Current status and propsects of R-packages for the design of experiments

by Emi Tanaka

Abstract Re-running an experiment is generally costly and in some cases impossible due to limited resources, so the design of experiment plays a critical role in increasing the quality of experimental data. In this article I describe the current state of the R-packages for the design of experiments through textual analysis and download trends. I discuss also the software design of widely utilised R packages in the field of experimental design and conclude with discussion of some future prospects for the field.

Introduction

The critical role of data collection is well captured in the expression "garbage in, garbage out" – in other words, if the collected data is rubbish then no analysis, however complex it may be, can make something out of it. Methods for data collection can be dichotomised by the type of data collected – namely, experimental or observational – or alternatively, categorised into experimental design or survey design. Survey designs often, although not always, aim to collect observational data whilst experimental designs exclusively center on experimental data. This dichotomisation, to a great extent, is seen in CRAN task views where R-packages in experimental design are in *ExperimentalDesign* and R-packages in survey designs are in *OfficialStatistics*.

In the CRAN task view of *ExperimentalDesign*, there are 113 R packages for experimental design and analysis of data from experiments, henceforth referred to as "DoE packages" in this paper. The sheer quantity and variation in the output experimental design in the R-packages are arguably unmatched with any other programming languages, e.g. in Python (van Rossum, 1995), only a handful of libraries that generate design of experiment exist (namely pyDOE, pyDOE2, dexpy, experimenter and GPdoemd) with limited outputs.

Explorative Analysis

- Not much difference between title and description.
- Just go with description and combine bi & tri.
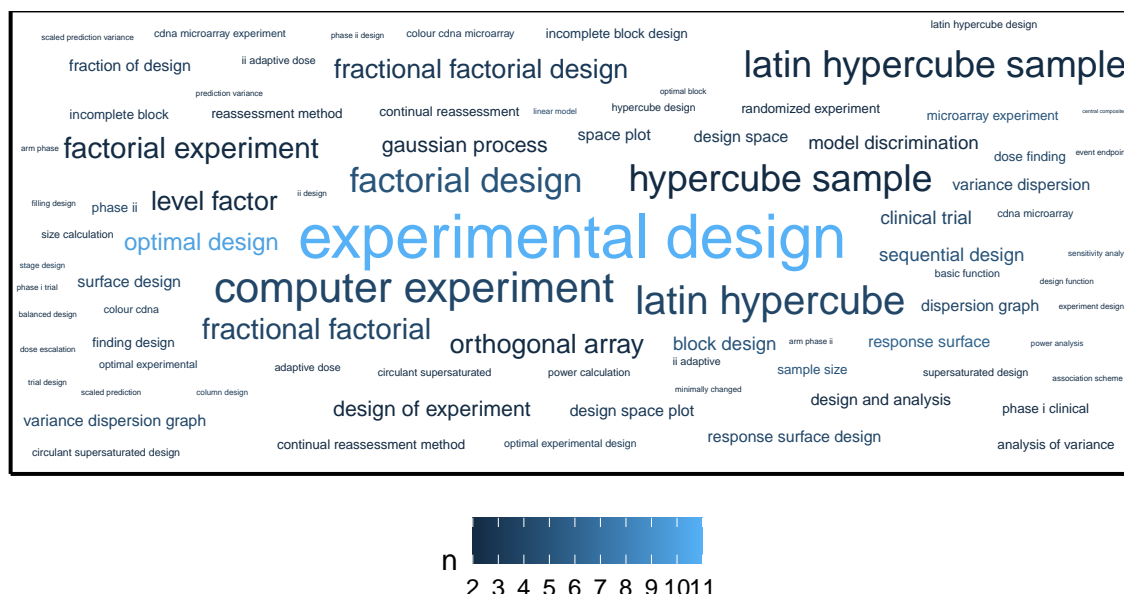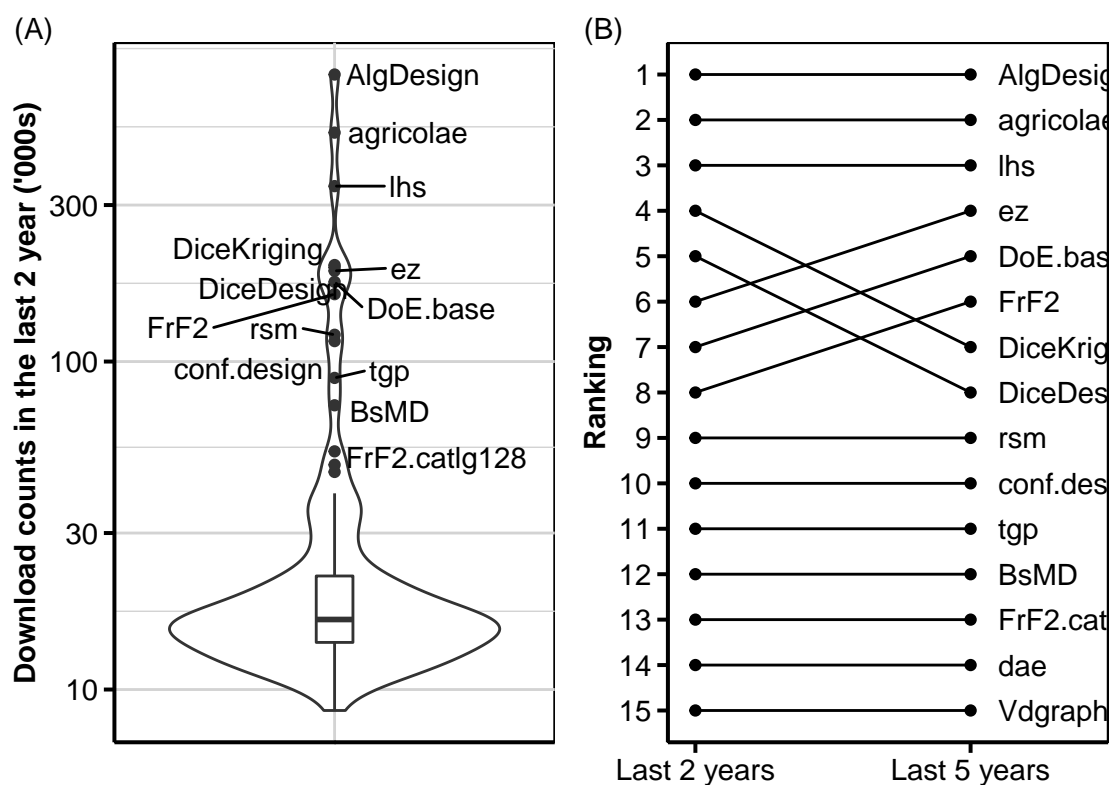
Figure 1: The above graph shows the total download counts from 2019-03-06 to 2021-03-04 of DoE packages.

