

# Package ‘simhelpers’

September 4, 2024

**Type** Package

**Title** Helper Functions for Simulation Studies

**Version** 0.3.0

**Description**

Calculates performance criteria measures and associated Monte Carlo standard errors for simulation results. Includes functions to help run simulation studies. Our derivation and explanation of formulas and our general simulation workflow is closely aligned with the approach described by Morris, White, and Crowther (2019) <[DOI:10.1002/sim.8086](https://doi.org/10.1002/sim.8086)>.

**URL** <https://meghapsimatrix.github.io/simhelpers/>

**BugReports** <https://github.com/meghapsimatrix/simhelpers/issues>

**Depends** R (>= 2.10)

**License** GPL-3

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.3.2

**SystemRequirements** RStudio

**Imports** stats, furr, tidyr, rstudioapi, Rdpack

**Suggests** dplyr, tibble, purrr, future, knitr, rmarkdown, pkgdown, covr, testthat, kableExtra, ggplot2, broom, boot

**RdMacros** Rdpack

**VignetteBuilder** knitr

**NeedsCompilation** no

**Author** Megha Joshi [aut, cre] (<<https://orcid.org/0000-0001-7936-076X>>),  
James Pustejovsky [aut] (<<https://orcid.org/0000-0003-0591-9465>>)

**Maintainer** Megha Joshi <megha.j456@gmail.com>

**Repository** CRAN

**Date/Publication** 2024-09-04 14:30:02 UTC

## Contents

alpha_res . . . . .	2
bootstrap_CIs . . . . .	3
bootstrap_pvals . . . . .	5
bundle_sim . . . . .	7
calc_absolute . . . . .	9
calc_coverage . . . . .	10
calc_rejection . . . . .	11
calc_relative . . . . .	12
calc_relative_var . . . . .	13
create_skeleton . . . . .	14
evaluate_by_row . . . . .	14
extrapolate_coverage . . . . .	15
extrapolate_rejection . . . . .	18
repeat_and_stack . . . . .	21
Tipton_Pusto . . . . .	22
t_res . . . . .	22
welch_res . . . . .	23
<b>Index</b>	<b>24</b>

---

alpha_res	<i>Cronbach's alpha simulation results</i>
-----------	--

---

### Description

A dataset containing simulation results from estimating Cronbach's alpha and its variance.

### Usage

```
alpha_res
```

### Format

A tibble with 1,000 rows and 3 variables:

**A** estimate of alpha.

**Var\_A** estimate of the variance of alpha.

**true\_param** true alpha used to generate the data.

bootstrap\_CIs

*Calculate one or multiple bootstrap confidence intervals***Description**

Calculate one or multiple bootstrap confidence intervals, given a sample of bootstrap replications.

**Usage**

```
bootstrap_CIs(
  boot_est,
  boot_se = NULL,
  est = NULL,
  se = NULL,
  CI_type = "percentile",
  level = 0.95,
  B_vals = length(boot_est),
  reps = 1L,
  format = "wide",
  seed = NULL
)
```

**Arguments**

<code>boot_est</code>	vector of bootstrap replications of an estimator.
<code>boot_se</code>	vector of estimated standard errors from each bootstrap replication.
<code>est</code>	numeric value of the estimate based on the original sample. Required for <code>CI_type = "normal"</code> , <code>CI_type = "basic"</code> , and <code>CI_type = "student"</code> .
<code>se</code>	numeric value of the estimated standard error based on the original sample. Required for <code>CI_type = "student"</code> .
<code>CI_type</code>	Character string or vector of character strings indicating types of confidence intervals to calculate. Options are "normal", "basic", "student", and "percentile" (the default).
<code>level</code>	numeric value between 0 and 1 for the desired coverage level, with a default of 0.95.
<code>B_vals</code>	vector of sub-sample sizes for which to calculate confidence intervals. Setting <code>B_vals = length(boot_est)</code> (the default) will return bootstrap confidence intervals calculated on the full set of bootstrap replications. For <code>B_vals &lt; length(boot_est)</code> , confidence intervals will be calculated after sub-sampling (without replacement) the bootstrap replications.
<code>reps</code>	integer value for the number of sub-sample confidence intervals to generate when <code>B_vals &lt; length(boot_est)</code> , with a default of <code>reps = 1</code> .

