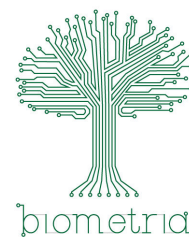


RESSALVA

Atendendo solicitação do(a)
autor(a), o texto completo desta tese
será disponibilizado somente a partir
de 01/08/2023.



Universidade Estadual Paulista “Júlio de Mesquita Filho”
Instituto de Biociências – Câmpus de Botucatu
Programa de Pós-graduação em Biometria



Modelagem conjunta e univariada: Uma aplicação para dados longitudinais de crescimento infantil

Elizabete de Jesus Pinto

Botucatu
2022

Elizabete de Jesus Pinto

Modelagem conjunta e univariada: Uma aplicação para dados longitudinais de crescimento infantil

Tese de Doutorado apresentada ao Curso de Programa de Pós-graduação em Biometria da Universidade Estadual Paulista “Júlio de Mesquita Filho” como parte dos requisitos necessários para a obtenção do título de Doutora em Biometria.

Orientador: Profa. Dra. Miriam Harumi Tsunemi

Botucatu
2022

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: ROSANGELA APARECIDA LOBO-CRB 8/7500

Pinto, Elizabete de Jesus.

Modelagem conjunta e univariada : uma aplicação para dados longitudinais de crescimento infantil / Elizabete de Jesus Pinto. - Botucatu, 2022

Tese (doutorado) - Universidade Estadual Paulista "Júlio de Mesquita Filho", Instituto de Biociências de Botucatu

Orientador: Miriam Harumi Tsunemi

Capes: 10203001

1. Estudos longitudinais. 2. Modelos lineares (Estatística). 3. Crianças - Crescimento. 4. Cópulas (Estatística matemática)

Palavras-chave: Brokenstick; Cópula bivariada; Estudo longitudinal; Modelo linear misto; SITAR.

ELIZABETE DE JESUS PINTO

**MODELAGEM CONJUNTA E UNIVARIADA: UMA APLICAÇÃO PARA
DADOS LONGITUDINAIS DE CRESCIMENTO INFANTIL**

Tese de doutorado defendida e aprovada em Botucatu-SP, no dia 01 de agosto, pela banca examinadora constituída pelos professores:

Profa. Dra. Miriam Harumi Tsunemi
Orientador

Profa. Dra. Liciane Vaz de Arruda Silveira
Examinador

Profa. Dra. Lídia Raquel de Carvalho
Examinador

Profa. Dra. Rosemeire Leovigildo Fiaccone
Examinador

Prof. Dr. Maurício Lima Barreto
Examinador

Dedicatória

A meus pais, Augusto e Helena, pelo amor e apoio incondicional.

A minha irmã Tânia, pela amizade e incentivo em todos os momentos.

Agradecimentos

A redação desta tese é fruto de esforço que é, ao mesmo tempo, solitário e coletivo. Assumindo a total responsabilidade pelas falhas nele contidas, expresso aqui o meu reconhecimento por toda a colaboração recebida.

Primeiro agradeço a Deus, o Caminho, a Verdade e a Vida. Agradeço pela Sua presença constante em minha vida, por me guiar e orientar. A Ele a minha eterna gratidão e fé!

A meus pais, Augusto e Helena, agradeço pelo incentivo, apoio e orações. Vocês são meus exemplos de vida, que apesar da pouca instrução, sempre me ensinaram valores importantes na minha formação pessoal e profissional.

As minhas queridas irmãs Tânia e Sandra pela amizade, incentivo e apoio na concretização desse sonho e aos meus irmãos José, Carlos, Ana, Vera e Ana Lúcia pelo carinho e apoio. Aos meus sobrinhos Yuri, Marília, Victor, Vanessa e Maria Clara por aquecerem minha vida de alegrias e, nos momentos de ausência a lembrança de cada sorriso me deu força para continuar. Aos meus cunhados agradeço pelo apoio e compreensão.

A Profa. Dra. Miriam Harumi Tsunemi, agradeço por sua orientação, confiança, apoio, incentivo, compreensão, motivação, amizade e contribuições ao longo da caminhada.

A Profa. Dra. Ana Marlúcia Oliveira (*in memoriam*) pelo carinho, amizade, contribuições, sugestões, pelas palavras de incentivo em meus momentos de angústias, medos e inquietações e por ter sempre uma palavra de afeto, encorajamento e sabedoria. Meu sincero agradecimento pela preciosa contribuição na minha formação. Obrigado por tudo!

A Profa. Dra. Rosemeire Leogevido Fiaccone, por sempre contribuir na minha trajetória acadêmica, pela amizade, confiança e motivação fundamentais para construir o conhecimento indispensável na conclusão desse trabalho.

Agradeço aos amigos Juliana, Bethina, Gustavo Tomanik, Thiago, Felipe Teles, Renato, Carlos e Felipe Camargo, pelo companheirismo, momentos especiais dentro e fora da Universidade. Vocês fizeram toda diferença! Aos amigos Mariese, Michele e Gustavo pela cooperação, pela troca de experiências, por compartilharem comigo momentos de angustias e alegrias.

Ao Prof. Dr. Maurício Barreto, Dra. Maria Yury e toda a equipe do Centro de Integração de Dados e Conhecimentos para Saúde (Cidacs/IGM/Fiocruz) pelo acolhimento e pela confiança no desenvolvimento deste estudo.

Ao Grupo de Estudos em Nutrição do CIDACS (GeNut-CIDACS), em especial a Profa. Dra. Rita Ribeiro da Silva, por sua valiosa amizade, pela confiança, por acreditar que daria certo, sua ajuda foi simplesmente essencial. Aos colegas Natanael, Juliana e Carolina pela troca de experiências e valiosas contribuições para o avanço deste trabalho.

Aos professores e equipe técnica do Programa de Pós-Graduação em Biometria, pelos ensinamentos, convivência e todo suporte.

À banca examinadora que tão gentilmente aceitaram participar e colaborar com esta tese.

Ao Prof. Dr. Paulo Henrique Ferreira da Silva por sua disponibilidade, contribuições e apoio.

Ao Centro de Ciências da Saúde da Universidade Federal do Recôncavo da Bahia pela liberação das atividades docentes durante o doutorado e por todo apoio.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES.

A todos que ajudaram de forma direta ou indireta na minha formação e compartilham dessa etapa da minha vida.

Resumo

O crescimento infantil é um importante indicador de saúde desde a primeira infância até a puberdade, pois pode determinar o estado de saúde na idade adulta. O monitoramento das medidas de peso e altura pode auxiliar nas ações da gestão de políticas públicas para a prevenção da desnutrição e obesidade, voltada à promoção da saúde e nutrição adequada ao crescimento na infância. O acompanhamento dessas medidas antropométricas, através de medidas coletadas repetidamente no mesmo indivíduo, torna possível responder as questões relativas às mudanças que ocorrem na dinâmica do crescimento, como a trajetória do crescimento de cada indivíduo ao longo do tempo e as diferenças que ocorrem no crescimento médio entre o grupo de indivíduos. Entendendo que o monitoramento dessas medidas são importantes índices de saúde na infância e a forma mais adequada de avaliar seu estado nutricional e de saúde geral, o objetivo principal desta tese foi estimar trajetórias de crescimento com modelagem conjunta e univariada (marginal) para as medidas antropométricas de peso e altura, na população menor de dois anos de idade de uma coorte eletrônica de base populacional linkados com dados de rotina de nascidos vivos (SINASC) e do Sistema de Vigilância Alimentar e Nutricional (SISVAN), entre os anos 2008 a 2017. Foram utilizadas duas abordagens para lidar com dados longitudinais multivariados. Na primeira abordagem sugerimos uma comparação de modelos comumente usados para caracterizar as trajetórias de crescimento infantil como os Modelos de Efeitos Mistos, *Brokenstick* e o SITAR. Na segunda abordagem, o principal interesse foi estudar a evolução conjunta de duas variáveis respostas para o mesmo indivíduo, que exige métodos mais complexos e desafiadores, pois além da existência da dependência entre as medidas repetidas no mesmo indivíduo considera-se a correlação entre as variáveis respostas que pode mudar com o tempo e, além disso, a correlação natural existente em dados longitudinais. Dessa forma, explorou-se a modelagem via funções de cópulas bivariadas, com intuito de melhorar a modelagem de dados bivariados com formas mais flexíveis de se estimar a dependência no padrão de crescimento infantil.

