

**Comparação de modelo clássico e Bayesiano para dados de óbitos perinatais no ISEA,
Campina Grande-PB**

**Comparison of classic and Bayesian model for data on perinatal deaths at ISEA,
Campina Grande-PB**

**Comparación del modelo clásico y bayesiano para datos sobre muertes perinatales en
ISEA, Campina Grande-PB**

Recebido: 10/06/2020 | Revisado: 23/06/2020 | Aceito: 01/07/2020 | Publicado: 18/07/2020

Mácio Augusto de Albuquerque

ORCID: <http://orcid.org/0000-0002-0113-9130>

Universidade Estadual da Paraíba, Brasil

E-mail: marcioaa@uepb.edu.br

Sandro Lins Lopes de Lucena

ORCID: <http://orcid.org/0000-0002-7327-9626>

Universidade Estadual da Paraíba, Brasil

E-mail: sandro.lins.lucena@hotmail.com

Kleber Napoleão Nunes de Oliveira Barros

ORCID: <http://orcid.org/0000-0003-2515-329>

Universidade Estadual da Paraíba, Brasil

E-mail: kleberbarros@cct.uepb.edu.br

Resumo

Modelos lineares generalizados são úteis, dentre outras situações, quando se quer ajustar modelos a dados que não seguem normalidade e não podem ser ajustados usando apenas a regressão linear simples. Outra ferramenta poderosa de estimação são os métodos Bayesianos, baseado em probabilidades condicionais. Neste trabalho apresenta-se um ajuste de modelos de regressão logístico com parâmetros estimado pelo método da máxima verossimilhança que é atualizado usando as técnicas da inferência Bayesiana. Tais métodos foram aplicados em dados obtidos no Instituto de Saúde Elpídio de Almeida que fica localizado no município de Campina Grande - PB. As informações referem-se a pacientes gestantes atendidas nesta unidade de saúde. Objetivou-se obter o melhor modelo possível que nos forneça informação sobre a chance de óbito de uma criança em função de variáveis maternas usando o método de estimação da máxima verossimilhança e o método Bayesiano. Os ajustes e diagnósticos dos

modelos foram realizados com auxílio do software R. Constatou-se que o modelo estimado pela máxima verossimilhança é muito próximo do modelo Bayesiano.

Palavras-chave: Regressão logística; Inferência Bayesiana; Óbito perinatal.

Abstract

Generalized linear models are useful, among other situations, when you want to fit models to data that do not follow normality and cannot be adjusted using only simple linear regression. Another powerful estimation tool is the Bayesian methods, based on conditional probabilities. This work presents an adjustment of logistic regression models with parameters estimated by the maximum likelihood method, which is updated using Bayesian inference techniques. Such methods were applied to data obtained at the Instituto de Saúde Elpídio de Almeida, which is located in the city of Campina Grande - PB. The information refers to pregnant patients seen at this health unit. The objective was to obtain the best possible model that provides us with information about the chance of death of a child due to maternal variables using the maximum likelihood estimation method and the Bayesian method. The adjustments and diagnostics of the models were performed with the aid of the R software. It was found that the model estimated by the maximum likelihood is very close to the Bayesian model.

Keywords: Logistic regression; Bayesian inference; Perinatal death.

Resumen

Los modelos lineales generalizados son útiles, entre otras situaciones, cuando se desea ajustar modelos a datos que no siguen la normalidad y no pueden ajustarse utilizando solo una regresión lineal simple. Otra poderosa herramienta de estimación son los métodos bayesianos, basados en probabilidades condicionales. Este trabajo presenta un ajuste de modelos de regresión logística con parámetros estimados por el método de máxima verosimilitud, que se actualiza utilizando técnicas de inferencia bayesianas. Dichos métodos se aplicaron a los datos obtenidos en el Instituto de Saúde Elpídio de Almeida, ubicado en la ciudad de Campina Grande - PB. La información se refiere a pacientes embarazadas atendidas en esta unidad de salud. El objetivo fue obtener el mejor modelo posible que nos brinde información sobre la posibilidad de muerte de un niño debido a variables maternas utilizando el método de estimación de máxima verosimilitud y el método bayesiano. Los ajustes y diagnósticos de los modelos se llevaron a cabo con la ayuda del software R. Se encontró que el modelo estimado por la máxima probabilidad está muy cerca del modelo bayesiano.

Palabras clave: Regresión logística; Inferencia bayesiana; Muerte perinatal..

1. Introdução

Nos últimos anos o Brasil vem enfrentando mudanças sociais, principalmente nas últimas três décadas. Por meio do desenvolvimento social, o país vem obtendo índices cada vez mais positivo, que quantificam a melhora de vida da população. Na saúde pública, por exemplo, há diversos estudos que constata a influência de fatores sociais na vida de menores recém-nascidos. Silva et al. (2006) relatam que os determinantes sociais repercutem na situação de saúde na população de menores de um ano de vida enquanto que Sarinho (1998) destaca sua atenção para a idade materna (abaixo de 20 anos e acima de 35 anos) está associada à maior probabilidade de morte neonatal. O maior risco de mortalidade perinatal em adolescentes tem sido atribuído a maior incidência de baixo peso ao nascer, parto prematuro e complicações na gestação (Sarinho, 1998).

A saúde infantil no Brasil vem apresentando avanços importantes principalmente a partir da década de 1990 e 2000, demonstrada por indicadores que mostram significativas melhorias nos determinantes sociais das doenças em conjunto a organização dos serviços de saúde (Victoria et al., 2011). Esses avanços são promovidos, em parte, por estudos científicos que podem detectar resultados e oferecer possíveis soluções para problemas que a saúde pública venha a enfrentar.

Dentre os avanços científicos destacamos o desenvolvimento dos Modelos Lineares Generalizados (MLG's). Estes modelos foram propostos por (Nelder & Wedderburn, 1972), e apresentam uma componente sistemática, uma componente aleatória e uma função de ligação. Trabalha-se com o modelo logístico, que é um caso especial de modelo linear generalizado quando se tem a função de probabilidade binomial para a variável resposta e a função de ligação logit. O modelo logístico é utilizado quando se tem variável resposta do tipo binária, por isso o uso da função binomial.