format	character string controlling the format of the output. If format = "wide" (the default), then different types of confidence intervals will be returned in separate columns. If format = "long", then confidence intervals of different types will appear on different rows of dataset. If format = "wide-list", then different types of confidence intervals will be returned in separate columns and the result will be wrapped in an unnamed list.
seed	Single numeric value to which the random number generator seed will be set. Default is NULL, which does not set a seed.

### Details

Confidence intervals are calculated following the methods described in Chapter 5 of Davison and Hinkley (1997). For basic non-parametric bootstraps, the methods are nearly identical to the implementation in `boot.ci` from the `boot` package.

### Value

If format = "wide", the function returns a `data.frame` with `reps` rows per entry of `B_vals`, where each row contains confidence intervals for one sub-sample replication.

If format = "long", the function returns a `data.frame` with one row for each `CI_type`, each replication, and each entry of `B_vals`, where each row contains a single confidence interval for one sub-sample replication.

If format = "wide-list", then the output will be structured as in format = "wide" but will be wrapped in an unnamed list, which makes it easier to store the output in a tibble, and will be assigned the class "bootstrap\_CIs".

### References

Davison, A.C. and Hinkley, D.V. (1997). *\_Bootstrap Methods and Their Application\_*, Chapter 5. Cambridge University Press.

### Examples

```
# generate t-distributed data
N <- 50
mu <- 2
nu <- 5
dat <- mu + rt(N, df = nu)

# create bootstrap replications
f <- \(x) {
  c(
    M = mean(x, trim = 0.1),
    SE = sd(x) / sqrt(length(x))
  )
}

booties <- replicate(399, {
  sample(dat, replace = TRUE, size = N) |>
  f()
})
```

```
  })

  res <- f(dat)

  # calculate bootstrap CIs from full set of bootstrap replicates
  bootstrap_CIs(
    boot_est = booties[1,],
    boot_se = booties[2,],
    est = res[1],
    se = res[2],
    CI_type = c("normal", "basic", "student", "percentile"),
    format = "long"
  )

  # calculate multiple bootstrap CIs using sub-sampling of replicates
  bootstrap_CIs(
    boot_est = booties[1,],
    boot_se = booties[2,],
    est = res[1],
    se = res[2],
    CI_type = c("normal", "basic", "student", "percentile"),
    B_vals = 199,
    reps = 4L,
    format = "long"
  )

  # calculate multiple bootstrap CIs using sub-sampling of replicates,
  # for each of several sub-sample sizes.
  bootstrap_CIs(
    boot_est = booties[1,],
    boot_se = booties[2,],
    est = res[1],
    se = res[2],
    CI_type = c("normal", "basic", "student", "percentile"),
    B_vals = c(49,99,199),
    reps = 4L,
    format = "long"
  )
)
```

---

bootstrap\_pvals

*Calculate one or multiple bootstrap p-values*

---

### **Description**

Calculate one or multiple bootstrap p-values, given a bootstrap sample of test statistics.

