreu com o restante, como os extremos.

Para medir o grau de relacionamento entre as variáveis quantitativas, foi calculada a correlação linear para as variáveis que são idade, escolaridade, renda familiar, tamanho da família, atividade física, IMC, responsável pela compra de alimentos na sua família, leitura de rótulos nutricionais, horas de acompanhamento das mídias e a influência das propagandas.

As variáveis em estudo foram avaliadas quanto à normalidade pelo teste de Shapiro-Wilk. O resultado caracterizou ausência de normalidade pelos dados ($P < 0,0001$) em todas as variáveis, o que significa que as variáveis não seguem a distribuição normal. Devido à ausência de normalidade, optou-se por se realizar a estimação da correlação pelo coeficiente de Spearman.

Na Tabela 2 será mostrado as estimativas do estudo e classificação por cores de acordo com o nível de significância dos coeficientes de correlação Spearman das variáveis

Tabela 2 – Estimativas do estudo e classificação por cores de acordo com o nível de significância dos coeficientes de correlação Spearman das variáveis.

Variáveis	1	2	3	4	5	6	7	8	9	10
1-IMC	1	0,29	-0,06	0,04	0,12	-0,09	-0,05	-0,02	0,11	-0,17
2-Idade		1	0,15	0,19	-0,01	0,06	-0,23	-0,07	0,31	-0,06
3-Escolaridade			1	0,46	-0,14	0,13	0,02	0,09	0,07	0,17
4-Renda familiar				1	0,17	0,09	-0,04	0,05	0	0,08
5-Tamanho da família					1	-0,01	-0,01	0,06	-0,13	-0,12
6-Leitura de rótulo						1	-0,03	-0,09	0,15	0,13
7-Horas passadas nas mídias							1	0,06	-0,02	0,01
8-Propaganda								1	-0,04	-0,01
9-Compra seu alimento									1	0,03
10-Atividade física										1

Fonte: Elaboração da autora.

Legenda: ■ nível de significância maior que 10 %
■ nível de significância 10%
■ nível de significância 5%
■ nível de significância %

Na Tabela 2 estão apresentados a correlação entre as variáveis, a relação mais forte foi verificada entre a escolaridade e a renda familiar, o que indica que existe uma relação positiva moderada entre as variáveis, outras variáveis como idade e IMC, responsável pela compra do seu alimento e idade, apresentaram relação positiva fraca moderada. A leitura de rótulos e o IMC tem uma correlação negativa (-0,09), uma relação bem fraca, porém significativa, caracterizando que quanto maior o IMC menos se faz a leitura de rótulos, o que leva a concordância com estudos como o de Bonnano *et al.* (2018) e o de Souza *et al.* (2019).

Na Figura 4 é apresentada o diagrama de dispersão e o box plot da variável IMC em função das outras variáveis incluídas nesse estudo. Entretanto, não é possível observar nenhuma relação clara entre as variáveis.

Figura 4 - Gráficos de dispersão e box plot da variável IMC em função das outras variáveis incluídas no estudo.

(Continua)

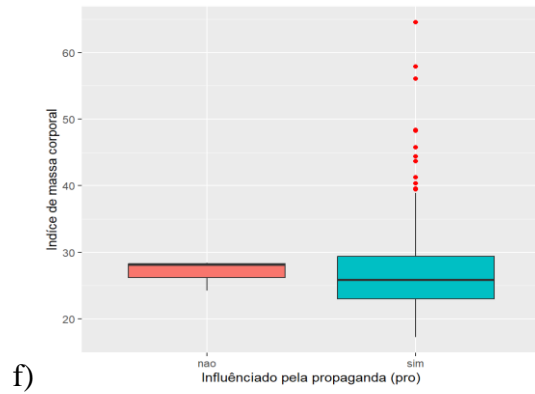
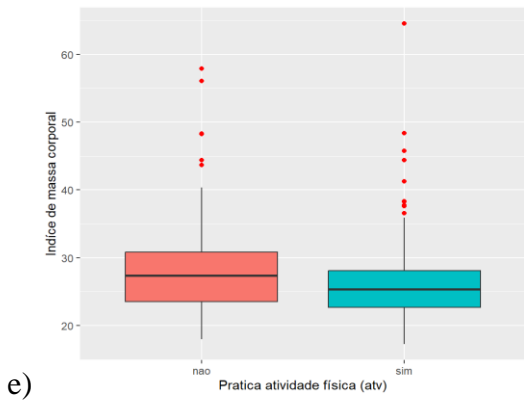
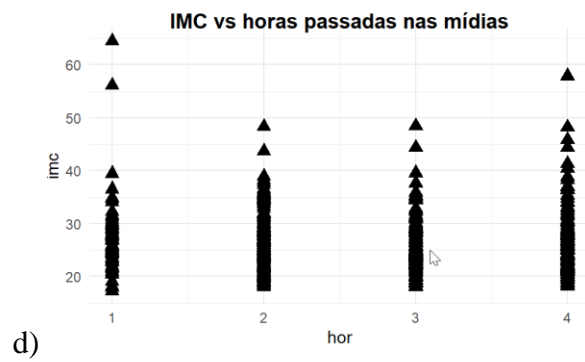
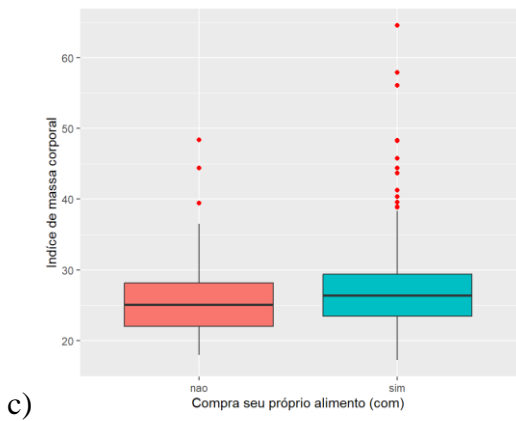
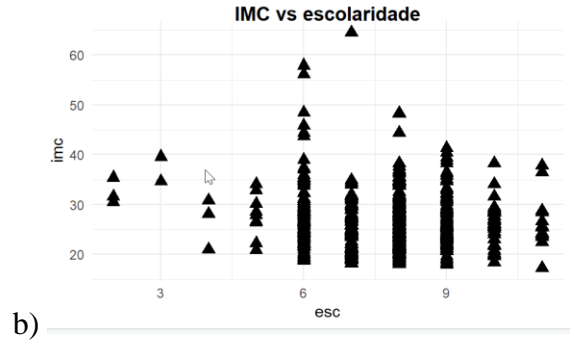
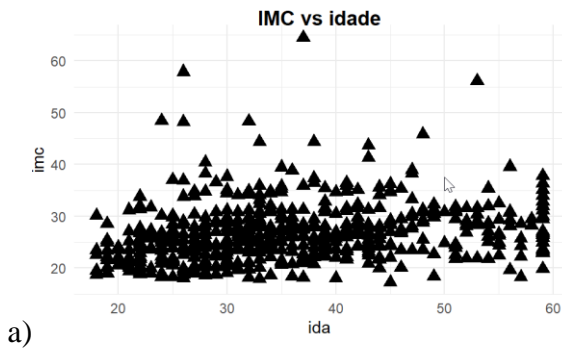
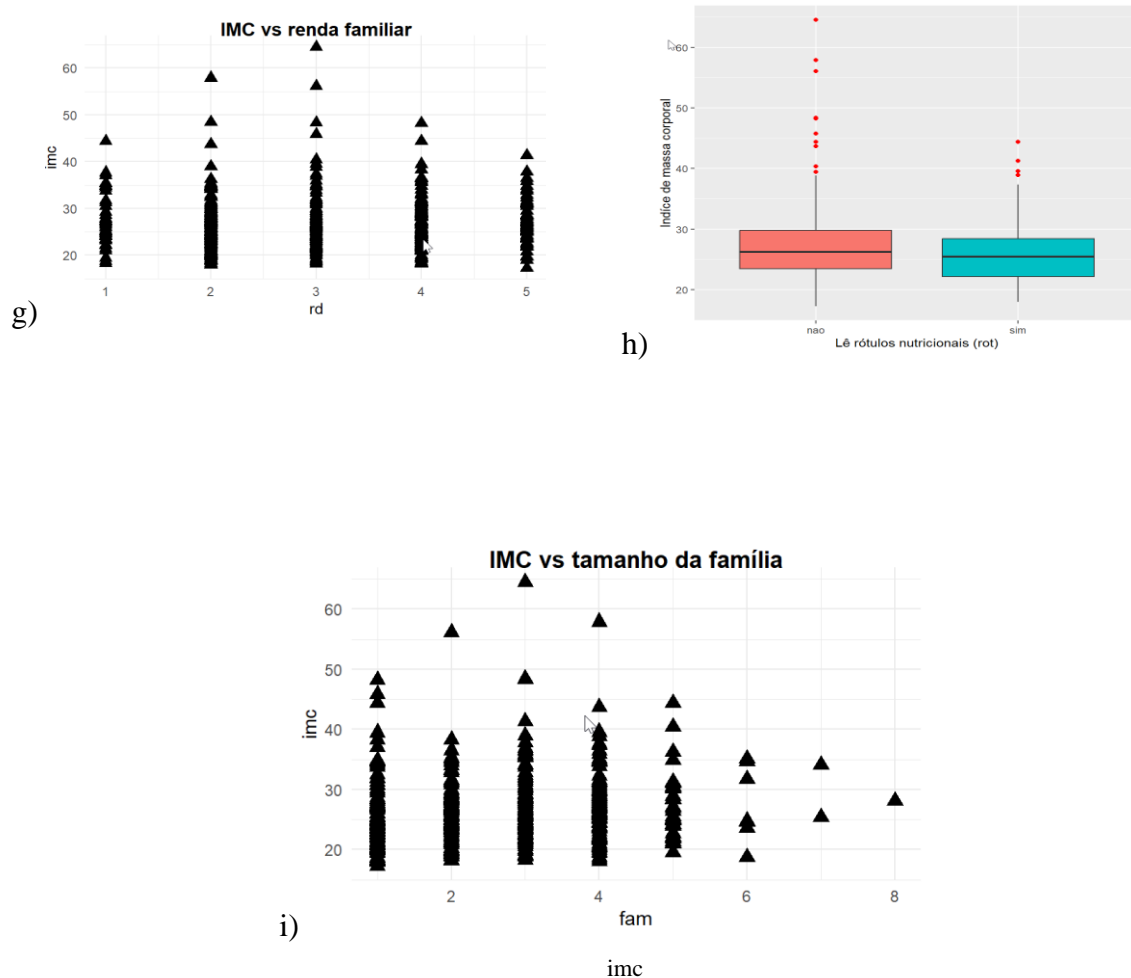


