

**INTERVALOS DE CONFIANÇA
PARA CURVAS PERCENTUAIS
DE PESO FETAL ESTIMADO
EM GESTAÇÕES GEMELARES**

Adriana Sañudo

Dissertação apresentada ao
Instituto de Matemática e Estatística
da Universidade de São Paulo
para obtenção do grau de
Mestre em Estatística

Área de concentração: **Estatística**
Orientador: **Prof. Dr. Julio da Motta Singer**

- São Paulo, novembro de 2000 -

INTERVALOS DE CONFIANÇA PARA CURVAS PERCENTUAIS
DE PESO FETAL ESTIMADO EM GESTAÇÕES GEMELARES

Este exemplar corresponde à redação final devidamente corrigida e defendida por Adriana Sañudo e aprovada pela comissão julgadora.

São Paulo, 27 de novembro de 2000.

Banca examinadora:

- Prof. Dr. Julio da Motta Singer
- Profa. Dra. Lúcia Pereira Barroso
- Profa. Dra. Renée Xavier de Menezes

*Aos meus pais, Enrique e Vera, cujo sacrifício e dedicação
em todos estes anos permitiram a minha realização profissional.
A vocês, a minha eterna gratidão pelo amor, carinho e apoio.*

Agradecimentos

Ao Prof. Julio da Motta Singer, pela orientação, amizade, confiança, apoio e extrema dedicação durante todas as fases deste trabalho, a quem devo gratidão e profundo respeito pessoal e profissional.

A todos os professores do Departamento de Estatística do IME e especialmente aos professores Dalton, Antonio Carlos e Lisbeth que despertaram o meu interesse pela pesquisa.

Aos amigos da Escola Paulista de Medicina Prof. Clóvis, Suely, Peninha, Alexandre, Rebeca, Graciela e Eloísa pelo apoio na fase final deste trabalho.

Aos amigos queridos do curso com os quais eu já estudei muito e não posso deixar de agradecer a amizade e paciência. Desses, não posso deixar de citar a Andréa, Ângela, Delhi, Karina, Liliam, Lílian, Luís, Pedro e a Raquel.

Aos meus familiares pelo apoio incondicional em todos os momentos.

Ao Centro de Estatística Aplicada pelo conjunto de dados analisados neste trabalho. Em especial à Dra. Mariza Marie Fujita do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo quem forneceu os dados e ao Alexandre Ryuzo Shinzato que executou a fase inicial da análise desse trabalho.

À Ângela Tavares Paes e ao Valdir Okano pela colaboração na parte computacional além do carinho e amizade.

Ao Edson Pavanello pelo seu companheirismo e apoio nesta fase final do trabalho.

Resumo

Consideramos a estimação de curvas percentuais para a avaliação do crescimento fetal ao longo do período gestacional presumido de 12 a 39 semanas. Utilizando modelos log-lineares de efeitos aleatórios com termos de primeiro e segundo graus, indicamos como bandas de confiança aproximadas podem ser construídas para as diversas curvas percentuais (3%, 10%, 25%, 50%, 75%, 90% e 97%) a partir dos elementos da matriz de covariância assintótica dos estimadores dos parâmetros do modelo. Ilustramos os resultados com dados de um estudo realizado no ambulatório de pré-natal da Clínica Obstétrica do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo. Esses dados correspondem a uma amostra de 80 fetos provenientes de gestações gemelares duplas avaliados ultra-sonograficamente em dois ou mais instantes ao longo do período gestacional. Confirmamos a validade da técnica proposta por intermédio da comparação dos resultados com aqueles provenientes de um estudo de simulação.

Abstract

We considered the estimation of percentual curves to evaluate fetal growth during the presumed pregnancy period of 12 to 39 weeks. Using random effects linear-log models with first and second-degree terms, we indicate how approximate confidence bands for different percentual curves (3%, 10%, 25%, 50%, 75%, 90% and 97%) may be obtained from assymptotic covariance matrix of the model parameters estimators. The results were illustrated with data obtained from a study carried out at the prenatal ambulatory of the Clínica Obstétrica do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo. The data correspond to a sample of 80 fetal from twin pregnancies, which were ultrasonographically evaluated at two or more instants throughout the pregnancy. We confirmed the validity of the proposed technique by comparing the results with those obtained from a simulation study.

Sumário

1	Introdução	2
2	Descrição dos Dados	6
3	Ajuste dos Modelos	9
4	Precisão das curvas percentuais	13
A	Tabelas	27
B	Gráficos	40
C	Expressão matricial do modelo de efeitos aleatórios	43

Capítulo 1

Introdução

Intervalos de referência e/ou Percentis populacionais de medidas antropométricas, concentração de algumas substâncias, etc. são importantes para vários ramos da Medicina. Por exemplo, nos laboratórios clínicos, intervalos de referência são ferramentas comuns para determinar medidas não usuais ou extremas de uma variável de interesse; se um valor individual estiver fora do intervalo correspondente, investigações clínicas adicionais são recomendadas para identificação de uma possível disfunção.

Um intervalo de referência de $\gamma\%$ é uma faixa de valores em torno da mediana que contém $\gamma\%$ dos dados. Por exemplo, um intervalo de referência de 90% contém observações que estão entre o 5° e o 95° percentis.

O termo ‘intervalo de referência’ foi sugerido por Gräsbeck and Saris (1969) e até o ano de 1970, a expressão ‘intervalos de normalidade’ era mais comumente usada em seu lugar; valores fora dessa faixa eram ditos ‘anormais’. Uma falha dessa terminologia é que, para o intervalo acima, por definição, 10% dos indivíduos ‘normais’ teriam valores fora do ‘intervalo de normalidade’. Além disso, esse termo sugere que a população de referência é constituída de ‘indivíduos normais, saudáveis’, o que pode nem sempre ser o caso. Também pode existir uma possível confusão com a distribuição Gaussiana (usualmente referida como ‘normal’); modelar os dados sob suposição de normalidade é uma alternativa comum mas nem sempre apropriada para a estimação dos limites de referência. Apesar disso, a expressão ‘intervalo de normalidade’ ainda é usada (Wright and Royston, 1999).

Em muitos casos a variável de interesse pode ser afetada por características da

população de referência como idade, sexo, hábito de fumar, peso, altura ou fatores genéticos. Nesses casos, convém considerar a construção de intervalos de referência estratificados, ou seja, intervalos para cada subpopulação gerada pelas combinações dos níveis das características intervenientes.

Supondo que a distribuição da variável de interesse (X) seja Gaussiana, o percentil populacional de ordem γ ($0\% < \gamma < 100\%$) pode ser estimado com base numa amostra (aleatória simples) de tamanho n por $X_\gamma = \bar{X} + z_\gamma S$, em que $\bar{X} = \sum_{i=1}^n X_i/n$ é a média amostral, z_γ é o quantil de ordem γ da distribuição Gaussiana padrão e $S = \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 / (n - 1)}$ é o desvio padrão amostral. Por exemplo, na construção do intervalo de referência de 90% da proteína IgM em uma população pediátrica na faixa etária de sete a nove meses, Jolliff et al. (1982) obtiveram $\bar{X} = 0,80g/L$ e $S = 0,23g/L$; utilizando, respectivamente, $z_{2,5\%} \cong -2$ e $z_{97,5\%} \cong 2$ como valores aproximados dos quantis de ordem 5% e 95% da distribuição Gaussiana padrão, o percentil de ordem 5% da distribuição da proteína IgM na subpopulação em estudo é estimado por $X_{5\%} = 0,34g/L$ e o de ordem 95% por $X_{95\%} = 1,26g/L$.

Se em vez de estratificar a população de referência há interesse em estudar uma variável (peso fetal, por exemplo) ao longo de uma determinada escala ordenada (como idade gestacional), a obtenção de percentis populacionais envolve conceitos e técnicas estatísticas mais sofisticadas. Nesse contexto, o que se deseja são **Curvas de referência** e/ou **Curvas percentuais**, cuja obtenção pode ser encarada como um problema de ajuste de curvas de crescimento. No exemplo citado, essas curvas são importantes para o acompanhamento do peso fetal durante o período gestacional, pois a partir delas podem-se propor condutas obstétricas e/ou neonatais adequadas.

Curvas de peso ao nascimento, isto é, curvas construídas através dos pesos dos recém-nascidos vivos para diferentes idades gestacionais, existem há algum tempo na literatura, tanto para gestações únicas (Lubchenco et al., 1963) quanto para gestações gemelares (Naeye et al., 1966). Sob o ponto de vista da aplicação prática, a limitação desses estudos pós-natais reside no fato de que crianças nascidas prematuramente são incluídas na base de dados e podem viesar os resultados (Fujita, 1998).

Com a introdução de técnicas obstétricas ultra-sonográficas, é possível avaliar o crescimento fetal por meio de múltiplos parâmetros biométricos e fazer estimativas de

peso intra-uterino (Fujita, 1998), eliminando assim a deficiência mencionada acima.

As estimativas de curvas de peso intra-uterino podem ser obtidas de estudos realizados sob dois tipos de planejamento para a coleta dos dados: o **transversal** (“cross-sectional”) e o **longitudinal**. Caso cada feto seja avaliado somente em uma única idade gestacional, o estudo é chamado de transversal, ao passo que se cada feto for avaliado em diferentes idades gestacionais, o estudo é do tipo longitudinal. Como na prática clínica é comum avaliar o feto mais do que uma vez durante a gestação, os dados assim gerados são freqüentemente longitudinais.

Deter et al. (1982) construíram curvas de crescimento médio para fetos únicos com base em estudos longitudinais. Hadlock et al. (1982) apresentam críticas às curvas de crescimento obtidas por análise transversal, uma vez que podem não refletir com precisão o crescimento fetal individual e sugerem que as curvas derivadas de estudos longitudinais são mais apropriadas para esse fim, além de representarem melhor o crescimento médio da população investigada. Wright and Royston (1997, 1999) discutem, comparam e fazem uma revisão dos métodos estatísticos apropriados para estimar tanto intervalos de referência como curvas de referência para dados provenientes de estudos transversais. Royston (1991) discute métodos para estimar curvas de referência para o crescimento fetal de gestações únicas em estudos longitudinais.

É importante distinguir dois objetivos da construção de curvas de crescimento: a estimação de uma curva média e a estimação de curvas percentuais. Para o primeiro caso, o interesse está em estimar uma curva que modele a relação entre a média da variável de interesse (por exemplo, peso fetal) e o tempo (idade gestacional). Quando a distribuição da variável de interesse é simétrica, a curva média é a curva mediana, ou seja, corresponde ao percentil 50%. Através das curvas percentuais podem-se avaliar trajetórias individuais com mais informação do que aquela conseguida apenas com a curva média.

A precisão das curvas percentuais (a média incluída) é um item importante nesse processo, embora raramente seja explicitada. Essa precisão pode nos dar informações sobre a quantidade e a qualidade dos dados. O principal objetivo deste trabalho é propor métodos de cálculo da precisão de curvas de crescimento percentuais para o peso fetal a partir de um conjunto de dados obtidos de um estudo longitudinal. Em

particular, consideramos gestações gemelares duplas dado que nestes casos os estudos são poucos e têm resultados e conclusões muitas vezes conflitantes (Fujita, 1998). O interesse obstétrico na avaliação do crescimento fetal para gestações gemelares reside na tendência de o peso médio dos recém-nascidos nesses casos ser menor do que o de gestações únicas, justificando um tratamento diferenciado.

Devido à impossibilidade de mensurar diretamente o peso do feto (PF), ele deve ser estimado por meio de outras características antropométricas, como diâmetro biparietal (DBP), circunferência cefálica (CC), circunferência abdominal (CA) e comprimento do fêmur (CF), obtidas a partir de exames ultra-sonográficos. Hadlock et al. (1985) sugerem a seguinte fórmula, largamente utilizada na prática, para obter uma estimativa do peso fetal

$$\log PF = 1,3596 - 0,00386 CA \cdot CF + 0,0064 CC + 0,00061 DBP \cdot CA + 0,0424 CA + 0,174 CF. \quad (1.1)$$

No capítulo 2, estão descritos o planejamento, os objetivos e o processo de coleta de dados do estudo de onde foram obtidos os dados. No capítulo 3, descrevemos a estratégia empregada por Singer e Shinzato (1999) para a estimação das curvas de crescimento e apresentamos o modelo de efeitos aleatórios sugerido por esses autores. Finalmente, no capítulo 4, propomos um método analítico para a construção dos intervalos de confiança para essas curvas e comparâmo-lo com os intervalos de confiança obtidos a partir de uma simulação similar àquela realizada por Royston (1991).

Capítulo 2

Descrição dos Dados

Os dados analisados neste trabalho foram provenientes de um estudo clínico, prospectivo e longitudinal realizado num período de três anos, durante o qual foram acompanhadas gestantes do Ambulatório de Pré-Natal da Clínica Obstétrica do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo. Desse acompanhamento, foi obtida uma amostra de 80 fetos, satisfazendo o seguinte protocolo:

- provenientes de gestações gemelares duplas;
- ausência de patologia materna intercorrente durante a gestação;
- ausência de malformação ou óbito fetal e
- concordância do peso dos gêmeos (diferença entre os pesos menor ou igual a 20% do peso do maior gêmeo ao nascimento).

Para efeito deste projeto, as características observadas nesses fetos foram:

- identificação da gestante;
- identificação do feto dentro de cada gestação;
- idade gestacional (semanas) e
- peso fetal estimado (gramas).

A idade gestacional foi determinada com base em um de três critérios diferentes:

1. identificação do primeiro dia após a última menstruação;
2. identificação do comprimento cabeça-nádega (Robinson and Fleming, 1975);
3. média aritmética da idade estimada através de vários parâmetros biométricos (Hadlock et al., 1985).

O critério 1 foi utilizado nos casos em que a data da última menstruação era conhecida, porém deveria ser coerente com a estimativa obtida por um dos outros dois critérios na primeira ultra-sonografia. O critério 2 foi utilizado para estimar as idades gestacionais inferiores ou iguais a 12 semanas e o critério 3 foi utilizado para estimar as idades gestacionais nos demais casos.