Palavras-chave: Crescimento infantil, Estudo Longitudinal, Modelo Linear Misto, *Brokenstick*, SITAR, Cópula Bivariada.

Abstract

Child growth is an important health indicator from early childhood to puberty, as childhood can determine health status in adulthood. The monitoring of weight and height measurements can help public policies management actions to prevent malnutrition, obesity, and help promoting adequate nutrition for childhood growth. The monitoring of these anthropometric measurements repeatedly in the same individual along time makes it possible to answer questions regarding the changes that occur in the dynamics of growth, such as the trajectory of each individual's growth over time and the differences that can occur in the average growth between the group of individuals. Understanding that monitoring these measures are important indices of childhood health and the most appropriate way to assess their nutritional status and general health, the main objective of this thesis was to estimate growth trajectories with joint and univariate (marginal) modeling for anthropometric measurements of weight and height, in the population under two years of age of a population-based electronic cohort linked with routine live birth data (SINASC) and the Food and Nutrition Surveillance System (SISVAN), between the years 2008 to 2017. Two approaches were used to deal with multivariate longitudinal data. In the first approach, we suggest a comparison of models commonly used to characterize child growth trajectories, such as the Mixed Effects Models, *Brokenstick* and SITAR. In the second approach, the main interest was to study the joint evolution of two response variables for the same individual, which requires more complex and challenging methods, as the existence of dependence between repeated measurements in the same individual, the correlation between the variables is considered responses that can change over time and, in addition, the natural correlation existing in longitudinal data. Thus, modeling via bivariate copula functions was explored, in order to improve the modeling of bivariate data with more flexible ways of estimating dependence on the pattern of child growth.

Keywords: Child growth, Longitudinal Study, Mixed Linear Model, *Brokenstick*, SITAR, Bivariate Copula.

Lista de figuras

Figura 1 – Gráfico de Perfil de peso e altura por sexo de crianças menores de 2 anos, Brasil, 2008-2017	35
Figura 2 – Tempos de observação para um subconjunto aleatório de 100 crianças entre 0 e 2 anos do Banco Completo (Brasil, 2008-2017)	36
Figura 3 – Tempos de observação para a proposta de análise para um subconjunto aleatório de 100 crianças entre 0 e 2 anos, Brasil, 2008-2017	38
Figura 4 – Curvas estimadas pelo modelo <i>Brokenstick</i> por sexo, para dados antropométricos peso e altura de crianças a termo 0 a 2 anos de idade, Brasil, 2008-2017	39
Figura 5 – Curvas estimadas pelo modelo <i>Brokenstick</i> por sexo, para dados antropométricos peso e altura de crianças a termo com baixo peso ao nascer de 0 a 2 anos de idade, Brasil, 2008-2017	40
Figura 6 – Trajetória média ajustada para peso e altura, pelo modelo de efeitos mistos para crianças de 0 a 2 anos de idade por sexo e classificação de peso ao nascer.	43
Figura 7 – Trajetória média ajustada e observada para peso e altura, pelo modelo de efeitos mistos para crianças PAN de 0 a 2 anos de idade por sexo.	44
Figura 8 – Trajetória média ajustada e observada para peso e altura, pelo modelo de efeitos mistos para crianças BPN de 0 a 2 anos de idade por sexo.	45
Figura 9 – Gráficos de dispersão e histograma dos resíduos condicionais padronizados e quantis em relação aos pesos estimados pelo Modelo de efeitos mistos	46
Figura 10 – Gráficos de dispersão e histograma dos resíduos condicionais padronizados e quantis em relação as alturas estimadas pelo Modelo de efeitos misto	47
Figura 11 – Trajetória média ajustada para peso e altura, pelo modelo <i>Brokenstick</i> para crianças PAN de 0 a 2 anos de idade por sexo.	49
Figura 12 – Trajetória média ajustada para peso e altura, pelo modelo <i>Brokenstick</i> para crianças BPN de 0 a 2 anos de idade por sexo	49
Figura 13 – Gráficos de dispersão dos resíduos condicionais padronizados pelo ajuste do modelo <i>Brokenstick</i> para o peso de crianças de 0 a 2 anos por sexo e classificação de peso ao nascer idade	50
Figura 14 – Gráficos de dispersão dos resíduos condicionais padronizados pelo ajuste do modelo <i>Brokenstick</i> para a altura de crianças de 0 a 2 anos por sexo e classificação de peso ao nascer	50
Figura 15 – Histograma dos resíduos condicionais padronizados pelo ajuste do modelo <i>Brokenstick</i> para o peso de crianças de 0 a 2 anos idade por sexo e classificação de peso ao nascer	51

Figura 16 – Histograma dos resíduos condicionais padronizados pelo ajuste do modelo <i>Brokenstick</i> para a altura de crianças de 0 a 2 anos idade por sexo e classificação de peso ao nascer	51
Figura 17 – Valor ajustado pelo modelo <i>Brokenstick</i> versus observado dos dados antropométricos brutos de crianças PAN do sexo masculino de 0 a 2 anos idade . . .	52
Figura 18 – Valor ajustado pelo modelo <i>Brokenstick</i> versus observado dos dados antropométricos brutos de crianças PAN do sexo feminino de 0 a 2 anos de idade . .	52
Figura 19 – Valor ajustado pelo modelo <i>Brokenstick</i> versus observado dos dados antropométricos brutos de crianças BPN do sexo masculino de 0 a 2 anos idade. . .	53
Figura 20 – Valor ajustado pelo modelo <i>Brokenstick</i> versus observado dos dados antropométricos brutos de crianças BPN do sexo feminino de 0 a 2 anos idade. . . .	53
Figura 21 – Trajetória média do peso ajustado pelo modelo SITAR (esquerda), e curva de velocidade de crescimento (direito) para peso de crianças entre 0 e 2 anos. .	59
Figura 22 – Trajetória média da altura ajustada pelo modelo SITAR (esquerda), e curva de velocidade de crescimento (direito) para altura de crianças entre 0 e 2 anos. .	60
Figura 23 – Trajetória média, ajustada e observada, para peso e altura, pelo modelo SITAR para crianças PAN do sexo masculino de 0 a 2 anos	61
Figura 24 – Trajetória média, ajustada e observada, para peso e altura, pelo modelo SITAR para crianças PAN do sexo feminino de 0 a 2 anos	61
Figura 25 – Trajetória média ajustada e observada para peso e altura, pelo modelo SITAR para crianças BPN do sexo masculino de 0 a 2 anos	62
Figura 26 – Trajetória média ajustada e observada para peso e altura, pelo modelo SITAR para crianças BPN do sexo feminino de 0 a 2 anos	62
Figura 27 – Gráficos de dispersão dos resíduos padronizados em relação aos valores ajustados pelo modelo SITAR para variáveis antropométricas em crianças PAN de 0 a 2 anos por sexo.	63
Figura 28 – Gráficos de dispersão dos resíduos padronizados em relação aos valores ajustados pelo modelo SITAR para variáveis antropométricas em crianças BPN de 0 a 2 anos por sexo.	63
Figura 29 – Gráficos de probabilidade normal a partir dos modelos SITAR ajustado para variáveis antropométricas, para crianças PAN entre 0 e 2 anos por sexo . . .	64
Figura 30 – Gráficos de probabilidade normal a partir dos modelos SITAR ajustado para variáveis antropométricas, para crianças BPN entre 0 e 2 anos por sexo . . .	64
Figura 31 – Histograma dos resíduos padronizados pelo ajuste do modelo SITAR para o peso de crianças de 0 a 2 anos, por sexo e classificação de peso ao nascer . .	65
Figura 32 – Histograma dos resíduos padronizados pelo ajuste do modelo SITAR para a altura de crianças de 0 a 2 anos, por sexo e classificação de peso ao nascer .	65
Figura 33 – Histogramas do peso e altura para crianças PAN de 0 a 2 anos. Brasil 2008-2017	66
Figura 34 – Boxplot do peso e altura para crianças PAN de 0 a 2 anos. Brasil 2008-2017	67