Figura 4 - Gráficos de dispersão e box plot da variável IMC em função das outras variáveis incluídas no estudo.

(Conclusão)



Fonte: Elaboração da autora.

Na Figura 4 é possível notar que a variável IMC em relação as outras variáveis não apresenta, de forma evidente, informações claras de uma possível relação. Fazer um ajuste de um modelo de regressão quantílica simples, para quantificar o valor esperado do IMC em todos seus quantis em função de outras variáveis permitirá explorar melhor.

4.2 ESPECIFICAÇÕES DO MODELO QUANTÍLICO

O modelo de regressão quantílica utilizado e já explicado anteriormente, possibilita se obter informações a respeito do τ -ésimo quantil condicional do IMC da população do sul de Minas Gerais.

$$IMC_j = \beta_0 + \beta_{rot}rot_j(\tau) + \beta_{ida}ida_j(\tau) + \beta_{rd}rd_j(\tau) + \beta_{esc}esc_j(\tau) + \beta_{fam}fam_j(\tau) + \beta_{com}com_j(\tau) + \beta_{pro}pro_j(\tau) + \beta_{hor}hor_j(\tau) + \beta_{atv}atv_{(\tau)j} + \epsilon_j \quad (41)$$

em que os erros ϵ_j são normalmente distribuídos com média zero e variância constante e as variáveis Rot é leitura de rótulo, Ida é idade, Rd é renda familiar, Esc é nível de escolaridade, Fam é tamanho da família, Com é compra seu próprio alimento, Pro é influenciado por propagandas, Hor é horas passadas nas mídias e Atv é atividade física regular.

A Tabela 3 apresenta as estimativas dos coeficientes de regressão quantílica para os quantis 0.15, 0.25, 0.5, 0.75, 0.85 e 0.9. Além das estimativas, são apresentados também os valores de p que cada estimativa apresentou e, para comparação, as estimativas dos respectivos parâmetros do modelo de regressão linear clássica usando o MMQO. Vale ressaltar que na literatura não existe um valor fixo de τ para ser avaliado. Dessa forma a escolha desses quantis busca analisar o efeito das covariáveis em diferentes pontos na cauda inferior, superior e no quantil mediano da distribuição condicional da variável resposta. Aqui, o valor de p é obtido assumindo que os resíduos são independentes e identicamente distribuídos (iid), tornando a análise mais robusta.

Para ser feito o MMQO foram realizados todos os testes propostos, apresentando que os dados não apresentam homocedasticidade e nem colinearidade de acordo com o Teste White e o VIF respectivamente.

Tabela 3- Estimativas para o modelo de regressão quantílica condicional, do MMQO e respectivas significâncias.

(Continua)

Variáveis	$\tau = 0.15$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.85$	$\tau = 0.90$	MMQO
<i>Intercepto</i>	19,17 0,0000***	19,33 0,0002***	24,04 0,0001***	25,1 0,0001***	24,67 0,0001***	23,86 0,0001***	23,49 0,0001***
<i>Idade</i>	0,09 0,0001***	0,11 0,0002***	0,11 0,0001***	0,11 0,0001***	0,13 0,0001***	0,15 0,0001***	0,1 0,0001***
<i>escolaridade</i>	-0,28 0,0976*	-0,33 0,1471	-0,26 0,1173	-0,11 0,6491	0,19 0,4847	-0,13 0,7463	-0,31 0,0862*
<i>Renda familiar</i>	0,25 0,2176	0,16 0,59	0,1 0,6403	0,15 0,6226	-0,39 0,2554	-0,52 0,2894	0,9 0,6846
<i>Tamanho da família</i>	0,45 0,0081***	0,35 0,1433	0,45 0,0083***	0,24 0,33	0,55 0,0405**	0,65 0,1058	0,39 0,0317**
<i>Atividade física</i>	-0,12 0,7791	0,02 0,9737	-1,4 0,0013***	-1,84 0,0031***	-1,53 0,0318**	-1,84 0,0711*	-1,19 0,0109**

Tabela 3- Estimativas para o modelo de regressão quantílica condicional, do MMQO e respectivas significâncias

(Conclusão)

Variáveis	$\tau = 0.15$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.85$	$\tau = 0.90$	MMQO
<i>propaganda</i>	-1,33 0,6454	0,7402 0,8557	-0,89 0,7623	-0,18 0,9667	2,2 0,6474	3,73 0,5877	0,24 0,9408
<i>Compra seu alimento</i>	0,86 0,0859*	0,77 0,1471	0,77 0,1318	0,73 0,3166	0,78 0,3497	1,93 0,1056	1,12 0,042**
<i>Leitura de rótulos</i>	-1,01 0,0201**	-1,33 0,1471	-0,61 0,0935*	-0,78 0,0178**	-1,88 0,0092***	-2,67 0,0097***	-1,32 0,0056***
<i>Horas passadas nas mídias</i>	0,26 0,2397	0,99 2,1471	-0,12 0,602	0,26 0,4189	0,91 0,0129**	0,69 0,1843	0,21 0,385

Fonte: Elaboração da autora.

Legenda: *** Nível de significância 1 %

** Nível de significância 5%

* Nível de significância 10%

Conforme apresentado na Tabela 3 a regressão quantílica ao longo de seus quantis estão ajustadas as faixas de valores, sendo os primeiros quantis responsáveis pelos valores mais baixos do IMC e assim por diante até que nos últimos quantis ficassem os maiores valores de IMC. Assim, as estimativas variam conforme esses valores. Portanto, é possível observar que na variável explicativa idade. As estimativas desta variável, em todos os quantis foi significativa e de efeito positivo. Em relação a variável resposta IMC, está situação sugere que quanto maior for a idade maior será o IMC, em toda sua distribuição e na sua média. Sendo um ganho de 0,11 kg/m² no IMC a cada unidade ano que aumenta na idade. Também é possível notar que ao longo dos quantis e na média (MMQO), a estimativa não variou muito, mostrando que a relação da idade com o IMC ocorre em toda sua distribuição. Em acordo com alguns estudos que identificam o envelhecimento como determinante de diversas modificações na composição corporal estando associado a aumento da massa gordurosa e mudanças no seu padrão de distribuição. Os riscos associados ao aumento da massa gordura ou decorrentes do baixo peso são diferentes nos idosos (SANTOS *et al.*, 2019).

Já a escolaridade apresentou significância apenas no primeiro quantil e no MMQO, além de apresentar uma variação em suas estimativas que tem efeito negativo que varia entre os valores de -0,33 até 0,11. Esta situação pode ser explicada de maneira empírica, devido que quanto melhor o nível de escolaridade melhor vai ser seu entendimento sobre alimentação e com isso provavelmente poderá ocorrer uma diminuição no IMC. De acordo com Bonanno *et al.* (2018) o nível de escolaridade também se encontra associado ao excesso de peso e

obesidade, dado que as pessoas com ensino superior apresentam um menor risco de desenvolver excesso de peso e obesidade, comparativamente às pessoas com ensino pré-escolar, mas não sendo observado nesse estudo.

