

Análise Completa do estudo de depressão de Riesby e do Crescimento de Árvores Sitka Spruce.

Cícero Hitzschky

cicero.hitzschky@alu.ufc.br

Departamento de Estatística e Matemática Aplicada
Universidade Federal do Ceará

21 de fevereiro de 2025



Sumário

- 1 Dataset Riesby
 - Desenho do Estudo

- 2 Dataset Sitka Spruce
 - Análise

Dataset Riesby

Sobre o conjunto de dados

- O dataset Riesby representa um ensaio clínico psiquiátrico longitudinal descrito em Reisby et al. (1977) para tratamento de depressão.
- O estudo focou na relação longitudinal entre os níveis plasmáticos de imipramina (IMI) e desipramina (DMI) e a resposta clínica em 66 pacientes internados com depressão é a mudança nas pontuações de depressão semana a semana.
- Como a imipramina se biotransforma no metabólito ativo desmetilimipramina (ou desipramina), a medição da desipramina também foi feita neste estudo.

Desenho do Estudo

Fase Inicial

Período de Placebo

Tratamento

Doses de 225mg/dia de imipramina por 4 semanas.

Avaliação

Escala de classificação de depressão de Hamilton (Hamilton, 1960).

Medições

Nível plasmático de imipramina (IMI) e seu metabólito desipramina (DMI) medidos no final de cada semana de tratamento.



Desenho do Estudo

Coleta de dados

- Sexo
- Diagnóstico de Depressão: Endógena ou Reativa (Não endógena).



Desenho do Estudo

Número de Participantes

Um total de 66 indivíduos sendo a variação por semana dada por:

- Semana 0: 61 participantes.
- Semana 1: 63 participantes.
- Semana 2: 65 participantes.
- Semana 3: 58 participantes.

O conjunto de dados

Tabela: Níveis plasmáticos de imipramina (IMI) e desipramina (DMI) e HDRS score em pacientes com depressão durante o tratamento psiquiátrico.

ID	Score (HDRS)	Semana	Sexo	Endógena	IMI(mg/L)	DMI(mg/L)
101	-8	0	0	0	4,043050	4,204690
101	-19	1	0	0	3,931830	4,812180
101	-22	2	0	0	4,330730	4,962840
101	-23	3	0	0	4,369450	4,962840
103	-18	0	1	0	2,772590	5,236440
⋮	⋮	⋮	⋮	⋮	⋮	⋮
360	12	3	0	1	3,637590	4,844190
361	-19	0	1	1	4,204690	3,784190
361	-22	1	1	1	4,584970	4,234110
361	-23	2	1	1	4,382030	4,189650
361	-11	3	1	1	4,624970	4,189650



O conjunto de dados

Questões de Interesse

- O tratamento obteve resultados satisfatórios?
- O tratamento em pacientes endógenos é mais eficaz?



O conjunto de dados

Tabela: Escore HDRS dos pacientes em cada semana de tratamento.

Semana ID	0	1	2	3
101	-8	-19	-22	-23
103	-18	-9	-18	-20
104	-11	-16	-10	-29
105	-6	-6	-9	-13
⋮	⋮	⋮	⋮	⋮
607	0	-3	-10	-26
608	-10	-12	-21	-20
609	-3	-11	-10	-23
610	-1	-11	N/A	-23

Medidas de Resumo

Tabela: Medidas de resumo do Score de Depressão HDRS para cada sexo nas diferentes semanas de tratamento.

		Quantidade	Média	DP	Min	Q1	Mediana	Q3	Max
Sexo	Semana								
Feminino	0	19	-2,79	5,24	-13	-6,00	-3,0	0,00	8
	1	19	-6,58	8,49	-24	-11,00	-5,0	-2,00	8
	2	19	-9,58	8,21	-23	-15,00	-9,0	-4,00	7
	3	18	-10,67	8,98	-23	-16,75	-14,5	-5,25	12
Masculino	0	45	-5,80	5,23	-19	-10,00	-4,0	-2,00	6
	1	46	-7,00	5,60	-22	-10,00	-7,0	-3,00	4
	2	44	-9,55	7,04	-23	-12,00	-10,0	-6,50	6
	3	40	-11,4	7,44	-29	-15,00	-11,0	-6,00	7

Gráfico de trajetórias individuais do Score HDRS para cada sexo.