Figura 35 – Gráficos de quantis para os resíduos de confundimento mínimo padronizados obtidos a partir dos modelos ajustados para peso e altura de crianças PAN de 0 a 2 anos de idade por sexo	70
Figura 36 – Gráficos de quantis para os resíduos de confundimento mínimo padronizados obtidos a partir dos modelos ajustados para peso e altura de crianças BPN de 0 a 2 anos de idade por sexo	71

Lista de tabelas

Tabela 1 – Característica sociodemográficas materna e do recém-nascido da população para crianças de 0 a 2 anos. Brasil 2008-2017	34
Tabela 2 – Proposta de análise para dados irregulares de crianças menores de 2 anos aninhada a coorte de 100 milhões de brasileiro. Brasil 2008 - 2017.	37
Tabela 3 – Comparação do modelo <i>Brokenstick</i> para variáveis antropométricas peso e altura em relação ao número de observações, tempo de processamento e variância residual para crianças a termo PAN de 0 a 2 anos e por sexo. . . .	39
Tabela 4 – Comparação do modelo <i>Brokenstick</i> para variáveis antropométricas peso e altura em relação ao número de observações, tempo de processamento e variância residual para crianças a termo BPN de 0 a 2 anos e por sexo. . . .	40
Tabela 5 – Número de medições das variáveis antropométricas brutas por intervalo de idade alvo para crianças de 0 a 2 anos, por sexo e classificação de peso ao nascer - Banco PropostaMed	41
Tabela 6 – Número de medições das variáveis antropométricas brutas por criança de 0 a 2 anos, por sexo e classificação de peso ao nascer - Banco PropostaMed . .	41
Tabela 7 – Modelos de efeitos mistos para variáveis antropométricas peso (kg) e altura (cm) de crianças menores de 2 anos de idade. Brasil, 2008-2017.	43
Tabela 8 – Estimativas e desvio padrão (DP) dos coeficientes do modelo <i>Brokenstick</i> para o peso das crianças menores de 2 anos de idade (nós nas idades alvo 0, 2,6,12 e 24 meses, por sexo e classificação de peso ao nascer)	48
Tabela 9 – Estimativas e desvio padrão (DP) dos coeficientes do modelo <i>Brokenstick</i> para a altura das crianças menores de 2 anos de idade (nós nas idades alvo 0, 2,6,12 e 24 meses, por sexo e classificação de peso ao nascer)	48
Tabela 10 – Comparação dos ajustes dos modelos SITAR aplicados aos dados antropométricos em crianças PAN de 0 a 2 anos, com diferentes parametrizações e transformações na variável idade em meses.	55
Tabela 11 – Comparação dos ajustes dos modelos SITAR aplicados aos dados antropométricos em crianças BPN de 0 a 2 anos, com diferentes parametrizações e transformações na variável idade em meses.	56
Tabela 12 – Modelo SITAR de melhor ajuste para o sexo masculino aplicados aos dados antropométricos de peso (kg) em crianças de 0 a 2 anos por classificação de peso ao nascer. Brasil 2008-2017.	57
Tabela 13 – Modelo SITAR de melhor ajuste para o sexo feminino aplicados aos dados antropométricos de peso (kg) em crianças de 0 a 2 anos por classificação de peso ao nascer. Brasil 2008-2017.	57

Tabela 14 – Modelo SITAR de melhor ajuste para o sexo masculino aplicados aos dados antropométricos de altura (cm) em crianças de 0 a 2 anos por classificação de peso ao nascer. Brasil 2008-2017.	58
Tabela 15 – Modelo SITAR de melhor ajuste para o sexo feminino aplicados aos dados antropométricos de altura (cm) em crianças de 0 a 2 anos por classificação de peso ao nascer. Brasil 2008-2017.	58
Tabela 16 – Distribuição em relação as características da base de dados e amostra de 10%. Brasil 2008-2017	59
Tabela 17 – Ajuste do modelo <i>Brokenstick</i> para uma amostra de crianças PAN, do sexo masculino entre de 0 a 2 anos de idade. Brasil 2008-2017	68
Tabela 18 – Ajuste do modelo <i>Brokenstick</i> para uma amostra de crianças PAN, do sexo feminino entre de 0 a 2 anos de idade. Brasil 2008-2017	68
Tabela 19 – Ajuste do modelo <i>Brokenstick</i> para crianças BPN, do sexo masculino de 0 a 2 anos de idade. Brasil 2008-2017	69
Tabela 20 – Ajuste do modelo <i>Brokenstick</i> para crianças com BPN, do sexo feminino de 0 a 2 anos de idade. Brasil 2008-2017	69
Tabela 21 – Principais resultados da estimativa IFM para os modelos <i>Brokenstick</i> baseados em cópulas bivariados ajustados aos dados de uma amostra de crianças PAN, e parâmetros de associação de cópulas por sexo	70
Tabela 22 – Principais resultados da estimativa para os modelos <i>Brokenstick</i> baseados em cópulas bivariados ajustados aos dados de uma amostra de crianças BPN, e parâmetros de associação de cópulas por sexo	71

Lista de abreviaturas e siglas

AIC	<i>Akaike information criteria</i>
BIC	<i>Bayesian information criteria</i>
BPN	Baixo peso ao nascer
CadÚnico	Cadastro Único para Programas Sociais do Governo Federal
CIDACS	Centro de Integração de Dados e Conhecimentos para Saúde
DP	Desvio Padrão
E/I	Estatura por idade
IFM	Método da inferência para as marginais
IMC	Índice de massa corpórea
MV	Máxima verossimilhança
MVR	Máxima verossimilhança restrita
PAN	Peso adequado ao nascer
P/E	Peso por estatura
P/I	Peso por idade
RSD	Desvio Padrão Residual
SINASC	Sistema de Informações sobre Nascidos Vivos
SISVAN	Sistema de Vigilância Alimentar e Nutricional
SITAR	<i>SuperImposition by Translation And Rotation</i>