### **Usage**

```
bootstrap_pvals(
```

```

boot_stat,
stat,
alternative = "two-sided",
B_vals = length(boot_stat),
reps = 1L,
enlist = FALSE,
seed = NULL
)

```

### Arguments

<code>boot_stat</code>	vector of bootstrap replications of a test statistic.
<code>stat</code>	numeric value of the test statistic based on the original sample.
<code>alternative</code>	a character string specifying the alternative hypothesis, must be one of "two-sided" (the default), "greater" or "less".
<code>B_vals</code>	vector of sub-sample sizes for which to calculate p-values. Setting <code>B_vals = length(boot_stat)</code> (the default) will return a single p-value calculated on the full set of bootstrap replications. For <code>B_vals &lt; length(boot_stat)</code> , p-values will be calculated after sub-sampling (without replacement) the bootstrap replications.
<code>reps</code>	integer value for the number of sub-sample p-values to generate when <code>B_vals &lt; length(boot_stat)</code> , with a default of <code>reps = 1</code> .
<code>enlist</code>	logical indicating whether to wrap the returned values in an unnamed list, with a default of <code>FALSE</code> . Setting <code>enlist = TRUE</code> makes it easier to store the output as a single entry in a tibble.
<code>seed</code>	Single numeric value to which the random number generator seed will be set. Default is <code>NULL</code> , which does not set a seed.

### Details

p-values are calculated by comparing `stat` to the distribution of `boot_stat`, which is taken to represent the null distribution of the test statistic. If `alternative = "two-sided"` (the default), then the p-value is the proportion of the bootstrap sample where the absolute value of the bootstrapped statistic exceeds the absolute value of the original statistic. If `alternative = "greater"`, then the p-value is the proportion of the bootstrap sample where the value of the bootstrapped statistic is larger than the original statistic. If `alternative = "less"`, then the p-value is the proportion of the bootstrap sample where the value of the bootstrapped statistic is less than the original statistic.

### Value

The format of the output depends on several contingencies. If only a single value of `B_vals` is specified and `reps = 1`, then the function returns a vector with a single p-value. If only a single value of `B_vals` is specified but `B_vals < length(boot_stat)` and `reps > 1`, then the function returns a vector p-values, with an entry for each sub-sample replication. If `B_vals` is a vector of multiple values, then the function returns a list with one entry per entry of `B_vals`, where each entry is a vector of length `reps` with entries for each sub-sample replication.

If `enlist = TRUE`, then results will be wrapped in an unnamed list, which makes it easier to store the output in a tibble.

## References

Davison, A.C. and Hinkley, D.V. (1997). *\_Bootstrap Methods and Their Application\_*, Chapter 4. Cambridge University Press.

## Examples

```
# generate data from two distinct populations
dat <- data.frame(
  group = rep(c("A", "B"), c(40, 50)),
  y = c(
    rgamma(40, shape = 7, scale = 2),
    rgamma(50, shape = 3, scale = 4)
  )
)
stat <- t.test(y ~ group, data = dat)$statistic

# create bootstrap replications under the null of no difference
boot_dat <- dat
booties <- replicate(399, {
  boot_dat$group <- sample(dat$group)
  t.test(y ~ group, data = boot_dat)$statistic
})

# calculate bootstrap p-values from full set of bootstrap replicates
bootstrap_pvals(boot_stat = booties, stat = stat)

# calculate multiple bootstrap p-values using sub-sampling of replicates
bootstrap_pvals(
  boot_stat = booties, stat = stat,
  B_vals = 199,
  reps = 4L
)

# calculate multiple bootstrap p-values using sub-sampling of replicates,
# for each of several sub-sample sizes.
bootstrap_pvals(
  boot_stat = booties, stat = stat,
  B_vals = c(49,99,199),
  reps = 4L
)
```

## Description

Bundle a data-generation function, a data-analysis function, and (optionally) a performance summary function into a simulation driver.

**Usage**

```
bundle_sim(
  f_generate,
  f_analyze,
  f_summarize = NULL,
  reps_name = "reps",
  seed_name = "seed",
  summarize_opt_name = "summarize",
  row_bind_reps = TRUE
)
```

**Arguments**

f_generate	function for data-generation
f_analyze	function for data-analysis. The first argument must be the data, in the format generated by f_analyze().
f_summarize	function for calculating performance summaries across replications. The first argument must be the replicated data analysis results. Default is NULL, so that no summary function is used.
reps_name	character string to set the name of the argument for the number of replications, with a default value of "reps".
seed_name	character string to set the name of the argument for the seed option, with a default value of "seed". Set to NULL to remove the argument from the simulation driver.
summarize_opt_name	character string to set the name of the argument for where to apply f_summarize to the simulation results, with a default value of TRUE. Ignored if no f_summarize function is specified. Set to NULL to remove the argument from the simulation driver.
row_bind_reps	logical indicating whether to combine the simulation results into a data frame using rbind(), with a default value of TRUE. If FALSE, then the function will return replications in a list and so f_summarize must be able to take a list as its first argument.

