

ANÁLISE QUANTITATIVA DE DADOS EM LINGUÍSTICA

Programa de Pós-Graduação em Linguística

2021.2

Prof. Dr. Ronaldo Lima Jr.

ronaldojr@letras.ufc.br

<http://ronaldolimajr.github.io>

Objetivo: Capacitar pesquisadores em formação em Linguística para o uso adequado de técnicas de inferência estatística para análise de dados linguísticos.

Ementa: Princípios básicos da estatística. Estatística descritiva. Noções de probabilidade. Estatística inferencial. Testes estatísticos. Modelos estatísticos.

Semana	Conteúdo
27/09–03/10	Introdução: o papel da análise quantitativa de dados na linguística; instalando o R e o RStudio
04/10–10/10	Conceitos básicos: população e amostra; tipos de variável (nominal, ordinal, discreta e contínua; dependente e independente); comandos básicos do R para análise quantitativa
11/10–17/10	Conceitos de estatística descritiva: medidas de tendência centrais (média, mediana e moda); medidas de dispersão (desvio-padrão, variância e erro padrão)
18/10–24/10	Noções de probabilidade: conceitos, distribuição normal, distribuição binomial, histogramas, valor de p
25/10–31/10	Introdução a Testes de Hipótese: hipótese nula e hipótese alternativa; testes unilaterais e bilaterais; erro de tipo I e erro de tipo II; intervalo de confiança; tamanho do efeito; teste de Shapiro
01/11–07/11	Teste de proporção e tabelas de frequência; Chi-quadrado e gráfico de barras
08/11–14/11	Teste- t , teste de Wilcoxon e gráfico de caixas; ANOVA, teste de Kruskal-Wallis e teste de Tukey
15/11–21/11	Correlação de Pearson e de Spearman e gráficos de dispersão
22/11–28/11	Semana do tira-dúvida e do <i>catching up</i>
29/11–05/12	Modelos estatísticos; regressão linear simples
06/12–12/12	Regressão linear múltipla
13/12–22/12	Interação entre variáveis
23/12–09/01	Regressão logística
10/01–16/01	Recesso
17/01–23/01	Regressão ordinal
24/01–30/01	Modelos hierárquicos (efeitos mistos)
31/01–06/02	Semana do tira-dúvida e do <i>catching up</i>
07/02–13/02	Finalização e entrega da análise de dados

Avaliação:

Exercícios: 60 pts

Análise de dados: 40 pts

Bibliografia:

Garcia, Guilherme D. **Data visualization and analysis in second language research**. NY: Routledge, 2021.

Garcia, Guilherme D. **Introduction to data analysis using R**. Web Tutorial, 2019–2021.

GOdoy, Mahayana Cristina. **Introdução aos modelos lineares mistos para os estudos da linguagem**. PsyArXiv, 2019.

- Godoy, Mahayana Cristina. [Análise e Visualização de Dados quantitativos em Linguística](#). Github.
- Gries, Stefan Th. **Statistics for linguistics with R : A practical introduction**. Berlin: Walter de Gruyter, 2013.
- Larson-Hall, Jenifer. **A guide to doing statistics in second language research using SPSS and R**. Routledge, 2015.
- Levshina, Natalia. **How to do linguistics with R: Data exploration and statistical analysis**. Amsterdam: John Benjamins Publishing Company, 2015.
- Loerts, Hanneke; Lowie, Wander; Seton, Bregtje. **Essential Statistics for Applied Linguistics: Using R Or JASP**. Amsterdam: Macmillan International, Red Globe Press, 2020.
- Oushiro, Livia. [Introdução à Estatística para Linguistas](#) (Version 2.0.3). Zenodo. 2021.
- Sonderegger, Morgan. [Regression modeling for linguistic data](#). OSF. 2021.
- Sonderegger, Morgan; Wagner, Michael; Torreira, Francisco. [Quantitative Methods for Linguistic Data](#). v 1.0 (out/2018), 2018.
- Winter, Bodo. **Statistics for linguists: An introduction using R**. Routledge, 2019.