Sumário

1	INTRODUÇÃO	1
1.1	Objetivos	5
2	MODELOS PARA TRAJETÓRIA DE CRESCIMENTO INFANTIL	6
2.1	Características dos estudos longitudinais	6
2.2	Abordagem Univariada	8
2.2.1	Modelo Linear de Efeitos Mistos - MLM	8
2.2.2	Especificação do modelo	9
2.2.3	Estimação Pontual	12
2.2.4	Modelo <i>Brokenstick</i> – <i>Splines</i> Lineares	13
2.2.5	Modelo SITAR	16
2.2.6	Análise de Resíduo	18
2.3	Abordagem Multivariada - Cópulas	19
2.3.1	Conceitos e Definições	19
2.3.2	Famílias Paramétricas de Cópulas	23
2.3.2.1	Clayton	23
2.3.2.2	Frank	24
2.3.2.3	Gumbel	24
2.3.2.4	Gaussiana	25
2.3.2.5	t-Student	25
2.3.3	Inferência	26
3	DESCRIÇÃO DO BANCO DE DADOS	28
3.1	Fonte de Dados	28
3.2	Integração e desidentificação das bases de dados	30
3.3	Descrição da abordagem metodológica e dos procedimentos	31
3.3.1	Critérios de exclusões	31
4	RESULTADOS	33
4.1	Estratégia de análise de dados para medições irregulares	35
4.2	Ajustes dos modelos univariados para as trajetórias de crescimento	42
4.2.1	Modelo Linear Misto	42
4.2.2	Modelo <i>Brokenstick</i>	44
4.2.3	Modelo SITAR	47

4.3	Ajuste do Modelo Bivariado - Cópula	57
5	DISCUSSÃO	72
6	CONCLUSÃO	75
6.1	Perspectivas Futuras	76
	Referências	78
	Apêndices	84
	APÊNDICE A – CÓDIGOS DO PROGRAMA R	85

1 Introdução

Há uma crescente ênfase nas pesquisas da área de saúde envolvendo a modelagem do crescimento ao longo da vida e na identificação de fatores associados ao crescimento (COLE, 2002; JOHNSON; BALAKRISHNA; GRIFFITHS, 2013; TILLING et al., 2014; HOWE et al., 2016; ANDERSON et al., 2019; DOBSON et al., 2020; CONFORTIN et al., 2021). O crescimento infantil é um importante indicador de saúde desde a primeira infância até a puberdade, pois pode determinar o estado de saúde na idade adulta. A investigação das curvas de crescimento através do peso e altura é a maneira mais apropriada de avaliar o estado nutricional e o crescimento infantil (COLE, 2002; HOSSEINI et al., 2014). Por sua própria natureza, o crescimento físico é um processo dinâmico longitudinal e, portanto, gera desafios conceituais e metodológicos (HOWE et al., 2016; ANDERSON et al., 2019).

Diversos fatores devem ser considerados na determinação dos padrões de crescimento infantil, como os aspectos nutricionais, culturais, ambientais, sociais, biológicos, genéticos, para além daqueles relacionados ao ambiente social e cultural em que vivem a criança e sua família (SPYRIDES et al., 2008; CONFORTIN et al., 2021). O crescimento e desenvolvimento adequados indicam condições plenas de saúde e vida da criança. Assim, o monitoramento do estado antropométrico pode garantir a promoção e manutenção da saúde, bem como indicar o momento de intervenção sobre fatores capazes de comprometê-la.

O crescimento é um dos principais indicadores de saúde na infância. A avaliação e monitoração do crescimento ponderal e linear são ações definidas nas políticas de atenção à saúde e nutrição para esta fase da vida (ANDERSON et al., 2019). Altura e índice de massa corporal (IMC) são medidas antropométricas da qualidade da nutrição e salubridade do ambiente de vida durante a infância e adolescência e preditivos de saúde e desenvolvimento ao longo da vida. (HORIMOTO et al., 2020). O acompanhamento dessas medidas vai ajudar numa ação precoce, prevenindo o desenvolvimento de morbidades futuras.

O acompanhamento das medidas antropométricas é fundamental desde o nascimento do indivíduo, pois existem diferenças no crescimento entre as nascidas a termo (idade gestacional igual ou maior que 37 semanas) com peso adequado ao nascer e aquelas nascidas prematuramente (idade gestacional inferior a 37 semanas), e ainda aquelas com baixo peso ao nascer, definido como inferior a 2500g, que pode ser uma consequência da prematuridade ou estar associada à restrição de crescimento intrauterino, ou uma combinação das duas (SULLIVAN et al., 2008). O

baixo peso ao nascer é um dos principais fatores que afetam a morbimortalidade infantil em todo o mundo, ((MANYEH et al., 2016) e, na vida adulta, está associado com ao aumento de várias doenças como a síndrome metabólica e diabetes mellitus.

O acompanhamento do peso e da altura nas escalas brutas ou padronizadas (altura ou peso relativo de uma criança em uma determinada idade na forma de escore Z), pode auxiliar as ações da gestão de política pública para a prevenção da desnutrição e obesidade voltadas à promoção da saúde e nutrição adequada ao crescimento na infância. O registro de tais medidas, do mesmo indivíduo, de forma repetida e periódica, durante um período de tempo, apresenta caráter longitudinal e melhora a consistência da informação, fornecendo dados sobre as variações global e individual ao longo do tempo. Esse registro torna possível responder as questões relativas às mudanças que ocorrem na dinâmica do crescimento, particularmente sobre duas questões que interessam diretamente a esse evento. A primeira delas diz respeito à trajetória do comportamento do crescimento de cada indivíduo ao longo do tempo. A segunda é a que busca prever as diferenças que ocorrem no crescimento médio entre o grupo de indivíduos (KOLETZKO; RZEHA; REGNAULT, 2014).

Uma característica importante em dados longitudinais é a sua estrutura balanceada ou desbalanceada, sendo a última mais comumente encontrada. O desbalanceamento é uma consequência natural em estudo de coorte, pois o tempo total de observações nem sempre é igual para todos os indivíduos, e ainda pode ocorrer a presença de padrões de observações desequilibradas entre os sujeitos devendo, portanto, ser considerado na análise dos dados (DIGGLE et al., 2002).

Existem muitos métodos para avaliar a trajetória de crescimento, mas há uma semelhança entre eles que leva a suposição de uma curva média geral para uma determinada população, e as diferenças entre as crianças podem ser explicadas como desvios dessa curva média (VERBEKE, 1997; ANDERSON et al., 2019). Esses métodos levam em consideração a correlação entre observações repetidas na mesma criança para obter erros padrões apropriados, por exemplo, os modelos multiníveis que são frequentemente usados para descrever trajetórias, os modelos não lineares, as análises de curva de crescimento latente e modelagem de mistura (ANDERSON et al., 2019; MCNEISH; MATTA, 2018; TILLING et al., 2014).