Cada um dos 40 pares de fetos foi avaliado ultra-sonograficamente em dois ou mais instantes ao longo do período gestacional presumido de 12 a 39 semanas. Na Tabela A.1 (Apêndice A) apresentamos os pesos estimados para cada um dos fetos. Como as ocasiões em que foram obtidas as estimativas do peso fetal são distintas para cada mãe, pode-se classificar este estudo como longitudinal desbalanceado em relação ao tempo. Vale ressaltar, que nem sempre é possível observar todas as medidas, e assim não há como obter uma estimativa para o peso fetal nas idades gestacionais avaliadas.

Para cada uma das idades gestacionais avaliadas encontram-se dispostos na Tabela A.2, a média, mediana, desvio padrão, primeiro e terceiro quartis e o coeficiente de variação dos pesos fetais estimados. Pode-se perceber que os valores do desvio padrão tendem a aumentar com a idade gestacional, sugerindo uma provável heterocedasticidade dos dados. Nota-se ainda que o coeficiente de variação parece estar estabilizado em torno de 10% a 15%, indicando que o desvio padrão pode se relacionar de forma linear com a média. Além disso, pode-se notar que os valores da média e mediana são próximos, sugerindo uma simetria das distribuições envolvidas.

O gráfico de perfis individuais observados para o peso fetal estimado está apresentado na Figura B.1 (Apêndice B). Uma análise desse gráfico sugere que curvas exponenciais ou polinomiais do segundo ou terceiro grau podem ser utilizadas para aproximar as curvas de crescimento de interesse dentro do intervalo gestacional considerado (12 a 39 semanas). Na Figura B.2 apresentamos gráficos de perfis individuais do peso fetal estimado para cada um dos quarenta pares de gêmeos de cada gestação. O

comportamento semelhante de ambos os fetos de cada par sugere a inclusão de gestante como um fator adicional nos modelos estatísticos a serem empregados na análise.

Para cada idade gestacional, a hipótese de que os dados provêm de uma distribuição Gaussiana foi avaliada por intermédio do teste Q-Q (Johnson and Wichern, 1992). Esse teste foi concretizado para as idades gestacionais com pelo menos 5 observações. Em todos os casos, a hipótese de que os pesos fetais estimados têm distribuição Gaussiana não foi rejeitada. Os coeficientes obtidos e os valores contra os quais eles são comparados podem ser observados na Tabela A.3.

Capítulo 3

Ajuste dos Modelos

O fato de várias observações terem sido realizadas em cada feto e o de que se lidou com dois fetos por gestante sugerem que as observações podem ser correlacionadas e, conseqüentemente, técnicas de análise de dados longitudinais são mais apropriadas para o ajuste das curvas (ver Singer and Andrade (2000), por exemplo).

Dentre os modelos estatísticos disponíveis para tal finalidade, selecionaram-se os chamados modelos de efeitos aleatórios, pois permitem modelar curvas sob diversas estruturas de dependência intra-individual e ainda permitem considerar variâncias heterogêneas para as diferentes idades gestacionais; a incorporação dessa característica parece necessária tendo em vista os perfis individuais apresentados na Figura B.1. Apesar de sua denominação, esses modelos incluem tanto os efeitos fixos como os aleatórios. Os efeitos fixos determinam o perfil médio populacional e estes acrescidos dos efeitos aleatórios modelam os perfis individuais.

A estratégia de modelagem incluiu as três etapas descritas a seguir.

1. Obtenção de modelos aceitáveis para o perfil médio populacional levando em consideração apenas a correlação entre as medidas realizadas nos mesmos fetos. Nesta etapa a avaliação do ajuste dos modelos foi baseada na significância estatística dos coeficientes e de critérios de ajuste como o Critério da Informação de Akaike (AIC) e o Critério Bayesiano da Informação de Schwarz (BIC); maiores detalhes sobre o uso de tais critérios podem ser obtidos em Jones (1993), por exemplo.

2. Avaliação dos modelos selecionados em 1 sob uma estrutura de covariância mais complexa, envolvendo um termo adicional de correlação entre as medidas realizadas em fetos irmãos. Nesta etapa, a avaliação do ajuste de modelos também incluiu análise de resíduos.
3. Construção das curvas de peso fetal estimado para os percentis populacionais 3%, 10%, 25%, 50%, 75%, 90% e 97% para cada modelo selecionado em 2 e escolha do melhor modelo em termos da compatibilidade dos percentis populacionais estimados com os amostrais.

A primeira etapa do processo de obtenção dos modelos se iniciou com 16 modelos, sendo 8 modelos polinomiais para peso fetal estimado e 8 modelos polinomiais para o logaritmo natural do peso fetal estimado. Todos eles continham efeitos fixos e aleatórios lineares, quadráticos e cúbicos para a idade gestacional (IdG , IdG^2 e IdG^3) e se diferenciavam pela presença ou não dos efeitos fixos e aleatórios do intercepto, e do tipo de estrutura da matriz de covariâncias para os efeitos aleatórios (não estruturada ou definida por componentes de variância). Os modelos foram ajustados por meio de métodos de máxima verossimilhança restrita (ver Diggle et al. (1994), por exemplo). Para os cálculos utilizou-se o procedimento MIXED do software SAS (SAS/STAT). Com base na significância estatística dos parâmetros fixos e/ou aleatórios desses modelos, outros mais parcimoniosos (com menor número de parâmetros) foram gerados e seu ajuste reavaliado.

A cada um dos seis modelos selecionados na primeira etapa foi adicionado um efeito aleatório para modelar a correlação entre as medidas realizadas em fetos irmãos. Com base em uma análise de resíduos, três modelos com ajuste adequado e resultados mais coerentes com as suposições dos modelos de efeitos aleatórios foram selecionados para a terceira etapa.

Após a construção das curvas de peso fetal estimado para os vários percentis populacionais foram obtidos o número e a porcentagem de fetos com pelo menos 50% das observações limitadas superiormente pelas curvas percentuais estimadas. Maiores detalhes podem ser obtidos em Singer e Shinzato (1999). Com base nisso pode-se concluir que o modelo com melhor ajuste é dado por

$$\log PF = 0,78386 + 0,32835 \cdot IdG - 0,00372 \cdot IdG^2 + \text{erro aleatório} \quad (3.1)$$

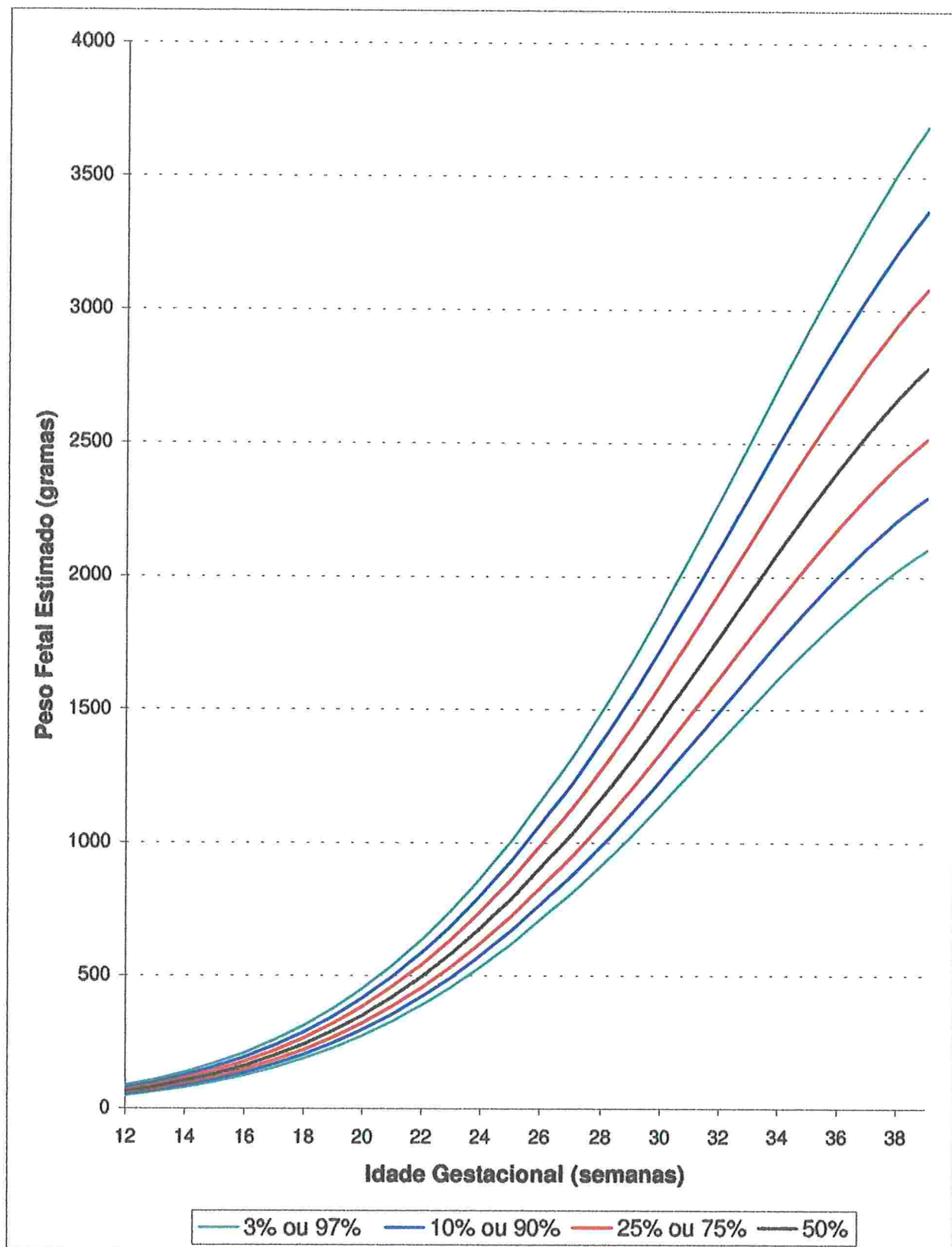
Pelo modelo ajustado, a relação entre o peso fetal estimado e a idade gestacional segue uma curva exponencial. Esta curva é semelhante àquela obtida por Hadlock et al. (1991) em um estudo do tipo transversal para fetos únicos, sugerindo a plausibilidade do modelo. A estrutura de covariância associada ao modelo inclui efeitos aleatórios de intercepto, idade gestacional e gestante.

Na Tabela 3.1 apresentamos as estimativas dos efeitos fixos e das variâncias dos efeitos aleatórios do modelo ajustado com os respectivos erros padrão. Na Figura 3.1 exibimos a curva média ajustada (curva correspondente ao percentil 50%) além de várias curvas percentuais populacionais estimadas sob o modelo descrito em (3.1).

Tabela 3.1: Estimativas e erros padrão para os efeitos fixos e aleatórios (Modelo 3.1).

	Parâmetro	Estimativa	Erro Padrão
Efeitos Fixos	Intercepto	0,78386274	0,06240590
	IdG	0,32835444	0,00435865
	IdG^2	-0,00371982	0,00007813
Variância dos Efeitos Aleatórios	Intercepto	0,02078325	0,00821534
	IdG	0,00002924	0,00000822
	Cov(Intercepto,IdG)	-0,00073119	0,00024583
	Gestante	0,00961932	0,00267804
	Resíduo	0,00453083	0,00032030

Figura 3.1: Curvas de percentis populacionais estimados.



Capítulo 4

Precisão das curvas percentuais

Este capítulo está dividido em duas seções; na primeira indicamos um método analítico para a construção dos intervalos de confiança para as curvas percentuais populacionais; na segunda seção, a construção dos intervalos de confiança é concretizada por intermédio de uma simulação similar àquela realizada por Royston (1991).

O modelo de efeitos aleatórios dado por (3.1) pode ser especificado genericamente como

$$y_{ij}(t) = \alpha + \beta_1 t + \beta_2 t^2 + a_{ij} + b_{ij} t + \epsilon_{ij}(t) \quad (4.1)$$

com $i = 1, 2$, $j = 1, \dots, 40$ e $t = 12, \dots, 39$ em que

- i) $y_{ij}(t)$ é o logaritmo natural do peso fetal estimado para o i -ésimo feto da j -ésima mãe na idade gestacional t ;
- ii) α , β_1 e β_2 são os coeficientes polinomiais (fixos) de grau zero, um e dois, respectivamente;
- iii) a_{ij} e b_{ij} são os coeficientes (aleatórios) associados aos termos de grau zero e um, respectivamente, para o i -ésimo feto da j -ésima mãe, e
- iv) $\epsilon_{ij}(t)$ é um termo aleatório associado ao i -ésimo feto da j -ésima mãe na idade gestacional t .

Assumimos também que os coeficientes aleatórios seguem uma distribuição normal

bivariada, ou seja, que

$$\begin{pmatrix} a_{ij} \\ b_{ij} \end{pmatrix} \sim N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix} \right]$$

com $\sigma_a^2 > 0$ e $\sigma_b^2 > 0$. Além disso, o termo aleatório é decomposto em

$$\epsilon_{ij}(t) = m_j + \delta_{ij}(t) \quad (4.2)$$

em que $m_j \sim N(0, \sigma_m^2)$ e $\delta_{ij}(t) \sim N(0, \sigma^2)$ são variáveis aleatórias independentes e denotam os efeitos de gestante e erro de medida, respectivamente.

