

# Metodi quantitativi per la ricerca in linguistica applicata

Docente: dott. Jacopo Saturno, Università di Bergamo

## *Date*

31 marzo 2023, ore 10 - 13

21 aprile 2023, ore 10 - 13

5 maggio 2023, ore 10 – 13

12 maggio 2023, ore 10 -13

*Aula:* 329, via Necchi 9

## *Descrizione*

Il seminario offre un'introduzione alla manipolazione e analisi statistica di dati quantitativi per la ricerca in linguistica applicata.

I partecipanti saranno accompagnati lungo l'intero percorso dalla formulazione delle domande di ricerca alla pubblicazione dei risultati, passando per le diverse fasi della manipolazione dei dati (formattazione, visualizzazione grafica, analisi statistica inferenziale). A questo scopo si replicheranno due studi sul tema dell'intercomprensione tra lingue romanze, facilmente accessibili anche a studenti senza competenze specifiche.

Laddove possibile, il corso presenta diversi approcci metodologici al medesimo problema. In particolare, nelle lezioni iniziali si mostrerà come affiancare l'interfaccia grafica del foglio di calcolo (es. Excel) alla flessibilità del linguaggio di programmazione R. Nelle fasi più avanzate del corso ci si concentrerà invece su quest'ultimo strumento, indispensabile per l'analisi statistica inferenziale.

## *Contenuti*

1. Fasi della ricerca
  - Domande di ricerca e impostazione del modello statistico
  - Raccolta e organizzazione dei dati
2. Fondamenti del linguaggio R e manipolazione di dati
  - Il linguaggio R: le basi
  - Manipolazione di dati in R: vettori, data frame, funzioni
3. Statistica descrittiva
  - Statistica descrittiva
  - Rappresentazione grafica dei dati
4. Statistica inferenziale
  - Cenni sui GLMM con esempi
  - Presentazione dei risultati

## *Requisiti*

Non è richiesta alcuna conoscenza pregressa di statistica né di programmazione.

I partecipanti sono pregati di installare preventivamente sul proprio PC i seguenti programmi:

R <https://cran.r-project.org>

R studio Desktop <https://posit.co>

Sublime Text <https://www.sublimetext.com>

MS Excel (preferibilmente) o equivalente (es. Libreoffice Calc, OpenOffice Calc)

### *Bibliografia*

Baayen, Harald. 2008. *Analyzing linguistic data: a practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Nicenboim, Bruno & Vasishth, Shravan. 2016. Statistical methods for linguistic research: Foundational Ideas-Part II. *Language and Linguistics Compass* 10(11). 591–613. (doi:10.1111/lnc3.12207)

Vasishth, Shravan & Nicenboim, Bruno. 2016. Statistical Methods for Linguistic Research: Foundational Ideas - Part I.: *Language and Linguistics Compass* 10(8). 349–369. (doi:10.1111/lnc3.12201)

Vasishth, Shravan & Schad, Daniel & Bürki, Audrey & Kliegl, Reinhold. *Linear Mixed Models in Linguistics and Psychology: A Comprehensive Introduction*. ([https://vasishth.github.io/Freq\\_CogSci/](https://vasishth.github.io/Freq_CogSci/))

Wickham, Hadley & Grolemund, Garrett. 2016. *R for data science: import, tidy, transform, visualize, and model data*. First edition. Sebastopol, CA: O'Reilly.

Winter, Bodo. 2020. *Statistics for linguists: an introduction using R*. New York: Routledge, Taylor & Francis Group.

**Jacopo Saturno** insegna “Linguistica Generale” e “Informatica e Dati Linguistici” presso l’Università di Bergamo, in cui si è addottorato nel 2017 con una tesi sulle fasi iniziali dell’acquisizione del polacco L2. I suoi principali interessi di ricerca riguardano l’acquisizione di lingue slave come L2 (specie russo e polacco) e l’intercomprensione in area slava e romanza, con particolare attenzione alle basi empiriche del fenomeno e alle sue applicazioni glottodidattiche e traduttive.