Dois métodos foram utilizados para estimar os parâmetros do modelo, o método da máxima verossimilhança, ver (Mardia & Marshall, 1984), e o método Bayesiano. Para método da máxima verossimilhança a seleção das variáveis é feita pelo método stepwise que se baseia na observação de que uma variável pode ser movida de um conjunto de variáveis para outro conjunto de variáveis que dá maior redução no resíduo. Em outras palavras, o método consiste em selecionar modelos onde as variáveis são adicionadas e/ou excluídas uma de cada vez para que o modelo se aproxime ao máximo dos dados.

A adequação do modelo pode ser verificada por meio de vários critérios. O Critério de Informação de Akaike e o desvio residual pode ser utilizado para verificar o quão bom é o

modelo ajustado pelo método da máxima verossimilhança e por meio da análise residual verifica-se a adequabilidade do modelo estimado. O estudo residual pode ser feito diretamente observando o envelope simulado dos resíduos conforme descreve (Moral, Hinde & Demétrio, 2017).

Com apoio teórico dos modelos lineares generalizados e com apoio dos programas computacionais há grande facilidade para a implementação das técnicas bayesianas para modelos complexos. Diante desse contexto, objetivamos obter o melhor modelo possível que nos forneça informação sobre a chance de óbito de uma criança em função de variáveis maternas usando o método de estimação da máxima verossimilhança e o método Bayesiano.

2 Materiais e Métodos

O referido estudo trata-se estudos descritivos que enfatizando também o método quantitativo e qualitativo. Os dados foram coletados no Instituto de Saúde Elpídio de Almeida e referem-se a características de gestantes atendidas na maternidade no ano de 2018. Inicialmente tinha-se informações de 5604 pacientes no banco de dados, e para cada paciente coletou-se vinte variáveis, dentre características da mãe e da criança. Devido à ausência de informações no preenchimento dos formulários, dados ausentes, e após minuciosa mineração do banco de dados, a amostra ficou restrita e composta por informações de 2807 mulheres que foram atendidas nesta instituição no ano citado.

Dos prontuários, com as características das gestantes, trabalhamos com a idade de cada paciente, o grau de instrução, a situação conjugal, a raça declarada e o número de gestações e o mês do atendimento. Depois da mineração dos dados, fez-se uma análise descritiva buscando verificar as frequências de cada variável e entender o perfil das mães gestantes que passaram pela maternidade.

Trata-se na modelagem, com a variável resposta binária em que 0 indica que a mãe não sofreu com óbitos de seu filho em nenhuma de suas gestações e 1 indica que a mãe já perdeu seu filho no período perinatal, ou seja, a criança veio a óbito. Na sequência considera-se o modelo de regressão linear generalizado saturado, com as variáveis idade, o número de gestação, o grau de instrução, a situação conjugal e a raça declarada, com distribuição binomial e função de ligação logit, ajustando parâmetros pelo método da máxima verossimilhança, para verificar quais das variáveis podem ter relação com a chance de óbito dos produtos gestacionais.

Aplicando-se os métodos de seleção de modelos descritos na seção anterior, ajustasse

um modelo logístico com tendência linear e outro modelo logístico com tendência quadrática na variável idade. Em seguida verifica-se, por meio dos critérios de bondade do ajuste e da figura de envelope simulado, a adequação de cada um dos modelos ajustados.

Ao encontrar o melhor modelo aplica-se a inferência Bayesiana nos parâmetros do mesmo, por meio de simulação de Monte Carlos via Cadeias de Markov, tendo como distribuição a priori para os parâmetros distribuições normais com médias iguais a suas respectivas estimativas dada pela estimação máxima verossimilhança e credibilidade 0,01, obtendo-se as estimativas dos parâmetros atualizadas e a distribuição a posteriori com a aplicação direta do Teorema de Bayes.

3 Resultado e Discussão

3.1 Análise Descritiva

Dos 2807 prontuários que compõem este estudo, 30,96% (IC95% = [29,25%;32,67%]) das mães sofreram com o óbito de pelo menos um produto de sua(s) gestação(ões), enquanto que 69,04% (IC95% = [67,33%;70,75%]) não registraram óbitos conforme observado na Tabela 1. Avaliando por variável, no tocante à situação conjugal das mães, há prevalência das solteiras, com percentual de 55,51% (IC95% = [53,67%;57,35%]), seguido de mães casadas com 26,00% (IC95% = [24,38%;27,62%]) na (Tabela 1). Das 1558 mães solteiras, 29,97% (IC95% = [27,69%;32,24%]) tiveram pelo menos um insucesso em alguma de sua (s) gravidez. 50% dos produtos gerados pelas mulheres viúvas vieram a óbito (Tabela 1).

A frequência, segundo o grau de instrução, foi maior para as mulheres que possuem apenas o Ensino Fundamental II com 1167 mães. Em segundo lugar registrou-se mãe com Ensino Médio, com frequência de 590 mães, e em terceiro lugar 430 mães com Ensino Fundamental I (Tabela 1). Constatou-se ainda 1,61% (IC95% = [1,14%;2,07%]) das mães não possuem escolaridade. Dessas 44,44% (IC95% = [29,92%;58,96%]) perderam pelo menos um de seu (s) filho (s). No caso de mães com ensino superior, completo ou incompleto, apenas 22,77% (IC95% = [14,59%;30,95%]) tiveram registros de óbitos para pelo menos um de seus filhos. No quesito racial mães pardas são ampla maioria. Registrou-se que 93,30% (IC95% = [92,37%;94,22%]) das mulheres são pardas. Mães declaradas brancas foram apenas 5,38 (IC95% = [4,54%;6,21%]), pretas 1,28%(IC95% = [1,07%;1,49%]) e indígena apenas uma (Tabela 1). Pode-se concluir que 31,16% (IC95% = [29,38%;32,93%]) das mães pardas não obtiveram êxito em pelo menos uma gestação. Esse número é de 38,89% das mães declaradas

pretas.

Quanto a idade verificou-se que, em média, as mães têm 27,63 anos (IC95% = [27,43;27,82]) com números variando dos 13 anos e 50 anos de idade. Os registros foram alocados em quatro faixas etária conforme observado na Tabela 1. A classe de mulheres com idade entre 23 e 27 anos obteve maior frequência, foram 814 pacientes, ou seja 29%(IC95% = [27,32%;30,68%]) do total, seguida de perto pelo grupo de mulheres que possuem entre 32 e 50 anos de idade, que registrou 769 casos, 27,40% (IC95% = [25,75%;29,04%]) do total (Tabela 1). Registrou-se 103 casos de mães menores de idade, ou seja, entre 13 e 17 anos, das quais 46 não tiveram êxito em pelo menos uma gestação.