Essas suposições nos permitem escrever

$$\text{Var}[\epsilon_{ij}(t)] = \sigma_m^2 + \sigma^2$$

e

$$\begin{aligned} \text{Cov}[\epsilon_{1j}(t), \epsilon_{2j}(t)] &= \mathbb{E}[(m_j + \delta_{1j}(t))(m_j + \delta_{2j}(t))] \\ &= \mathbb{E}[m_j^2 + m_j\delta_{2j}(t) + m_j\delta_{1j}(t) + \delta_{1j}(t)\delta_{2j}(t)] \\ &= \mathbb{E}[m_j^2] \\ &= \sigma_m^2. \end{aligned}$$

Assumindo independência entre todos os componentes aleatórios do modelo, ou sejam, a_{ij} , b_{ij} , m_j e $\delta_{ij}(t)$, temos

$$\mathbb{E}[y_{ij}(t)] = \mathbb{E}[y(t)] = \alpha + \beta_1 t + \beta_2 t^2 \quad (4.3)$$

e

$$\begin{aligned} \text{Var}[y_{ij}(t)] &= \text{Var}[y(t)] = \text{Var}[a_{ij} + b_{ij}t + \epsilon_{ij}(t)] \\ &= \text{Var}[a_{ij}] + t^2 \text{Var}[b_{ij}] + \text{Var}[\epsilon_{ij}(t)] \\ &\quad + 2t \text{Cov}[a_{ij}, b_{ij}] + 2 \text{Cov}[a_{ij}, \epsilon_{ij}(t)] \\ &\quad + 2t \text{Cov}[b_{ij}, \epsilon_{ij}(t)] \\ &= \sigma_a^2 + t^2 \sigma_b^2 + 2t \sigma_{ab} + \sigma_m^2 + \sigma^2. \end{aligned} \quad (4.4)$$

O percentil populacional de ordem γ ($0\% < \gamma < 100\%$) para a idade gestacional t é dado por

$$y_\gamma(t) = \mathbb{E}[y(t)] + z_\gamma \sqrt{\text{Var}[y(t)]} \quad (4.5)$$

em que z_γ é o quantil de ordem γ da distribuição Gaussiana padrão.

Para estimar esse percentil precisamos de estimadores de $\mathbb{E}[y(t)]$ e $\text{Var}[y(t)]$. Utilizando a notação $\hat{\theta}$ para denotar um estimador de θ , temos de (4.3) e (4.4) que

$$\hat{\mathbb{E}}[y(t)] = \hat{y}(t) = \hat{\alpha} + \hat{\beta}_1 t + \hat{\beta}_2 t^2 \quad (4.6)$$

com $\hat{\alpha}$, $\hat{\beta}_1$ e $\hat{\beta}_2$ denotando estimadores de α , β_1 e β_2 , respectivamente, e

$$\widehat{\text{Var}}[y(t)] = \hat{\sigma}_a^2 + t^2 \hat{\sigma}_b^2 + 2 t \hat{\sigma}_{ab} + \hat{\sigma}_m^2 + \hat{\sigma}^2 \quad (4.7)$$

com $\hat{\sigma}_a^2$, $\hat{\sigma}_b^2$, $\hat{\sigma}_{ab}$, $\hat{\sigma}_m^2$ e $\hat{\sigma}^2$, denotando estimadores de σ_a^2 , σ_b^2 , σ_{ab} , σ_m^2 e σ^2 , respectivamente.

Com base em (4.6) e (4.7) um estimador para (4.5) é dado por

$$\hat{y}_\gamma(t) = \hat{y}(t) + z_\gamma \sqrt{\widehat{\text{Var}}[y(t)]}. \quad (4.8)$$

Variando γ , esse resultado permite construir as curvas de crescimento associadas aos percentis de interesse. A escolha dos estimadores utilizados acima merece alguma discussão. No caso dos parâmetros de regressão (α , β_1 e β_2) o método de estimação mais utilizado é o de máxima verossimilhança (MV), pois sob condições de regularidade (ver Sen and Singer, (1993), por exemplo) satisfeitas nos casos que examinamos aqui, os estimadores gerados são não viciados e assintoticamente eficientes. Para os componentes da estrutura de covariância σ_a^2 , σ_b^2 , σ_{ab} , σ_m^2 e σ^2 , o método de MV produz estimadores viciados (ver Searle et al., (1992), por exemplo) pelo fato de não levar em consideração a perda de graus de liberdade resultantes da estimação dos parâmetros de regressão (α , β_1 e β_2). Neste caso recomenda-se utilizar o método de máxima verossimilhança restrita (MVR) proposto originalmente por Patterson and Thompson (1971) e particularizado por Laird and Ware (1982) para modelos de efeitos aleatórios como aqueles que investigamos neste trabalho. Segundo o método de MVR, a estimação dos componentes de variância baseia-se nos resíduos calculados depois do ajuste por mínimos

quadrados ordinários dos efeitos fixos. Em vez de se usar as observações diretamente, o método de MVR utiliza uma combinação linear dos dados, escolhida de tal modo que não contenha qualquer efeito fixo, não importando qual seja o seu valor. Esse procedimento é equivalente à obtenção dos estimadores de MV dos componentes de variância a partir dos resíduos gerados pelo método de mínimos quadrados ordinários. Por exemplo, quando os dados são provenientes de uma amostra aleatória simples, X_1, \dots, X_n , independente e identicamente distribuída segundo uma distribuição $N(\mu, \sigma^2)$, o estimador de MV de σ^2 é dado por $\hat{\sigma}^2 = \sum_{i=1}^n (X_i - \bar{X})^2/n$ e é viciado para σ^2 , pois $E(\hat{\sigma}^2) = \sigma^2(1 - 1/n)$ ao passo que o estimador de MVR, $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2/(n - 1)$, é não viciado para σ^2 . Em geral a estimação pelo método de MVR produz estimativas menos viciadas dos componentes de variância do que aquelas geradas pelo método de MV (Jones, 1993). Utilizando um estudo de simulação, Fellingham and Raghunathan (1995) discutem as suposições distribucionais sobre os efeitos aleatórios em análise de dados longitudinais e concluem que quando os efeitos aleatórios não seguem uma distribuição Gaussiana, mas são simétricos, os estimadores de MVR dos efeitos fixos não são afetadas, mas os estimadores dos componentes de variância são altamente variáveis. Se os efeitos aleatórios provêm de uma distribuição assimétrica, o método de MVR produz estimativas bastante deficientes. Nesse trabalho utilizaremos estimadores gerados pelo método de MVR mesmo sabendo que eles podem ser viciados. O resultado obtido é avaliado por intermédio da comparação com outro obtido de um estudo de simulação como aquele discutido em Royston (1991).

Para avaliar a precisão das curvas definidas por (4.8) é necessário obter os correspondentes intervalos de confiança (IC) cujos limites inferior e superior para um coeficiente de confiança λ ($0\% < \lambda < 100\%$) são respectivamente

$$\hat{y}_\gamma^-(t) = \hat{y}_\gamma(t) - z_\lambda \sqrt{\widehat{\text{Var}}[\hat{y}_\gamma(t)]}$$

$$\hat{y}_\gamma^+(t) = \hat{y}_\gamma(t) + z_\lambda \sqrt{\widehat{\text{Var}}[\hat{y}_\gamma(t)]}$$

em que z_λ é o quantil de ordem $(1+\lambda)/2$ da distribuição Gaussiana padrão, ($0 < \lambda < 1$).

Para determinar os valores dos limites do IC, precisamos obter um estimador de

$\text{Var}[\hat{y}_\gamma(t)]$. Nesse sentido, observemos que

$$\begin{aligned}\text{Var}[\hat{y}_\gamma(t)] &= \text{Var}\left[\hat{y}(t) + z_\gamma \sqrt{\widehat{\text{Var}}[y(t)]}\right] \\ &= \text{Var}[\hat{y}(t)] + z_\gamma^2 \text{Var}\left[\sqrt{\widehat{\text{Var}}[y(t)]}\right] + 2z_\gamma \text{Cov}\left[\hat{y}(t), \sqrt{\widehat{\text{Var}}[y(t)]}\right].\end{aligned}\quad (4.9)$$

Desenvolvamos então expressões para o cálculo de cada parcela do segundo membro de (4.9). Em primeiro lugar, notemos que

$$\begin{aligned}\text{Var}[\hat{y}(t)] &= \text{Var}\left[\hat{\alpha} + \hat{\beta}_1 t + \hat{\beta}_2 t^2\right] \\ &= \text{Var}[\hat{\alpha}] + t^2 \text{Var}[\hat{\beta}_1] + t^4 \text{Var}[\hat{\beta}_2] + 2t \text{Cov}[\hat{\alpha}, \hat{\beta}_1] \\ &\quad + 2t^2 \text{Cov}[\hat{\alpha}, \hat{\beta}_2] + 2t^3 \text{Cov}[\hat{\beta}_1, \hat{\beta}_2].\end{aligned}$$

Agora, para calcular $\text{Var}\left[\sqrt{\widehat{\text{Var}}[y(t)]}\right]$ primeiramente observemos que, pelo Método Delta (ver Sen and Singer (1993), por exemplo) temos

$$\text{Var}\left[\sqrt{\widehat{\text{Var}}[y(t)]}\right] \cong \frac{1}{4 \cdot \mathbb{E}\left[\widehat{\text{Var}}[y(t)]\right]} \cdot \text{Var}\left[\widehat{\text{Var}}[y(t)]\right]$$

com

$$\begin{aligned}\mathbb{E}\left[\widehat{\text{Var}}[y(t)]\right] &= \mathbb{E}\left[\hat{\sigma}_a^2 + t^2 \hat{\sigma}_b^2 + 2t \hat{\sigma}_{ab} + \hat{\sigma}^2 + \hat{\sigma}_m^2\right] \\ &= \mathbb{E}\left[\hat{\sigma}_a^2\right] + t^2 \mathbb{E}\left[\hat{\sigma}_b^2\right] + 2t \mathbb{E}\left[\hat{\sigma}_{ab}\right] + \mathbb{E}\left[\hat{\sigma}^2\right] + \mathbb{E}\left[\hat{\sigma}_m^2\right] \\ &\cong \sigma_a^2 + t^2 \sigma_b^2 + 2t \sigma_{ab} + \sigma^2 + \sigma_m^2.\end{aligned}$$

Se utilizarmos estimadores de MVR de σ_a^2 , σ_b^2 , σ_{ab} , σ_m^2 e σ^2 , um estimador aproximadamente não viciado de $\mathbb{E}\left[\widehat{\text{Var}}[y(t)]\right]$ é

$$\hat{\mathbb{E}}\left[\widehat{\text{Var}}[y(t)]\right] = \hat{\sigma}_a^2 + t^2 \hat{\sigma}_b^2 + 2t \hat{\sigma}_{ab} + \hat{\sigma}_m^2 + \hat{\sigma}^2.$$

A magnitude de seu possível vício será avaliada através de um estudo de simulação .

Por outro lado, utilizando (4.7) observemos que

$$\begin{aligned}
\text{Var} \left[\widehat{\text{Var}}[y(t)] \right] &= \text{Var} \left[\hat{\sigma}_a^2 + t^2 \hat{\sigma}_b^2 + 2 t \hat{\sigma}_{ab} + \hat{\sigma}^2 + \hat{\sigma}_m^2 \right] \\
&= \text{Var} \left[\hat{\sigma}_a^2 \right] + t^4 \text{Var} \left[\hat{\sigma}_b^2 \right] + 4 t^2 \text{Var} \left[\hat{\sigma}_{ab} \right] \\
&\quad + \text{Var} \left[\hat{\sigma}^2 \right] + \text{Var} \left[\hat{\sigma}_m^2 \right] + 2 t^2 \text{Cov} \left[\hat{\sigma}_a^2, \hat{\sigma}_b^2 \right] \\
&\quad + 4 t \text{Cov} \left[\hat{\sigma}_a^2, \hat{\sigma}_{ab} \right] + 2 \text{Cov} \left[\hat{\sigma}_a^2, \hat{\sigma}_m^2 \right] \\
&\quad + 2 \text{Cov} \left[\hat{\sigma}_a^2, \hat{\sigma}^2 \right] + 4 t^3 \text{Cov} \left[\hat{\sigma}_b^2, \hat{\sigma}_{ab} \right] \\
&\quad + 2 t^2 \text{Cov} \left[\hat{\sigma}_b^2, \hat{\sigma}^2 \right] + 2 t^2 \text{Cov} \left[\hat{\sigma}_b^2, \hat{\sigma}_m^2 \right] \\
&\quad + 4 t \text{Cov} \left[\hat{\sigma}_{ab}, \hat{\sigma}^2 \right] + 4 t \text{Cov} \left[\hat{\sigma}_{ab}, \hat{\sigma}_m^2 \right] \\
&\quad + 2 \text{Cov} \left[\hat{\sigma}^2, \hat{\sigma}_m^2 \right].
\end{aligned}$$

Um estimador de $\text{Var} \left[\widehat{\text{Var}}[y(t)] \right]$ é obtido substituindo as variâncias e co-variâncias na expressão pelos componentes da matriz de covariância assintótica de $(\hat{\sigma}_a^2, \hat{\sigma}_b^2, \hat{\sigma}_{ab}, \hat{\sigma}_m^2, \hat{\sigma}^2)$. A expressão desse estimador pode ser obtida do resultado (C.4) do Apêndice C.

Finalmente, notemos que no último termo de (4.9), $\hat{y}(t)$ é uma função linear do vetor α definido em (C.1) do Apêndice C, e que $\widehat{\text{Var}}[y(t)]$ é uma função de uma transformação do vetor de dados que não depende de α , dado que corresponde ao estimador de MVR. Então, $\text{Cov} \left[\hat{y}(t), \sqrt{\widehat{\text{Var}}[y(t)]} \right] = \mathbf{0}$ pelo resultado (C.8).

Os resultados da aplicação ao conjunto de dados sob investigação estão dispostos nas Tabelas A.5 a A.7 (Apêndice A) e Figuras 4.1 a 4.3.

Figura 4.1: Percentis populacionais 3% e 97% estimados e intervalos de confiança (95%).

