

US EPA ARCHIVE DOCUMENT

Version 2.12 (Released October 30, 2006)

Improved the help documentation for functions `psurvey.analysis`, `cat.analysis`, and `cont.analysis`.

Version 2.11 (Released June 30, 2006)

Modified functions in the library to remove use of the global variable named "software.program" that was used to switch between sections of code used in R and sections of code used in S-PLUS.

Version 2.10 (Released March 31, 2006)

Modified function `cont.analysis` to combine the `Pct` and `Tot` output data frames. Also, removed percentile estimates based on the cdf for the total from the `Pct` and `Pct.D` output data frames for `cont.analysis`.

Version 2.9 (Released December 31, 2005)

Modified function `relrisk` so that estimates are calculated correctly when cell counts or margin totals equal zero.

Modified function `cont.analysis` to remove the total estimate from the output data frame named `Tot`.

Version 2.8 (Released September 30, 2005)

Modified all functions in the library that generate warning messages so that the messages are stored in a data frame named `warn.df` rather than being written to the console. Also, created a function named `warnprnt` that prints warning messages stored in the `warn.df` data frame.

Modified functions `psurvey.analysis`, `cat.analysis`, `cont.analysis`, `category.est`, `cdf.est`, `cdf.decon`, and `total.est` to support the addition of variables for support values and size-weights in the design data frame, which is an argument used by the `psurvey.analysis`, `cat.analysis`, and `cont.analysis` functions.

Modified functions `psurvey.analysis`, `cat.analysis`, `cont.analysis`, `category.est`, `cdf.est`, `cdf.decon`, and `total.est` to ensure that a two-stage sampling design has a sufficient number of sites in each stage one sampling unit to allow variance calculation.

Version 2.7 (Released June 30, 2005)

Modified functions `psurvey.analysis`, `cat.analysis`, and `cont.analysis` to check for repeat occurrences of site IDs after subsetting input data frames by the logical variable in the sites data frame. Previously, checking for repeat occurrences of site IDs was conducted prior to subsetting by the logical variable in the sites data frame. Also, modified the `dframe.check` function to remove checking for repeat occurrences of site IDs.

Modified function `total.est` to eliminate the possibility of a negative value for the estimated population variance.

Modified documentation for function `relrisk` to clarify the relationship between arguments to the function that specify levels of the response and stressor categorical variables and calculation of the relative risk estimate.

Modified function `input.check` so that, when the survey design is stratified, arguments `popsize` and `unitsize` are only checked to ensure that values are provided for the stratum codes that actually are used by a population or by a subpopulation of a population. Also, modified functions `cat.analysis` and `cont.analysis` to reflect the changes to function `input.check`.

Version 2.6 (Released March 31, 2005)

Created a function (`relrisk`) that computes the relative risk estimate for a 2x2 table of cell counts defined by a categorical response variable and a categorical explanatory (`stressor`) variable for an unequal probability design.

Modified functions `cat.analysis` and `cont.analysis` so that no analysis is attempted when a subpopulation of a population type contains either no data or only a single value for an indicator.

Version 2.5 (Released December 31, 2004)

Renamed functions in the library to remove the suffix `".fcf"` from function names.

Removed use of the argument named `"type.cat"` (the type of categorical response variable) from the `psurvey.analysis`, `cat.analysis`, `category.est`, and `dframe.check` functions.

Version 2.4 (Released September 30, 2004)

Modified functions `psurvey.analysis.fcf`, `cat.analysis.fcf`, `cont.analysis.fcf`, and `input.check.fcf` to allow known size of the resource and known sum of the size-weights for subpopulations.

Modified the function `uniqueID.fcf` so that site IDs that are stored as a factor are processed without error.

Modified functions `cdf.test.fcf` and `cdfvar.test.fcf` so that a warning message is printed whenever the combined number of values in any of the classes used to define the CDFs is less than five.

Version 2.3 (Released June 30, 2004)

Created a file named `"UserGuide.pdf"` that provides a narrative discussion regarding use of the library.

Modified functions `category.est.fcf`, `cdf.est.fcf`, and `total.fcf` so that requisite values are removed from the vector of values in `popsize` (the known size of the resource) and `unitsize` (the known sum of the size-weights of the resource) whenever removal of missing values from a stratified sample results in strata that no longer contain any values and, therefore, were removed from the analysis.

Version 2.2 (Released March 31, 2004)

In function `dframe.check.fcf`, changed stop messages to warning messages when repeat site IDs are encountered.

Created a function named `uniqueID.fcn` that creates unique site IDs by appending a unique number to each occurrence of a site ID.

Created a function named `vecprint.fcn` that takes an input vector and outputs a character string with line breaks inserted.

In functions `cat.analysis.fcn` and `cont.analysis.fcn`, corrected an error that resulted in loss of the input values for `N.cluster`, `popsiz`, `stagelsiz`, and `unitsiz` in unstratified samples.

In function `cont.analysis.fcn`, added a call to the CDF deconvolution function (`cdf.decon.fcn`) when values are provided for measurement error variance. For that case, two additional data frames named "CDF.D" and "Pct.D" containing the deconvoluted CDF estimates and deconvoluted percentile estimates, respectively, were added to the output list.

In function `input.check.fcn`, added error checking for measurement error variance and variance of measurement error variance.

