Revision history for misig repository

Nathan (Nat) Goodman

May 1, 2019

Release 1.20a 2019-04-29

Support posted version of `ovrht` blog post. "Your P-Values are Too Small! And So Are Your Confidence Intervals!"

Major code changes

- Refactored data generation code into multiple files analogous to `R/doc...` files
- Added support for d2ht (aka 'het') distribution

New files

- `ovrht.Rmd`, `ovrht.TRN.Rmd` and related
  - Source files for `ovrht` blog post. `TRN` is modified to generate Word document that can be ported to TRN more easily
- `R/dat.R`
  - Top-level and generic data generation functions, analogous to `R/doc.R`
- `R/dat_mndht.R`, `R/dat_ovrfx.R`, `R/dat_ovrht.R`, `R/dat_readme.R`
  - Data generation function for each document
- `R/doc_mndht.R`
  - Document generation functions for `mndht` blog post. Runs, but still a rough sketch. Expect name to change before post goes public
- `R/doc_ovrht.R`
  - Document generation functions for `ovrht` blog post

Changed files

- `R/datman.R`
  - Add functions for top level data stored in `datadir`
  - Add functions for `hetd` simulations
  - Extend `save_tbl` to save nonflat objects. These aren't tables, of course, but it's sometimes useful to store them in table directories
- `R/doc.R`
  - Add param to `dotbl` for saving nonflat objects
- `R/doc_ovrfx.R`
  - Replace `n.meand`, `d.meand` params by `n.fixd`, `d.fixd`
- `R/doc_readme`
  - Replace `n.meand`, `d.meand` params by `n.fixd`, `d.fixd`
  - More TBD
- `R/init.R`
  - Abandon `n.meand`, `d.meand` params - redundant with `n.fixd`, `d.fixd`
  - Add `outdir` param to drive creation of output directories. Needed now that documents have become so varied
  - Add params for `ovrht` document
  - Add params for saving top level data
  - Remove obsolete params - `fpr.cutoff`, `fnr.cutoff`
  - Stop setting params for doc types that don't use those params
- `R/plot.R`
  - `plothist`: add `add`, color params
  - `plotpvsd`: add support for d2ht (aka 'het') distribution

- plotm: add `spar` param for `spline` smoother
    - `hline`, `vline`: add `cex` param
- `R/run.R`
    - Add `source` commands for new `R/dat.R`, `dat_<document>.R` files
    - Remove obsolete `source` command for `R/dodata.R`
    - Replace calls to `dosim`, `domeand` by more general call to `dodat`
- `R/stats.R`
    - Add many functions for d2ht (aka 'het') distribution
    - Replace obsolete calls to `parent` by `param`
    - Vectorize `ci_d2t` function
- `R/util.R`
    - Add `spar` param to `splinem` smoothing function
    - Add `ignore` param to `nvq` to ignore NULL and non-existant names

Known bugs

1. Functions in the various `doc_` files are clobbering each other. Could also happen with `dat_` files
    - Suggests I've reached the limit of my ersatz object-like programming
    - One solution is to move sourcing of doc specific files into `init`. Be aware this will break my current debugging practice of sourcing `run.R` to get 'em all

Release 1.10 2019-02-19

First real version of README. Renamed and refactored top-level file. See below.

Changed, new, and deleted files

- `README.Rmd`

    - First version of README doc

- Rename top level `R/ovrfx.R` to `R/run.R` and move data generation code to `R/dodata.R`. When writing README, I saw that old scheme made it hard to explain how to run program

- `R/doc.R`

    - Improve management of figure and table labels - numbers, prefixes, suffixes
    - Add `sect_start` function to adjust figure and table labels at start of new section
    - Change `need.init` default in `dodoc` to F - works better with current `init` scheme

- `R/doc_readme.R`

    - Support first version of README doc

- `R/init.R`

    - Update defaults for `readme` doc
    - Change figure and table labels per new scheme in R/doc.R

- `R/plot.R`

    - Fix xlab, ylab defaults when x, y are not defaults

- `R/util.R`

    - Fix bug in loess smoother when data column names are not valid formula symbols

Release 1.01 2019-02-16

Support posted version of `ovrfx` blog post.