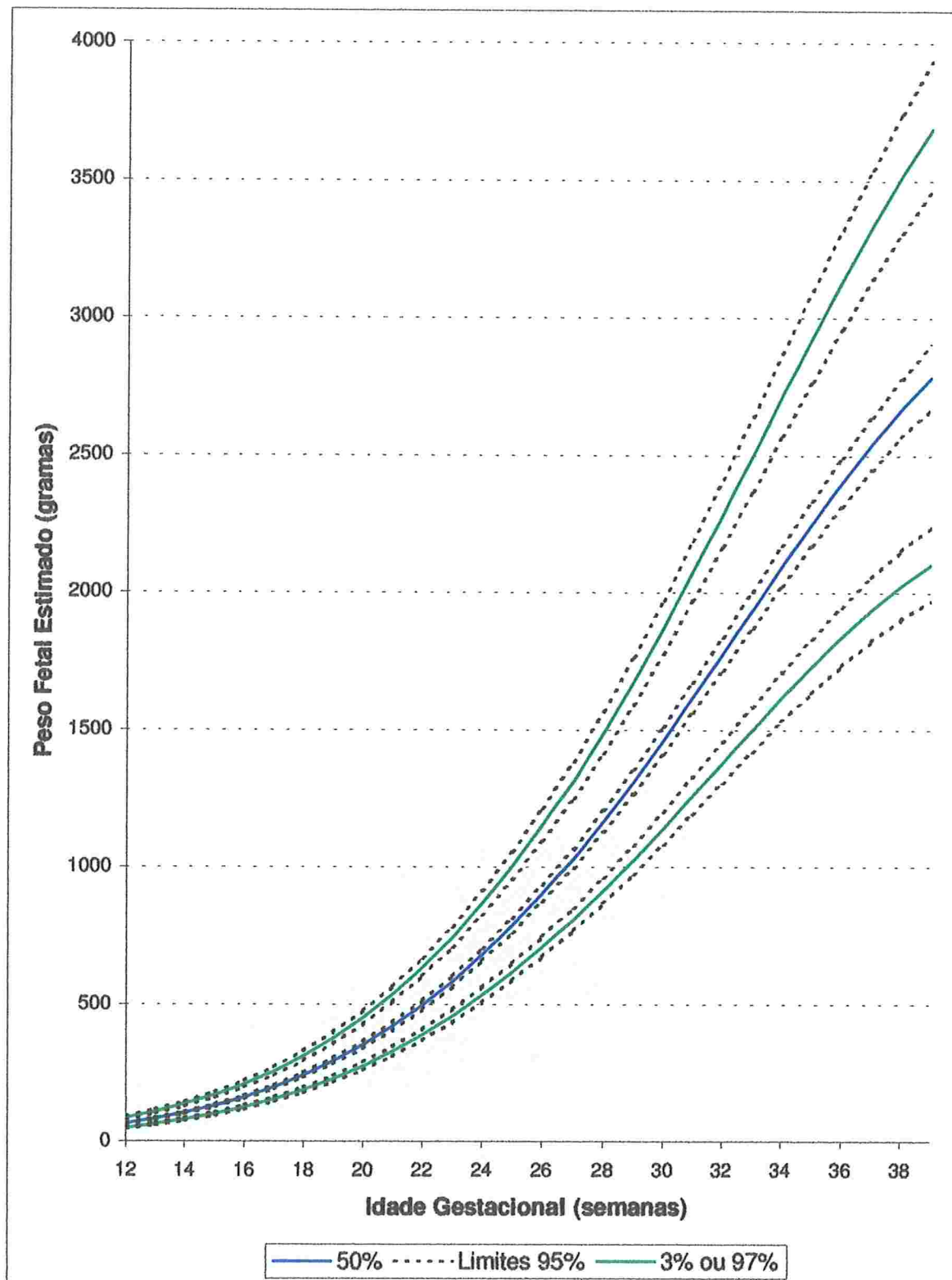


Figura 4.2: Percentis populacionais 10% e 90% estimados e intervalos de confiança (95%).

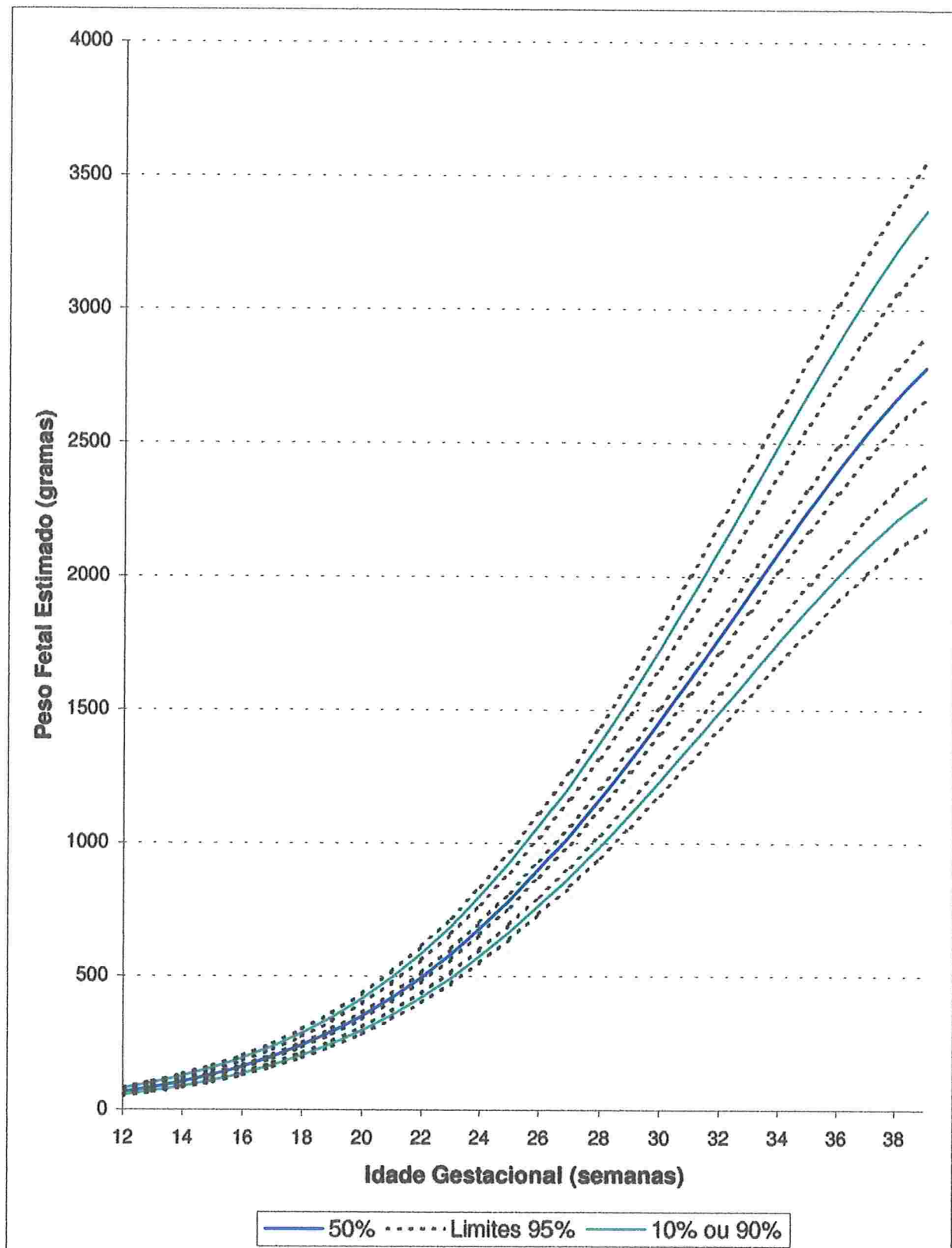
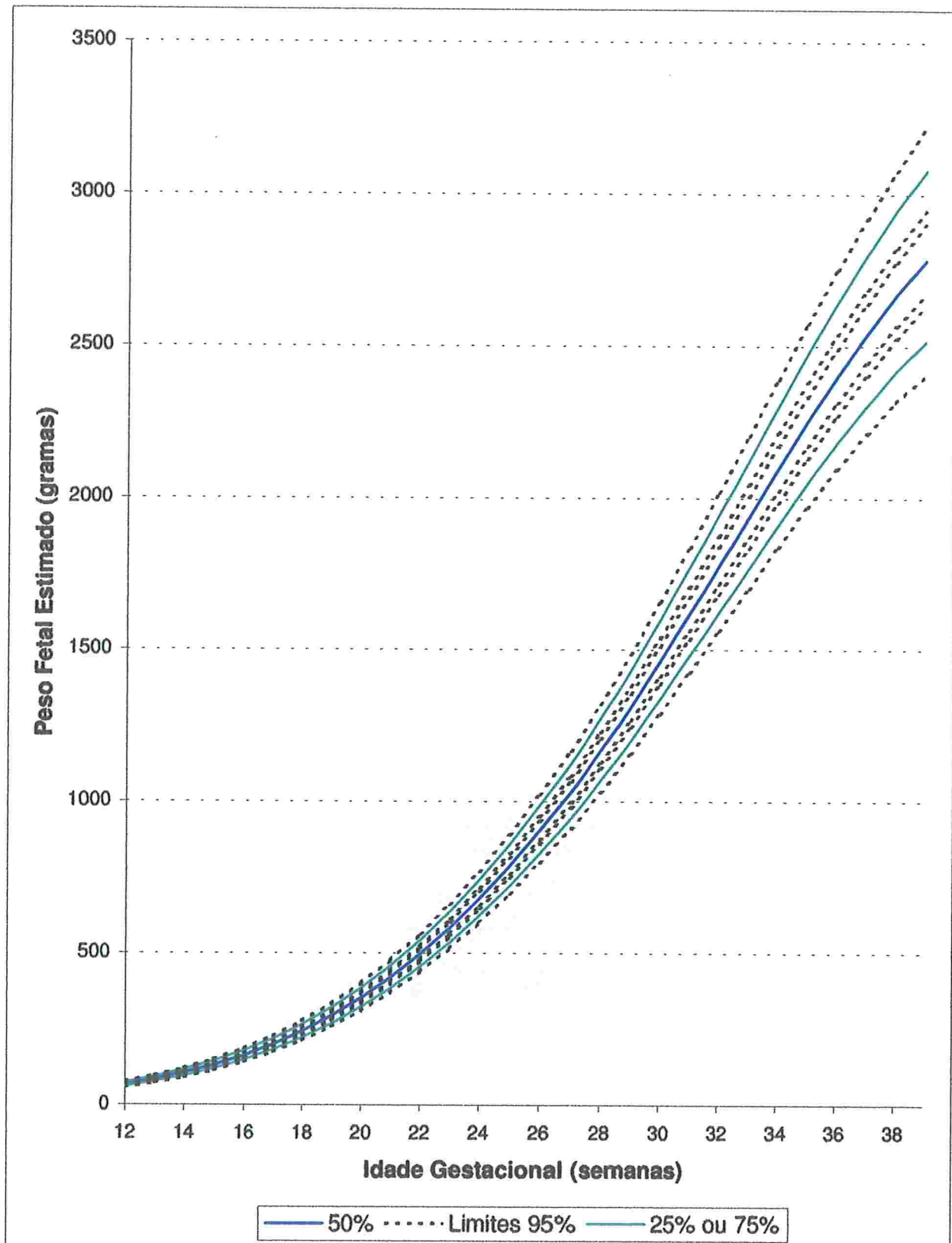


Figura 4.3: Percentis populacionais 25% e 75% estimados e intervalos de confiança (95%).



As estimativas utilizadas para o cálculo dos limites de confiança para os percentis populacionais no método analítico foram obtidas pelo método de MVR. Com o objetivo de verificar qual o efeito do possível vício dessas estimativas nos limites de confiança, consideramos um estudo de simulação.

Foram gerados 200 conjuntos de dados para o logaritmo natural do peso fetal estimado, levando-se em consideração a estrutura de incidência apresentada na Tabela A.1. Para a construção de cada um desses conjuntos de dados, utilizamos as estimativas dispostas na Tabela 3.1 como valores dos parâmetros. Mais especificamente, foram gerados 80 valores (um para cada feto das 40 mães) de acordo com a distribuição

$$\begin{pmatrix} a_{ij} \\ b_{ij} \end{pmatrix} \sim N_2 \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0,02078325 & -0,00073119 \\ -0,00073119 & 0,00002924 \end{pmatrix} \right].$$

Para o efeito de gestante, foram gerados 40 valores de acordo com a distribuição

$$m_j \sim N(0, 0.00961932).$$

Para o efeito residual foram gerados 552 valores, ou seja, um valor em cada uma das idades gestacionais avaliadas para cada um dos fetos de acordo com a distribuição

$$\delta_{ij}(t) \sim N(0, 0.00453083).$$

O *software* utilizado para a geração destas variáveis aleatórias foi o **S-PLUS** versão 4.5. Para cada combinação dos valores \hat{a}_{ij} , \hat{b}_{ij} , \hat{m}_j e $\hat{\delta}_{ij}(t)$ gerados, utilizamos o modelo (3.1) para obter o logaritmo natural do peso fetal estimado para cada feto em cada idade gestacional efetivamente observada no conjunto de dados original. Um dos 200 conjuntos de dados assim obtidos está apresentado na Tabela A.4 (Apêndice A).

O modelo descrito em (4.1) foi ajustado a cada um dos 200 conjuntos de dados, obtendo-se assim estimativas dos parâmetros fixos e das variâncias dos efeitos aleatórios. Com a média dessas estimativas construíram-se os percentis populacionais simulados, e então calcularam-se os limites de confiança (95%) desses percentis simulados. Estes resultados estão apresentados nas Tabelas A.5 a A.7 (Apêndice A) e Figuras 4.4 a 4.6.

Figura 4.4: Percentis populacionais 3% e 97% empíricos e intervalos de confiança (95%).