**Value**

A function to repeatedly run the 'f\_generate' and 'f\_analyze' functions and (optionally) apply 'f\_summarize' to the resulting replications.

**Examples**

```
f_G <- rnorm
f_A <- function(x, trim = 0) data.frame(y_bar = mean(x, trim = trim))
f_S <- function(x, calc_sd = FALSE) {
  if (calc_sd) {
    res_SD <- apply(x, 2, sd)
    res <- data.frame(M = colMeans(x), SD = res_SD)
  } else {
```



```
    res <- data.frame(M = colMeans(x))
  }
  res
}

# bundle data-generation and data-analysis functions
sim1 <- bundle_sim(f_generate = f_G, f_analyze = f_A)
args(sim1)
res1 <- sim1(4, n = 70, mean = 0.5, sd = 1, trim = 0.2)
res1

# bundle data-generation, data-analysis, and performance summary functions
sim2 <- bundle_sim(f_generate = f_G, f_analyze = f_A, f_summarize = f_S)
args(sim2)
res2 <- sim2(24, n = 7, mean = 0, sd = 1, trim = 0.2, calc_sd = TRUE)
res2

# bundle data-generation and data-analysis functions, returning results as a list
sim3 <- bundle_sim(f_generate = f_G, f_analyze = f_A, row_bind_reps = FALSE)
args(sim3)
res3 <- sim3(4, n = 70, mean = 0.5, sd = 3, trim = 0.2)
res3
```

---

calc\_absolute

*Calculate absolute performance criteria and MCSE*

---

## Description

Calculates absolute bias, variance, mean squared error (mse) and root mean squared error (rmse). The function also calculates the associated Monte Carlo standard errors.

## Usage

```
calc_absolute(
  data,
  estimates,
  true_param,
  criteria = c("bias", "variance", "stddev", "mse", "rmse"),
  winz = Inf
)
```

## Arguments

data	data frame or tibble containing the simulation results.
estimates	vector or name of column from data containing point estimates.
true_param	vector or name of column from data containing corresponding true parameters.

criteria	character or character vector indicating the performance criteria to be calculated, with possible options "bias", "variance", "stddev", "mse", and "rmse".
winz	numeric value for winsorization constant. If set to a finite value, estimates will be winsorized at the constant multiple of the inter-quartile range below the 25th percentile or above the 75th percentile of the distribution. For instance, setting winz = 3 will truncate estimates that fall below $P25 - 3 * IQR$ or above $P75 + 3 * IQR$ .

**Value**

A tibble containing the number of simulation iterations, performance criteria estimate(s) and the associated MCSE.

**Examples**

```
calc_absolute(data = t_res, estimates = est, true_param = true_param)
```

---

calc_coverage	<i>Calculate confidence interval coverage, width and MCSE</i>
---------------	---

---

**Description**

Calculates confidence interval coverage and width. The function also calculates the associated Monte Carlo standard errors. The confidence interval percentage is based on how you calculated the lower and upper bounds.