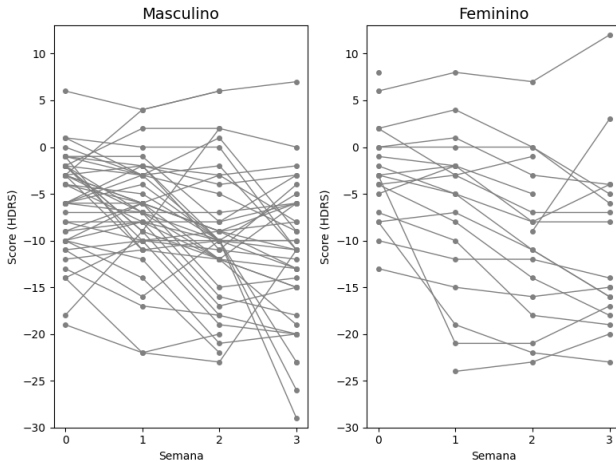


Gráfico de trajetórias individuais do Score HDRS com perfil médio para cada sexo.

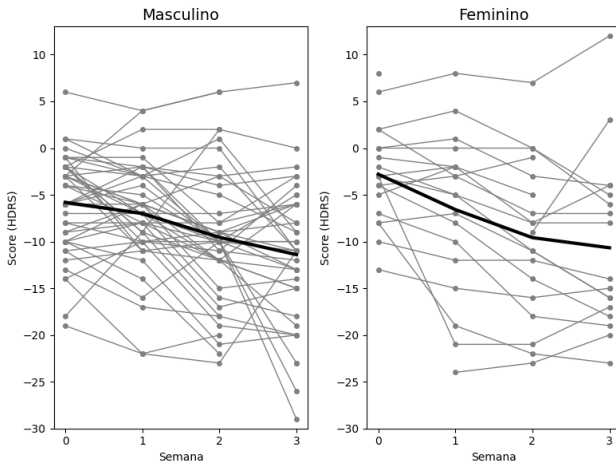
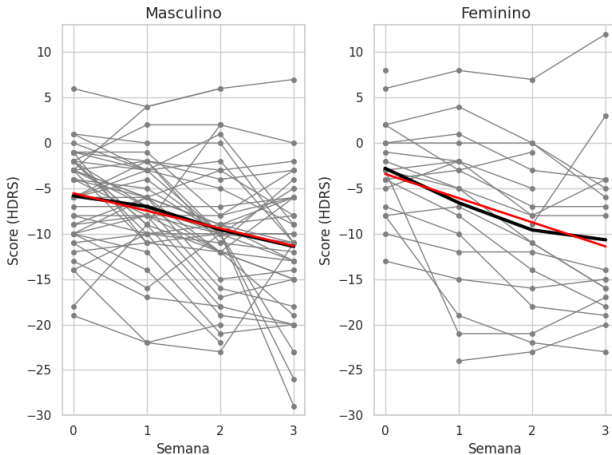


Gráfico de trajetórias individuais do Score HDRS com perfil médio e regressão para cada sexo.



Modelos de Regressão simples

Modelo de Regressão para o sexo feminino

$$y = -3.41 - 2.66x + \varepsilon$$

Modelo de Regressão para o sexo masculino

$$y = -5.53 - 1.93x + \varepsilon$$



Gráfico de trajetórias individuais do Score HDRS para o sexo feminino com perfil médio e regressão simples e regressão polinomial do 2nd grau ajustada.