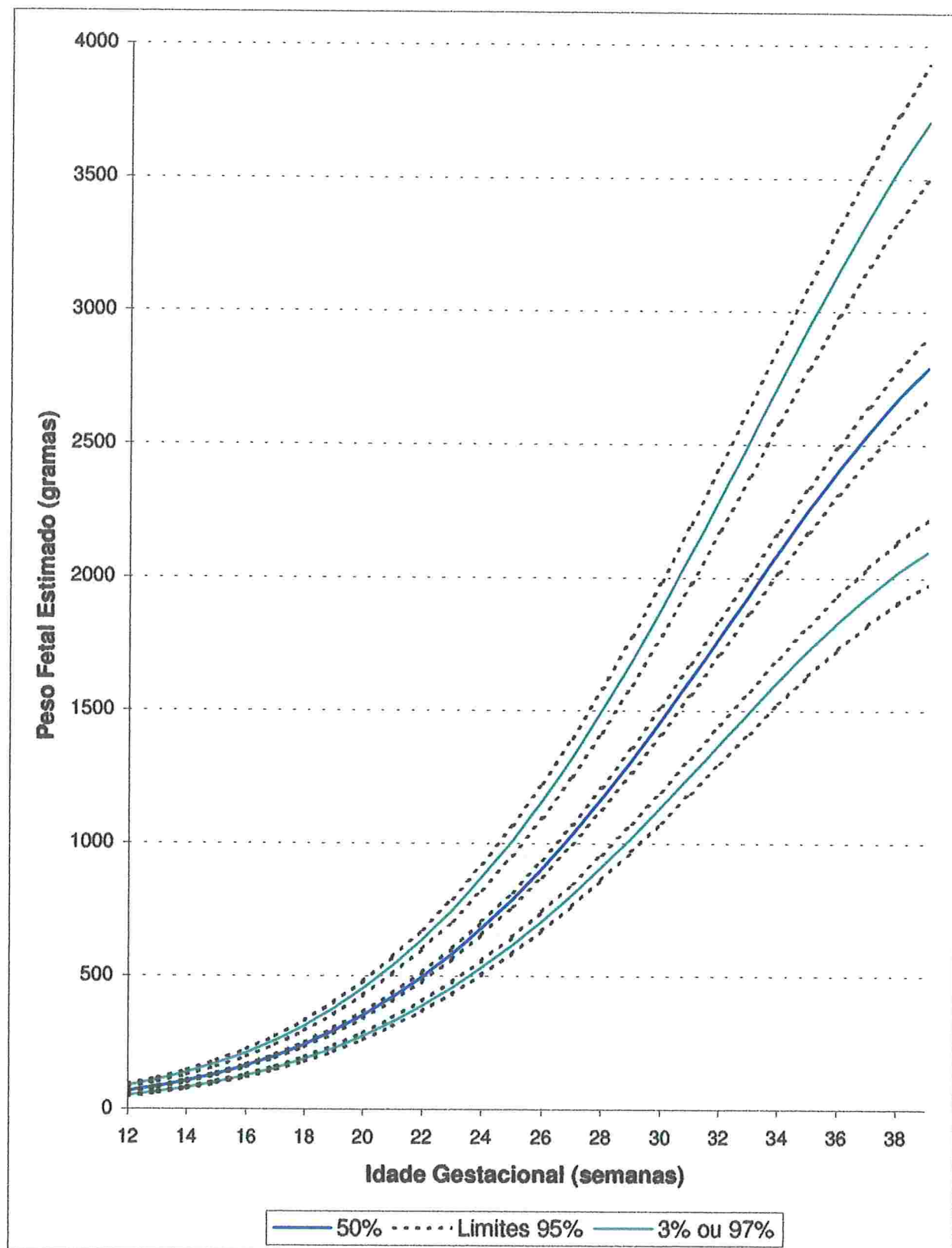


Figura 4.5: Percentis populacionais 10% e 90% empíricos e intervalos de confiança (95%).

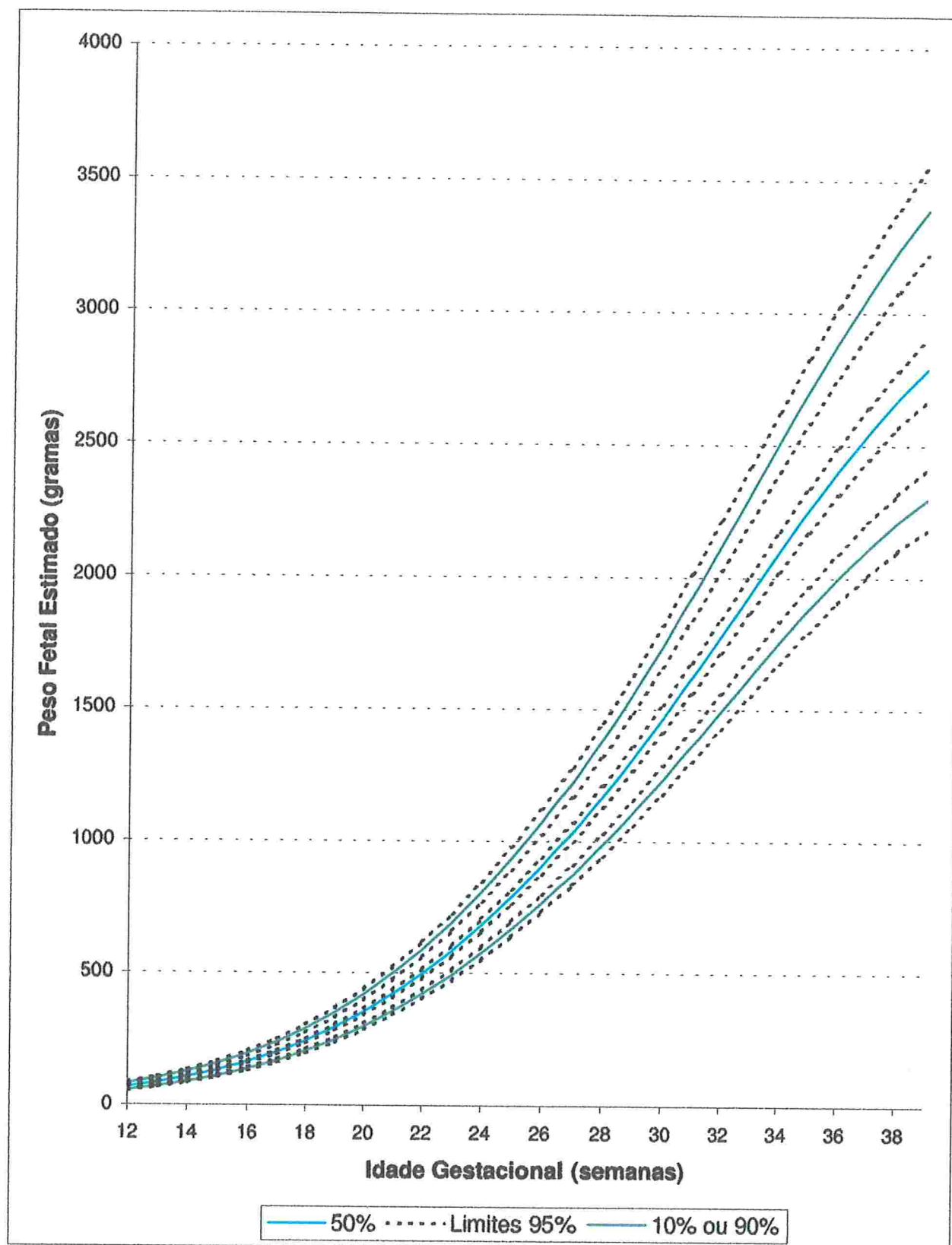
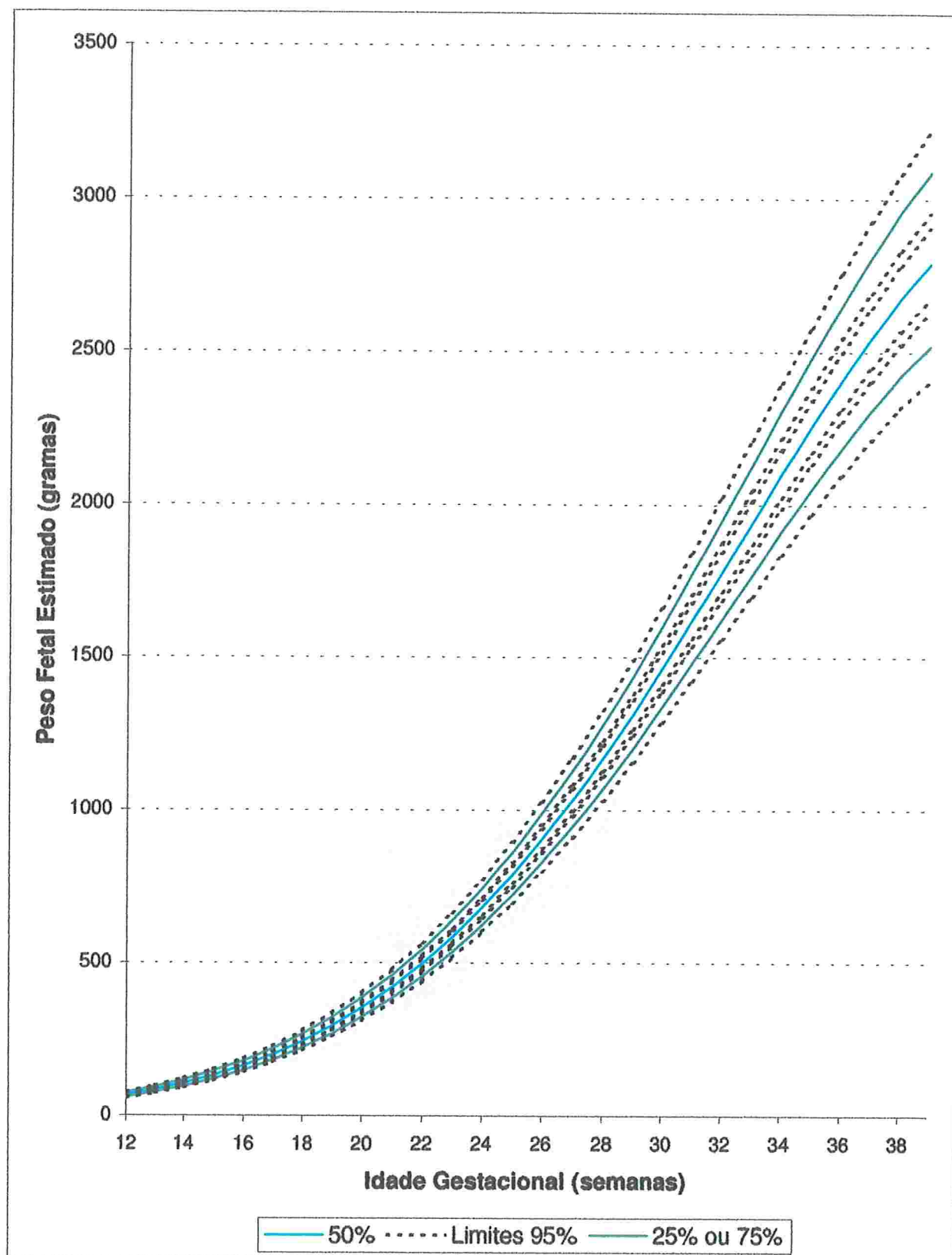


Figura 4.6: Percentis populacionais 25% e 75% empíricos e intervalos de confiança (95%).



Analisando as Tabelas A.5 a A.7 e os Gráficos 4.1 a 4.6 podemos notar que os intervalos de confiança de 95% apresentam um padrão muito parecido. Para quantificar a magnitude da discrepância entre os dois métodos calculamos as diferenças percentuais

$$\frac{|\text{valor simulado} - \text{valor calculado analiticamente}|}{\text{valor simulado}} \cdot 100\%$$

para cada um dos percentis populacionais em cada uma das idades gestacionais avaliadas. A maior diferença percentual obtida foi de 1.4%; além disso, em 98% dos casos a diferença percentual foi menor do que 1% e em 73% dos casos a diferença percentual foi menor do que 0.5%.

Com base nos resultados descritos acima podemos concluir que o método analítico, apesar de usar os valores obtidos pelo método de MVR, e possivelmente viciados, pode ser utilizado em substituição à simulação, uma vez que ele é mais prático e rápido.

A inclusão desta informação em gráficos de curvas de crescimento do peso fetal estimado permite uma avaliação da precisão das estimativas dos percentis populacionais.

Apêndice A

Tabelas

Tabela A.2: Estatísticas descritivas para peso fetal estimado (g)

Idade Gestacional (semanas)	Nº fetos	Peso Fetal Estimado (gramas)					
		Média	Desvio Padrão	Coef. de Variação	1º Quartil	Mediana	3º Quartil
12	2	68	8	12%	-	-	-
13	6	76	3	3%	74	75	79
14	2	92	6	6%	-	-	-
15	2	153	15	10%	-	-	-
16	17	167	28	17%	146	162	189
17	2	205	10	5%	-	-	-
18	12	253	20	8%	242	256	261
19	16	318	49	16%	290	326	341
20	16	350	50	14%	316	356	370
21	8	394	41	10%	361	399	424
22	16	475	52	11%	435	479	509
23	26	584	90	15%	508	574	648
24	14	709	98	14%	642	697	769
25	36	785	112	14%	700	786	832
26	9	998	106	11%	914	994	1074
27	30	993	117	12%	940	978	1054
28	22	1201	132	11%	1092	1182	1262
29	38	1314	157	12%	1220	1315	1407
30	19	1522	189	12%	1398	1506	1642
31	43	1605	218	14%	1436	1570	1759
32	29	1746	244	14%	1613	1739	1878
33	40	1956	236	12%	1806	1964	2105
34	31	2134	257	12%	1966	2091	2352
35	37	2294	241	11%	2192	2319	2434
36	38	2440	266	11%	2200	2423	2654
37	18	2540	372	15%	2242	2621	2794
38	21	2619	322	12%	2394	2525	2901
39	2	2638	444	17%	-	-	-

Tabela A.3: Teste de normalidade dos dados

Id. Gestacional (semanas)	Nº fetos	Coef. de Correlação para Gráfico Q-Q	Pontos Críticos nível de sig. de 10%
12	2	-	-
13	6	0.9460	0.9032
14	2	-	-
15	2	-	-
16	17	0.9900	0.9503
17	2	-	-
18	12	0.9760	0.9351
19	16	0.9800	0.9503
20	16	0.9700	0.9503
21	8	0.9700	0.9351
22	16	0.9870	0.9503
23	26	0.9870	0.9665
24	14	0.9850	0.9503
25	36	0.9880	0.9740
26	9	0.9630	0.9351
27	30	0.9830	0.9715
28	22	0.9700	0.9604
29	38	0.9920	0.9771
30	19	0.9870	0.9604
31	43	0.9930	0.9792
32	29	0.9870	0.9715
33	40	0.9930	0.9771
34	31	0.9780	0.9715
35	37	0.9910	0.9740
36	38	0.9890	0.9771
37	18	0.9770	0.9604
38	21	0.9760	0.9604
39	2	-	-

Tabela A.5: Intervalos de confiança (95%) para os percentis 3% e 97%.