Release 1.00 2019-02-14

Support third version of `ovrfx` blog post. Final or near-final version.

Changed files

- `R/doc_ovrfx.R`
  - Remove the 'zoom' figure (old figure 2)
  - Add 'sampling distribution' figure (new figure 3)
  - Use theoretical meand calculations for 'average significant effect size' (figure 4) and numbers in text
- `R/init.R`
  - Unify parameters for fixd simulation (used for empirical meand calculation) and theoretical meand calculation
- `R/ovrfx.R`
  - Turn theoretical meand calculation on again
- `R/plot.R`
  - Minor cleanup of plotpvsd
- `R/stats.R`
  - Fix blunder in d_d2t. Have to scale by sqrt(n/2) to account for difference in x-density
  - Mistake also in repwr/R/stats.R, which I will fix separately
- `css/html_document.css`
  - Add css for blockquote. Rmd's built-in blockquote style sets the font too big
  - Solution from https://stackoverflow.com/questions/44191576. Thanks!

Release 0.90 2019-02-13

Support second version of `ovrfx` blog post. Current title "When You Select Significant Findings, You're Selecting Inflated Estimates". Not final version but close.

Major code changes

- Abandon file caching. Way more trouble than it's worth
- Handle multiple types of simulation
  - `rand` - the original siumulation scheme: randomly choose d and run one simulation per choice
  - `fixd` - more typical simulation scheme: set `d` to certain fixed values and run many simulations for each value
- Handle two ways of calculating mean significant effect sizes, 'meand'
  - `empi` - estimate empirically from simulation data
  - `theo` - theoretical calculation from sampling distribution

Changed files

- `R/datman.R`
  - Add functions to `save` and `load` new file types: `sim_rand`, `sim_fixd`, `meand_empi`, `meand_theo`
  - Replace generic functions, eg, `save_nd`, by ones specific to each file type. Though seemingly less factored, it's way simpler now that I've abandoned caching, esp. since we have so few file types at present
  - Remove `keep` functions. These supported now-abandoned file caching
  - Remove caching related code from `save` and `load` functions
  - Redefine `get` functions to be synonyms for `load`. Previous distinction was needed for now-abandoned file caching
- `R/doc_ovrfx.R`
  - Near complete rewrite for second version of blog post
- `R/init.R`
  - Add params for multiple types of simulation: `sim_rand`, `sim_fixd`

- Add params for calculation of mean significant effect sizes: `meand`
  - Remove params for caching: `load`, `keep`, `clean.cache`
  - Remove code for initializing cache
  - `datadir` no longer has `m` component. Makes no sense with multiple simulation types
- `R/ovrfx.R`
  - Add function for `fixd` simulation and rename function for original `rand` simulation: `dosim_fixd`, `dosim_rand`
  - Add function for empirical calculation of mean significant effect size and rename function for theoretical calculation: `domeand_empi`, `domeand_theo`
  - Incorporate new `R/datman.R` functions for saving files
  - Remove code for getting data from cache if possible
- `R/plot.R`
  - Add function to plot histogram: `plothist`
  - Fiddle with 'extra' line params and code. Removed specialized params like `d.crit` and `d.pop` in favor of the more general `vline` and `hline`
  - Add `x0` param to `pval_legend` function to adjust location. Crude but adequate for immediate needs
  - Add convenience function `d2col` that maps `d` values to colors by combining `d2pval` and `pval2col`

Release 0.50 2019-02-03

Support first real version of `ovrfx` blog post ready for external review. Current title "Significance Testing Makes Bad Choices"

- Rename repo from `effit` to `misig`. Better reflects the content I expect for this repo
- Rename first blog post from `siglo` to `ovrfx` - stands for 'overestimate with fixed effect size'

Release 0.10 2019-01-01

First version. Contains boilerplate files such as this one and the document `sigbi.Rmd` (presently just a stub). Will contain code to support the document when released.

Copyright & License

Copyright (c) 2019 Nathan Goodman

The software is **open source and free**, released under the MIT License. The documentation is **open access**, released under the Creative Commons Attribution 4.0 International License.