Tabela 1 – Número de mães atendidas no ISEA segundo a situação conjugal, o grau de instrução, raça e grupo de idade em 2018.

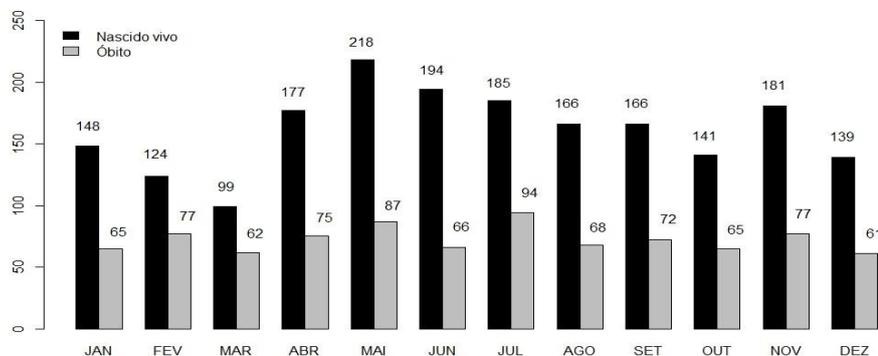
	Óbito		Total N(%)	Valor p
	Não n(%)	Sim n(%)		
Situação Conjugal				0,367
Solteira	1091 (38,87)	467 (16,64)	1558 (55,51)	
Casada	498 (17,74)	232 (8,26)	730 (26,00)	
Viúva	5 (0,18)	5 (0,18)	10 (0,36)	
Divorciada	7 (0,25)	6 (0,21)	13 (0,46)	
União Estável	337 (12,01)	159 (5,66)	496 (17,67)	
Grau de Instrução				0,018
Sem Escolaridade	25 (0,89)	20 (0,71)	45 (1,61)	
Fundamental I	430 (15,32)	225 (8,02)	655 (23,33)	
Fundamental II	815 (29,03)	352 (12,54)	1167 (41,57)	
Ensino Médio	590 (21,02)	249 (8,87)	839 (29,89)	
Superior Incompleto	31 (1,10)	5 (0,18)	36 (1,28)	
Superior Completo	47 (1,67)	18 (0,64)	65 (2,32)	
Raça				0,324
Branca	112 (3,99)	39 (1,39)	151 (5,38)	
Preta	22 (0,78)	14 (0,50)	36 (1,28)	
Parda	1803 (64,23)	816 (29,07)	2619 (93,30)	
Indígena	1 (0,04)	0 (0,00)	1 (0,04)	
Faixa Etária				0,001
[13 – 22]	477 (16,99)	190 (6,77)	667 (23,76)	
[23 – 27]	593 (21,13)	221 (7,87)	814 (29,00)	
[28 – 31]	365 (13,00)	192 (6,84)	557 (19,84)	
[32 – 50]	503 (17,92)	266 (9,48)	769 (27,40)	
Total	1938 (69,04)	869 (30,96)	2807 (100,00)	

Fonte: Autores.

O teste para a normalidade de Anderson Darling aponta para a não normalidade para todas as variáveis incluídas no estudo, a 0,05% de significância, dado que o valor p é menor que 0,001. O teste de qui-quadrado foi realizado para verificar a hipótese de que as variáveis estejam interferindo diretamente no óbito do produto das gestações. Observando a Tabela 1, quanto a situação conjugal, de acordo com o teste, não pode-se rejeitar a hipótese, a 0,05% de significância dado valor $p = 0,367$, de que a taxa de óbito independe da situação conjugal da mãe. Conclusão semelhante pode-se ter com relação a variável raça da mãe. De acordo com teste, com valor $p = 0,324$, não há indícios para afirmar, a 0,05 de significância, que a chance de óbito tenha dependência com a variável raça. Para o grau de instrução da mãe, pode-se afirmar que não há indícios para rejeitar a hipótese de que esta variável interfira diretamente na chance de a criança nascer viva a 0,05 de significância, o teste aponta valor $p = 0,018$. A idade da mãe interfere diretamente na chance de a criança vir a óbito, a 5% significância. O teste qui-quadrado obteve valor $p = 0,001$ na Tabela 1.

Observa-se na Figura 1 que o mês de Maio com 305 atendimentos dos quais 87 produtos de gestações vieram a óbito. Em seguida tem-se Julho com 279 casos registrados dos quais 66 foram óbito. Junho foi o terceiro mês com mais registros, foram 260 atendimentos registrados sendo 194 partos bem-sucedidos e 66 casos de óbito. O mês com menor número de atendimento foi Março, com 161 registros. Desses 62 resultaram em óbito. Dezembro e Fevereiro foram os outros dois meses com menor número de partos, com frequências 200 e 201, respectivamente. No mês de dezembro registrou-se 61 óbitos e no mês de fevereiro foram registrados 77 óbitos.

Figura 1 – Número de partos em que a criança nasceu viva e casos de óbitos por mês em 2018.



Fonte: Autores.

2.2 Escolha do Modelo

O vetor de respostas, Y , neste caso é binária e assume valor 0 quando a mãe atendida declara não ter perdido seu filho em nenhuma de suas gestações e valor 1 quando a mãe declara que não obteve sucesso em alguma de suas gestações. As variáveis explicativas são: a idade da mãe, a situação conjugal nos níveis solteira, casada, viúva, e união estável. A raça autodeclarada pela paciente nos níveis branca, preta, parda e indígena. O grau de instrução nos níveis sem escolaridade, fundamental I, fundamental II, ensino médio, ensino superior incompleto, e superior completo, também utilizamos a variável número de gestações.

Ajusta-se inicialmente o modelo saturado com tendência linear na variável idade, usando o método da máxima verossimilhança obtém-se as estimativas dos parâmetros. Observa-se na Tabela 2 os resultados encontrados, a significância para as estimativas é de 0,05. Observa-se que a variável raça não foi significativa. O modelo saturado registrou $AIC = 3121,3$ e desvio residual de 3091,3.

Tabela 2 – Estimativas de máxima verossimilhança dos parâmetros do modelo saturado com tendência linear na variável idade e intervalo de confiança.