Id. Gestacional	Limite	Limites de 95% para os percentis 3% e 97%			
		(3%) Método		(97%) Método	
		Analítico	Simulação	Analítico	Simulação
12	inferior	47	47	81	81
	superior	54	53	93	94
13	inferior	59	60	102	102
	superior	68	67	117	118
14	inferior	75	76	128	128
	superior	86	85	146	147
15	inferior	94	95	160	160
	superior	107	106	181	182
16	inferior	117	118	197	197
	superior	132	131	222	224
17	inferior	145	146	242	242
	superior	163	161	272	274
18	inferior	178	179	295	295
	superior	199	197	329	332
19	inferior	216	217	356	356
	superior	241	239	397	400
20	inferior	260	261	428	427
	superior	289	288	475	479
21	inferior	311	312	509	509
	superior	345	344	565	570
22	inferior	369	369	602	602
	superior	409	407	667	673
23	inferior	434	434	706	707
	superior	481	479	783	789
24	inferior	507	506	824	824
	superior	561	559	912	919
25	inferior	587	586	953	954
	superior	649	647	1055	1063
26	inferior	674	672	1095	1097
	superior	746	743	1212	1221
27	inferior	768	765	1249	1252
	superior	850	847	1383	1392
28	inferior	868	865	1415	1419
	superior	961	958	1567	1577
29	inferior	973	970	1591	1597
	superior	1078	1075	1764	1774
30	inferior	1082	1078	1777	1784
	superior	1201	1196	1972	1981
31	inferior	1194	1190	1970	1980
	superior	1327	1321	2188	2197
32	inferior	1307	1303	2168	2180
	superior	1455	1447	2412	2420

Tabela A.5:(Cont.) Intervalos de confiança (95%) para os percentis 3% e 97%.

Id. Gestacional	Limite	Limites de 95% para os percentis 3% e 97%			
		(3%) Método		(97%) Método	
		Analítico	Simulação	Analítico	Simulação
33	inferior	1420	1416	2369	2385
	superior	1583	1574	2641	2647
34	inferior	1530	1526	2570	2589
	superior	1709	1698	2871	2875
35	inferior	1635	1631	2768	2790
	superior	1832	1818	3100	3102
36	inferior	1735	1730	2960	2984
	superior	1948	1932	3324	3323
37	inferior	1825	1820	3141	3167
	superior	2056	2038	3540	3536
38	inferior	1904	1900	3309	3338
	superior	2154	2133	3744	3737
39	inferior	1972	1968	3461	3490
	superior	2241	2217	3933	3924

Tabela A.6: Intervalos de confiança (95%) para os percentis 10% e 90%.

Id. Gestacional	Limite	Limites de 95% para os percentis 10% e 90%			
		(10%) Analítico	Método Simulação	(90%) Analítico	Método Simulação
12	inferior	51	52	75	75
	superior	58	58	85	85
13	inferior	65	66	95	95
	superior	73	73	106	107
14	inferior	83	83	119	119
	superior	92	92	133	133
15	inferior	104	104	148	148
	superior	115	114	165	166
16	inferior	129	130	183	183
	superior	142	141	203	204
17	inferior	159	160	225	225
	superior	175	174	248	249
18	inferior	194	195	274	275
	superior	213	212	301	303
19	inferior	236	237	332	332
	superior	258	257	363	366
20	inferior	284	285	399	399
	superior	311	309	435	438
21	inferior	340	340	475	475
	superior	370	369	518	521
22	inferior	402	403	562	562
	superior	439	437	612	616
23	inferior	473	473	660	660
	superior	515	514	718	723
24	inferior	552	552	769	770
	superior	601	600	837	842
25	inferior	639	639	890	891
	superior	696	694	968	974
26	inferior	734	734	1022	1023
	superior	799	797	1112	1119
27	inferior	837	836	1166	1168
	superior	911	909	1269	1275
28	inferior	946	945	1320	1323
	superior	1030	1028	1438	1444
29	inferior	1061	1060	1484	1488
	superior	1156	1154	1617	1624
30	inferior	1181	1179	1656	1662
	superior	1288	1285	1806	1812
31	inferior	1305	1303	1835	1842
	superior	1424	1421	2003	2009
32	inferior	1430	1427	2019	2027
	superior	1562	1558	2205	2210

Tabela A.6:(Cont.) Intervalos de confiança (95%) para os percentis 10% e 90%.

Id. Gestacional	Limite	Limites de 95% para os percentis 10% e 90%			
		(10%) Método		(90%) Método	
		Análítico	Simulação	Análítico	Simulação
33	inferior	1555	1552	2204	2214
	superior	1701	1696	2411	2415
34	inferior	1678	1675	2390	2401
	superior	1838	1832	2618	2621
35	inferior	1797	1792	2571	2585
	superior	1972	1963	2822	2824
36	inferior	1908	1904	2746	2760
	superior	2100	2089	3022	3021
37	inferior	2011	2007	2911	2927
	superior	2219	2207	3213	3211
38	inferior	2102	2098	3063	3080
	superior	2328	2313	3392	3388
39	inferior	2180	2176	3199	3215
	superior	2424	2408	3557	3551

Tabela A.7: Intervalos de confiança (95%) para os percentis 25% e 75%.

Id. Gestacional	Limite	Limites de 95% para os percentis 25% e 75%			
		(25%) Método		(75%) Método	
		Analítico	Simulação	Analítico	Simulação
12	inferior	57	57	69	69
	superior	63	63	77	77
13	inferior	72	72	87	87
	superior	80	79	97	97
14	inferior	91	91	110	110
	superior	100	99	121	121
15	inferior	114	114	137	137
	superior	124	124	150	151
16	inferior	141	142	170	170
	superior	154	153	185	186
17	inferior	174	175	209	209
	superior	189	188	227	227
18	inferior	212	213	255	255
	superior	230	229	276	277
19	inferior	258	259	308	308
	superior	278	278	333	334
20	inferior	310	311	370	371
	superior	334	334	399	401
21	inferior	370	371	441	442
	superior	399	398	475	477
22	inferior	438	440	522	523
	superior	472	471	562	564
23	inferior	515	517	613	614
	superior	554	554	660	662
24	inferior	601	602	715	716
	superior	646	646	769	772
25	inferior	695	697	828	829
	superior	748	748	890	893
26	inferior	799	800	951	952
	superior	859	859	1022	1026
27	inferior	911	911	1084	1086
	superior	979	979	1166	1169
28	inferior	1030	1031	1228	1229
	superior	1108	1108	1320	1324
29	inferior	1156	1157	1380	1382
	superior	1244	1244	1484	1487
30	inferior	1288	1288	1539	1542
	superior	1386	1386	1656	1659
31	inferior	1424	1424	1704	1708
	superior	1533	1533	1835	1838
32	inferior	1563	1562	1873	1877
	superior	1684	1683	2018	2021

Tabela A.7:(Cont.) Intervalos de confiança (95%) para os percentis 25% e 75%.

Id. Gestacional	Limite	Limites de 95% para os percentis 25% e 75%			
		(25%) Analítico	Método Simulação	(75%) Analítico	Método Simulação
33	inferior	1701	1701	2044	2049
	superior	1835	1833	2204	2206
34	inferior	1838	1837	2214	2219
	superior	1984	1983	2390	2391
35	inferior	1971	1969	2380	2385
	superior	2131	2128	2573	2573
36	inferior	2096	2095	2539	2544
	superior	2271	2267	2751	2750
37	inferior	2212	2210	2688	2694
	superior	2403	2398	2920	2917
38	inferior	2317	2314	2824	2830
	superior	2525	2518	3078	3074
39	inferior	2407	2404	2945	2950
	superior	2633	2625	3222	3216

Apêndice B

Gráficos

Figura B.1: Perfis individuais do peso fetal estimado de 80 fetos gêmeos

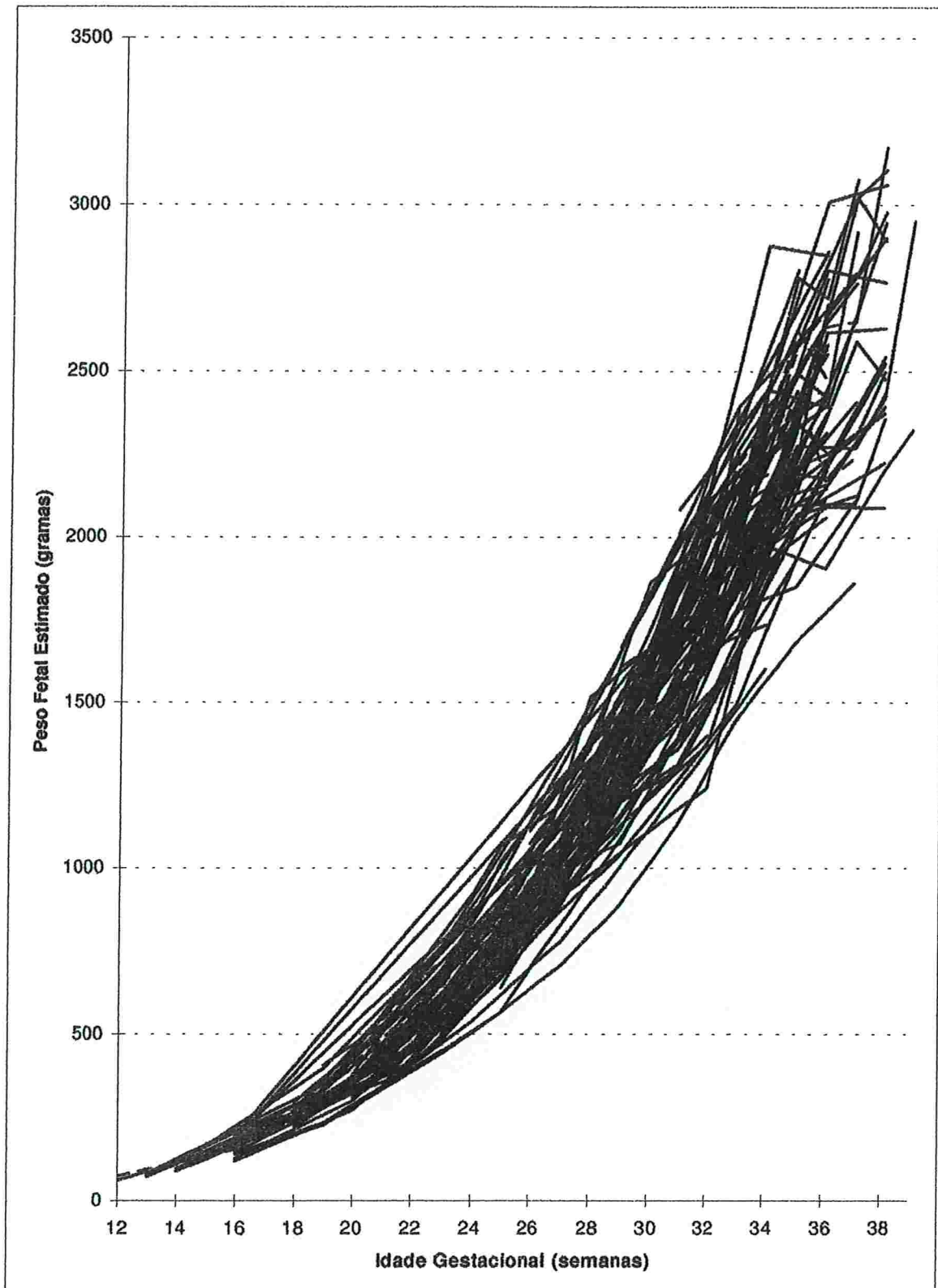
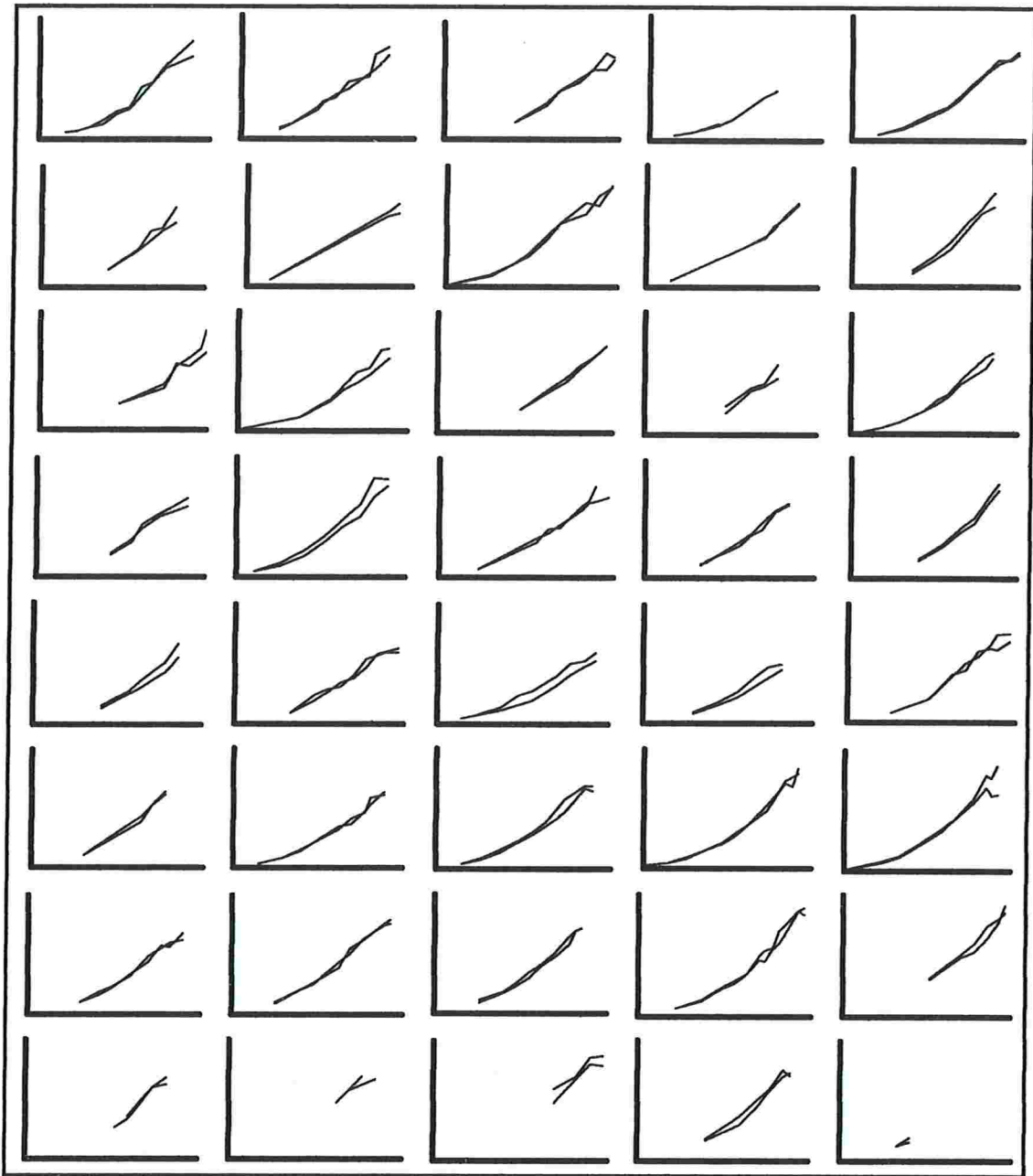