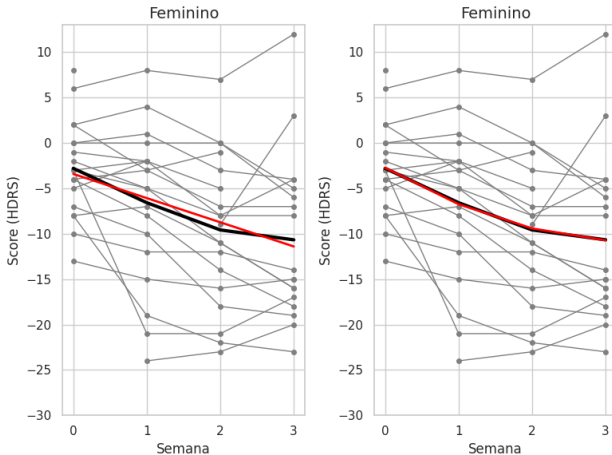
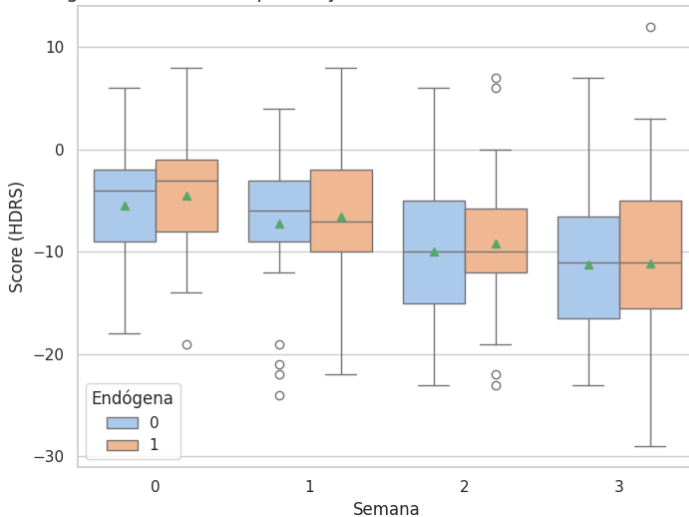
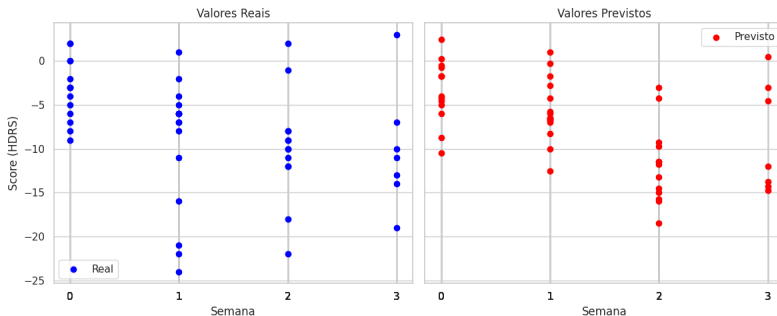