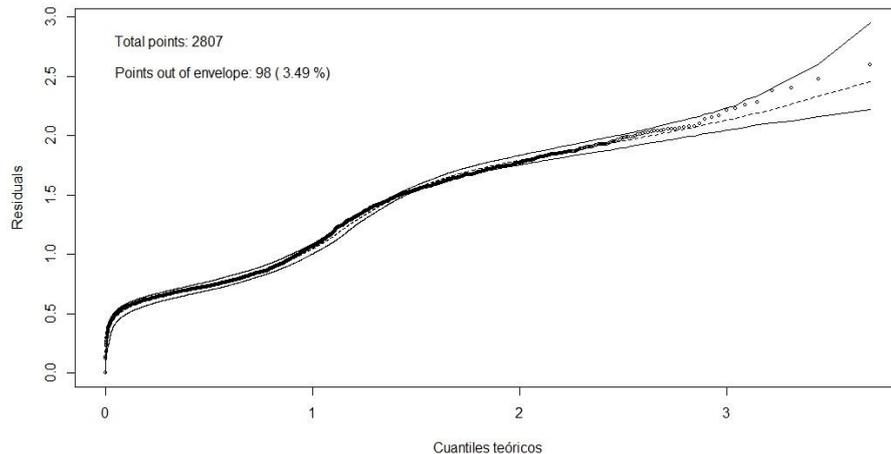
Parâmetro	Estimativa	E. P.	valor t	IC _{2,5%}	IC _{97,5%}	Pr(>/t/)
Intercepto	-2,0500	0,509	-4,003	-3,060	-1,06	0,001***
Idade (I)	-0,0476	0,008	-5,590	-0,060	-0,03	0,001***
Número de gestações (ν)	0,5973	0,036	16,440	0,530	0,670	0,001***
Situação Conjugal						
Casada (τ_2)	0,3514	0,111	-0,880	0,130	0,570	0,002**
Viúva (τ_3)	0,3234	0,708	0,460	-1,120	1,720	0,648
Divorciada (τ_4)	0,3925	0,629	0,620	-0,880	1,630	0,533
União estável (τ_5)	0,1493	0,120	1,250	-0,090	0,380	0,213
Raça/Cor						
Preta (ω_2)	0,2078	0,438	0,470	-0,680	1,050	0,635
Parda (ω_3)	0,0690	0,199	0,350	-0,310	0,470	0,729
Indígena (ω_4)	-11,4744	324,744	-0,040	-	54,700	0,972
Grau de escolaridade						
E. fundamental I (γ_2)	0,6910	0,397	1,740	-0,080	1,490	0,082.
E. fundamental II (γ_3)	0,9788	0,401	2,440	0,200	1,780	0,015*
E. Médio (γ_4)	1,2468	0,408	3,060	0,460	2,060	0,002**
E. S. Incompleto (γ_5)	0,4190	0,633	0,660	-0,90	1,620	0,508
E. S. Completo (γ_6)	1,3616	0,496	2,750	0,390	2,340	0,006**

Fonte: Autores.

*** Significante a 0,001 ** Significante a 0,01 * Significante a 0,05 . Significante a 0,10

A Figura 2 observa-se o gráfico de envelope simulado dos cálculos teóricos versus os resíduos. Observe que mesmo a variável raça não sendo significativa para o modelo, a 0,05 de significância, o gráfico mostra apenas 3,49% dos pontos fora do envelope.

Figura 2 – Envelope simulado para os resíduos do modelo saturado com tendência linear na variável idade.



Fonte: Autores.

Usando o método stepwise, observando os critérios AIC e o desvio residual, o modelo foi reajustado para a função

$$\log\left(\frac{\theta_i}{1-\theta_i}\right) = 1,2695 - 0,0479I + 0,5914v + 0,3063\tau_2 + 0,3278\tau_3 + 0,5906\tau_4 + 0,7203\tau_5 \quad (4.1)$$

O modelo (4.1) registrou $AIC = 3116,1$ e desvio residual 3092,1. Houve, portando, uma diminuição no valor de AIC e um pequeno aumento no valor do desvio residual. Na Tabela 3 observa-se as estimativas dos parâmetros e seus respectivos intervalos de confiança. Detecta-se que, com relação aos níveis de situação conjugal, mulheres casadas se diferencia estatisticamente dos demais níveis quanto a chance de óbito do filho. Em relação aos graus de instrução, mulheres com ensino fundamental II, ensino médio e ensino superior completo se diferenciam das mulheres analfabetas, com ensino fundamental I e/ou com ensino superior completo.

Tabela 3 – Estimativas de máxima verossimilhança dos parâmetros do modelo reduzido com tendência linear na variável idade e respectivos intervalos de confiança.

Parâmetro	Estimativa	E. P.	valor z	IC _{2,5%}	IC _{97,5%}	Pr(> z)
Intercepto	-1,2695	0,241	-5,27	-1,741	-0,797	0,001***
Idade (I)	-0,0479	0,008	-5,65	-0,064	-0,031	0,001***
Número de Gestações (ν)	0,5914	0,036	16,60	0,522	0,661	0,001***
Situação Conjugal						
Casada (τ_2)	0,3063	0,106	2,88	0,098	0,515	0,004**
Grau de Instrução						
E. Fundamental II (γ_3)	0,3278	0,119	2,75	0,094	0,561	0,006**
E. Médio (γ_4)	0,5906	0,130	4,53	0,335	0,846	0,001***
E. S. Completo (γ_6)	0,7203	0,311	2,32	0,111	1,329	0,020*

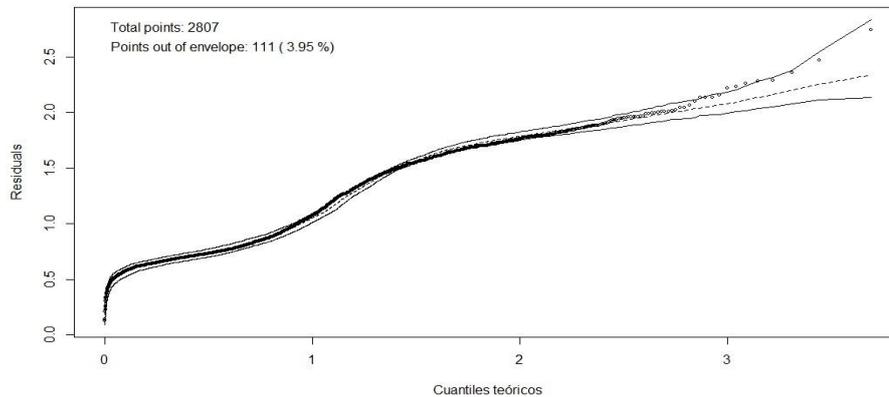
Fonte: Autores.