A variável renda familiar e a influência da propaganda não apresentou nenhum tipo de significância em todos seu quantis e no MMQO, ambas apresentaram estimativas variáveis e com efeitos positivos e negativos em relação ao IMC. Já a variável, tamanho da família apresentou significância nos quantis 0.15, 0.50 e a 0.85 e no MMQO. Na variável compra seu próprio alimento, foi verificado a significância no quantil 0.15 e no MMQO, além de que essas variáveis apresentaram um efeito positivo em todas as suas estimativas ao longo dos quantis e na média. Ou seja, quanto maior for a família e compra seu próprio alimento, provavelmente seu IMC pode ser maior. O que leva a pesquisa de Anjos (2016) que relata que quanto mais pessoas convivem em um ambiente mais complicada fica o monitoramento dos responsáveis do que a criança come levando a uma propensão para a obesidade, além de que um local com muitas pessoas pode elevar o consumo de alimentos processados, pela necessidade de maior rapidez, praticidade e por ser produtos mais baratos.

A variável atividade física regular apresentou significância em quase todos os quantis, menos no quantis 0.15 e no 0.25, além de apresentar efeito negativo. Ou seja, quanto maior for atividade física maior será a propensão do indivíduo participante da pesquisa ter um IMC menor, o que fica de acordo com o trabalho científico de Anjos (2016) que mostra através de dados estatísticos que as práticas de atividade física se apresentam como um fator de proteção para a obesidade, reduzindo o acúmulo de calorias ao longo do dia.

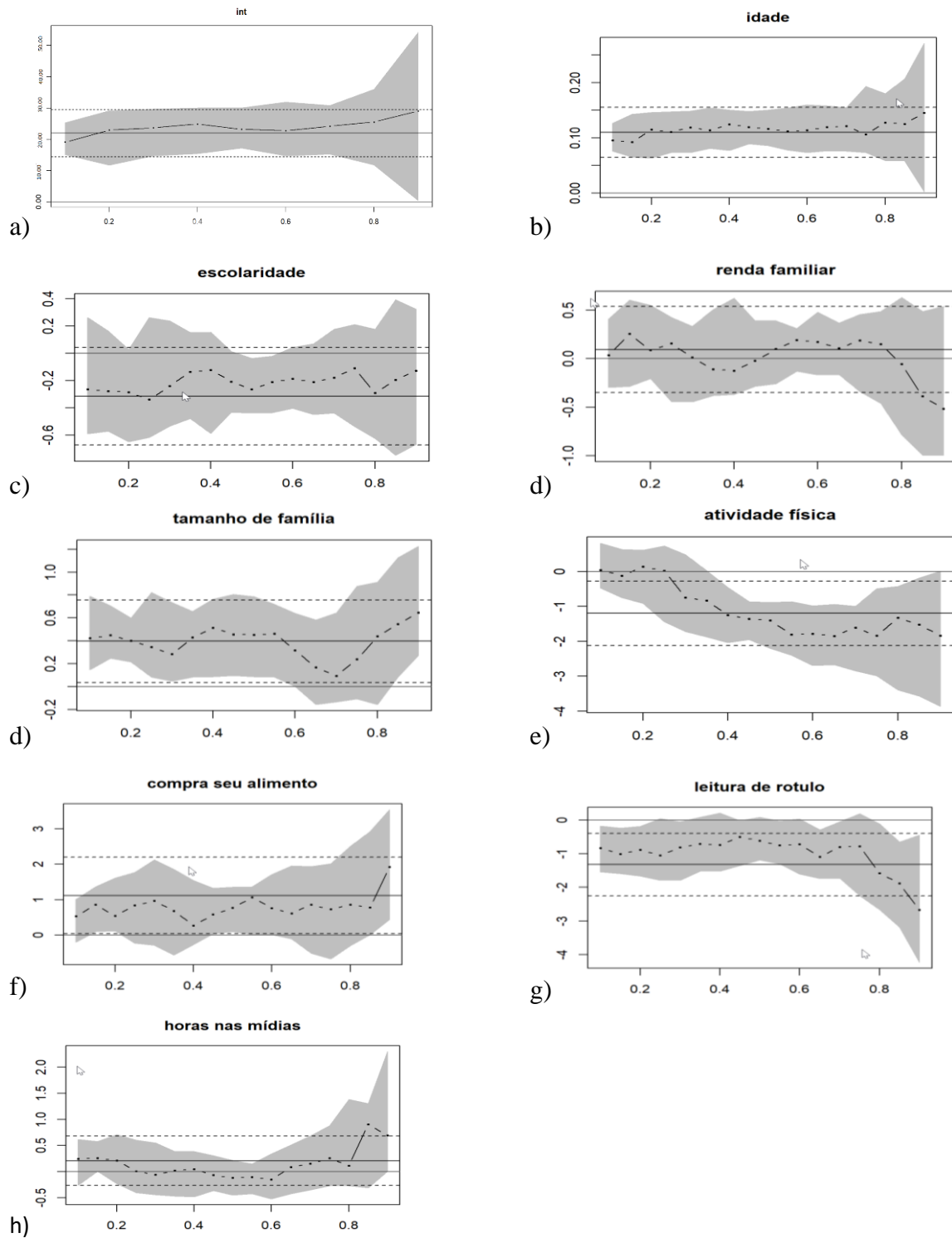
Já as horas passada nas mídias apresentou significância apenas no quantil 0.85, sendo de efeito positivo. Esta situação pode ser explicada empiricamente que quanto mais tempo se passa nas mídias ou na internet, menos tempo tem para cuidar do corpo e com isso um possível aumento do IMC será mais provável. De acordo com Anacleto *et al.* (2011) o desenvolvimento tecnológico começa a influenciar hábitos rotineiros do ser humano, funções simples como ir ao supermercado, pagar as contas, e até mesmo fazer compras de roupas e sapatos podem ser resolvidos pela internet. O uso excessivo de *smartphones* e outros dispositivos eletrônicos já vem sendo associado a inúmeros problemas de saúde física e mental. Segundo a pesquisa de Morrón (2013), indivíduos jovens que passam cinco ou mais horas diárias usando o celular estão 43% mais propensos a se tornarem obesos. O risco é ainda maior para as mulheres; elas apresentam o dobro do risco de ficar com excesso de peso em comparação com os homens. Os dados revelaram que 25% das pessoas acima do peso e 5% das obesas costumam utilizar o celular por muito tempo (MORRÓN,2013). A pesquisa ainda sugere que indivíduos que usam

o celular de forma excessiva estão duas vezes mais propensos a ingerir bebidas açucaradas, como refrigerante, *fast-food* e doces (MORRÓN,2013). Além disso, eles apresentam menor propensão a se exercitar. Esses comportamentos estão relacionados ao desenvolvimento de obesidade.

A variável leitura de rótulo apontou ter significância em quase todas as suas estimativas de quantis e no MMQO, menos no quantil 0.25, além de apresentar resultados parecidos com estudos como Kim, Nayga e Capps Jr. (2000), Loureiro *et al.* (2012), Bonnano *et al.* (2018) e Souza *et al.* (2019). A variável apresentou em suas estimativas efeito negativo em todos seus quantis e no MMQO, o que sugere que quanto mais as pessoas se dispuserem a lerem mais os rótulos e terem conhecimentos sobre os ingredientes que estão ingerindo, isto poderá influenciar seu IMC de modo que ele possa diminuir, por causa do efeito de uma alimentação mais conscientizado. Segundo Bonnano *et al.* (2018) a rotulagem dos alimentos, ao orientar o consumidor sobre a qualidade e a quantidade dos constituintes nutricionais dos produtos, auxilia escolhas alimentares apropriadas, sendo indispensável, no entanto, a fidedignidade das informações. Em seu estudo também foi apresentado estimativa negativa nos quantis da variável leitura de rótulos nutricionais, mostrando que quando aumenta a variável leitura de rótulos o IMC diminui, ou seja, podendo relacionar que o uso dos rótulos nutricionais podem ser essenciais para o combate a obesidade.

A Figura 5 permite visualizar as estimativas dos parâmetros da regressão quantílica e compará-las com as que seriam obtidas caso se ajustasse um modelo de regressão normal linear em função dos quantis. Ou seja, é possível visualizar de forma mais geral o comportamento das estimativas de cada parâmetro em diferentes quantis da distribuição condicional da variável resposta. Para comparação, as linhas pontilhadas são as estimativas e o intervalo com 90% de confiança do estimador de mínimos quadrados ordinários (MQO). Por outro lado, a área hachurada refere-se ao intervalo com 90% de confiança das estimativas da regressão quantílica. Dessa forma, para obter um gráfico mais detalhado definiu-se um conjunto de quantis mais amplo.

