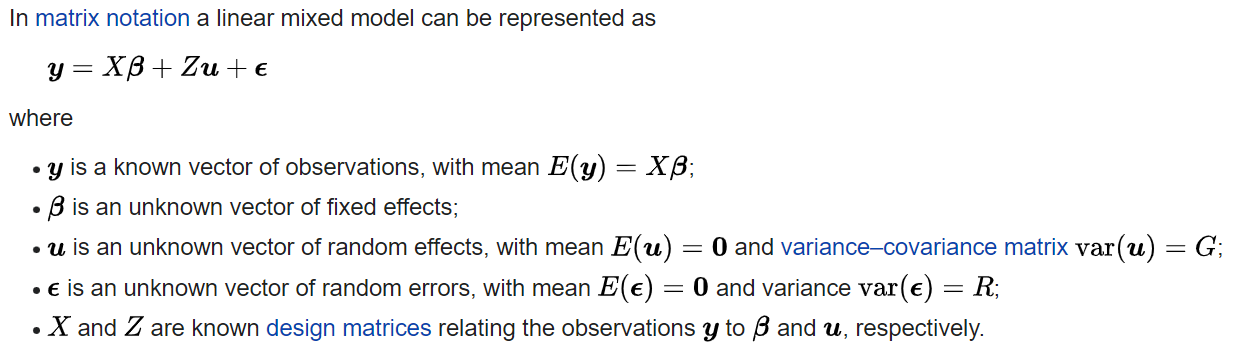
**Mixed-effects Model**

Wikipedia

* Contains fixed effects and random effects.
* Formula (by Wikipedia) **y = X.B + Z.u + e**



where:

* + B (vector) represents fixed effects – unknown

u (vector) represents random effects – unknown

e (vector) represents random errors – unknown

* + X, Z: “design matrices” relating y to B and u
    - In regression analysis
    - Contain values of explanatory variables of a set of objects, which are collected from experiments/measurements.

*[Note]*

*Explanatory variable – on x-axis.*

*Response variable – on y-axis.*

* + - Is a form of writing only?
    - Example of how a design matrix relates parameters in a multiple linear regression model in red:

Table

Description automatically generated

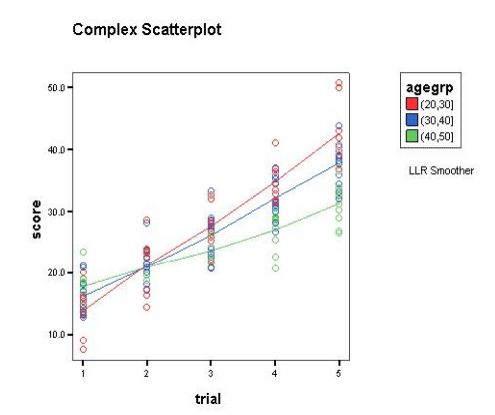
Russell’s resource

* Mixed models are an extremely useful modeling tool for situations in which there is some dependency among observations in the data, where the *correlation typically arises from the observations being clustered in some way* – ok, then a mixed model might be useful
* Packages used:
  + lme4
  + tidyverse: for data processing
  + merTools: optional
  + glmmTMB: optional
  + brms: optional
  + modelr: optional
  + nlme: part of base R, no need for install
* As for applications, ANOVA methods can be seen as special cases of a mixed model. Mixed models can be seen as a first step in expanding one’s tool set beyond the generalized linear model.
* Terminologies:
  + Random effects:
    - Have levels that are not of primary interest, but rather are thought of as a random selection from a much larger set of levels ([Ref1](https://www.stat.cmu.edu/~hseltman/309/Book/chapter15.pdf)).
  + Fixed effects:
    - The typical main effects one would see in a linear regression model in a similar manner ([Ref2](https://m-clark.github.io/mixed-models-with-R/random_intercepts.html)).
    - Have levels that are of primary interest and would be used again if the experiment were repeated ([Ref1](https://www.stat.cmu.edu/~hseltman/309/Book/chapter15.pdf)).
    - Subject effects are almost always random effects, while treatment levels are almost always fixed effects ([Ref1](https://www.stat.cmu.edu/~hseltman/309/Book/chapter15.pdf)).

*[Note]*

*Exploratory Data Analysis (EDA) is an (mostly graphical) approach/philosophy for*

*data analysis that employs a variety of techniques to.*

* (*[*Ref1*](https://www.stat.cmu.edu/~hseltman/309/Book/chapter15.pdf)*)*

“By default commits and tags are marked "Verified" if they are signed with a GPG, SSH, or S/MIME key that was successfully verified. If a commit or tag has a signature that can't be verified by GitHub, we mark the commit or tag "Unverified." In all other cases no verification status is displayed.” – fix?

References

[1] Seltman, H.J., 2012. Experimental design and analysis.

@misc{seltman2012experimental,

title={Experimental design and analysis},

author={Seltman, Howard J},

year={2012},

publisher={Carnegie Mellon University Pittsburgh}

}

[2]

*Backlog Russell’s email on Mixed-Effects:*

*“This includes a mixed effect model, with a fixed effect for year and distance to Sydney/Parramatta, and a random effect for the SA2.*

*Both random intercept only and random slope models are compared.*

*It looks like a mixed effects model is necessary.”*

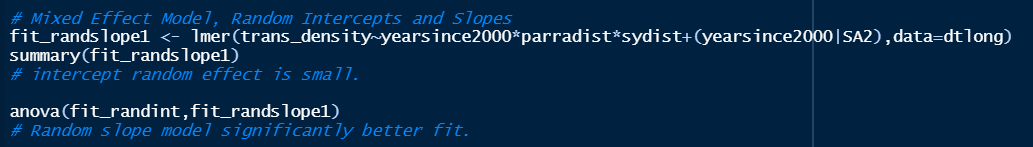
[*Workshop*](https://cloudstor.aarnet.edu.au/plus/s/IC41PpY8ZymLtmU) *and* [*Weblink*](https://m-clark.github.io/mixed-models-with-R/introduction.html)

--

Details from our Project:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| (Population Density) | ~ | (Year) | \* | (Distance from centroids) | “SA2”? |
|  |  | *fixed effect* |  | *fixed effect* | *random effect* |

Note that fixed effects are always going to happen in an experiment. Random effects are factors that influence the observations.



*Russell’s code*