*** Significante a 0,001 ** Significante a 0,01 * Significante a 0,05

A Figura 3 mostra o envelope simulado dos resíduos para o modelo reduzido com tendência linear na variável idade. Observa-se que o modelo é significativo, a 0,05 de significância, pois há apenas 3,95% dos pontos simulados fora do envelope. Em outras palavras, o modelo se ajusta aos dados e todas as suas variáveis são significativas. Neste caso, pode-se conservar a hipótese de que o modelo se ajusta aos dados.

Espera-se que, naturalmente por motivos biológicos, a mulher ainda com pouca idade tenha chance zero de êxito na gestação e que atingido o período fértil, geralmente no início da adolescência, essa chance seja positiva e aumente consideravelmente com o passar dos anos atingindo um valor máximo. Posteriormente, quando a carga hormonal feminina tende a baixar, a chance de a mulher ter sucesso no processo de gestação tende a diminuir. Esse comportamento não é linear.

Figura 3 – Envelope simulado para os resíduos do modelo reduzido com tendência linear na variável idade.



Fonte: Autores.

Adicionando-se a tendência quadrática no modelo tem-se novas estimativas, que estão descritas na Tabela 4. O modelo possui $AIC = 3112,3$ e desvio residual de $3086,3$. Com relação ao modelo reduzido com tendência linear na variável idade verifica-se uma queda de $3,8$ no valor do AIC e uma redução de $5,8$ no valor do desvio residual. Essa diminuição no valor do desvio é um indicativo de melhoramento no ajuste do modelo aos dados. O modelo quadrático ajustado tem a equação

$$\log \frac{\theta_i}{1-\theta_i} = -0,1781I + 0,0023I^2 + 0,6007\nu + 0,3583\tau_2 + 1,0268\gamma_3 + 1,3180\gamma_4 + 1,3901\gamma_6.$$

O modelo apresentou o menor AIC e o menor valor para o desvio. Aparentemente é o modelo que melhor explica a relação dos dados com a chance de óbito. As estimativas dos parâmetros do modelo podem ser vistas na Tabela 4.

Tabela 4 – Estimativas de máxima verossimilhança dos parâmetros do modelo reduzido com tendência quadrática na variável idade e intervalo de confiança.

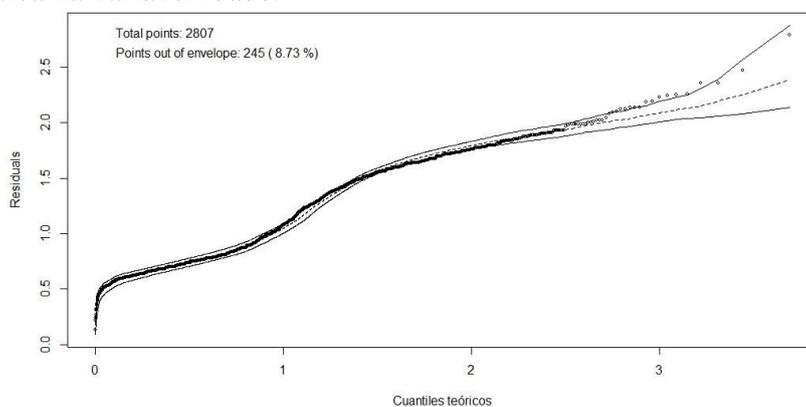
Parâmetro	Estimativa	E. P.	valor z	IC _{2,5%}	IC _{97,5%}	Pr(> z)
Intercepto	-0,2805	0,843	-0,330	-1,940	1,370	0,739
Idade I	-0,1781	0,054	-3,270	-0,280	-0,070	0,001**
Idade ²	0,0023	0,001	2,430	0,001	0,004	0,015*
Número de Gestações (ν)	0,6007	0,036	16,470	0,5300	0,670	0,001***
Situação Conjugal						
Casada (τ_2)	0,3583	0,112	3,210	0,140	0,580	0,002**
Viúva (τ_3)	0,3311	0,708	0,470	-1,120	1,730	0,640
Divorciada (τ_4)	0,4299	0,629	0,680	-0,840	1,660	0,494
União Estável (τ_5)	0,1503	0,120	1,250	-0,090	0,380	0,209
Grau de Instrução						
E. Fundamental I (γ_2)	0,7372	0,404	1,820	-0,040	1,550	0,068.
E. Fundamental II (γ_3)	1,0268	0,408	2,520	0,240	1,850	0,012*
E. Médio (γ_4)	1,3180	0,415	3,180	0,520	2,150	0,002**
E. S. Incompleto (γ_5)	0,5043	0,638	0,790	-0,820	1,720	0,429
E. S. Completo (γ_6)	1,3901	0,501	2,770	0,4100	2,380	0,006**

Fonte: Autores.

*** Significante a 0,001, ** Significante a 0,01, * Significante a 0,05, Significante a 0,10.

A Figura 4 mostra o envelope dos quantis teóricos versus os resíduos. Neste modelo, reduzido com tendência quadrática na variável idade, a porcentagem de pontos fora do envelope aumentou quando comparado com o modelo reduzido e apenas tendência linear na idade. Foram registrados 245, ou seja, 8,73% dos pontos fora do envelope. Significa que não se pode admitir que o modelo é significativo, a 0,05 de significância, para explicar a relação entre os dados e a chance de óbito.

Figura 4 – Envelope simulado para os resíduos do modelo reduzido com tendência quadrática na variável idade.



Fonte: Autores.

Conforme a Tabela 5, nos ajuda a compreender a seleção do modelo, nela podemos ver os valores do AIC, dos desvios residuais e da porcentagem de pontos fora do envelope simulado para cada modelo.

Selecionado o modelo que tem as melhores condições para modelar a relação entre a chance de óbito e as variáveis explicativas fizemos a atualização dos parâmetros por meio do Teorema de Bayes.

Tabela 5 – Tabela resumo dos valores de AIC, desvio residual e percentual de pontos fora do envelope simulado.

Modelo	AIC	Desvio Residual	Percentual(%)
Linear Saturado	3121,3	3091,3	3,49%
Linear Reduzido	3116,1	3092,1	3,95%
Quadrático Reduzido	3112,3	3086,3	8,73%

Fonte: Autores.