Figura B.2: Peso fetal estimado para 40 pares de gêmeos



Eixo vertical: peso fetal estimado (0 a 3500g)

Eixo horizontal: idade gestacional (12 a 39 semanas)

Apêndice C

Expressão matricial do modelo de efeitos aleatórios

O modelo definido em (4.1) pode ser escrito na forma matricial como

$$\mathbf{y}_{ij} = \mathbf{X}_{ij} \boldsymbol{\alpha} + \mathbf{Z}_{ij} \mathbf{b}_{ij} + \mathbf{e}_{ij}, \quad i = 1, 2, j = 1, \dots, 40 \quad (\text{C.1})$$

em que \mathbf{y}_{ij} ($p_{ij} \times 1$) é o vetor dos logaritmos dos pesos fetais estimados para o i -ésimo feto da j -ésima mãe nas p_{ij} idades gestacionais avaliadas, \mathbf{X}_{ij} ($p_{ij} \times p$) é a matriz (conhecida) de especificação do modelo populacional, $\boldsymbol{\alpha}$ ($p \times 1$) é o vetor de parâmetros populacionais desconhecidos, \mathbf{Z}_{ij} ($p_{ij} \times q$) é a matriz (conhecida) de especificação associada aos efeitos aleatórios, \mathbf{b}_{ij} ($q \times 1$) é o vetor de efeitos aleatórios associados ao i -ésimo feto da j -ésima mãe e \mathbf{e}_{ij} ($p_{ij} \times 1$) é um vetor de erros aleatórios. Além disso, supomos que \mathbf{e}_{ij} e \mathbf{b}_{ij} são independentes com $\mathbf{b}_{ij} \sim N(\mathbf{0}, \mathbf{G})$ e $\mathbf{e}_{ij} \sim N(\mathbf{0}, \mathbf{R}_{p_{ij}})$ para $i = 1, 2, j = 1, \dots, 40$, em que \mathbf{G} ($q \times q$) e $\mathbf{R}_{p_{ij}}$ ($p_{ij} \times p_{ij}$) são matrizes de covariâncias, e $\mathbf{R}_{p_{ij}}$ depende de i, j somente através de sua dimensão p_{ij} . Note que para o modelo adotado (4.1), temos $p = 3, q = 2$,

$$\mathbf{G} = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix} \quad \text{e} \quad \mathbf{R}_{p_{ij}} = \mathbf{1}_{p_{ij}} \mathbf{1}_{p_{ij}}^t \sigma_m^2 + \mathbf{I}_{p_{ij}} \sigma^2$$

em que $\mathbf{1}_{p_{ij}}$ ($p_{ij} \times 1$) é um vetor de dimensão p_{ij} com todos elementos iguais a 1 e $\mathbf{I}_{p_{ij}}$ é a matriz identidade de ordem p_{ij} . Por exemplo, para o feto 1 da mãe 1, com observações realizadas nas idades gestacionais 16, 18, 20, 22, 24, 26, 28, 30, 32 e 36 semanas temos

$$\mathbf{X}_{11} = \begin{pmatrix} 1 & 16 & 16^2 \\ 1 & 18 & 18^2 \\ 1 & 20 & 20^2 \\ 1 & 22 & 22^2 \\ 1 & 24 & 24^2 \\ 1 & 26 & 26^2 \\ 1 & 28 & 28^2 \\ 1 & 30 & 30^2 \\ 1 & 32 & 32^2 \\ 1 & 36 & 36^2 \end{pmatrix} \quad \text{e} \quad \mathbf{Z}_{11} = \begin{pmatrix} 1 & 16 \\ 1 & 18 \\ 1 & 20 \\ 1 & 22 \\ 1 & 24 \\ 1 & 26 \\ 1 & 28 \\ 1 & 30 \\ 1 & 32 \\ 1 & 36 \end{pmatrix}.$$

Conseqüentemente,

$$\mathbf{y}_{ij} \sim N(\mathbf{X}_{ij} \boldsymbol{\alpha}, \boldsymbol{\Sigma}_{ij})$$

com $\boldsymbol{\Sigma}_{ij} = \mathbf{R}_{p_{ij}} + \mathbf{Z}_{ij} \mathbf{G} \mathbf{Z}_{ij}^t$ e $\text{Cov}(\mathbf{y}_{ij}, \mathbf{y}_{i'j}) = \sigma_m^2 \mathbf{1}_{p_{ij}} \mathbf{1}_{p_{i'j}}^t$ se $i \neq i'$ e $\text{Cov}(\mathbf{y}_{ij}, \mathbf{y}_{i'j'}) = \mathbf{0}$ se $i \neq i'$. A matriz $\boldsymbol{\Sigma}_{ij}$, $i = 1, 2, j = 1, \dots, 40$ é uma função conhecida de um vetor

$\boldsymbol{\theta}$ ($k \times 1$) de parâmetros desconhecidos que não dependem de $\boldsymbol{\alpha}$, isto é, $\boldsymbol{\Sigma}_{ij} = \boldsymbol{\Sigma}_{ij}(\boldsymbol{\theta})$ para $i = 1, 2, j = 1, \dots, 40$. Notemos que no caso investigado aqui $k = 5$, e assim temos $\boldsymbol{\theta} = (\theta_1, \dots, \theta_5)^t$ com $\theta_1 = \sigma_a^2$, $\theta_2 = \sigma_{ab}$, $\theta_3 = \sigma_b^2$, $\theta_4 = \sigma_m^2$ e $\theta_5 = \sigma^2$.

Escrevendo $\mathbf{G} = \sum_{d=1}^3 \theta_d \mathbf{A}_d$, com $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $A_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ e $A_3 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ obtemos

$$\begin{aligned} \boldsymbol{\Sigma}_{ij}(\hat{\boldsymbol{\theta}}) &= \mathbf{Z}_{ij} \mathbf{G} \mathbf{Z}_{ij}^t + \mathbf{1}_{p_{ij}} \mathbf{1}_{p_{ij}}^t \sigma_m^2 + \mathbf{I}_{p_{ij}} \sigma^2 \\ &= \mathbf{Z}_{ij} \left(\sum_{d=1}^3 \theta_d \mathbf{A}_d \right) \mathbf{Z}_{ij}^t + \mathbf{1}_{p_{ij}} \mathbf{1}_{p_{ij}}^t \sigma_m^2 + \mathbf{I}_{p_{ij}} \sigma^2 \\ &= \sum_{d=1}^3 \theta_d \mathbf{Z}_{ij} \mathbf{A}_d \mathbf{Z}_{ij}^t + \mathbf{1}_{p_{ij}} \mathbf{1}_{p_{ij}}^t \sigma_m^2 + \mathbf{I}_{p_{ij}} \sigma^2 \\ &= \sum_{d=1}^5 \theta_d \mathbf{D}_{ijd} \end{aligned}$$

em que $\mathbf{D}_{ijd} = \mathbf{Z}_{ij} \mathbf{A}_d \mathbf{Z}_{ij}^t$ para $d = 1, 2$ e 3 , $\mathbf{D}_{ijd} = \mathbf{1}_{p_{ij}} \mathbf{1}_{p_{ij}}^t$ para $d = 4$ e $\mathbf{D}_{ijd} = \mathbf{I}_{p_{ij}}$ para $d = 5$.

De (C.1) temos que, para $i = 1, 2, j = 1, \dots, 40$,

$$\mathbb{E}(\mathbf{y}_{ij}) = \mathbf{X}_{ij} \boldsymbol{\alpha} \quad (\text{C.2})$$

$$\text{Var}(\mathbf{y}_{ij}) = \boldsymbol{\Sigma}_{ij} = \mathbf{Z}_{ij} \mathbf{G} \mathbf{Z}_{ij}^t + \mathbf{1}_{p_{ij}} \mathbf{1}_{p_{ij}}^t \sigma_m^2 + \mathbf{I}_{p_{ij}} \sigma^2. \quad (\text{C.3})$$

Neste caso, os parâmetros desconhecidos do modelo (C.1) são $\boldsymbol{\alpha}$ e $\boldsymbol{\theta}$, com $\boldsymbol{\alpha}$ representando o vetor (3×1) constituído pelos parâmetros de regressão (α , β_1 e β_2) e $\boldsymbol{\theta}$ representando o vetor (5×1) constituído pelos componentes de variância (σ_a^2 , σ_{ab} , σ_b^2 , σ_m^2 e σ^2). Para obtermos estimadores de $\mathbb{E}(\mathbf{y}_{ij})$ e $\text{Var}(\mathbf{y}_{ij})$ precisamos estimar os parâmetros $\boldsymbol{\alpha}$ e $\boldsymbol{\theta}$ do modelo (C.1). Esses estimadores podem ser obtidos através do método de MV ou de MVR. Para obtermos estimativas de MVR de $\boldsymbol{\theta}$ podemos maximizar a função de verossimilhança restrita ou, equivalentemente, seu logaritmo natural,

$$\begin{aligned} l_R(\mathbf{y}_{ij}; \boldsymbol{\theta}) &= K - \frac{1}{2} \sum_{i=1}^2 \sum_{j=1}^{40} \ln |\boldsymbol{\Sigma}_{ij}| - \frac{1}{2} \ln \left| \sum_{i=1}^2 \sum_{j=1}^{40} \mathbf{X}_{ij}^t \boldsymbol{\Sigma}_{ij}^{-1} \mathbf{X}_{ij} \right| \\ &\quad - \frac{1}{2} \sum_{i=1}^2 \sum_{j=1}^{40} (\mathbf{y}_{ij} - \mathbf{X}_{ij} \tilde{\boldsymbol{\alpha}})^t \boldsymbol{\Sigma}_{ij}^{-1} (\mathbf{y}_{ij} - \mathbf{X}_{ij} \tilde{\boldsymbol{\alpha}}) \end{aligned}$$

em que $\tilde{\alpha} = (\sum_{i=1}^2 \sum_{j=1}^{40} \mathbf{X}_{ij}^t \Sigma_{ij}^{-1} \mathbf{X}_{ij})^{-1} (\sum_{i=1}^2 \sum_{j=1}^{40} \mathbf{X}_{ij}^t \Sigma_{ij}^{-1} \mathbf{y}_{ij})$ e K é uma constante que não depende de θ .

Nesse sentido, necessitamos calcular as primeiras e segundas derivadas parciais de $l_R(\mathbf{y}_{ij}; \theta)$ em relação a $\theta = (\theta_1, \dots, \theta_5)^t$, que são dadas por

$$\mathbf{s}_R = \left[\left\langle \frac{\partial l_R}{\partial \theta_d} \right\rangle_d \right]$$

e

$$\mathbf{H}_R = \left[\left\langle \frac{\partial^2 l_R}{\partial \theta_d \partial \theta_h} \right\rangle_{dh} \right]$$

em que expressões do tipo $[\langle \rangle_a]$ denotam vetores cujo a -ésimo elemento é dado pela expressão entre $\langle \rangle$ e expressões do tipo $[\langle \rangle_{ab}]$ denotam matrizes cujo elemento de ordem (a, b) é dado pela expressão entre $\langle \rangle$.

O estimador de MVR de θ é a solução de

$$\mathbf{s}_R = [\partial l_R / \partial \theta] = \mathbf{0}.$$

A matriz de informação de Fisher associada à função de verossimilhança restrita é dada por

$$E(\theta) = -E[\mathbf{H}_R(\theta)].$$

Com base nisso, um estimador da matriz de covariância assintótica do estimador de $\hat{\theta}_R$ pode ser obtida através da inversa da matriz de informação de Fisher; mais especificamente:

$$\widehat{\text{Var}}_A(\hat{\theta}_R) = 2 \left[\left\langle \sum_{i=1}^2 \sum_{j=1}^{40} \text{tr} \mathbf{P}_{ij} \mathbf{D}_{ijd} \mathbf{P}_{ij} \mathbf{D}_{ijh} \right\rangle_{dh} \right]^{-1} \quad (\text{C.4})$$

em que $\mathbf{P}_{ij} = \Sigma_{ij}^{-1} - \Sigma_{ij}^{-1} \mathbf{X}_{ij} (\sum_{i=1}^2 \sum_{j=1}^{40} \mathbf{X}_{ij}^t \Sigma_{ij}^{-1} \mathbf{X}_{ij})^{-1} \mathbf{X}_{ij}^t \Sigma_{ij}^{-1}$.