Na atualidade, muitos modelos estatísticos de crescimento têm sido explorados, especialmente por considerar diferenças inter e intra-individuais. O modelo linear de efeitos mistos é frequentemente usado, pois permite descrever a tendência temporal considerando a correlação existente entre as medidas repetidas do crescimento no mesmo indivíduo, e possui diferencial de se adequar quando o número de medições difere entre os indivíduos, ou quando as medições são feitas em tempos diferentes, tendo o tempo como uma covariável do modelo e atinge a convergência a grande volume de dados ((TILLING et al., 2014; MCNEISH; MATTA, 2018; ANDERSON et al., 2019). Uma abordagem é modelar as variáveis dependentes como uma simples função da idade (que pode ser linear ou pode incluir polinômios). Outro método denominado modelo *Brokenstick* é um caso especial do modelo linear de efeitos misto, que considera o tempo como

uma função B-spline linear e o indivíduo como o fator de agrupamento. Ele modela a trajetória de crescimento como uma série de linhas conectadas ("*splines* lineares", também chamados de modelos "quebrado" ou "por partes") unidos em nós. Este modelo também é flexível a dados desbalanceados e com falta de padrão regular entre as medidas (HOWE et al., 2016; BUUREN, 2020).

Outra abordagem que tem sido muito usada na atualidade para estimar trajetória de crescimento, foi proposto por Cole (2010), e é conhecido como *Superimposition by Translation and Rotation* (SITAR), que envolve cada indivíduo com uma curva que é uma versão deslocada e transformada da curva média de crescimento. Deslocar a curva verticalmente para cima ou para baixo corresponde às mudanças médias, deslocá-la horizontalmente para a esquerda ou direita corresponde a diferentes tempos de crescimento.

Os modelos descritos acima consideram apenas uma variável resposta antropométrica. Entretanto, é muito comum em estudos longitudinais a observação de um conjunto de variáveis respostas medidas repetidamente para um mesmo indivíduo, e desperta o interesse em estudar a relação de dependência entre as variáveis respostas no tempo nos padrões de crescimento infantil. A abordagem de cópula, que é a distribuição conjunta de duas ou mais variáveis respostas escrita como produto das distribuições marginais e condicionais das variáveis, possibilita estudar a evolução conjunta das variáveis peso e altura.

A investigação em grandes bases de dados com registros eletrônicos na área da saúde, por meio dos modelos descritos acima, permite auxiliar no monitoramento do crescimento e desenvolvimento infantil que são ações essenciais para a prevenção da desnutrição, obesidade e promoção da saúde na infância. Portanto, estimar as trajetórias de crescimento permite a caracterização de seus diferentes padrões, como, por exemplo, identificar as crianças que crescem com sucesso e aquelas cujo crescimento é menos saudável. Para as crianças que estão com crescimento menos saudável, é possível identificar a necessidade de intervenção e quando ocorre o momento de recuperação (HOSSEINI et al., 2014).

O papel das trajetórias de saúde ao longo da vida é um dos temas que destaca a necessidade de dados capazes de subsidiar as próximas etapas da pesquisa em saúde pública e saúde da população, tendo como ferramenta os estudos longitudinais (JUTTE; ROOS; BROWNELL, 2011; HARRON et al., 2017). No entanto, a coleta longitudinal de dados é difícil de coordenar, demorada e cara, além das questões de perda de seguimento, bem como limitações resultantes de critérios de exclusão e requisitos de consentimento do sujeito que reduzem a generalização para a população em geral (TU et al., 2004). Dessa forma, os dados administrativos vinculados compreendem informações já coletadas de forma ampla e diligente em grandes populações para outros fins, depois mescladas no nível individual usando identificadores únicos e anônimos e disponibilizadas para pesquisa acadêmica (HARRON et al., 2017). Esses ambientes ricos em informações podem, com custo relativamente baixo, fornecer amostras muito grandes e observações de longo prazo para futuras pesquisas em saúde pública que a coleta de dados

primários geralmente não consegue (JUTTE; ROOS; BROWNELL, 2011; HARRON et al., 2017).

A coorte eletrônica de base populacional linkados com dados de rotina de nascidos vivos (SINASC) e do Sistema de Vigilância Alimentar e Nutricional (SISVAN) do Ministério da Saúde, é um exemplo de grande base de dados administrativos que coleta informações contínuas sobre condições e monitoramento do estado nutricional da população. Essa grande base de dados, com a diversidade de populações, tem potencial para ser explorada para estimar trajetórias de crescimento na primeira infância. Esses registros são de demanda espontânea e são encontrados nos estabelecimentos de Assistência de Saúde. No entanto, as informações sobre as medições antropométricas das crianças, que muitas vezes permanecem nos serviços, são subutilizadas pela dificuldade metodológica não somente no desenho do estudo, mas, particularmente, na condução das análises estatísticas de medidas biológicas, especialmente aquelas computadas ao longo do tempo.

Este estudo tem o objetivo de modelar a trajetória do peso e altura de crianças menores de 2 anos da coorte eletrônica de base populacional linkados com dados de rotina de nascidos vivos (SINASC) e do Sistema de Vigilância Alimentar e Nutricional (SISVAN), uma base com grande volume de dados ainda pouco explorada na literatura. Mais especificamente, este estudo visa comparar três técnicas de modelagem marginal diferentes: os Modelos de Efeitos Mistos, *Brokenstick* e SITAR para as medidas de peso e altura na sua forma bruta longitudinalmente e em segundo explorar a relação de dependência dessas variáveis conjuntamente no padrão de crescimento infantil através da abordagem de Cópula.

6 Conclusão

Esta tese teve como objetivo geral estimar trajetórias de crescimento com modelagem conjunta e univariada (marginal) para as medidas antropométricas de peso e altura, na população menor de dois anos de idade, em uma coorte eletrônica de base populacional linkados com dados de rotina de nascidos vivos (SINASC) e do Sistema de Vigilância Alimentar e Nutricional (SISVAN) do Ministério da Saúde, no período de 2008 a 2017. O acompanhamento desta trajetória permite auxiliar no monitoramento do crescimento e desenvolvimento infantil que são ações essenciais para a prevenção da desnutrição, obesidade e promoção da saúde na infância, permitindo caracterizar diferentes padrões de crescimento, como, por exemplo, as crianças que crescem com sucesso e aquelas cujo crescimento é menos saudável.

Uma característica de dados longitudinais observacionais é a presença de observações que variam no tempo entre os indivíduos. Esses dados também podem apresentar tempos de acompanhamento irregulares, a exemplo do ocorrido na base de dados em estudo, que são desafios significativos para a análise quantitativa. Dentre as estratégias disponíveis na literatura para análise dos tempos de observações irregulares, adotamos a sugerida por Buuren (2020) e Lukko *et al* (2020) que foi a de criar uma grade de tempo pré-especificada, denominada de “idade alvo”, com o objetivo de padronizar a idade e de colocar o espaçamento de maneira igualitária para todos. No entanto, o banco de dados ainda apresentou desbalanceamento entre as medidas.

Para diminuir as incertezas em relação à estratégia sugerida foram criados três bancos de dados "Completo", "Proposta" e "PropostaMed", com o objetivo de compará-los e escolher aquele que viabilizaria os ajustes dos modelos de trajetória pelo modelo *Brokenstick*. Adotou-se o banco de dados PropostaMed para ajuste dos modelos univariados e bivariados, pois nele o tempo de processamento foi otimizado, houve convergência dos modelos e não foi observado uma diferença significativa entre as curvas ajustadas e variâncias residuais.

Neste estudo, foram utilizadas duas abordagens para lidar com dados longitudinais multivariados. Na primeira abordagem sugerimos uma comparação de modelos comumente usados para caracterizar as trajetórias de crescimento infantil, como os Modelos de Efeitos Mistos, *Brokenstick* e o SITAR, para modelar as medidas de peso e altura na sua forma bruta. Na segunda abordagem estudamos a relação de dependência dessas variáveis conjuntamente no padrão de crescimento infantil através da abordagem de Cópula.