Cada parâmetro tem distribuição a priori $N(\mu_{MV}, 100)$, em que μ_{MV} é sua respectiva estimativa de máxima verossimilhança dada no modelo clássico. Isso reflete a informação a pouca informação que temos sobre os valores da variabilidade de cada parâmetro. Simulando 11 mil vezes, com saltos de 10 estimativas, sendo que as primeiras mil simulações são descartadas obtemos as atualizações dos parâmetros do modelo. A Tabela 6 apresenta a média e o desvio padrão e o intervalo de credibilidade de cada parâmetro.

Tabela 6 – Estimativas dos parâmetros do modelo bayesiano e intervalos de credibilidade.

Parâmetro	Média	Desvio Padrão	Ic2,5%	Ic97,5%
Intercepto	-1,26011	0,245500	-1,73867	-0,79012
Idade	-0,04864	0,008742	-0,06544	-0,03154
Número de Gestações (ν)	0,59535	0,036691	0,52452	0,66571
Situação Conjugal				
Casada (τ_2)	0,30285	0,107853	0,07748	0,50499
Grau de Instrução				
E. Fundamental II (γ_3)	0,32936	0,124712	0,06613	0,58345
E. Médio (γ_4)	0,59074	0,134722	0,33933	0,85858
E. Superior Completo (γ_6)	0,72752	0,312674	0,09130	1,31497

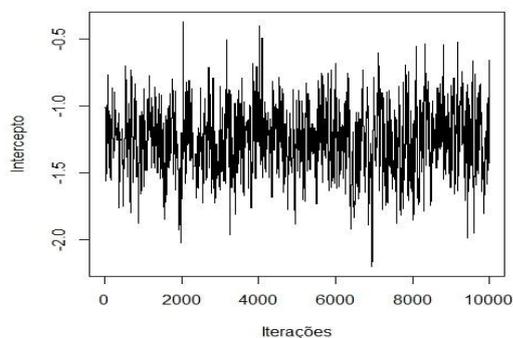
Fonte: Autores

O modelo bayesiano é dado pela equação (4.2)

$$\log\left(\frac{\theta_i}{1-\theta_i}\right) = 1,2601 - 0,0486I + 0,5953v + 0,3028\tau_2 + 0,3294\tau_3 + 0,5907\tau_4 + 0,7275\tau_5 \quad (4.2)$$

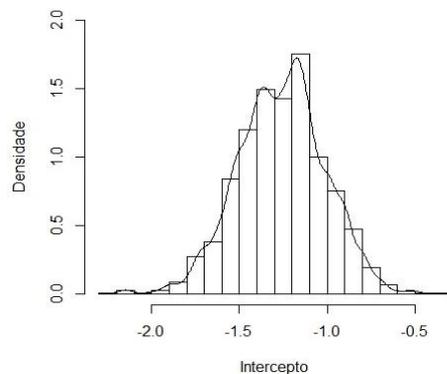
A sequência de figuras a seguir, Figura 5 até a Figura 18, mostra as séries temporais dos parâmetros e as suas respectivas densidades a posterioris. Podemos ver a convergência dos parâmetros nas figuras.

Figura 5 – Valores estimados nas iterações Intercepto



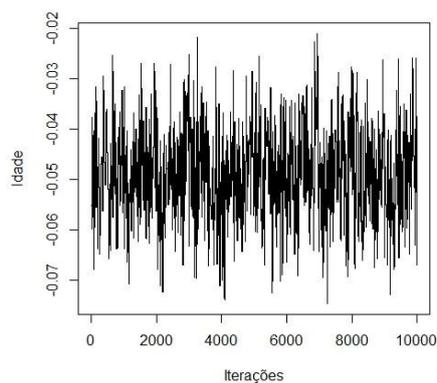
Fonte: Autores.

Figura 6 - Densidade a posteriori do Intercepto



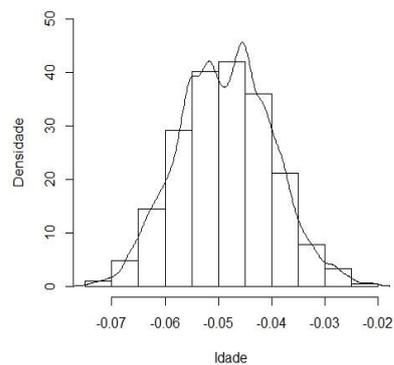
Fonte: Autores.

Figura 7 – Valores estimados nas iterações do parâmetro (I)



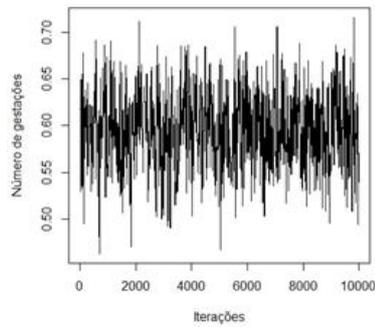
Fonte: Autores.

Figura 8 – Densidade a posteriori do iterações do parâmetro (I).



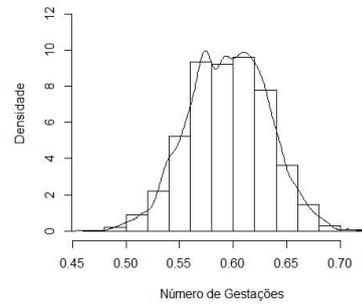
Fonte: Autores.

Figura 9 – Valores estimados nas iterações do parâmetro ν .



Fonte: Autores.

Figura 10 – Densidade a posteriori parâmetro ν .



Fonte: Autores.

Figura 11 – Valores estimados nas iterações do parâmetro τ_2 .

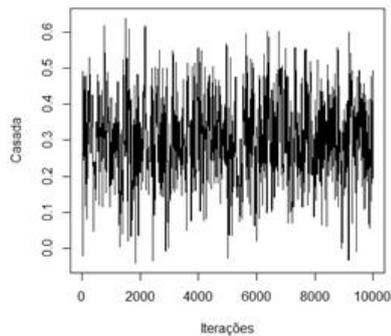
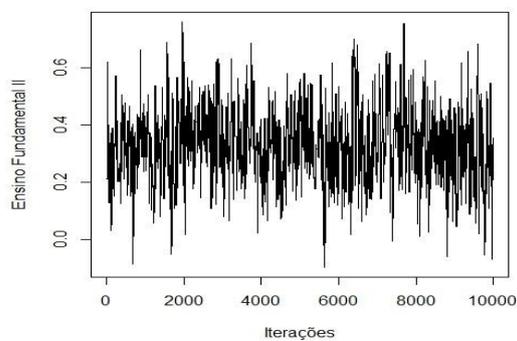


Figura 13 – Valores estimados nas iterações do parâmetro γ_3 .