**Usage**

```
calc_coverage(
  data,
  lower_bound,
  upper_bound,
  true_param,
  criteria = c("coverage", "width"),
  winz = Inf
)
```

**Arguments**

data	data frame or tibble containing the simulation results.
lower_bound	vector or name of column from data containing lower bounds of confidence intervals.
upper_bound	vector or name of column from data containing upper bounds of confidence intervals.
true_param	vector or name of column from data containing corresponding true parameters.

criteria	character or character vector indicating the performance criteria to be calculated, with possible options "coverage" and "width".
winz	numeric value for winsorization constant. If set to a finite value, estimates will be winsorized at the constant multiple of the inter-quartile range below the 25th percentile or above the 75th percentile of the distribution. For instance, setting winz = 3 will truncate estimates that fall below $P25 - 3 * IQR$ or above $P75 + 3 * IQR$ .

**Value**

A tibble containing the number of simulation iterations, performance criteria estimate(s) and the associated MCSE.

**Examples**

```
calc_coverage(data = t_res, lower_bound = lower_bound,
              upper_bound = upper_bound, true_param = true_param)
```

---

 calc\_rejection

*Calculate rejection rate and MCSE*


---

**Description**

Calculates rejection rate. The function also calculates the associated Monte Carlo standard error.

**Usage**

```
calc_rejection(data, p_values, alpha = 0.05, format = "wide")
```

**Arguments**

data	data frame or tibble containing the simulation results.
p_values	vector or name of column from data containing p-values.
alpha	scalar or vector indicating the nominal alpha level(s). Default value is set to the conventional .05.
format	option "wide" (the default) will produce a tibble with one row, with separate variables for each specified alpha. Option "long" will produce a tibble with one row per specified alpha.

**Value**

A tibble containing the number of simulation iterations, performance criteria estimate and the associated MCSE.

**Examples**

```
calc_rejection(data = t_res, p_values = p_val)
```

---

 calc\_relative

*Calculate relative performance criteria and MCSE*


---

**Description**

Calculates relative bias, mean squared error (relative mse), and root mean squared error (relative rmse). The function also calculates the associated Monte Carlo standard errors.

**Usage**

```
calc_relative(
  data,
  estimates,
  true_param,
  criteria = c("relative bias", "relative mse", "relative rmse"),
  winz = Inf
)
```

**Arguments**

data	data frame or tibble containing the simulation results.
estimates	vector or name of column from data containing point estimates.
true_param	vector or name of column from data containing corresponding true parameters.
criteria	character or character vector indicating the performance criteria to be calculated, with possible options "relative bias", "relative mse", and "relative rmse".
winz	numeric value for winsorization constant. If set to a finite value, estimates will be winsorized at the constant multiple of the inter-quartile range below the 25th percentile or above the 75th percentile of the distribution. For instance, setting winz = 3 will truncate estimates that fall below $P25 - 3 * IQR$ or above $P75 + 3 * IQR$ .

**Value**

A tibble containing the number of simulation iterations, performance criteria estimate(s) and the associated MCSE.

**Examples**

```
calc_relative(data = t_res, estimates = est, true_param = true_param)
```

---

calc_relative_var	<i>Calculate jack-knife Monte Carlo SE for variance estimators</i>
-------------------	--

---

### Description

Calculates relative bias, mean squared error (relative mse), and root mean squared error (relative rmse) of variance estimators. The function also calculates the associated jack-knife Monte Carlo standard errors.

### Usage

```
calc_relative_var(
  data,
  estimates,
  var_estimates,
  criteria = c("relative bias", "relative mse", "relative rmse"),
  winz = Inf,
  var_winz = winz
)
```

### Arguments

data	data frame or tibble containing the simulation results.
estimates	vector or name of column from data containing point estimates.
var_estimates	vector or name of column from data containing variance estimates for point estimator in estimates.
criteria	character or character vector indicating the performance criteria to be calculated, with possible options "relative bias", "relative mse", and "relative rmse".
winz	numeric value for winsorization constant. If set to a finite value, estimates will be winsorized at the constant multiple of the inter-quartile range below the 25th percentile or above the 75th percentile of the distribution. For instance, setting winz = 3 will truncate estimates that fall below $P25 - 3 * IQR$ or above $P75 + 3 * IQR$ .
var_winz	numeric value for winsorization constant for the variance estimates. If set to a finite value, variance estimates will be winsorized at the constant multiple of the inter-quartile range below the 25th percentile or above the 75th percentile of the distribution. For instance, setting var_winz = 3 will truncate variance estimates that fall below $P25 - 3 * IQR$ or above $P75 + 3 * IQR$ . By default var_winz is set to the same constant as winsorize.