Figura 5 - Estimativas (linhas contínuas) e intervalo de confiança de 90% (área hachurada) para os coeficientes de regressão considerando um conjunto denso de quantis, $\tau = 0.05, 0.10, \dots, 0.90$.



Fonte: Elaboração da autora.

Nota: A linha horizontal em zero é marcada como referência. As linhas pontilhadas referem-se ao intervalo de confiança de mínimos quadrados ordinários com intervalo de confiança de 90%.

Como resultado é verificado que as estimativas do intercepto, denotado por int , crescem gradativamente em função dos quantis. Além disso, $\dot{\beta}_0(\tau)$ demonstra ser significativo para todos os quantis avaliados, uma vez que o valor zero não está contido no intervalo com 90% de confiança da regressão quantílica.

As estimativas dos coeficientes relacionados a ida mantêm valores próximos dentro do intervalo de confiança via MMQO em boa parte dos quantis, especificamente nos quantis inferiores e mediano. É observado também que suas estimativas crescem lentamente para $\tau = 0.75$. As estimativas relacionadas as variáveis esc , rd , atv , pro , rot e hor apresentam um comportamento oscilatório. Além disso, a maior parte de suas estimativas estão contidas no intervalo via MMQO. As estimativas relacionadas as variáveis fam e com assumem um padrão oscilatório positivo para os quantis. As estimativas relacionadas a variável rot assumem um padrão oscilatório negativo para o quantis. É visto também que em diversos pontos as estimativas da regressão quantílica se mantêm contidas dentro do intervalo de confiança de mínimos quadrados ordinários, não demonstrado haver diferença estatística entre as estimativas obtidas por meio das duas abordagens.

Após estimar os parâmetros do modelos um ponto importante é verificar se as diferenças observadas nos diferentes quantis é realmente significativa. Para isso, pode-se utilizar o teste de Wald (KOENKER; BASSETT, 1982), cuja hipótese nula é:

$$H_0 = \beta_1(\tau) = \beta_2(\tau) = \beta_3(\tau) = \beta_4(\tau) = \beta_5(\tau) = \beta_7(\tau) = \beta_8(\tau) = \beta_9(\tau) = 0,$$

contra a hipótese de que pelo menos um desses parâmetros é diferente de zero. Ou seja, é possível testar de forma simultânea múltiplos coeficientes e verificar se a função do τ -ésimo e θ -ésimo quantil condicional são diferentes. Este teste permite verificar a real necessidade de se utilizar a regressão quantílica visto que não é preciso ajustar diferentes modelos se o efeito das covariáveis é uniforme ao longo dos quantis (SANTOS, 2012).

Como resultado obteve-se um valor-p 0,001 ao testar todos os parâmetros de regressão nos quantis $\tau = (0.15, 0.25, 0.50, 0.75, 0.85, 0.90)$ de forma simultânea. Portanto, a rejeição da hipótese nula sugere que pelo menos um dos coeficientes é estatisticamente diferente dos demais, justificando o uso da regressão quantílica.

Com o objetivo de avaliar a qualidade do ajuste dos modelos utilizou-se o teste da falta de ajuste para verificar a hipótese nula de linearidade do modelo contra a hipótese alternativa de não linearidade (HE; ZHU, 2003). Outra maneira de verificar a qualidade do ajuste é a partir de uma medida similar ao pseudo- R^2 , introduzida por Koenker e Machado (1999), denotada por

$R^1(\tau)$. Na Tabela 4 é apresentado os valores da medida de bondade de ajuste e os valores de p referente ao teste da falta de ajuste.

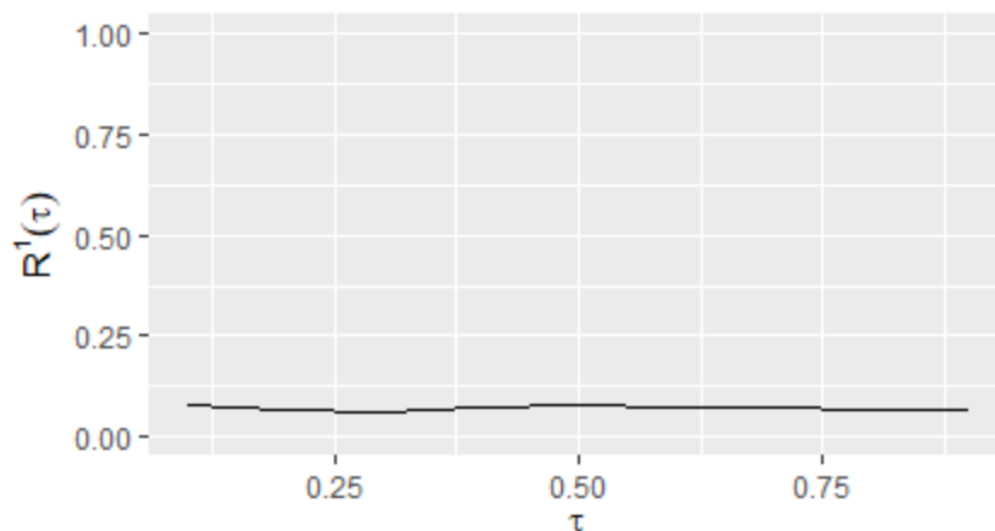
Tabela 4– Medida de ajuste e Valor-p referente ao teste da falta de ajuste.

τ -ésimo quantil	Medida de ajuste pseudo- R^2	Valor-p do teste de linearidade
0.15	0,098	0,064
0.25	0,085	0,132
0.50	0,094	0,002
0.75	0,078	0,075
0.85	0,075	0,032
0.90	0,077	0,230

Fonte: Elaboração da autora.

Na tabela 4 pode-se verificar que os modelos ajustados nos quantis mediano e 0.85 rejeitaram a hipótese de linearidade, uma vez que seu valor-p é menor do que o nível de significância de 5%. Em relação a medida de qualidade de ajuste, $R^1(\tau)$, é possível notar um decrescimento dos seus valores ao longo dos quantis, nos levando a entender que os modelos explicam menos da variável dependente, nos levando a entender que a contribuição não é a mesma, embora os valores sejam próximos. Ou seja, suas contribuições são maiores na cauda inferior da distribuição. Vale ressaltar que a maior contribuição é avaliada no quantil 0.15 apresentando um valor de 0,098 Na Figura 6 é apresentado o gráfico da medida do pseudo- R^2 ao longo dos quantis ajustados.

Figura 6 - Medida da bondade de ajuste, $R^1(\tau)$, para o modelo de regressão quantílica

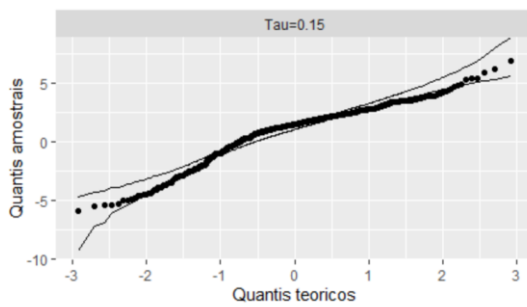


Fonte: Elaboração da autora.

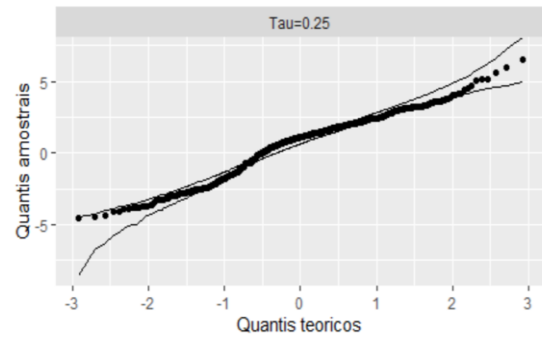
Para avaliar a qualidade do ajuste por meio da análise dos resíduos pode-se utilizar o gráfico de envelope simulado exibido na Figura 7.

Figura 7 - Gráfico de envelope simulado para modelos de regressão quantílica considerando os resíduos quantílicos e os quantis condicionais 0.15, 0.25, 0.50, 0.75, 0.85 e 0.90 de IMC.