Diagrama de Caixa da pontuação HDRS em cada semana de tratamento



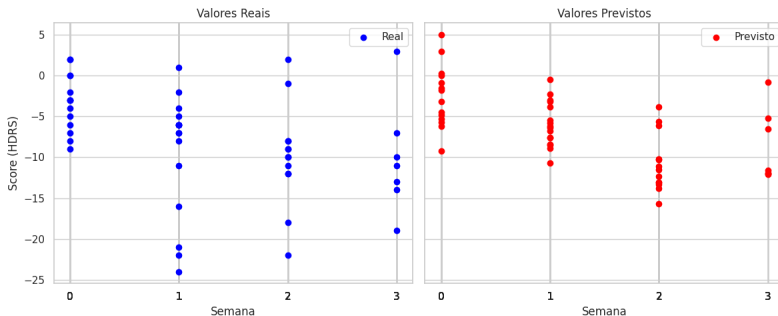
Regressão via Florestas Aleatórias

Para 4 árvores



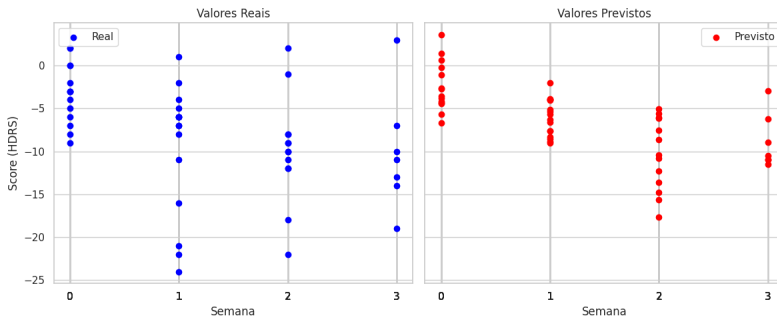
Regressão via Florestas Aleatórias

Para 10 árvores



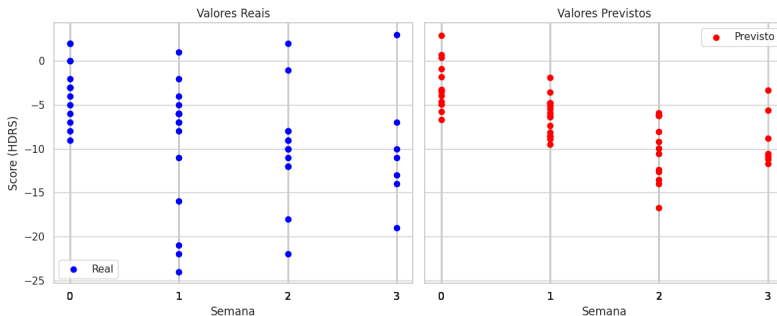
Regressão via Florestas Aleatórias

Para 50 árvores



Regressão via Florestas Aleatórias

Para 100 árvores



Melhor Escolha de Hiperparametro

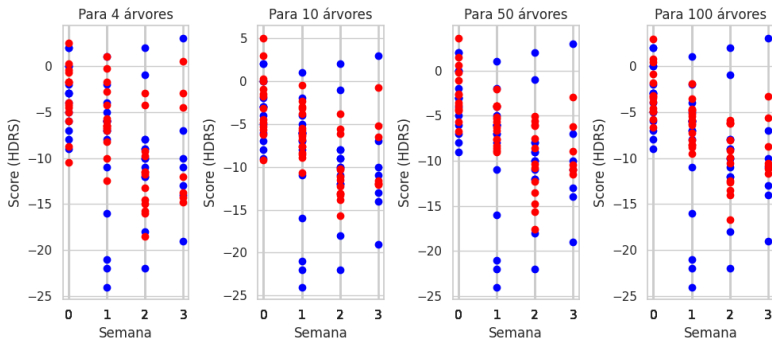
Tabela: Comparação de Hiperparâmetros do Random Forest

Num de Árvores	EQM	R^2 Score	EMA
4	44.06	-0.0309	5.06
10	31.50	0.2630	4.33
50	33.30	0.2203	4.42
100	32.70	0.2300	4.27



Regressão via Florestas Aleatórias

Comparação de Valores Reais e Previstos



Dataset Sitka Spruce

Sobre o conjunto de dados

- O dataset Sitka Spruce apresenta medições longitudinais do crescimento de árvores da espécie *Picea sitchensis*.
- A variável de resposta é o logaritmo do tamanho das árvores, onde o tamanho é calculado como o produto da altura da árvore e o quadrado do diâmetro.
- O estudo acompanha o desenvolvimento das árvores ao longo do tempo, analisando o crescimento com base nas medições repetidas.
- Como a poluição por ozônio é comum em áreas urbanas, o impacto do aumento das concentrações de ozônio no crescimento das árvores é de grande interesse.

Desenho do Estudo

Objetivo

Avaliar o efeito da poluição por ozônio no crescimento das árvore

Medições

- Logaritmo do tamanho das árvores, calculado como o produto da altura da árvore e do quadrado do diâmetro (*log inch*).
- Idade das árvores (em anos).



Desenho do Estudo

De 79 árvores, um total de 54 árvores foi cultivado com exposição ao ozônio a 70 ppb; 25 foram cultivadas em condições de controle.