A comparação entre os parâmetros estimados e análise de resíduos a partir de modelos

de regressão linear de efeitos mistos, *Brokenstick* e SITAR, utilizados para análise de dados longitudinais, permitiu concluir que o modelo linear de efeitos mistos não foi flexível o suficiente para modelar a curva de crescimento e SITAR não apresentou um bom ajuste aos dados. Dessa forma estes modelos não se ajustaram na avaliação do crescimento da criança nos dois primeiros anos de vida, quando comparado ao modelo *Brokenstick*, para a base de dados em estudo.

O modelo multinível de *spline* linear forneceu um método útil para resumir as trajetórias de crescimento, reduzindo a dimensionalidade dos dados em medições de crescimento e fornecendo resumos de crescimento interpretáveis. Os resultados sugerem que as incertezas associadas à transformação de vários tempos de observação em medidas repetidas menos irregulares foram pequenas. É provável que a abordagem *Brokenstick* funcione com sucesso em uma ampla gama de conjuntos de dados. A falta de avaliação do diagnóstico do modelo final com a análise de resíduos é uma limitação dessa abordagem, pois ainda não está implementada no pacote *Brokenstick* do R. O critério de escolha da posição dos nós foi respaldado pela literatura.

A abordagem univariada traz como limitação a impossibilidade de modelar conjuntamente algumas variáveis ao longo do tempo, levando também em consideração a correlação entre as variáveis respostas. Entretanto, na segunda abordagem, o principal interesse é a modelagem conjunta dessas variáveis dada a correlação natural entre as mesmas. Dessa forma, explorou-se a modelagem via funções de cópulas bivariadas, com intuito de melhorar a modelagem de dados bivariados com formas mais flexíveis de se estimar essa dependência. Levando em consideração os resultados dos critérios de seleção, as cópulas Gumbel (crianças do sexo masculino (PAN)) e Frank (demais grupos estudados), foram escolhidas devido à maior precisão quanto à representação da dependência entre as variáveis peso e altura.

Por fim, na prática, o entendimento destas trajetórias de crescimento, poderá apoiar as discussões acerca das ações de Vigilância Alimentar e Nutricional (VAN) no âmbito da Atenção Primária a Saúde no SUS. A médio prazo, subsidiar o planejamento de ações para prevenção e enfrentamento dos agravos relacionados aos desvios das trajetórias de crescimento infantil, onde é preciso intensificar ações de prevenção e promoção da saúde focalizando os fatores associados ao excesso de peso em idade mais tenras e assim reduzir a gravidade dessa condição e, com isso, reduzir o risco de desfechos adversos crônicos ao longo da vida. Ao longo prazo, assegurar a segurança alimentar e nutricional desta e de novas gerações.

6.1 Perspectivas Futuras

- Ressalta-se que no Brasil são escassos os estudos longitudinais que avaliam a dinâmica de crescimento físico e composição corporal na primeira infância, o que pode ser devido às dificuldades relacionadas a logística complexa e ao alto custo da pesquisa longitudinal. A base de dados em estudo apresenta a potencialidade de responder a essas questões sem limitações amostrais e estratificada por grupos de interesse, como características perinatal.

Destaca-se a necessidade de trabalhos futuros ampliando a faixa etária estudada, para crianças de 0 a 5 anos e considerando o sexo, peso ao nascer e prematuridade, entendendo que o monitoramento das medidas antropométricas são importantes índices de saúde na infância e a forma mais adequada de avaliar o estado nutricional e de saúde geral.

- Os pacotes do programa R, SITAR e *Brokenstick* foram desenvolvidos para admitir apenas a variável idade como covariável. Um desafio é poder implementar a inclusão de mais covariáveis no modelo permitindo que outras variáveis preditoras possam contribuir na explicação dessa trajetória, e ainda implementar a avaliação das suposições de validação do modelo através da análise gráfica dos resíduos.
- Uma limitação encontrada neste trabalho e que nos motiva a buscar soluções é a estimação dos erros padrões das estimativas para os modelos *Brokenstick* derivados da cópula bivariada. Esforços têm sido empregados buscando novas estratégias que permitam a estimação desse parâmetro, inclusive levando em consideração o seu tempo de processamento.
- Em estudos futuros pode ser interessante testar a relação de dependências em indicadores antropométricos padronizados, explorando novas famílias de cópula que possam se ajustar melhor aos dados e verificar se a relação de dependência entre as medidas antropométricas varia ao longo do tempo. Além disso, buscar estratégia computacional, como por exemplo, a paralelização como meio a viabilizar a análise de cópula envolvendo dados massivos, reduzindo o tempo computacional que foi um dos grandes desafios encontrados.

Referências

- AKAIKE, H. On entropy maximization principle. *Application of statistics*, North-Holland Publishing Company, p. 27–41, 1977. 13
- ALI, M. S. et al. Administrative data linkage in brazil: potentials for health technology assessment. *Frontiers in pharmacology*, Frontiers, v. 10, p. 984, 2019. 30
- ALMEIDA, D. et al. Examining the quality of record linkage process using nationwide brazilian administrative databases to build a large birth cohort. *BMC Medical Informatics and Decision Making*, BioMed Central, v. 20, n. 1, p. 1–9, 2020. 30
- ANDERSON, C. et al. Comparing predictive abilities of longitudinal child growth models. *Statistics in medicine*, Wiley Online Library, v. 38, n. 19, p. 3555–3570, 2019. 1, 2, 9, 14, 15, 16, 44, 47, 73
- BARBOSA, G. C. et al. Cidacs-rl: a novel indexing search and scoring-based record linkage system for huge datasets with high accuracy and scalability. *BMC medical informatics and decision making*, BioMed Central, v. 20, n. 1, p. 1–13, 2020. 30
- BARRETO, M. L. et al. The centre for data and knowledge integration for health (cidacs): linking health and social data in brazil. *International Journal of Population Data Science*, Swansea University, v. 4, n. 2, 2019. 31
- BEATH, K. J. Infant growth modelling using a shape invariant model with random effects. *Statistics in medicine*, v. 26, n. 12, p. 2547–2564, 2007. 16, 72
- BERKEY, C. S.; REED, R. B. A model for describing normal and abnormal growth in early childhood. *Human biology*, JSTOR, p. 973–987, 1987. 9
- BOOR, C. D.; BOOR, C. D. *A practical guide to splines*. [S.l.]: springer-verlag New York, 1978. v. 27. 15
- BOTTON, J. et al. Postnatal weight and height growth velocities at different ages between birth and 5 y and body composition in adolescent boys and girls. *The American journal of clinical nutrition*, Oxford University Press, v. 87, n. 6, p. 1760–1768, 2008. 9
- BRASIL. *Manual Operacional para uso do Sistema de Vigilância Alimentar e Nutricional: SISVAN VERSÃO 3.0*. [S.l.]: Ministério da Saúde, 2017. Disponível em: <<http://sisaps.saude.gov.br/sisvan/documentos/index>>, acesso em 09/04/2021. 29
- BRASIL. *Secretaria de Vigilância da Saúde. Sistema de Informações sobre Nascidos Vivos (SINASC): Coordenação Geral de Informações e Análises Epidemiológicas - CGIAE*. 2018. Disponível em: <<http://svs.aids.gov.br/dantps/cgiae/sinasc/>>, acesso em 09/04/2021. 29, 30
- BUUREN, S. van. Broken stick model for irregular longitudinal data. *Journal of Statistical Software*, Foundation for Open Access Statistics, Submitted for publication, p. 1–47, 2020. 3, 7, 8, 13, 14, 15, 16, 32, 44, 45, 73
- CAMERON, N. The human growth curve, canalization and catch-up growth. In: *Human growth and development*. [S.l.]: Elsevier, 2012. p. 1–22. 42, 73