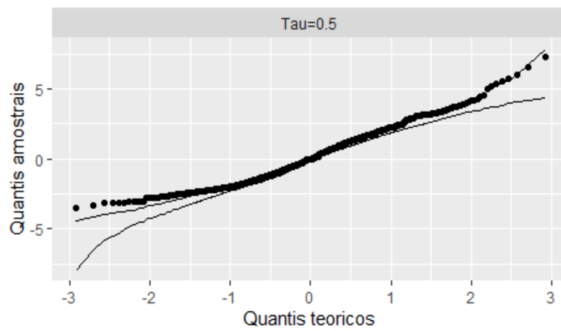
a)



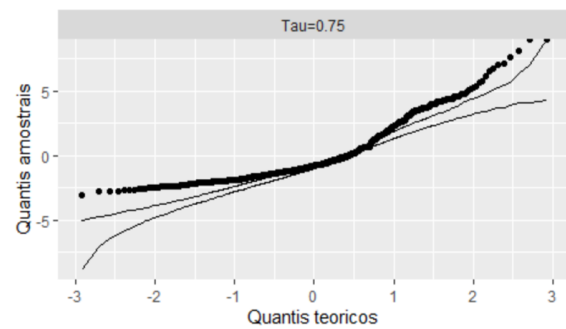
b)



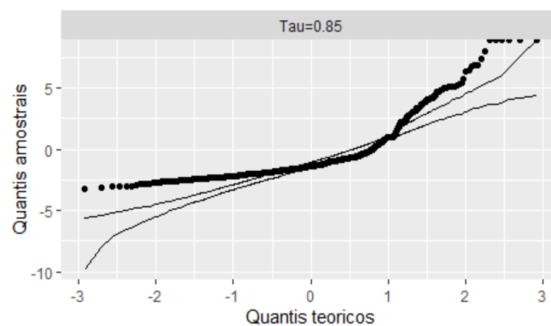
c)



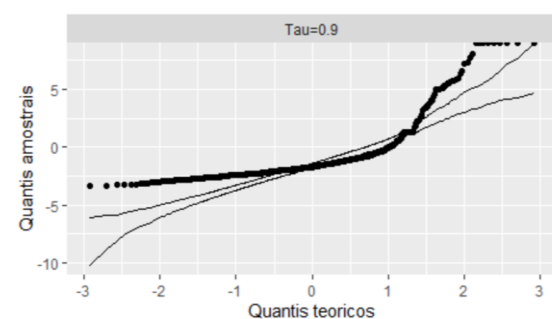
d)



e)



f)



Fonte: Elaboração da autora.

Como apresentado em Santos (2012) o uso dessa técnica requer a suposição de alguma distribuição probabilidade para a variável resposta. Um resultado bastante conhecido na literatura diz que o estimador de mínimos quadrados ordinários coincide com o estimador de

máxima verossimilhança quando os erros seguem distribuição normal padrão. De forma similar, para modelos de regressão quantílica tem-se que o estimador dos desvios absolutos ponderados coincide com o estimador de máxima verossimilhança quando erros seguem distribuição Laplace Assimétrica, $LA(\mu, \sigma, \tau)$, com parâmetros μ , σ e τ (YU; ZHANG, 2005). Infelizmente se observa uma falta de ajuste nos quantis superiores mas qualifica o ajuste dos quantis inferiores.

5 CONSIDERAÇÕES FINAIS

Neste estudo, avaliou-se a relação entre a leitura dos ingredientes nos rótulos dos alimentos e os níveis de IMC entre adultos. Usando um estimador de regressão quantílica condicional (RQC), destaca-se a existência de uma relação negativa entre a leitura de rótulos de alimentos e o IMC de adultos. Foi descoberto que os coeficientes que capturam a relação entre a leitura de rótulos e o IMC são maiores em níveis de IMC mais altos, os resultados apoiam políticas para promover escolhas alimentares informadas, pois a leitura de rótulos pode ter um efeito mais acentuado na redução do IMC para indivíduos com sobrepeso e obesidade. Além disso, nossos resultados se somam à literatura anterior, destacando como a relação entre a leitura dos ingredientes nos rótulos e o IMC adulto é mais acentuada entre os indivíduos com maior probabilidade de sobrepeso e obesidade, e que são menos propensos a usar rótulos de alimentos, como menor rendimento, com menor escolaridade e que não praticam práticas um estilo de vida saudável (prática esportiva). Assim, nossos resultados fornecem suporte para políticas que promovam o uso de rótulos de alimentos e o fornecimento de informações adicionais e de fácil compreensão, pois o uso de rótulos nutricionais desempenha um papel importante

As evidências encontradas apontam que ainda existem obstáculos entre o consumidor e as informações nutricionais contidas nos rótulos de alimentos, e um desses obstáculos é, principalmente, a falta de entendimento e compreensão dos dados disponíveis nos rótulos. Ainda, a análise desses estudos mostrou a importância de políticas públicas para implantação de rótulos nutricionais que estejam vinculados às mais recentes recomendações nutricionais, além da linguagem propícia a cada população, pois grupos populacionais distintos comportam-se de maneira diferente, dependendo do modelo e da rotulagem nutricional.

A partir dos resultados explicitados nesta dissertação, constata-se que a rotulagem dos alimentos é uma ferramenta potencialmente útil no manejo de como pessoas com DCNT podem fazer escolhas alimentares mais saudáveis e de acordo com suas limitações. O uso da rotulagem na frente da embalagem em forma de advertência e semáforo são modelos de rótulos mais acessíveis às diferentes classes sociais. No entanto, uma reformulação de uma linguagem alinhada ao conhecimento da população brasileira é uma medida que poderia ser amplamente adotada, a fim de beneficiar todos os consumidores

REFERÊNCIAS

- ASSOCIAÇÃO BRASILEIRA DE AGÊNCIAS DE PUBLICIDADE. **Números oficiais da indústria da comunicação e seu impacto na economia brasileira: 2008**. São Paulo, 2008. Disponível em: http://www.abapnacional.com.br/images/publicacoes/abap_ibge_web.pdf. Acesso em: 12 nov. 2022.
- AGHA, M.; AGHA, R. The rising prevalence of obesity: part A: impact on public health. **International Journal Of Surgery. Oncology**, v. 2, n. 7, p. e17, 2017.
- ALMEIDA, A. T. C. de. Determinantes dos piores e melhores resultados educacionais dos alunos da rede pública de ensino fundamental no Brasil. **Planejamento e Políticas Públicas**, n. 42, p. 42, 2014.
- AMEMIYA, T. **Advanced econometrics**. Cambridge: Harvard University, 1985.
- ANJOS, L. **Obesidade e saúde pública**. São Paulo, FIOCRUZ, 2016.
- ANTIPORTA, A. *et al.* Health and nutrition of indigenous and nonindigenous children in the Peruvian Amazon/Situacion de salud y nutricion de ninos indigenas y ninos no indigenas de la Amazonia peruana. **Revista Panamericana de Salud Pública**, v. 38, n. 1, p. 49-57, 2015.
- FIESP/CIESP**. Federação das Indústrias do Estado de São Paulo/ Centro de Indústrias do Estado de São Paulo. “A mesa dos brasileiros: transformação, confirmações e condições”. São Paulo, 2017. Disponível em: <http://hotsite.fiesp.com.br/amesadosbrasil/>. Acesso: 20 jul. 2020.
- ANVISA. Consulta Pública no 71, de 10 de nov. de 2006. Proposta de regulamento técnico sobre oferta, propaganda, publicidade, informação e outras práticas correlatas cujo objeto seja a divulgação ou promoção de alimentos com quantidades elevadas de açúcar, de gordura saturada, de gordura trans, de sódio e de bebidas com baixo teor nutricional, quaisquer que sejam as formas e meios de sua veiculação. **Diário Oficial da União, Brasília 2006**. 2006.
- ANVISA. Resolução RDC nº 24, de 15 de junho de 2010. Dispõe sobre a oferta, propaganda, publicidade, informação e outras práticas correlatas cujo objetivo seja a divulgação e a promoção comercial de alimentos considerados com quantidades elevadas de açúcar, gordura saturada, gordura trans, sódio e de bebidas com baixo teor nutricional. **Diário Oficial da União 2010**; 14 jul. 2010.
- ANVISA. Gerência-Geral de Alimentos. **Relatório preliminar de análise de impacto regulatório sobre rotulagem nutricional**. Brasília. Maio de 2018. Disponível em: <http://portal.anvisa.gov.br/documents/219201/5197494/An%C3%A1lise+de+Impacto+Regulatório+sobre+Rotulagem+Nutricional.pdf/c63f2471-4343-481d-80cb-00f4b2f72118?version=1.0>. Acesso em: 19 jul. 2020.
- ANVISA. **Anvisa aprova norma sobre rotulagem nutricional**. Brasília-DF, 2020. 07 de out. de 2020. Disponível em: <https://www.gov.br/anvisa/pt-br/assuntos/noticias-anvisa/2020/aprovada-norma-sobre-rotulagem-nutricional>. Acesso em: 01 dez. 2020.

BAIRD, M. F. O lobby na regulação da publicidade de alimentos da Agência Nacional de Vigilância Sanitária. **Revista de Sociologia e Política**, v. 24, p. 67-91, 2016.

BANDEIRA, L. M. *et al.* Desempenho e percepção sobre modelos de rotulagem nutricional frontal no Brasil. **Revista de Saúde Pública**, v. 55, p. 19, 2021.

BANTERLE, A.; CAVALIERE, A. Is there a relationship between product attributes, nutrition labels and excess weight? Evidence from an Italian region. **Food Policy**, v. 49, p. 241-249, 2014.