Fonte: Autores.

Figura 12 – Densidade do parâmetro a posteriori τ_2 .

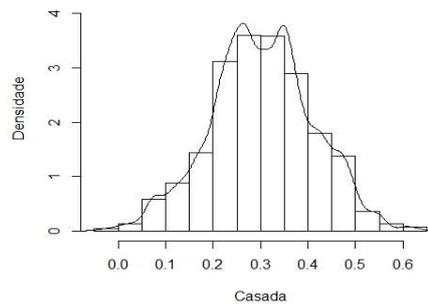
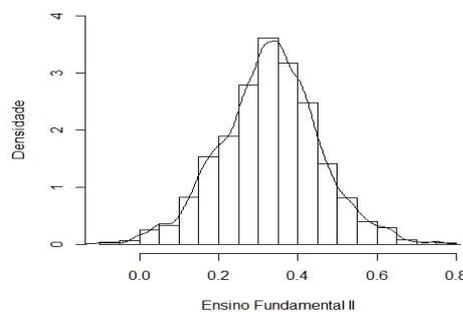
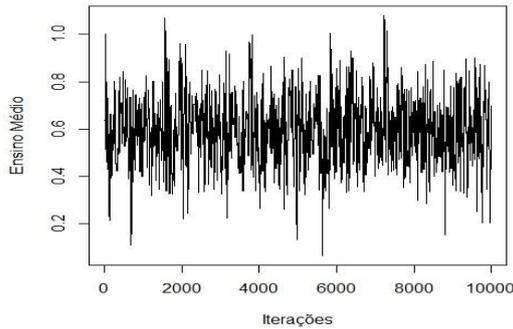


Figura 14 – Densidade a posteriori do parâmetro γ_3



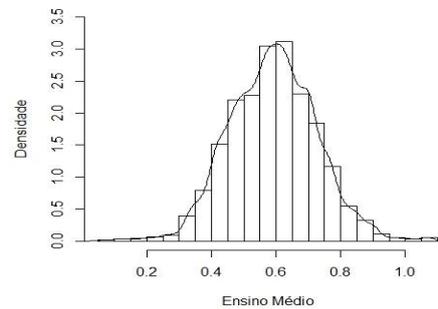
Fonte: Autores.

Figura 15 – Valores estimados nas iterações do parâmetro γ_4 .



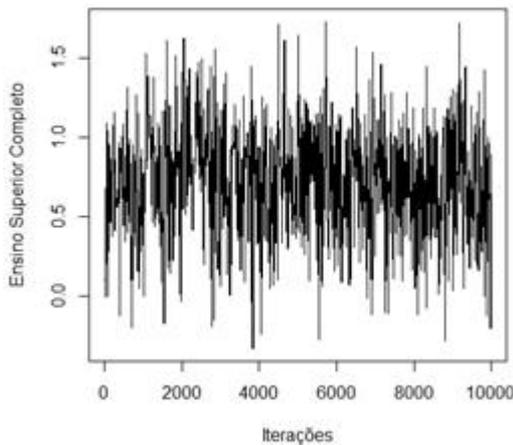
Fonte: Autores.

Figura 16 – Densidade a posteriori do parâmetro γ_4 .



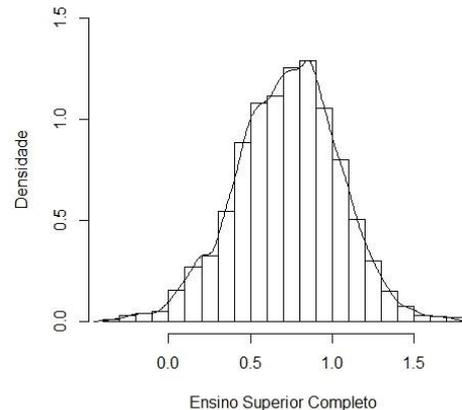
Fonte: Autores.

Figura 17 – Valores estimados nas iterações parâmetro γ_6



Fonte: Autores.

Figura 18 – Densidade a posteriori do parâmetro γ_6



Fonte: Autores.

2.1.1 Interpretação dos modelos e previsões

O modelo ajustado por verossimilhança, (4.1), é bastante semelhante ao modelo Bayesiano, (4.2). Observa-se essa semelhança pela proximidade dos valores dos parâmetros.

Os valores preditos pelos dois modelos também se assemelham, conforme mostra a Tabela 7 e a Tabela 8. O modelo (4.1) mostra 1881 casos de crianças nascidas vivas, confirmado pelos dados, e 127 casos de óbitos dados que nasceram vivas. O mesmo modelo também nos diz que 649 casos de crianças nascidas vivas quando na verdade elas foram casos de óbito e 220 casos de óbitos confirmados pelos dados.

Tabela 7 – Matriz de confusão do modelo (4.1).

		Valores Preditos		Total
		Vivo	Óbito	
Nascido	Vivo	1811	127	1938
	Óbito	649	220	869
Total		2460	347	2807

Fonte: Autores.

O modelo Bayesiano, (4.2), nos mostra resultado bastante semelhante. Esse modelo nos mostra 1810 casos de crianças que nasceram vivas, casos estes confirmados pelos dados, enquanto que 128 casos foram registrados como óbito erroneamente. Portanto o modelo (4.1) obteve 1 acerto a mais que o modelo Bayesiano. O modelo (4.2) ainda detecta 649 casos de nascidos vivos quando na verdade os indivíduos foram a óbito e 220 casos de óbitos quando realmente os indivíduos foram a óbito.

Tabela 8 – Matriz de confusão do modelo (4.2).

		Valores Preditos		Total
		Vivo	Óbito	
Nascido	Vivo	1810	128	1938
	Óbito	649	220	869
Total		2459	348	2807

Fonte: Autores.

Na Tabela 9 tem-se a acurácia, a taxa positiva verdadeira (TP), a taxa de falsos positivos (FP), a taxa negativa verdadeira (TN), a taxa de falsos negativos (FN), e precisão para os dois modelos ajustados. Os valores são praticamente iguais, alguns diferem muito pouco. O modelo (4.2) tem uma acurácia um pouco maior enquanto que o modelo (4.1) tem maior precisão.