- CEF. *Caixa Econômica Federal. Cadastro Único*. 2018. Disponível em: <<http://www.caixa.gov.br/cadastros/cadastro-unico/Paginas/default.aspx>>, acesso em 09/04/2021. 29
- CHEN, Y. et al. An epidemiological survey on low birth weight infants in china and analysis of outcomes of full-term low birth weight infants. *BMC pregnancy and childbirth*, BioMed Central, v. 13, n. 1, p. 1–9, 2013. 31
- CHO, H. The analysis of multivariate longitudinal data using multivariate marginal models. *Journal of Multivariate Analysis*, Elsevier, v. 143, p. 481–491, 2016. 19, 73
- COLE, T.; COLE, M. T. Package ‘sitar’. 2021. 17, 18, 32
- COLE, T. J. Assessment of growth. *Best Practice & Research Clinical Endocrinology & Metabolism*, Elsevier, v. 16, n. 3, p. 383–398, 2002. 1, 72
- COLE, T. J. Optimal design for longitudinal studies to estimate pubertal height growth in individuals. *Annals of human biology*, v. 45, n. 4, p. 314–320, 2018. 16, 47, 73
- COLE, T. J.; DONALDSON, M. D.; BEN-SHLOMO, Y. Sitar—a useful instrument for growth curve analysis. *International journal of epidemiology*, Oxford University Press, v. 39, n. 6, p. 1558–1566, 2010. 16, 47, 52
- CONFORTIN, S. C. et al. Rps brazilian birth cohorts consortium (ribeirão preto, pelotas and são luís): history, objectives and methods. *Cadernos de Saúde Pública*, SciELO Brasil, v. 37, 2021. 1
- CROZIER, S. R. et al. A discussion of statistical methods to characterise early growth and its impact on bone mineral content later in childhood. *Annals of human biology*, Taylor & Francis, v. 46, n. 1, p. 17–26, 2019. 16
- DIGGLE, P. et al. *Analysis of longitudinal data*. [S.l.]: Oxford university press, 2002. 2, 7, 42, 72
- DOBSON, A. et al. Flexible age-period-cohort modelling illustrated using obesity prevalence data. *BMC medical research methodology*, Springer, v. 20, n. 1, p. 1–9, 2020. 1
- EILERS, P. H.; MARX, B. D. Flexible smoothing with b-splines and penalties. *Statistical science*, Institute of Mathematical Statistics, v. 11, n. 2, p. 89–121, 1996. 13
- ESCARELA, G.; HERNÁNDEZ, A. Modelado de parejas aleatorias usando cópulas. *Revista Colombiana de Estadística*, Universidad Nacional de Colombia, v. 32, n. 1, p. 33–58, 2009. 21
- FAUSTO, M. A. et al. O modelo de regressão linear misto para dados longitudinais: uma aplicação na análise de dados antropométricos desbalanceados. *Cadernos de Saúde Pública*, SciELO Public Health, v. 24, p. 513–524, 2008. 72
- FERREIRA, P. H. et al. Bivariate copula-based linear mixed-effects models: An application to longitudinal child growth data. *TEMA (São Carlos)*, SciELO Brasil, v. 20, p. 37–59, 2019. 19, 21, 24, 25, 66, 73, 74
- FILTEAU, S. et al. Steady growth in early infancy is associated with greater anthropometry in indian children born low birth weight at term. *The Journal of nutrition*, Oxford University Press, v. 149, n. 9, p. 1633–1641, 2019. 16, 72
- FITZMAURICE, G. et al. *Longitudinal data analysis*. [S.l.]: CRC press, 2008. 7

- FITZMAURICE, G. M.; LAIRD, N. M.; WARE, J. H. *Applied longitudinal analysis*. [S.l.]: John Wiley & Sons, 2012. v. 998. 9
- FRAHM, G.; JUNKER, M.; SZIMAYER, A. Elliptical copulas: applicability and limitations. *Statistics & Probability Letters*, Elsevier, v. 63, n. 3, p. 275–286, 2003. 23
- GOLDSTEIN, H. et al. Multilevel growth curve models that incorporate a random coefficient model for the level 1 variance function. *Statistical methods in medical research*, SAGE Publications Sage UK: London, England, v. 27, n. 11, p. 3478–3491, 2018. 6, 16
- HARRON, K. et al. Challenges in administrative data linkage for research. *Big data & society*, Sage Publications Sage UK: London, England, v. 4, n. 2, p. 2053951717745678, 2017. 3, 4
- HAUSPIE, R. C.; CAMERON, N.; MOLINARI, L. *Methods in human growth research*. [S.l.]: Cambridge University Press, 2004. v. 39. 6, 18
- HORIMOTO, A. et al. Height and body-mass index trajectories of school-aged children and adolescents from 1985 to 2019 in 200 countries and territories: a pooled analysis of 2181 population-based studies with 65 million participants. 2020. 1
- HOSSEINI, S.-M. et al. Child weight growth trajectory and its determinants in a sample of iranian children from birth until 2 years of age. *International Journal of Preventive Medicine*, Wolters Kluwer–Medknow Publications, v. 5, n. 3, p. 348, 2014. 1, 3
- HOWE, L. D. et al. Linear spline multilevel models for summarising childhood growth trajectories: a guide to their application using examples from five birth cohorts. *Statistical methods in medical research*, SAGE Publications Sage UK: London, England, v. 25, n. 5, p. 1854–1874, 2016. 1, 3, 14, 15, 44, 72
- IDA, A.; ISHIMURA, N.; NAKAMURA, M. Note on the measures of dependence in terms of copulas. *Procedia Economics and Finance*, Elsevier, v. 14, p. 273–279, 2014. 21
- JENSS, R. M.; BAYLEY, N. A mathematical method for studying the growth of a child. *Human Biology*, Johns Hopkins Press, v. 9, n. 4, p. 556, 1937. 9
- JOE, H.; XU, J. J. The estimation method of inference functions for margins for multivariate models. *Technical Report 166*, University of British Columbia, n. 1, 1996. 26
- JOHNSON, W.; BALAKRISHNA, N.; GRIFFITHS, P. L. Modeling physical growth using mixed effects models. *American journal of physical anthropology*, Wiley Online Library, v. 150, n. 1, p. 58–67, 2013. 1
- JUTTE, D. P.; ROOS, L. L.; BROWNELL, M. D. Administrative record linkage as a tool for public health research. *Annual review of public health*, v. 32, n. 1, p. 91–108, 2011. 3, 4
- KOLETZKO, B.; RZEHA, P.; REGNAULT, N. *Analysis of Growth Trajectories*. [S.l.]: Karger, 2014. 2
- LAIRD, N. M.; WARE, J. H. Random-effects models for longitudinal data. *Biometrics*, JSTOR, p. 963–974, 1982. 10
- LIANG, Y.; LU, W.; YING, Z. Joint modeling and analysis of longitudinal data with informative observation times. *Biometrics*, Wiley Online Library, v. 65, n. 2, p. 377–384, 2009. 7