BERDEGUÉ, J; AGUIRRE, P. Obesidade que mata. **O Globo**, Rio de Janeiro, 12 fev. 2018. Disponível em: <https://oglobo.globo.com/opiniao/obesidade-que-mata-22386691>. Acesso em: 18 jul. 2020.

BLITSTEIN, J. L.; EVANS, W. D. Use of nutrition facts panels among adults who make household food purchasing decisions. **Journal of Nutrition Education and Behavior**, v. 38, n. 6, p. 360-364, 2016.

BRANDÃO, I. S.; SOARES, D. J. **A obesidade, suas causas e consequências para a saúde**. 2018. 17f. Monografia (Especialização em Saúde da Família) – Faculdade de Enfermagem. Universidade da Integração Internacional da Lusofonia Afro-Brasileira, São Francisco do Conde, 2018.

BRASIL. Ministério da Saúde. Decreto-Lei nº 986, de 21 de outubro de 1969. Institui normas básicas sobre alimentos. **Diário Oficial da União 1969**. Brasília, 21 out. 1969, 6p.

BRASIL. Presidência da República. Lei no 9.782, de 26 de janeiro de 1999. Define o Sistema Nacional de Vigilância Sanitária, cria a Agência Nacional de Vigilância Sanitária, e dá outras providências. **Diário Oficial da União 1999**. Brasília, 27 jan. 1999, 5p.

BRASIL. Câmara dos Deputados. **Projeto Lei 4328/2016**. Brasília, 04 fev. 2016, 15p. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2076868>. Acesso em: 19 jul. 2020.

BRASIL. Congresso Nacional. **Lei Nº 13.666/2018**. Brasília, 16 de maio 2018, 2p. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13666.htm. Acesso em: 20 jul. 2020.

BRASIL. Ministério da Saúde. Estimativas sobre frequência e distribuição sociodemográfica de fatores de risco e proteção para doenças crônicas nas capitais dos 26 estados brasileiros e no Distrito Federal em 2018. **Vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico (Vigitel) Brasil 2018**. Brasília-DF. 2019, 139p. Disponível em: <http://portalarquivos2.saude.gov.br/images/pdf/2019/jul./25/vigitel-brasil-2018.pdf>. Acesso em: 18 jul. 2020.

BRASIL. Ministério da Saúde. e-Gestor atenção Básica, Informação e gestão da Atenção Básica. **Históricos – ACS- Macrorregião de 2011-2019**. Secretaria de Atenção Primária à Saúde (SAPS). Brasília 2021, 160 p. Disponível em:

<https://egestorab.saude.gov.br/paginas/ acessoPublico/relatorios/relHistoricoCoberturaConsolidado.xhtml>. Acesso em: 20 jul. 2022.

BRASIL. Ministério da Educação. Programa Nacional de Alimentação Escolar. **Educação alimentar e nutricional – EAN**. Brasília, 2020. 189p. Disponível em: <https://www.fn-de.gov.br/programas/pnae/pnae-eixos-de-atuacao/pnae-educacao-alimentar-nutricional>. Acesso em: 20 jul. 2020.

BONANNO, A. *et al.* Food labels and adult BMI in Italy—An unconditional quantile regression approach. **Food Policy**, v. 74, p. 199-211, 2018.

BORAH, B. J.; BASU, A. Highlighting differences between conditional and unconditional quantile regression approaches through an application to assess medication adherence. **Health Econ**, v. 22, p. 1052-1070, 2013.

BOTTAI, M.; CAI, B.; MCKEOWN, R. Logistic quantile regression for bounded outcomes. **Statistics in Medicine**, v. 29, n. 2, p. 309–317, 2010.

BUCHINSKY, M. The dynamics of changes in the female wage distribution in the USA: a quantile regression approach. **Journal of Applied Econometrics**, v. 13, n. 1, p. 1-30, 1998.

CAIVANO, Simone *et al.* Conflitos de interesses nas estratégias da indústria alimentícia para aumento do consumo de alimentos ultraprocessados e os efeitos sobre a saúde da população brasileira. **DEMETRA: Alimentação, Nutrição & Saúde**, v. 12, n. 2, p. 349-360, 2017.

CAMERON, A. C. *et al.* Modelling the differences in counted outcomes using bivariate copula models with application to mismeasured counts. **The Econometrics Journal**, v. 7, n. 2, p. 566-584, 2004.

CAMPBELL-SCHERER, D.; SHARMA, A. M. Improving obesity prevention and Management in Primary Care in Canada. **Current Obesity Reports**, v. 5, n. 3, p. 327-332, 2016.

CAMPOS, S *et al.* Nutrition labels on pre-packaged foods: a systematic review. **Public Health Nutrition**, v. 14, n. 8, p. 1496-1506, 2011.

CANCER RESEARCH UK. Millennials top obesity chart before reaching middle age. **Cancer News**, United Kingdom, 2018. Disponível em: <https://www.cancerresearchuk.org/about-us/cancer-news/press-release/2018-02-26-millennials-top-obesity-chart-before-reaching-middle-age> . Acesso em: 18 jul. 2020.

CHAER, G.; DINIZ, R. R. P.; RIBEIRO, E. A. A técnica do questionário na pesquisa educacional. **Revista Evidência**, v. 7, n. 7, 2012.

CHEN, K., *et al.* Analysis of least absolute deviation. **Biometrika**, v. 95, p.107–122. 2008.

COUTINHO, J. G. *et al.* A organização da vigilância alimentar e nutricional no sistema único de saúde: histórico e desafios atuais. **Revista Brasileira de Epidemiologia**, v. 12, p. 688-699, 2009.

DRAPER, N. R.; SMITH, H. **Applied regression analysis**. Hoboken: John Wiley & Sons, 1998.

DRICHOUTIS, A. C. *et al.* Would consumers value food-away-from-home products with nutritional labels?. **Agribusiness: An International Journal**, v. 25, n. 4, p. 550-575, 2009.

DRICHOUTIS, A. C. *et al.* Nutritional labeling. **The Oxford Handbook of The Economics of Food Consumption And Policy**, p. 520-545, 2011.

FIRPO, S. *et al.* Unconditional quantile regressions. **Econometrica**, v. 77, n.3, p. 953-973, 2009.

FLEGAL, K. M. *et al.* Trends in obesity among adults in the United States, 2005 to 2014. **Jama**, v. 315, n. 21, p. 2284-2291, 2016.

FLORIDO, L. M. *et al.* Combate à obesidade: estratégias comportamentais e alimentares. **Cadernos da Medicina-UNIFESO**, v. 2, n. 2, 2019.

FUNDAÇÃO INSTITUTO OSWALDO CRUZ. Centro Especializado em Obesidade e Diabetes. **Tratamento da obesidade**. São Paulo, 2020. Disponível em: <https://centrodeobesidadeediabetes.org.br/tudo-sobre-obesidade/tratamento-da-obesidade/#:~:text=IMC%20entre%2025%2C0%20e,%2Fm2%3A%20obesidade%20grau%20III..> Acesso em: 18 nov. 2022.

GIBBONS, J. D.; CHAKRABORTI, S. **Nonparametric statistical inference**. Boca Raton: CRC Press, 2014.

GRUPO DE MÍDIA SÃO PAULO. **Mídia dados Brasil 2011**. São Paulo, 2011. Disponível em: <http://midiadados.digitalpages.com.br/home.aspx?edicao=3>. Acesso em: 12 nov. 2022.

GUTENBRUNNER, C.; JURECKOVA, J. Regression rank scores and regression quantiles, **The Annals of Statistics**, v. 20, p. 305–330, 1992.

GUTENBRUNNER, C., *et al.* Tests of linear hypotheses based on regression rank scores. **Journal of Nonparametric Statistics**, v. 2, p. 307–331, 1993.

HAO, L.; NAIMAN, D. **Quantile regression**. [S.l.]: Sage Publications, 2007.

HARDLE, W.; MANSKI, C. F. Nonparametric and semiparametric approaches to discrete response analysis. **Journal of Econometrics**, v. 58, n. 1-2, p. 1-2, 1993.

HECKMAN, J.; TOBIAS, J. L.; VYTLACIL, E. Simple estimators for treatment parameters in a latent-variable framework. **Review of Economics and Statistics**, v. 85, n. 3, p. 748-755, 2003.

HE, X.; ZHU, L. A lack-of-fit test for quantile regression. **Journal of the American Statistical Association**, v. 98, n. 464, p. 1013–1022, 2003.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Pesquisa de orçamentos familiares 2008-2009**: antropometria e estado nutricional de crianças, adolescentes e adultos no Brasil. Rio de Janeiro, 2010.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Prática de esporte e atividade física no Brasil 2018-2019**. Rio de Janeiro, 2019.

INSTITUTO NACIONAL DE CÂNCER (BRASIL). Posicionamento do Instituto Nacional de Câncer José Alencar Gomes da Silva acerca do sobrepeso e obesidade. **Revista Brasileira de Cancerologia**, Rio de Janeiro, v. 63, n. 1, p. 7-12, 2017.