What are the common types of experimental designs?

How do packages interplay with each other?

Which packages are widely utilised?

Software Design

Propsects and Discussion

word	n	word	n
optimal design	10	experimental design	11
experimental design	8	optimal design	10
clinical trial	5	package provide	7
dose finding	5	graphical user	6
sequential design	5	response surface	6
block design	4	user interface	6
microarray experiment	4	block design	5
response surface	4	contour plot	5
		design based	5
		effect model	5
		factorial design	5
		microarray experiment	5
		mixed effect	5
		provide function	5
		sample size	5
		sequential design	5

Table 1: My tables

agricolae

Helping info to get started

Introductory section which may include references in parentheses (R Core Team, 2012), or cite a reference such as R Core Team (2012) in the text.

Section title in sentence case

Let's check fi this works Figure @ref(fig:Rlogo).

This section may contain a figure such as Figure **??**.

Summary

This file is only a basic article template. For full details of *The R Journal* style and information on how to prepare your article for submission, see the Instructions for Authors.

About this format and the R Journal requirements

`rticles::rjournal_article` will help you build the correct files requirements:

- A R file will be generated automatically using `knitr::purl` - see https://bookdown.org/yihui/rmarkdown-cookbook/purl.html for more information.
- A tex file will be generated from this Rmd file and correctly included in `RJwapper.tex` as expected to build `RJwrapper.pdf`.
- All figure files will be kept in the default rmarkdown `*_files` folder. This happens because `keep_tex = TRUE` by default in `rticles::rjournal_article`
- Only the bib filename is to modifed. An example bib file is included in the template (`RJreferences.bib`) and you will have to name your bib file as the tex, R, and pdf files.

Bibliography

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2012. URL http://www.R-project.org/. ISBN 3-900051-07-0. [p3]

G. van Rossum. Python tutorial. Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, May 1995. [p1]

Emi Tanaka
Monash University
Monash University
Clayton campus, VIC 3800, Australia
http://emitanaka.org/
ORCiD: 0000-0002-1455-259X
emi.tanaka@monash.edu