É importante notar que o logaritmo da função de verossimilhança restrita, ou equivalentemente a função de verossimilhança restrita não envolve α ; então, uma vez encontrado o estimador de MV de θ ($\hat{\theta}_R$), α pode ser estimado através do método de mínimos quadrados generalizados, isto é

$$\hat{\alpha} = \left(\sum_{i=1}^2 \sum_{j=1}^{40} \mathbf{X}_{ij}^t \Sigma_{ij}^{-1}(\hat{\theta}_R) \mathbf{X}_{ij} \right)^{-1} \left(\sum_{i=1}^2 \sum_{j=1}^{40} \mathbf{X}_{ij}^t \Sigma_{ij}^{-1}(\hat{\theta}_R) \mathbf{y}_{ij} \right).$$

A matriz de covariância assintótica de $\hat{\alpha}$ pode ser estimada por

$$\widehat{\text{Var}}_A(\hat{\alpha}) = \left(\sum_{i=1}^2 \sum_{j=1}^{40} \mathbf{X}_{ij}^t \Sigma_{ij}^{-1}(\hat{\theta}_R) \mathbf{X}_{ij} \right)^{-1}$$

Detalhes sobre todos estes cálculos estão demonstrados em Andreoni (1989), por exemplo.

O modelo (C.1) também pode ser escrito na forma ainda mais compacta

$$\mathbf{y} = \mathbf{X} \boldsymbol{\alpha} + \mathbf{Z} \mathbf{b} + \mathbf{e} \quad (\text{C.5})$$

com $\mathbf{y} = (\mathbf{y}_{11}^t, \dots, \mathbf{y}_{2,40}^t)^t (\sum_{i=1}^2 \sum_{j=1}^{40} p_{ij} \times 1)$, $\mathbf{X} = (\mathbf{X}_{11}^t, \dots, \mathbf{X}_{2,40}^t)^t (\sum_{i=1}^2 \sum_{j=1}^{40} p_{ij} \times p)$, $\mathbf{Z} = (\mathbf{Z}_{11}^t, \dots, \mathbf{Z}_{2,40}^t)^t (\sum_{i=1}^2 \sum_{j=1}^{40} p_{ij} \times q)$ e $\mathbf{e} = (\mathbf{e}_{11}^t, \dots, \mathbf{e}_{2,40}^t)^t (\sum_{i=1}^2 \sum_{j=1}^{40} p_{ij} \times 1)$.

Conseqüentemente,

$$\begin{aligned} E(\mathbf{y}) &= \mathbf{X} \boldsymbol{\alpha} \\ \text{Var}(\mathbf{y}) &= \mathbf{V}(\boldsymbol{\theta}) = \text{diag}(\Sigma_{11}(\boldsymbol{\theta}), \dots, \Sigma_{2,40}(\boldsymbol{\theta})) \end{aligned}$$

onde $\text{diag}(\Sigma_{11}(\boldsymbol{\theta}), \dots, \Sigma_{2,40}(\boldsymbol{\theta}))$ representa uma matriz diagonal em blocos. Ou seja, $\mathbf{y} \sim N(\mathbf{X} \boldsymbol{\alpha}, \mathbf{V})$.

O estimador de MVR é essencialmente um estimador de MV baseado em um conjunto de dados linearmente transformados $\mathbf{y}^* = \mathbf{A} \mathbf{y}$ tal que a distribuição de \mathbf{y}^* não depende de $\boldsymbol{\alpha}$. Uma maneira de obter esta transformação é tomando \mathbf{A} como a matriz projeção

$$\mathbf{A} = \mathbf{I} - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t. \quad (\text{C.6})$$

Então, \mathbf{y}^* tem distribuição normal com média zero, não importando qual seja o valor de $\boldsymbol{\alpha}$. Do ponto de vista operacional não há necessidade de fazermos a transformação de \mathbf{y} para \mathbf{y}^* , como veremos a seguir.

Seja \mathbf{A} a matriz definida em (C.6) e \mathbf{B} a matriz de dimensão $(\sum_{i=1}^2 \sum_{j=1}^{40} p_{ij} \times (\sum_{i=1}^2 \sum_{j=1}^{40} p_{ij} - p))$ definida por $\mathbf{B} \mathbf{B}^t = \mathbf{A}$ e $\mathbf{B}^t \mathbf{B} = \mathbf{I}$. Finalmente, seja $\mathbf{Z} = \mathbf{B}^t \mathbf{y}$. Agora, para $\boldsymbol{\theta}$ fixo, o estimador de MV de $\boldsymbol{\alpha}$ é o estimador de mínimos quadrados generalizados,

$$\hat{\alpha} = (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} = \mathbf{H} \mathbf{y}.$$

com $\mathbf{H} = (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1}$.

Além disso, $\mathbb{E}(\mathbf{Z}) = \mathbf{0}$ e \mathbf{Z} e $\hat{\boldsymbol{\alpha}}$ são independentes, não importando o valor de $\boldsymbol{\alpha}$, como mostraremos a seguir. Primeiramente, notemos que

$$\mathbb{E}(\mathbf{Z}) = \mathbf{B}^t \mathbb{E}(\mathbf{y}) = \mathbf{B}^t \mathbf{X} \boldsymbol{\alpha} = \mathbf{B}^t \mathbf{B} \mathbf{B}^t \mathbf{X} \boldsymbol{\alpha}$$

pois $\mathbf{B}^t \mathbf{B} = \mathbf{I}$. Mas como $\mathbf{B} \mathbf{B}^t = \mathbf{A}$, podemos escrever

$$\mathbb{E}(\mathbf{Z}) = \mathbf{B}^t \mathbf{A} \mathbf{X} \boldsymbol{\alpha}$$

e

$$\mathbf{A} \mathbf{X} = \{\mathbf{I} - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t\} \mathbf{X} = \mathbf{X} - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{X} = \mathbf{0},$$

conseqüentemente $\mathbb{E}(\mathbf{Z}) = \mathbf{0}$.

Em segundo lugar,

$$\begin{aligned} \text{Cov}(\mathbf{Z}, \hat{\boldsymbol{\alpha}}) &= \mathbb{E}\{\mathbf{Z}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})^t\} \\ &= \mathbb{E}\{\mathbf{B}^t \mathbf{y}(\mathbf{y}^t \mathbf{H}^t - \boldsymbol{\alpha}^t)\} \\ &= \mathbf{B}^t \mathbb{E}(\mathbf{y} \mathbf{y}^t) \mathbf{H}^t - \mathbf{B}^t \mathbb{E}(\mathbf{y}) \boldsymbol{\alpha}^t \\ &= \mathbf{B}^t \{\text{Var}(\mathbf{y}) + \mathbb{E}(\mathbf{y}) \mathbb{E}(\mathbf{y})^t\} \mathbf{H}^t - \mathbf{B}^t \mathbb{E}(\mathbf{y}) \boldsymbol{\alpha}^t. \end{aligned}$$

Substituindo $\mathbb{E}(\mathbf{y}) = \mathbf{X} \boldsymbol{\alpha}$ nessa última expressão, temos

$$\begin{aligned} \text{Cov}(\mathbf{Z}, \hat{\boldsymbol{\alpha}}) &= \mathbf{B}^t (\mathbf{V} + \mathbf{X} \boldsymbol{\alpha} \boldsymbol{\alpha}^t \mathbf{X}^t) \mathbf{H}^t - \mathbf{B}^t \mathbf{X} \boldsymbol{\alpha} \boldsymbol{\alpha}^t \\ &= \mathbf{B}^t \mathbf{V} \mathbf{H}^t + \mathbf{B}^t \mathbf{X} \boldsymbol{\alpha} \boldsymbol{\alpha}^t \mathbf{X}^t \mathbf{H}^t - \mathbf{B}^t \mathbf{X} \boldsymbol{\alpha} \boldsymbol{\alpha}^t \end{aligned} \quad (\text{C.7})$$

Além disso, notemos que

$$\mathbf{X}^t \mathbf{H}^t = \mathbf{X}^t \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} = \mathbf{I}$$

e

$$\begin{aligned} \mathbf{B}^t \mathbf{V} \mathbf{H}^t &= \mathbf{B}^t \mathbf{V} \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \\ &= \mathbf{B}^t \mathbf{X} (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} = \mathbf{0} \end{aligned}$$

porque $\mathbf{B}^t \mathbf{X} = \mathbf{B}^t \mathbf{A} \mathbf{X} = \mathbf{0}$ como na demonstração de $\mathbb{E}(\mathbf{Z}) = \mathbf{0}$. Logo de (C.7) segue que

$$\begin{aligned} \text{Cov}(\mathbf{Z}, \hat{\boldsymbol{\alpha}}) &= \mathbf{0} + \mathbf{B}^t \mathbf{X} \boldsymbol{\alpha} \boldsymbol{\alpha}^t - \mathbf{B}^t \mathbf{X} \boldsymbol{\alpha} \boldsymbol{\alpha}^t \\ &= \mathbf{0}. \end{aligned} \quad (\text{C.8})$$

Referências Bibliográficas

- [1] Andreoni, S. (1989). Modelo de Efeitos Aleatórios para Análise de Dados Longitudinais Não Balanceados em Relação ao Tempo. Dissertação de Mestrado. São Paulo: Instituto de Matemática e Estatística. Universidade de São Paulo, Brasil.
- [2] Deter, R.L., Harrist, R.B., Hadlock, F.P. and Poindexter, A.N. (1982). Longitudinal studies of fetal growth with the use of dynamic image ultrasonography. **American Journal of Obstetrics and Gynecology**, 143: 545-554.
- [3] Diggle, P.J., Liang, K-Y. and Zeger, S.L. (1994). **Analysis of Longitudinal Data**. Oxford: Oxford Science Publications.
- [4] Fellingham, G.W. and Raghunathan, T.E. (1995). Sensitivity of point and interval estimates to distributional assumptions in longitudinal data analysis of small samples. **Communications in Statistics - Simulation**, 24: 617-630.
- [5] Fujita, M. (1998). Avaliação ultra-sonográfica do crescimento fetal em gestações gêmeares. Dissertação de Doutorado. São Paulo: Faculdade de Medicina. Universidade de São Paulo, Brasil.
- [6] Gräsbeck, R. and Saris, N-E. (1969). Establishment and use of normal values. **Scandinavian Journal of Clinical Laboratory Investigation**, 24: 62-63.
- [7] Hadlock, F.P., Deter, R., Harrist, R.B. and Park, S.K. (1982). Fetal biparietal diameter: a critical re-evaluation of the relation to menstrual age by means of realtime ultrasound. **Journal of Ultrasound in Medicine**, 1: 97-104.
- [8] Hadlock, F.P., Harrist, R.B. and Martinez-Poyer, J. (1991). In utero analysis of fetal growth: a sonographic weight standard. **Radiology**, 181: 129-133.

- [9] Hadlock, F.P., Harrist, R.B., Sharman, R.S., Deter, R.L. and Park, S.K. (1985). Estimation of fetal weight with the use of head, body, and femur measurements - a prospective study. **American Journal of Obstetrics and Gynecology**, 151: 333-337.
- [10] Johnson, R.A., and Wichern, D.W. (1992). **Applied Multivariate Statistical Analysis**. New Jersey: Prentice-Hall.
- [11] Jolliff, C.R., Cost, K.M., Stivrins, P.C., Grossman, P.P., Nolte, C.R., Franco, S.M., Fijan, K.J., Fletcher, L.L. and Shriner, H.C. (1982). Reference Intervals for Serum IgG, IgA, IgM, C3, and C4 as Determined by Rate Nephelometry. **Clinical Chemistry**, 28: 126-128.
- [12] Jones, R.H. (1993). **Longitudinal Data with Serial Correlation: a State-space Approach**. London: Chapman and Hall.
- [13] Laird, N.M. and Ware, J.H. (1982). Random-effects models for longitudinal data. **Biometrics**, 38: 963-974.
- [14] Lubchenco, L.O., Hansman, C., Dressler, M. and Boyd, E. (1963). Intrauterine growth as estimated from liveborn birth-weight data at 24-42 weeks of gestacion. **Pediatrics**, 32: 793-800.
- [15] Naeye, R.L., Benirschke, K., Hagstrom, J.W.C. and Marcus, C.C. (1966). Intrauterine growth of twins as estimated from liveborn birth-weight data. **Pediatrics**, 37: 409-416.
- [16] Patterson, H.D. and Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. **Biometrika**, 58: 545-554.
- [17] Robinson, H.P. and Fleming, J.E.E. (1975). A critical evaluation of sonar "crown-rump length" measurements. **British Journal of Obstetrics and Gynaecology**, 82: 702-710.
- [18] Royston, P. (1991). Calculation of unconditional and conditional reference intervals for foetal size and growth from longitudinal measurements. **Statistics in Medicine**, 14: 1417-1436.

- [19] SAS/STAT Software: (1997). **Changes and Enhancements through Release 6.12**. Cary, NC: SAS Institute Inc.
- [20] Searle, S.R., Casella, G. and McCulloch, C.E. (1992). **Variance Components**. New York: John Wiley & Sons, Inc.
- [21] Sen, P.K. and Singer, J.M. (1993). **Large Sample Methods in Statistics: An Introduction with Applications**. New York: Chapman and Hall.
- [22] Singer, J.M. and Andrade, D.F. (2000). Analysis of Longitudinal Data. In **Handbook of Statistics, Volume 18: Bio-Environmental and Public Health Statistics**, P.K.Sen and C.R. Rao (eds). Amsterdam: North Holland. (p. 115 a 160).
- [23] Singer, J.M. e Shinzato, A. (1999). Curva de peso fetal estimado em gestações gemelares. Relatório de Análise Estatística. São Paulo: Centro de Estatística Aplicada. Departamento de Estatística. Instituto de Matemática e Estatística. Universidade de São Paulo, Brasil.
- [24] Wright, E. M. and Royston, P. (1997). A comparison of statistical methods for age-related reference intervals. **Journal of the Royal Statistical Society**, 160: 47-69.
- [25] Wright, E. M. and Royston, P. (1999). Calculating reference intervals for laboratory measurements. **Statistical Methods in Medical Research**, 8: 93-112.