KEERY, H. *et al.* The impact of appearance-related teasing by family members. **Journal of Adolescent Health**, v. 37, n. 2, p. 120-127, 2015.

KIM, S. Y.; NAYGA JR, R. M.; CAPPS JR, O. The effect of food label use on nutrient intakes: an endogenous switching regression analysis. **Journal of Agricultural and Resource Economics**, p. 215-231, 2000.

KOENKER, R.; BASSETT, G. Regression quantiles. **Econometrica: Journal of The Econometric Society**, p. 33-50, 1978.

KOENKER, R.; BASSETT, G. Robust tests for heteroscedasticity based on regression quantiles. **Econometrica: Journal of The Econometric Society**, p. 43-61, 1982.

KOENKER, R.; MACHADO, J. A. F. Goodness of fit and related inference processes for quantile regression. **Journal of The American Statistical Association**, v. 94, n. 448, p. 1296-1310, 1999.

KOENKER, R. Quantile regression for longitudinal data. **Journal of Multivariate Analysis**, v. 91, n. 1, p. 74-89, 2004.

KOENKER, R. Inequality constrained quantile regression. **Sankhyā: The Indian Journal of Statistics**, p. 418-440, 2005.

KOENKER, R.; HALLOCK, K. F. Quantile regression. **Journal of Economic Perspectives**, v. 15, n. 4, p. 143-156, 2001.

LEDO, A.; CHAUD, D. M. A.; ABREU, E. S. Análise dos produtos alimentícios apresentados nas propagandas dos canais infantis de televisão por assinatura. **Revista Univap**, v. 25, n. 47, p. 70-85, 2019.

LOUREIRO, M. L. *et al.* Efeitos dos rótulos nutricionais na obesidade. **Economia Agrícola**, v. 43, n. 3, p. 333-342, 2012.

MACLEAN, J. C.; WEBBER, D. A.; MARTI, J. An application of unconditional quantile regression to cigarette taxes. **Journal of Policy Analysis and management**, v. 33, n. 1, p. 188-210, 2014.

MARTINELLI, S. S.; CAVALLI, S. B. Alimentação saudável e sustentável: uma revisão narrativa sobre desafios e perspectivas. **Ciência & Saúde Coletiva**, v. 24, p. 4251-4262, 2019.

MARTINS, A. P. B. É preciso tratar a obesidade como um problema de saúde pública. **Revista de Administração de Empresas**, v. 58, p. 337-341, 2018.

MARTÍNEZ-GONZÁLEZ, M. Á. *et al.* Physical inactivity, sedentary lifestyle and obesity in the European Union. **International Journal of Obesity**, v. 23, n. 11, p. 1192-1201, 1999.

MCCULLAGH, P.; NELDER, J. **Generalized linear models**. London: Chapman and Hall, 1989.

MENDONÇA, V. F. A relação entre o sedentarismo, sobrepeso e obesidade com as doenças cardiovasculares em jovens adultos: uma revisão da literatura. **Saúde e Desenvolvimento Humano**, v. 4, n. 1, p. 79-90, 2016.

MICHAIL, N. **Italy raises red flag once more over UK's**. London: Traffic Light Label. 2016. Disponível em: <https://www.foodnavigator.com/Article/2016/03/16/Italy-raises-red-flag-oncemore-over-UK-s-traffic-light-label>. Acesso em: 10 ago. 2022.

MICHAIL, N. **Six industry giants to launch UK-style traffic**. London: Light Labels in Europe. 2017a. Disponível em: <http://www.foodnavigator.com/Policy/Six-industry-giants-to-launch-UK-styletraffic-light-labels-in-Europe>. Acesso em: 10 ago. 2022.

MICHAIL, N. **Nutriscore to be France's**. London: Official Nutrition Label. 2017b. Disponível em: <https://www.foodnavigator.com/Article/2017/03/16/5-C-NutriScore-to-be-France-sofficial-nutrition-label>. Acesso em: 10 ago. 2022.

MINAS GERAIS. Secretaria de estado de saúde de Minas Gerais. **SES-MG inicia programas de combate a obesidade**. Belo Horizonte, 2018. 240p. Disponível em: <https://www.saude.mg.gov.br/>. Acesso em: 21 jul. 2020.

MOSTELLER, F.; TUKEY, J. W. **Data analysis and regression: a second course in statistics**. Boston: Addison-Wesley, 1977.

MORRÓN, M. M. Frecuencia de obesidad y sobrepeso en una institución de salud de la ciudad de Barranquilla. **Ciencia e Innovación en Salud**, v.1, n.1, 2013.

MUA, Y.; HEA, X. Power transformation toward a linear regression quantile. **Journal of the American Statistical Association**, v. 102, n. 477, p. 37-41, 2007.

NASCIMENTO, A. C. C. *et al.* Eficiência técnica da atividade leiteira em Minas Gerais: uma aplicação de regressão quantílica. **Revista Brasileira de Zootecnia**, v. 41, p. 783-789, 2012.

NORTHSTONE, K. *et al.* Are dietary patterns in childhood associated with IQ at 8 years of age? A population-based cohort study. **J Epidemiol Community Health**, v. 66, n. 7, p. 624-628, 2012.

- ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Obesity and overweight**. Rio de Janeiro, 2018. Disponível em: <https://www.who.int/en/news-room/fact-sheets/detail/obesity-and-overweight>. Acesso: 19 jul. 2020.
- PARQUE, T. Marketing direto e a estrutura das vendas agrícolas: uma abordagem de regressão quantílica incondicional. **Journal of Agricultural and Resource Economics**, v. 40, p. 266-284. 2015.
- PORTUGAL, M. S. **Modelos de parâmetros variáveis**: uma resenha crítica. [S.l: s.n], 1993.
- POWELL, J. Censored regression quantiles. **Journal of Econometrics**, v. 32, p. 143–155, 1986.
- R CORE TEAM. **R: A language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2022. Disponível em: <https://www.R-project.org>. Acesso em: 10 nov. 2022.
- RIO DE JANEIRO (ESTADO). Legislação Municipal do Rio de Janeiro/RJ. **Lei nº 5766**. Rio de Janeiro, 30 jun. 2014. 6 p. Disponível em: <https://leismunicipais.com.br/a1/rj/r/rio-de-janeiro/lei-ordinaria/2014/576/5766/lei-ordinaria-n-5766-2014-institui-o-estatuto-dos-portadores-de-obesidade-no-ambito-do-municipio-e-da-outras-providencias?r=p>. Acesso: 19 jul. 2020.
- RODRIGUES, K. A. S. **Diagnóstico em regressão L_1** . 2019. 100f. Dissertação (Mestrado em Estatística) – Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo - SP, 2019.
- ROQUE, R. C. **Estudo sobre a empregabilidade da previsão do índice BOVESPA usando Redes Neurais Artificiais**. 2009. 102f. Monografia (Trabalho de Conclusão de Curso em Eletrônica e Computação) – Instituto de Eletrônica e de Computação, Universidade Federal do Rio de Janeiro, Rio de Janeiro-RJ, 2009.
- ROSSI, F.; REIS, G. M. F.; FREITAS, W. J. **Cirurgiões debatem cirurgia bariátrica no SUS**. São Paulo: Sociedade Brasileira de Cirurgia Bariátrica e Metabólica, 2019. Disponível em: <https://www.scbcm.org.br/cirurgioes-debatem-cirurgia-bariatrica-no-sus/#:~:text=O%20tempo%20de%20espera%20para,que%20o%20tempo%20seja%20menor>. Acesso em: 19 jul. 2020.
- SANTOS, B. **Modelos de regressão quantílica**. 2012. 90f. Dissertação (Mestrado em Estatística) – Instituto de Matemática Estatística, Universidade de São Paulo, São Paulo - SP, 2012.
- SANTOS, T. N. R. *et al.* Prevalência do excesso de peso e fatores de risco para obesidade em adultos. **Revista Eletrônica Acervo Saúde**, n. 22, p. e308-e308, 2019.
- SERRA, G. M. A.; SANTOS, E. M. Saúde e mídia na construção da obesidade e do corpo perfeito. **Ciência & Saúde Coletiva**, v. 8, n. 3, p. 691-701, 2003.