Desenho do Estudo

Coleta de Dados

- Medições repetidas ao longo do tempo para acompanhar o crescimento.
- 79 árvores foram medidas do dia 469 até o dia 674



O conjunto de dados

Tabela: Comprimento de árvores *Picea sitchensis* ao longo de 205 dias

	linha	log pol	dias	arvore	tratamento
0	1	6,16	469	1	ozônio
1	2	6,18	496	1	ozônio
2	3	6,48	528	1	ozônio
3	4	6,65	556	1	ozônio
4	5	6,87	579	1	ozônio
⋮	⋮	⋮	⋮	⋮	⋮
627	628	5,95	556	79	controle
628	629	5,80	579	79	controle
629	630	6,21	613	79	controle
630	631	6,28	639	79	controle
631	632	6,34	674	79	controle



O conjunto de dados

Questões de Interesse

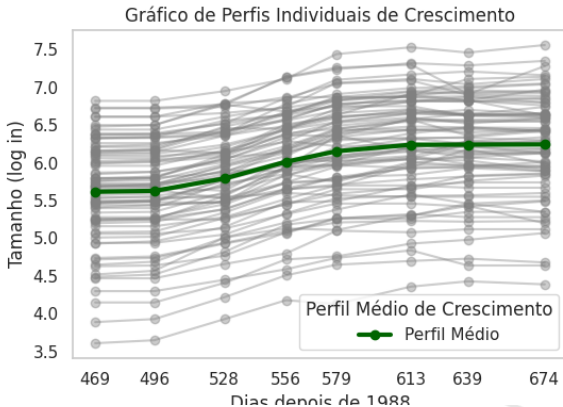
- A concentração de ozônio realmente influencia no crescimento das árvores?

Formato Longitudinal

Tabela: Comprimento de árvores *Picea sitchensis* ao longo de 205 dias

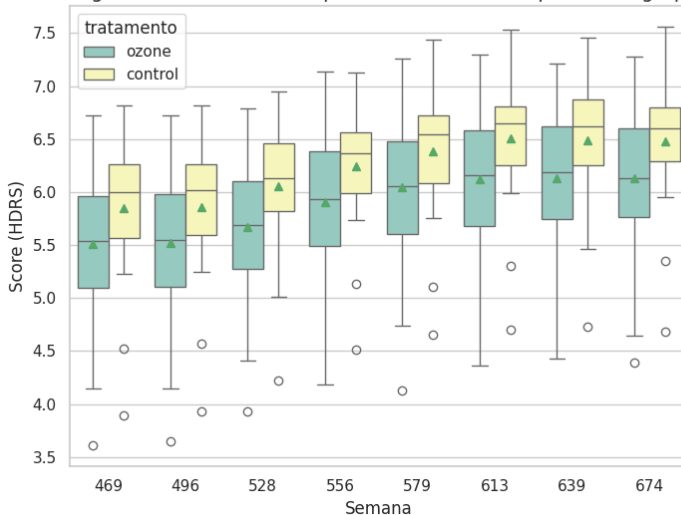
dias arvore	469	496	528	556	579	613	639	674
1	6,16	6,18	6,48	6,65	6,87	6,95	6,99	7,04
2	5,20	5,22	5,39	5,65	5,71	5,78	5,82	5,85
3	5,87	5,88	6,04	6,34	6,49	6,58	6,65	6,61
4	5,53	5,56	5,68	5,93	6,21	6,26	6,20	6,19
5	6,50	6,50	6,79	6,83	7,10	7,17	7,21	7,16
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
75	5,79	5,82	6,05	6,29	6,22	6,39	6,47	6,42
76	5,40	5,40	5,73	5,85	5,75	5,99	6,10	6,15
77	4,52	4,57	5,01	5,13	5,11	5,30	5,46	5,35
78	6,33	6,34	6,56	6,63	6,75	6,89	6,96	6,94
79	5,23	5,25	5,56	5,95	5,98	6,21	6,28	6,34