Version 2.1 (Released December 31, 2003)

Modified the functions for deconvolution of CDFs to allow usage of finite and continuous population correction factors in variance estimation.

Version 2.0 (Released September 30, 2003)

Created a new top-level function named `psurvey.analysis.fcn` that creates an object of class `psurvey.analysis`. The object created by the function is a list that contains all of the information necessary to use the analysis functions in the `psurvey.analysis` library.

Modified functions `cat.analysis.fcn` and `cont.analysis.fcn` to allow input of an object of class `psurvey.analysis`.

Created a function named `dframe.check.fcn` that checks site IDs, the sites data frame, the subpop data frame, the `data.cat` data frame, the `type.cat` vector, and the `data.cont` data frame to assure valid contents. The function is called by functions `psurvey.analysis.fcn`, `cat.analysis.fcn`, and `cont.analysis.fcn`.

Created a function named `input.check.fcn` that checks for compatibility of input values for functions `psurvey.analysis.fcn`, `cat.analysis.fcn`, `cont.analysis.fcn`, `caterory.est.fcn`, `cdf.est.fcn`, `cdf.decon.fcn`, and `total.est.fcn`.

Modified the functions for estimation of categories, estimation of CDFs, and estimation of population totals to allow usage of finite and continuous population correction factors in variance estimation.

For all functions, replaced usage of `R` (the known size of the resource) and `W` (the known sum of the size-weights for the resource) with `popsiz` and `unitsiz`, respectively.

For all functions, replaced usage of `T` and `F` with `TRUE` and `FALSE`, respectively.

Version 1.6 (Released June 30, 2003)

Created a function named `localmean.df.fcn` that calculates the degrees of freedom of the local mean variance-covariance estimator and modified function `cdfvar.test.fcn` to utilize the degrees of freedom for the F distribution versions of the test statistics.

Since library `mva` is loaded by R version 1.7.0 at startup, specific loading of that library was removed from the functions for estimation of categories (`category.est.fcn`), estimation of CDFs (`cdf.est.fcn`), deconvolution of CDFs (`cdf.decon.fcn`), testing differences between CDFs (`cdf.test.fcn`), and estimation of population totals (`total.est.fcn`).

Version 1.5 (Released March 31, 2003)

Modified the variance estimation functions associated with estimation of categories (`category.est.fcn`), estimation of CDFs (`cdf.est.fcn`), deconvolution of CDFs (`cdf.decon.fcn`), and estimation of population totals, etc. (`total.est.fcn`) to correct an error that resulted in underestimation of variance for two-stage sampling designs.

Modified the function for testing differences between CDFs (`cdf.test.fcn`) to include F distribution versions of the test statistics.

The functions for estimation of categories (`category.est.fcn`), estimation of CDFs (`cdf.est.fcn`), deconvolution of CDFs (`cdf.decon.fcn`), and estimation of population totals, etc. (`total.est.fcn`) were modified to allow a stratified sample to be reduced to a single stratum following removal of strata that either contain only missing values of the response variable or contain only a single value of the response variable.

Version 1.4 (Released December 31, 2002)

Added a function (`cdf.decon.fcn`) and support functions for deconvolution of a CDF.

Modified the top-level functions (`cat.analysis.fcn` and `cont.analysis.fcn`) to allow additional levels of a factor when creating the output data frame.

Version 1.3 (Released September 30, 2002)

Added top-level functions that allow input of more than one response variable at a time. One function (`cat.analysis.fcn`) performs categorical data analysis, and the other function (`cont.analysis.fcn`) performs continuous data analysis.

The functions for estimation of categories (`category.est.fcn`), estimation of CDFs (`cdf.est.fcn`), and estimation of population totals, etc. (`total.est.fcn`) and their support functions were modified to accommodate two-stage samples for both stratified and unstratified sampling designs.

Added a function (`adjwgt.fcn`) that adjusts initial survey design weights.

Added a function (`marinus`) that converts coordinates from latitude/longitude to the equidistant, cylindrical map projection.

Modified the category estimation function (`category.est.fcn`) to ensure that, when the known size of the resource is provided, the size estimates for the

categories sum to the known size of the resource.

Version 1.2 (Released June 30, 2002)

The functions for estimation of categories (`category.est.fcn`), estimation of CDFs (`cdf.est.fcn`), and estimation of population totals, etc. (`total.est.fcn`) and their support functions were modified to implement use of the known size of the resource (including know size of each stratum for a stratified sample) in the estimation algorithms.

Standardized output data frame format and naming conventions for all estimation functions and for both stratified and unstratified sampling designs within each estimation function.

Modified the CDF estimation function (`cdf.est.fcn`) so that both the CDF for the proportion (expressed as percent) and the CDF for the total are included in the output.

Combined the functions for category estimation and site status estimation into a single function named `category.est.fcn`.

Combined the functions for estimation of population total, population mean, and population variance into a single function named `total.est.fcn`.

Version 1.1 (Released March 31, 2002)

For a stratified sample, modified implementation of the stratum variable to allow the variable to be composed of character values in addition to numeric values (internally within the functions the stratum variable is converted to a factor).

For a stratified sample, implemented removal of strata containing a single value prior to conducting analysis.

For a stratified sample, eliminated output of estimates for each stratum.

Modified the missing value function to distinguish among numeric, factor, and character data and to remove an expanded set of values for each case.

Version 1.0 (Released January 31, 2002)

This is the original version of the library.