- SHELTON, N. J.; KNOTT, C. S. Association between alcohol calorie intake and overweight and obesity in English adults. **American Journal of Public Health**, v. 104, n. 4, p. 629-631, 2014.
- SILVA, T. B. C. *et al.* **Modelos de rotulagem nutricional frontal de alto conteúdo de nutrientes críticos comparado a outros modelos de rotulagem frontal nutricional: uma revisão sistemática.** Brasília- DF, 2019. 52 p.
- SILVA, B. L. *et al.* Rotulagem nutricional de alimentos: utilização e compreensão entre estudantes. **Brazilian Journal of Development**, v. 6, n. 11, p. 90688-90699, 2020.
- SOUZA, S. V.C.; JUNQUEIRA, R. G. A procedure to assess linearity by ordinary least squares method. **Analytica Chimica Acta**, v. 552, n.1-2, p. 25-30, 2005.
- SOUZA, S. A. **Modelagem da obesidade adulta nas nações: uma análise via modelos de regressão beta e quantílica.** 2017. 75f. Dissertação (Mestrado em Modelos de Decisão e Saúde) - Universidade Federal da Paraíba, João Pessoa, 2017.
- SOUZA, S. *et al.* **Modelagem da obesidade adulta nas nações adultas via modelo de regressão quantílica.** RBONE-Revista brasileira de obesidade, nutrição e emagrecimento, v. 13, n. 82, p. 992-1004, 2019.
- SPANOS, A. **Probability theory and statistical inference: econometric modeling with observational data.** Cambridge,UK : Cambridge University Press , 1999.
- STUTE, W., Nonparametric Model Checks for Regression, **The Annals of Statistics**, v. 25, p. 613–641. 1997
- VARIYAM, J.N.; CAWLEY, J. **Nutrition Labels and Obesity.** NBER Working Paper No. 11956. Disponível em: <http://www.nber.org/papers/w11956>. Acesso em: 10 ago. 2022.
- VALLGÅRDA, S. *et al.* Europe follow the US and declare obesity a disease?: a discussion of the so-called utilitarian argument. **European Journal of Clinical Nutrition**, v. 71, n. 11, p. 1263-1267, 2017.
- VILLAREAL, D. T. *et al.* Obesity in older adults: technical review and position statement of the American Society for Nutrition and NAASO, The Obesity Society. **The American Journal Of Clinical Nutrition**, v. 82, n. 5, p. 923-934, 2015.
- YEN, S. T.; SHAW, W. D.; YUAN, Y. Cigarette smoking and self-reported health in China. **China Economic Review**, v. 21, n. 4, p. 532-543, 2010.
- YEN, S. T.; BRUCE, D. J.; JAHNS, L. Supplemental nutrition assistance program participation and health: Evidence from low-income individuals in Tennessee. **Contemporary Economic Policy**, v. 30, n. 1, p. 1-12, 2012.
- YU, K.; MOYEED, R. A. Bayesian quantile regression. **Statistics & Probability Letters**, v. 54, n. 4, p. 437-447, 2001.
- YU, K.; ZHANG, J. A three-parameter asymmetric laplace distribution and its extension. **Communications in Statistics - Theory and Methods**, v. 34, p. 1867–1879, 2005.

APÊNDICE I

Seção 2 de 5

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Você está sendo convidado (a) a participar, como voluntário(a), da pesquisa – AVALIAÇÃO DO CUSTO DA OBESIDADE PARA O SISTEMA ÚNICO DE SAÚDE, COM DESENVOLVIMENTO DE FERRAMENTA DE MONITORAMENTO, no caso de você concordar em participar, favor marcar com "aceito" ao final do documento. Sua contribuição compreenderá responder a um instrumento para avaliar o impacto econômico da obesidade no seu município

PESQUISADOR(A) RESPONSÁVEL: Luciana Azevedo Cristina Garcia Lopes Alves, Marcelo Lacerda Rezende
ENDEREÇO: Universidade Federal de Alfenas – Unifal-MG, Rua Gabriel Monteiro da Silva, 700. Alfenas, MG. CEP 37130-000

TELEFONE: (35) 3701-9742,

PESQUISADORES PARTICIPANTES: Cristina Garcia Lopes Alves, Marcelo Lacerda Rezende, Daniela Braga Lima, Desnimar Alves Nogueira

PATROCINADOR: MS-SCTIE-Decit/CNPq Nº 26/2019 – Pesquisa em Alimentação e Nutrição

OBJETIVOS: Elaboração de diagnóstico situacional de uma macrorregião de saúde do estado de Minas Gerais, com identificação do impacto econômico da obesidade, e construção e validação de um instrumento de monitoramento que possa apoiar as ações de enfrentamento da obesidade na região estudada, com foco na atenção básica.

JUSTIFICATIVA: O estado nutricional da população brasileira indica que o Brasil vem enfrentando o aumento expressivo do sobrepeso e da obesidade. Em relação às políticas de saúde vigentes, vale ressaltar a importância da atenção básica, especialmente a atuação das equipes de saúde da família, na identificação dos problemas de saúde presentes nos territórios de saúde onde atuam, e no planejamento, execução e avaliação de ações que possam dar resolutividade às necessidades de saúde. Assim, faz-se necessário a construção de um instrumento de monitoramento dos fatores de risco e das condicionalidades referentes à obesidade, assim como o mapeamento da capacidade instalada nos sistemas locais e regionais de saúde. No sentido, de apoiar e fortalecer a organização das ações na rede pública de saúde, tanto relativas ao cuidado ao paciente obeso quanto às ações de prevenção e promoção da saúde na lógica do enfrentamento à obesidade, onde se destacam as ações de alimentação e nutrição.

PROCEDIMENTOS DO ESTUDO: Trata-se de uma pesquisa avaliativa, descritiva, transversal, de abordagem quantitativa, cujos atores/agentes sociais serão os profissionais da saúde nas Unidades Básicas de Saúde (UBS) da Macrorregião Sul.

A pesquisa será realizada por meio de aplicação de um questionário online (formulário google), na qual serão coletados informações sobre o peso, a estatura e a circunferência da cintura, todos autoreferidos, além de perguntas direcionadas sobre compreensão dos rótulos e publicidades de alimentos. A partir desses resultados iremos avaliar o estado nutricional relacionar com a compreensão dos rótulos e comportamento alimentar dos indivíduos entrevistados.

RISCOS E DESCONFORTOS E MEDIDAS: Sua participação nesse estudo implicará na disponibilidade de tempo para responder ao instrumento e tem como riscos possíveis desconfortos e/ou constrangimentos decorrentes do preenchimento do questionário. Como medidas preventivas de tais riscos garantimos a você abordagem cautelosa, bem como respeito ao seus valores, cultura e crenças. Ainda como medida preventiva do desconforto o orientamos que a sua concordância ou não em participar da pesquisa em nada irá alterar sua condição no local de trabalho. Em relação às medidas minimizadoras dos riscos, esclarecemos que sua identificação no questionário se dará apenas pelas iniciais do nome a fim de garantir o anonimato; que haverá sigilo em relação as respostas, tidas como confidenciais e utilizadas apenas para fins científicos. Reiteramos que será assegurado a você total sigilo sobre sua identidade e que você tem o direito de deixar de participar da pesquisa, em qualquer momento, sem que isso lhe traga qualquer prejuízo. Você poderá solicitar esclarecimentos quando sentir necessidade e sua resposta será respeitosamente utilizada em trabalhos e eventos científicos da área da saúde, sem restrições de prazo e citações, desde a presente data. Além disso, você poderá proceder à avaliação como juiz em ambiente seguro e privativo, uma vez que o acesso ao link do instrumento de coleta poderá ser acessado de qualquer dispositivo com acesso à internet. Você será beneficiado com a oportunidade de expressar sua opinião sobre um importante tema de saúde pública e participar da construção de um conhecimento mais aprofundado sobre o assunto. Ainda, os benefícios levantados para esta pesquisa está no potencial de novas medidas regulatórias, como a rotulagem e a publicidade de alimentos frente ao seu efeito sobre o comportamento alimentar, e com estimativa da redução da obesidade. Investigações como essa podem ser usadas como base para amparar novas propostas sobre a regulação de alimentos no país, ou mesmo revisão das medidas regulatórias vigentes.

GARANTIA DE ESCLARECIMENTO, LIBERDADE DE RECUSA E GARANTIA DE SIGILO: Você será esclarecido (a) sobre a pesquisa em qualquer aspecto que desejar. Você é livre para recusar-se a participar, retirar seu consentimento ou interromper a participação a qualquer momento. A sua participação é voluntária e a recusa em participar não irá acarretar qualquer penalidade ou perda de benefícios. O pesquisador irá tratar a sua identidade com padrões profissionais de sigilo. Os resultados da pesquisa serão enviados para você e permanecerão confidenciais. Seu nome ou o material que indique a sua participação não será liberado sem a sua permissão. Você não será identificado (a) em nenhuma publicação que possa resultar deste estudo. O pesquisador declara que irá cumprir todos os termos das Diretrizes e Normas Regulamentadoras de Pesquisa envolvendo Seres Humanos previstas na Resolução 466/2012 do Conselho Nacional de Saúde.

Ativar
Acesse

Declaro que li o Termo de Consentimento Livre e Esclarecido (TCLE) e que estou de acordo com * a participação na pesquisa

- Sim, aceito participar como voluntário na pesquisa
- Não, não aceito participar