Análise



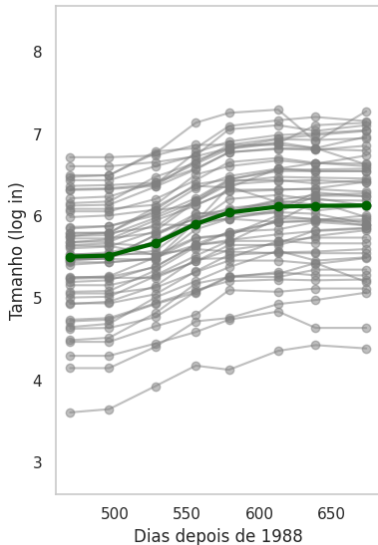
Análise

Diagrama de Caixa do comprimento das arvores para cada grupo

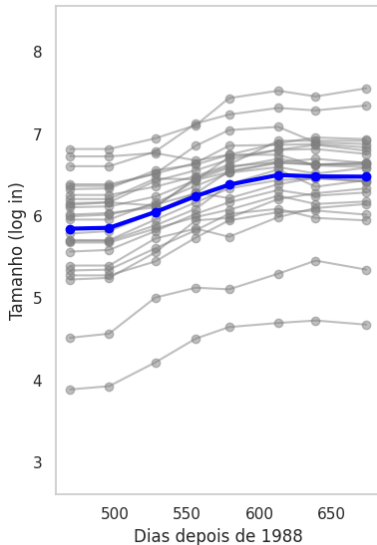


Análise

Perfis Individuais - Ozônio



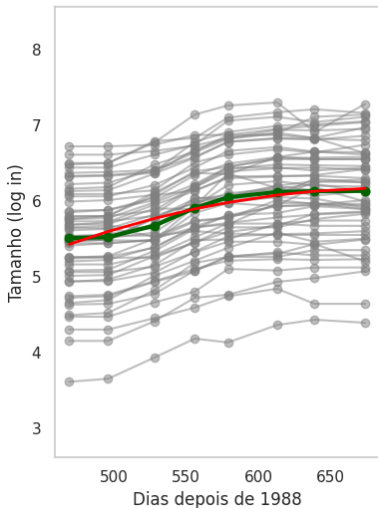
Perfis Individuais - Controle



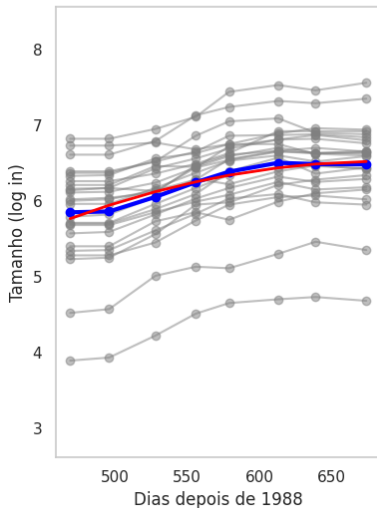
Análise

Ajuste via Polinômio de Grau 2

Perfis Individuais - Ozônio



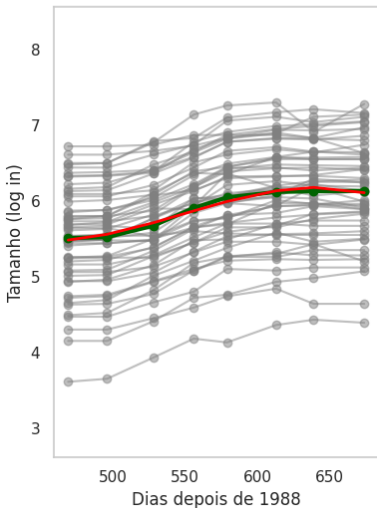
Perfis Individuais - Controle



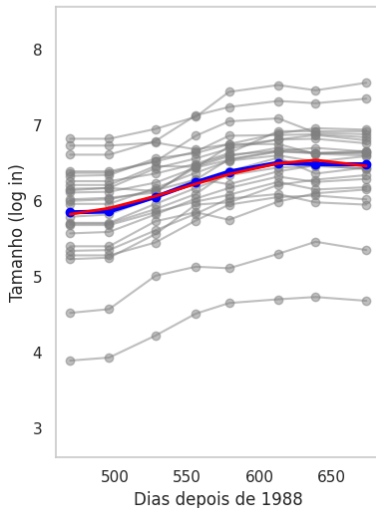
Análise

Ajuste via Polinômio de Grau 3

Perfis Individuais - Ozônio



Perfis Individuais - Controle



Tentativa de Modelo Misto

Mixed Linear Model Regression Results

```
=====
Model:                MixedLM Dependent Variable: log_in
No. Observations: 632    Method:                REML
No. Groups:           79    Scale:                0.0175
Min. group size:      8    Log-Likelihood:    163.5584
Max. group size:      8    Converged:         Yes
Mean group size:     8.0
=====
```

```
-----
                Coef. Std.Err.   z    P>|z| [0.025 0.975]
-----
Intercept      3.912      0.086 45.329 0.000   3.743   4.081
dias           0.004      0.000 45.943 0.000   0.003   0.004
arvore Var     0.425      0.551
=====
```



