

Revision history for misig repository

Nathan (Nat) Goodman

December 1, 2019

Release 1.40 2019-12-01

Support posted version of `confi` blog post. ‘Ladies And Gentlemen, I Introduce To You, “Plausibility Limits”’.

CAVEAT: TRN needs Word documents, but the Rmd-to-Word converter doesn’t handle output scaling of figures. The workaround I devised for the `effit` post didn’t work here: the scaling messed up the in-plot text. Since the scaling didn’t work, I didn’t bother learning how to adjust the file layout for for 1 row, 2 columns. Instead, I used the same full-size figures as in the html and pdf versions, and manually scaled the figures in Word to 40% scale.

Major code changes

- Added `withrows` function to iterate over rows of `cases` data frame. Replaces `withcase`

New files

- `confi.Rmd` and related
 - Source files for `confi` blog post.
- `dat_conf.R`, `doc_conf.R`
 - Data and document generation functions for `confi` blog post

Changed files

- `R/dat.R`
 - Do simulations in chunks. Much more efficient when `m` is big
 - Use `withrows` instead of `withcase` to iterate over cases
- `dat_ovrfx.R`, `doc_ovrfx.R`, `dat_ovrht.R`, `doc_ovrht.R`, `dat_readme.R`, `doc_readme.R`
 - Minor parameter changes mainly for stylistic consistency
- `dat_readmesupp.R`
 - Minor parameter changes mainly for stylistic consistency
- `doc_readmesupp.R`
 - Use `withrows` instead of `withcase` to iterate over cases
 - Minor parameter changes mainly for stylistic consistency
- `R/datman.R`
 - Use `nvq_file` to construct filenames more flexibly. Especially useful for simulation chunks
 - Add function to concatenate simulation chunks
- `R/doc.R`
 - Add support for multiple figures per plot file (`figdev`). Used to generate nicer Word output for TRN
 - Continue replacing deprecated `parent` function by `param`
- `R/init.R`
 - Add `tmpdir`. Used mainly for simulation chunks
 - Only create output directories needed by specific doc
 - Remove params for obsolete `supp` documents
 - Add ability to clean specific data types: subdirs and top level files
- `R/plot.R`
 - Add `legend.label` to top-level plot functions to control label for `pval_legend`. Used to display plausibility legends in `confi` blog post

- Add `plot_empty` function. Was used in now obsolete `ovrhtsupp`. Though not used now, seems worth keeping
- Add `fill_area` function to fill arbitrary areas
- `plotpvsd`: simplify setting of line properties for `confi` blog post
- `plotm`: let `x` be matrix; add `linear` smoothing method
- `vhline`: generalize `hlab`, `vlab` controls for displaying labels
- `pval_legend`: add `label` param to control legend title
- `fill-tail`: continue removing deprecated `parent` function
- `R/source.R`
 - Add `source_ifexists` function to `source` optional files
 - Only `source` doc-specific files if `param` environment setup; fixes bug in *de novo* startup
- `R/stats.R`
 - Change `ci.lo`, `ci.hi` variables to `lo`, `hi` for style
- `R/util.R`
 - Upper-case argument names in functions like `nvq` to avoid conflicts with user-supplied arguments
 - Add `nvq_file` function to more flexibly generate filenames
 - Add `i_pretty` function to pretty print `i` (used for simulation chunks)
 - Add `pretty` wrapper for pretty printing functions
 - Continue trying to create `parent` function that works :). Not sure I have it yet...
 - Add `linear` smoothing method
 - Add `smooth` wrapper for smoothing functions
 - Add `span` parameter to `loessm` smoothing function
 - Add `withrows` function to iterate over `cases` data frame; replaces `withcase`
 - Add rounding functions
 - Add `pick` function to pick `n` items approximately evenly spaced
 - Add `tol` argument to `between` to handle imprecise decimals
 - Add `tback` function: wrapper for `traceback` with arguments set the way I like them

Release 1.30 2019-06-12

First real version of README supplement

Major code changes

- Improve code for figure titles: generic `figtitle` function to generate figure titles instead of doc-specific functions; auto-scaling instead of setting `cex.main`
- Start cleanup of figure/table numbering scheme

New files

- `READMEsupp.Rmd`
 - Source file for README supplement
- `R/dat_readmesupp.R`, `R/doc_readmesupp.R`
 - Data and document generation functions for README supplement
- `R/source.R`
 - Handle problem of name collisions among doc-specific functions through `source_all` function that sources correct doc-specific files for current doc

Changed files

- `README.Rmd`
 - Incorporate new scheme for programmatically inserting figures
 - Add doc codes for README and README supplement in table
 - Change order of simulations in document to match order in code
- `R/dat.R`
 - Clean up setting of `verbose` in data generation functions

