

Package ‘ez.combat’

June 7, 2021

Type Package

Title Easy ComBat Harmonization

Version 1.0.0

Author Timothy R. Koscik [aut, cre, cph]

Maintainer Timothy Koscik <timothy-koscik@uiowa.edu>

Depends R (>= 3.5.0)

Imports stats

Description A dataframe-friendly implementation of ComBat Harmonization which uses an empirical Bayesian framework to remove batch effects.

Johnson WE & Li C (2007) <doi:10.1093/biostatistics/kxj037> ``Adjusting batch effects in microarray expression data using empirical Bayes methods."''

Fortin J-P, Cullen N, Sheline YI, Taylor WD, Aselcioglu I, Cook PA, Adams P, Cooper C, Fava M, McGrath PJ, McInnes M, Phillips ML, Trivedi MH, Weissman MM, & Shinhara RT (2017) <doi:10.1016/j.neuroimage.2017.11.024> ``Harmonization of cortical thickness measurements across scanners and sites."''

Fortin J-P, Parker D, Tun<e7> B, Watanabe T, Elliott MA, Ruparel K, Roalf DR, Satterthwaite TD, Gur RC, Gur RE, Schultz RT, Verma R, & Shinhara RT (2017) <doi:10.1016/j.neuroimage.2017.08.047> ``Harmonization of multi-site diffusion tensor imaging data."''

License Artistic-2.0

Encoding UTF-8

NeedsCompilation no

Repository CRAN

Date/Publication 2021-06-07 07:10:02 UTC

R topics documented:

| | |
|---------------------|----------|
| ez.combat | 2 |
| Index | 4 |

ez.combat

*ComBat Harmonization for Dataframe Objects***Description**

A easy to use function for applying ComBat Harmonization for batch effects on dataframe objects.

Usage

```
ez.combat(df,
          batch.var,
          adjust.var = "all",
          exclude.var = NULL,
          model = NULL,
          output = c("overwrite", "append"),
          error = c("abort", "continue"),
          use.eb = TRUE,
          verbose = FALSE)
```

Arguments

| | |
|-------------|--|
| df | Dataframe containing variables to be harmonized as well as the batch variable to harmonize for. |
| batch.var | A character string or integer indicating the dataframe column containing the batch variable. |
| adjust.var | A character vector indicating the names of the variables in the dataframe to be harmonized. Default value is "all", indicating that all variables should be harmonized (excluding the batch variable and any independent variables given in a model) |
| exclude.var | An optional character vector indicating which variables to exclude. Best used in conjunction with adjust.var = "all" if the number of variables to be excluded is fewer than the number to be included |
| model | a character vector specifying the model to be considered during harmonization. Models do not need to specify the dependent variable, and should follow R's standard modelling syntax, e.g., "~ group * age" |
| output | Select the behaviour of the output dataframe in the output list. "overwrite" will replace the unadjusted variables with adjusted ones, "append" will add the adjusted variables to the end of the output dataframe with ".cb" appended to the variable name. |
| error | Changes behaviour for handling certain input errors. "continue" harmonization will continue if variables have to be excluded, "abort" will discontinue harmonization |
| use.eb | Logical, Perform ComBat Harmonization with empirical bayes or without |
| verbose | Logical, print progress messages to console |

Value

A list containing: `df`: a dataframe with adjusted values `gamma.hat` and `delta.hat`: Estimated location and shift (L/S) parameters before empirical Bayes. `gamma.star` and `delta.star`: Empirical Bayes estimated L/S parameters. `gamma.bar`, `t2`, `a.prior` and `b.prior`: estimated prior distributions parameters.

Author(s)

Timothy R. Kosciak <timothy-kosciak@uiowa.edu>

Examples

```
# Harmonize 'iris' data by removing species effects:
cb <- ez.combat(df = iris, batch.var = "Species")

## plots of raw iris data
plot(iris$Sepal.Length, iris$Petal.Length, col = iris$Species, pch = 16,
     xlab = "Sepal Length", ylab = "Petal Length",
     main = "Flower Characteristics by Species")
legend(x = 4.5, y = 7, legend = levels(iris$Species), col = c(1:3), pch = 16)
boxplot(iris$Sepal.Length ~ iris$Species, notch = TRUE, las = 1,
       xlab = "Species", ylab = "Sepal Length", main = "Sepal Length by Species",
       cex.lab = 1.5, cex.axis = 1.5, cex.main = 2)

## plots after dataset is harmonized across species, i.e., species effects are removed
plot(cb$df$Sepal.Length, cb$df$Petal.Length, col = cb$df$Species,
     pch = 16,
     xlab = "Sepal Length", ylab = "Petal Length",
     main = "Flower Characteristics by Species - Harmonized")
legend(x = 4.5, y = 7, legend = levels(iris$Species), col = c(1:3), pch = 16)
boxplot(cb$df$Sepal.Length ~ cb$df$Species, notch = TRUE, las = 1,
       xlab = "Species", ylab = "Sepal Length",
       main = "Sepal Length by Species - Harmonized",
       cex.lab = 1.5, cex.axis = 1.5, cex.main = 2)
```

Index

ez.combat, [2](#)