tentativa de Equação de Estimação Generalizada

GEE Regression Results						
=====						
Dep. Variable:	log_in	No. Observations:	632			
Model:	GEE	No. clusters:	79			
Method:	Generalized	Min. cluster size:	8			
	Estimating Equations	Max. cluster size:	8			
Family:	Gaussian	Mean cluster size:	8.0			
Dependence structure:	Exchangeable	Num. iterations:	2			
Date:	Fri, 21 Feb 2025	Scale:	0.438			
Covariance type:	robust	Time:	15:42:16			
=====						
	coef	std err	z	P> z	[0.025	0.975]

Intercept	3.9121	0.106	36.949	0.000	3.705	4.120
dias	0.0037	0.000	30.905	0.000	0.003	0.004
=====						
Skew:	-0.5811	Kurtosis:	0.0652			
Centered skew:	-0.1805	Centered kurtosis:	0.3765			
=====						



Referências (1/5)

- [1] RIESBY, N. *Depression Study*. R Data Repository, v. 1.0, Depression of Riesby dataset. Disponível em: <https://www.rdocumentation.org/packages/datasets/versions/3.6.2/topics/Riesby>. Acesso em: 4 nov. 2024.
- [2] R CORE TEAM. *multilevelmod: Multilevel models and model diagnostics*. CRAN. Acesso em: 9 de fevereiro de 2025, de <https://cran.r-project.org/web/packages/multilevelmod/multilevelmod.pdf>.

Referências (2/5)

- [1] SITKA, M.; HAYWARD, B. *Growth of Sitka Trees*. R Data Repository, v. 1.0, Growth of Sitka Trees dataset. Disponível em: <https://www.rdocumentation.org/packages/datasets/versions/3.6.2/topics/sitka>. Acesso em: 4 nov. 2024.
- [2] HEDEKER, Donald; GIBBONS, Robert D. *Applied Longitudinal Data Analysis*. Hoboken: Wiley-Interscience, 2005. (Wiley Series in Probability and Statistics).

Referências (3/5)

- [1] DIGGLE, Peter; HEAGERTY, Patrick; LIANG, Kung-Yee; ZEGER, Scott. *Analysis of Longitudinal Data*. 2. ed. Oxford: Oxford University Press, 2002.
- [2] MONETTE, G. *Riesby*. RDrr.io, 2025. Disponível em: <https://rdrr.io/github/gmonette/spida15/man/Riesby.html>. Acesso em: 19 fev. 2025.

Referências (4/5)

- [1] MATPLOTLIB. *Matplotlib Documentation*. Disponível em: <https://matplotlib.org/stable/contents.html>. Acesso em: 21 fev. 2025.
- [2] NUMPY. *NumPy Documentation*. Disponível em: <https://numpy.org/doc/stable/>. Acesso em: 21 fev. 2025.

Referências (5/5)

- [1] PANDAS. *Pandas Documentation*. Disponível em: <https://pandas.pydata.org/pandas-docs/stable/>. Acesso em: 21 fev. 2025.
- [2] STATSMODELS. *Statsmodels Documentation*. Disponível em: <https://www.statsmodels.org/stable/index.html>. Acesso em: 21 fev. 2025.
- [3] SCIKIT-LEARN. *Scikit-learn Documentation*. Disponível em: <https://scikit-learn.org/stable/>. Acesso em: 21 fev. 2025.
- [4] SEABORN. *Seaborn Documentation*. Disponível em: <https://seaborn.pydata.org/>. Acesso em: 21 fev. 2025.

Obrigado!!!

Contato:

`cicero.hitzschky@alu.ufc.br`