- Implement data generation functions for all four statistics of interest, all simulation types, and both data generation models
- Start incorporating new `withcase` function to `lapply` loops; expect more later
- `R/dat_ovrht.R`
 - Minor changes for stylistic consistency
- `R/dat_readme.R`
 - Remove data generation functions not mentioned in README document, now that README supplement exists
- `R/doc.R`
 - Implement ‘section block’ figure/table numbering scheme, in which each section is a figure or table block; a step toward cleaning up figure/table numbering scheme
 - Add `figtitle` function to generate figure titles; replaces previous doc-specific functions
- `R/doc_ovrft.R`, `R/doc_ovrht.R`
 - Remove `cex.main` from function args in favor of auto-scaling
 - Replace doc-specific `title_ovrft`, `title_ovrht` by `figtitle`
- `R/doc_readme.R`
 - Remove `cex.main` from function args in favor of auto-scaling
 - Replace doc-specific `title_readme` by `figtitle`
 - Incorporate new ‘section block’ figure/table numbering scheme
- `R/init.R`
 - Add params for README supplement
 - Add call to new `source_doc` function to source doc-specific files
 - Add `sect` param in `init_doc` function to support new ‘section block’ figure/table numbering scheme
- `R/plot.R`
 - Implement auto-scaling of titles via `cex_title` function
 - Add checks for empty plot inputs in `plotdvst`, `plothist` functions
 - Add checks in `vhline` function for lines and labels outside plot area to avoid writing labels in the margins
 - Replace `T`, `F` by full words `TRUE`, `FALSE` after discovering that R doesn’t always treat these as synonyms
 - Add `vhdigits` arg to top level plot functions and propagate to `vhline` function
- `R/run.R`
 - Move sourcing of other files to `R/source.R` to address name collisions among doc-specific functions
- `R/util.R`
 - Extend `pmatch_choice` function to (optionally) match patterns in the middle of the text; I thought I’d use this in `doc_readmesupp` but didn’t
 - Add `withcase` function to improve `lapply` loops operating on rows of `cases` data frame; still not happy with it; expect further changes

Release 1.21 2019-05-22

Shortened README to contain only two figures per plot type; also added material for heterogeneous effect sizes

Changed and new files

- `README.Rmd`
 - Shorten README to only two figures per plot type
 - Add material for heterogeneous effect size model
 - Change words about how many simulations are run, performance, and so forth
- `R/dat.R`
 - Finish transition from special purpose `save` and `load` (aka `get`) functions for top level data to general purpose `save_data` and `load_data` (aka `get_data`) functions

- Add `dopower_hetd` function to compute power from `hetd` simulation
- Add `dopval_hetd` function to compute p-values from `hetd` simulation
- Add `dopval_d2ht` to compute p-values from `d2ht` distribution
- Add `doci_d2ht` to compute confidence intervals from `d2ht` distribution
- `R/dat_ovrht.R`
 - Replace inline computation of `pval` and `ci` tables with call to new functions in `R/dat.R`
- `R/dat_ovrhtsupp.R`, `R/doc_ovrhtsupp.R`
 - Renamed from `R/dat_mndht.R`, `R/doc_mndht.R`
 - Change `mndht` to `ovrhtsupp` in function names
 - Replace `meand` and `power` params by `hetd` params
- `R/dat_readme.R`
 - Complete rewrite to use functions in `R/dat.R`
- `R/datman.R`
 - Remove obsolete special purpose `save` and `load` (aka `get`) functions for top level data files; code now uses general purpose `save_data` and `load_data` (aka `get_data`) functions
 - Add `base` argument to `save_data` and `load_data` (aka `get_data`) functions to make it easier to specify file names for top level data files
- `R/doc_ovrfix.R`
 - Remove obsolete computation of `meand` from simulation
- `R/doc_ovrht.R`
 - Change `get_data` calls to use new names of data files; see `R/dat_ovrht.R`
 - Change figure and function names from `infl` to `over` for stylistic consistency
- `R/doc_readme.R`
 - Extensive rewrite for new `README.Rmd`
 - Only plot two figures per plot type
 - Add code for heterogeneous effect size model
- `R/init.R`
 - Increase `m.fixd` for `README` from `1e2` to `1e3`
 - Add `hetd params` for `README`
 - Remove params for theoretical mean significant effect size and power; code now uses simulation params
 - Add params for `ovrht pval` and `ci` tables
 - Change `clean` and `sectpx` params for `README` to reflect current usage
 - Edit some comments for stylistic consistency
- `R/plot.R`
 - Improve handling of “extra” vertical and horizontal lines and labels: only render if within plot region
- `R/run.R`
 - Change `source` statements to reflect that `mndht` document renamed to `ovrhtsupp`
- `R/stats.R`
 - Fix typo in comment

Release 1.20 2019-05-01

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`
 - `plohist`: 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-existent 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/ovrfix.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 `ovrfix` blog post.

Release 1.00 2019-02-14

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

Changed files

- `R/doc_ovrfix.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/ovrfix.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 simulation 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](#).