- LIN, H.; SCHARFSTEIN, D. O.; ROSENHECK, R. A. Analysis of longitudinal data with irregular, outcome-dependent follow-up. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Wiley Online Library, v. 66, n. 3, p. 791–813, 2004. 7
- LITTELL, R. C.; PENDERGAST, J.; NATARAJAN, R. Modelling covariance structure in the analysis of repeated measures data. *Statistics in medicine*, Wiley Online Library, v. 19, n. 13, p. 1793–1819, 2000. 10, 11
- LOKKU, A. et al. Summarizing the extent of visit irregularity in longitudinal data. *BMC Medical Research Methodology*, Springer, v. 20, p. 1–9, 2020. 7
- LU, Y.; PEARCE, A.; LI, L. Weight gain in early years and subsequent body mass index trajectories across birth weight groups: a prospective longitudinal study. *European journal of public health*, Oxford University Press, v. 30, n. 2, p. 316–322, 2020. 6
- MANYEH, A. K. et al. Socioeconomic and demographic determinants of birth weight in southern rural ghana: evidence from dodowa health and demographic surveillance system. *BMC pregnancy and childbirth*, BioMed Central, v. 16, n. 1, p. 1–9, 2016. 2
- MCNEISH, D.; MATTA, T. Differentiating between mixed-effects and latent-curve approaches to growth modeling. *Behavior research methods*, Springer, v. 50, n. 4, p. 1398–1414, 2018. 2, 9
- MD. Ministério da Saúde. Secretaria de Políticas Públicas. Saúde da Criança - Acompanhamento do crescimento e desenvolvimento infantil. 2002. Disponível em: <https://bvsmms.saude.gov.br/bvs/publicacoes/acompanhamento_crescimento_desenvolvimento_infantil_cab11.pdf> ,acesso em 22/11/2021. 35, 44
- MDS. Ministério do Desenvolvimento Social. Cadastro Único: O que é e para que serve. 2018. Disponível em: <<http://mds.gov.br/assuntos/cadastro-unico/o-que-e-e-para-que-serve>> , acesso em 09/04/2021. 29
- NELSEN, R. B. *An introduction to copulas*. [S.l.]: Springer Science & Business Media, 2007. 19, 20, 21, 22, 23, 73
- NOBRE, J. S.; SINGER, J. da M. Residual analysis for linear mixed models. *Biometrical Journal: Journal of Mathematical Methods in Biosciences*, Wiley Online Library, v. 49, n. 6, p. 863–875, 2007. 18, 19
- OHUMA, E. O. et al. A novel development indicator based on population-average height trajectories of children aged 0–5 years modelled using 145 surveys in 64 countries, 2000–2018. *BMJ global health*, BMJ Specialist Journals, v. 6, n. 3, p. 1–11, 2021. 16, 47, 72, 73
- PINHEIRO, J. et al. Package ‘nlme’. *Linear and nonlinear mixed effects models, version*, v. 3, n. 1, 2017. 13, 32, 42
- PINHEIRO, J. C.; BATES, D. M. Linear mixed-effects models: basic concepts and examples. *Mixed-effects models in S and S-Plus*, Springer, p. 3–56, 2000. 9
- PULLENAYEGUM, E. M.; LIM, L. S. Longitudinal data subject to irregular observation: A review of methods with a focus on visit processes, assumptions, and study design. *Statistical methods in medical research*, SAGE Publications Sage UK: London, England, v. 25, n. 6, p. 2992–3014, 2016. 8

- REHFELD, K. et al. Comparison of correlation analysis techniques for irregularly sampled time series. *Nonlinear Processes in Geophysics*, Copernicus GmbH, v. 18, n. 3, p. 389–404, 2011. [8](#)
- RIDDELL, C. A. et al. Classifying gestational weight gain trajectories using the sitar growth model. *Paediatric and perinatal epidemiology*, Wiley Online Library, v. 31, n. 2, p. 116–125, 2017. [47](#)
- RIDDELL, C. A. et al. Classifying gestational weight gain trajectories using the sitar growth model. *Paediatric and perinatal epidemiology*, Wiley Online Library, v. 31, n. 2, p. 116–125, 2017. [73](#)
- SILVERWOOD, R. J. *Issues in modelling growth data within a life course framework*. Tese (Doutorado) — London School of Hygiene and Tropical Medicine, 2008. [6](#)
- SIMPKIN, A. J. et al. Modelling height in adolescence: a comparison of methods for estimating the age at peak height velocity. *Annals of human biology*, Taylor & Francis, v. 44, n. 8, p. 715–722, 2017. [16](#)
- SINGER, J. M.; ANDRADE, D. d. Análise de dados longitudinais. *Simpósio Nacional de Probabilidade e Estatística*, Embrapa São Paulo, v. 7, 1986. [33](#)
- SPYRIDES, M. H. C. et al. Effect of predominant breastfeeding duration on infant growth: Prospective study using nonlinear mixed effect models. *Jornal de pediatria*, SciELO Brasil, v. 84, n. 3, p. 237–243, 2008. [1](#)
- STÜTZLE, W. et al. Shape-invariant modelling of human growth. *Annals of Human Biology*, Taylor & Francis, v. 7, n. 6, p. 507–528, 1980. [16](#)
- SULLIVAN, M. C. et al. Growth trajectories of preterm infants: birth to 12 years. *Journal of Pediatric Health Care*, Elsevier, v. 22, n. 2, p. 83–93, 2008. [1](#)
- SUZUKI, A. K. Modelos de sobrevivência bivariados baseados na cópula fgm: uma abordagem bayesiana. Universidade Federal de São Carlos, 2012. [24](#)
- TILLING, K. et al. Modelling childhood growth using fractional polynomials and linear splines. *Annals of Nutrition and Metabolism*, Karger Publishers, v. 65, n. 2-3, p. 129–138, 2014. [1](#), [2](#), [6](#), [9](#), [14](#), [15](#), [44](#)
- TOWERS, S. Potential fitting biases resulting from grouping data into variable width bins. *Physics Letters B*, Elsevier, v. 735, p. 146–148, 2014. [8](#)
- TU, J. V. et al. Impracticability of informed consent in the registry of the canadian stroke network. *New England Journal of Medicine*, Mass Medical Soc, v. 350, n. 14, p. 1414–1421, 2004. [3](#)
- VERBEKE, G. Linear mixed models for longitudinal data. In: *Linear mixed models in practice*. [S.l.]: Springer, 1997. p. 63–153. [2](#), [8](#)
- VILLAR, J. et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the newborn cross-sectional study of the intergrowth-21st project. *The Lancet*, Elsevier, v. 384, n. 9946, p. 857–868, 2014. [31](#)
- WADHAWAN, N. *Comparison of different weight growth models in a sample of children from 6 to 15 years*. Tese (Doutorado) — Concordia University, 2018. [16](#)

WEGMAN, E. J.; WRIGHT, I. W. Splines in statistics. *Journal of the American Statistical Association*, Taylor & Francis, v. 78, n. 382, p. 351–365, 1983. 13, 14

WHO. *World Health Organization. WHO Anthro Survey Analyser. World Health Organization*. 2011. Disponível em: <https://www.who.int/childgrowth/software/anthro_pc_manual_v322.pdf>, acesso em 22/11/2021. 44