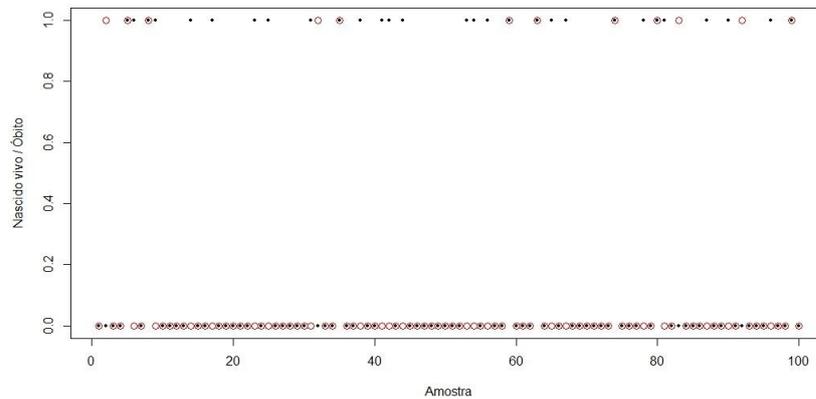
Tabela 9 – Indicadores dos modelos ajustados.

	Acurácia	TP	FP	TN	FN	Precisão
Modelo (4.1)	0,7232	0,2532	0,0655	0,9344	0,7468	0,6340
Modelo (4.2)	0,7235	0,2532	0,0660	0,9339	0,7468	0,6322

Fonte: Autores.

A Figura 19, mostra em preto uma amostra observada de 100 desfechos (nascidos vivos ou óbitos), em vermelho a previsão Bayesiana. Os valores, em vermelho, são obtidos probabilisticamente.

Figura 19 – Pontos preditos pelo modelo Bayesiano para uma amostra de 100 mulheres.



Fonte: Autores.

Simulando alguns resultados possíveis valores para as variáveis do modelo, e adotando o modelo Bayesianos, calculou-se a probabilidade estimada ($\hat{\theta}$) de óbito para cada situação. Observe a Tabela 10, destaca-se três situações. Na linha três, tem-se a situação em que a mãe de primeira viagem tem 13 anos de idade, não é casada e não tem ensino possui o fundamental II, a probabilidade de óbito da criança é de 0,215. Na linha sete, as mães têm 45 anos de idade, está em sua sétima gestação, não é casada e possui ensino médio. Nesta situação a probabilidade de desfecho trágico de sua gestação é de 0,740. Por fim, na linha nove, tem-se uma mãe que possui ensino superior completo, não é casada e está gerando seu primeiro filho com 50 anos de idade, a probabilidade de óbito da criança é de apenas 0,085.

Tabela 10 – Cálculo da probabilidade de óbito para 10 situações aleatórios possíveis.

Idade (<i>I</i>)	Número de Gestação (<i>v</i>)	Situação Conjugal	Grau de Instrução	θ
26	2	Outros	E. Fundamental II	0,268
29	2	Outros	Ensino Médio	0,291
13	1	Outros	Outros	0,215
38	1	Outros	E. Fundamental II	0,101
28	1	Casada	Outros	0,199
40	4	Outros	Outros	0,305
45	7	Outros	E. Médio	0,740
26	3	Outros	Outros	0,323
50	1	Outros	E. S. Completo	0,085
24	1	Outros	Outros	0,138

Fonte: Autores.

4. Considerações Finais

De todas as características apresentadas neste trabalho apenas mães casadas, referente à situação conjugal, e mães com ensino médio, referente ao grau de escolaridade podem explicar as chances óbitos perinatais ocorridos com pacientes atendidas em 2018 no ISEA. Mulheres casadas costumam fazer tratamentos para suprir as dificuldades em engravidar, nem sempre são bem-sucedidos, gerando um acréscimo no índice de óbito perinatal. Mulheres com ensino superior costumam ter mais acesso a informação, maior poder aquisitivo e acesso ao serviço de saúde com melhor qualidade e maior rapidez quando comparadas com mulheres que possuem apenas ensino médio, conseqüentemente há maior índice de sucesso no processo de gestação. Neste trabalho pôde-se constatar que a idade da mãe é um fator que interfere diretamente na chance de uma criança vir a óbito no período perinatal assim como o número de gestações. As outras variáveis que interferem são o fato da mãe ser casada e ter ensino fundamental II, ter ensino médio ou curso superior completo.

Referências

- Bolfarine, H., & Sandoval, M. C. (2001). Introdução à inferência estatística. Rio de Janeiro, 2.
- Mardia, K. V., & Marshall, R. J. (1984). Maximum likelihood estimation of models for Residual covariance in spatial regression. *Biometrika, Oxford University Press*, 71(1), 135–146.

Ministério da Saúde (2016). Entendendo o SUS. Brasília. Disponível em: <<http://portalsaude.saude.gov.br/index.php/cidadao/entenda-o-sus>>.

Moral, R. A., Hinde, J., & Demetrio, C. G. (2017). Half-normal plots and overdispersed models in r: The hnp package. *Journal of Statistical Software*, 81 (10), 1–23.

Nelder, J. A., & Baker, R. J. (1972). *Generalized linear models*. New York: Wiley Online Library.

Paraíba, J. da. Mais de 5 mil partos feitos no Isea-CG são de outras cidades e caso vai ao MFP. Campina Grande: Reportagem, 2018.

Sarinho S. W., Silva G. A. P., Melo F. D. A., & Guimarães M. J. B. (1998). Causas de óbitos neonatais na Cidade do Recife segundo critério de evitabilidade. *An Fac Med Univ Fed Pernambuco*; 12 (43), 112-5.

Silva C. F., Leite A. J. M., Almeida N. M. G. S., & Gondim R. C. (2006). Fatores de risco para mortalidade infantil em município do nordeste do Brasil: linkage entre bancos de dados de nascidos vivos e óbitos infantis-2000 a 2002. *Revista Brasileira de Epidemiologia*; 9 (1), 69–80.

Victora, C. G., Aquino, E. M. M. L. L. de, Leal, M. do C., Monteiro, C. A., Barros, F. C. L. F. de, & Szwarcwald, C. L. (2011). Saúde de mães e crianças no Brasil: progressos e desafios. *The Lancet*, 32-46. doi:10.1016/S0140-6736(11)60138-4

Porcentagem de contribuição de cada autor no manuscrito

Sandro Lins Lopes de Lucena - 30 %

Kleber Napoleão Nunes de Oliveira Barros - 35 %

Mácio Augusto de Albuquerque – 35%