### Value

A tibble containing the number of simulation iterations, performance criteria estimate(s) and the associated MCSE.

**Examples**

```
calc_relative_var(data = alpha_res, estimates = A, var_estimates = Var_A)
```

---

```
create_skeleton      Open a simulation skeleton
```

---

**Description**

Creates and opens a .R file containing a skeleton for writing a Monte Carlo simulation study.

**Usage**

```
create_skeleton()
```

**Examples**

```
## Not run:
create_skeleton()

## End(Not run)
```

---

```
evaluate_by_row      Evaluate a simulation function on each row of a data frame or tibble
```

---

**Description**

Evaluates a simulation function on each row of a data frame or tibble containing parameter values. Returns a single tibble with parameters and simulation results. The function uses `furrr::future_pmap`, which allows for easy parallelization.

**Usage**

```
evaluate_by_row(
  params,
  sim_function,
  ...,
  results_name = ".results",
  .progress = FALSE,
  .options = furrr::furrr_options(),
  system_time = TRUE
)
```

**Arguments**

params	data frame or tibble containing simulation parameter values. Each row should represent a separate set of parameter values.
sim_function	function to be evaluated, with argument names matching the variable names in params. The function must return a <code>data.frame</code> , <code>tibble</code> , or vector.
...	additional arguments passed to <code>sim_function</code> .
results_name	character string to set the name of the column storing the results of the simulation. Default is <code>".results"</code> .
.progress	A single logical. Should a progress bar be displayed? Only works with <code>multisession</code> , <code>multicore</code> , and <code>multiprocess</code> futures. Note that if a <code>multicore/multisession</code> future falls back to sequential, then a progress bar will not be displayed. <b>Warning:</b> The <code>.progress</code> argument will be deprecated and removed in a future version of <code>furrr</code> in favor of using the more robust <code>progressr</code> package.
.options	The future specific options to use with the workers. This must be the result from a call to <code>furrr_options()</code> .
system_time	logical indicating whether to print computation time. TRUE by default.

**Value**

A tibble containing parameter values and simulation results.

**Examples**

```
df <- data.frame(
  n = 3:5,
  lambda = seq(8, 16, 4)
)

evaluate_by_row(df, rpois)
```

---

`extrapolate_coverage` *Extrapolate coverage and width using sub-sampled bootstrap confidence intervals.*

---

**Description**

Given a set of bootstrap confidence intervals calculated across sub-samples with different numbers of replications, extrapolates confidence interval coverage and width of bootstrap confidence intervals to a specified (larger) number of bootstraps. The function also calculates the associated Monte Carlo standard errors. The confidence interval percentage is based on how you calculated the lower and upper bounds.

**Usage**

```

extrapolate_coverage(
  data,
  CI_subsamples,
  true_param,
  B_target = Inf,
  criteria = c("coverage", "width"),
  winz = Inf,
  nested = FALSE,
  format = "wide",
  width_trim = 0,
  cover_na_val = NA,
  width_na_val = NA
)

```

**Arguments**

<code>data</code>	data frame or tibble containing the simulation results.
<code>CI_subsamples</code>	list or name of column from data containing list of confidence intervals calculated based on sub-samples with different numbers of replications.
<code>true_param</code>	vector or name of column from data containing corresponding true parameters.
<code>B_target</code>	number of bootstrap replications to which the criteria should be extrapolated, with a default of $B = \text{Inf}$ .
<code>criteria</code>	character or character vector indicating the performance criteria to be calculated, with possible options "coverage" and "width".
<code>winz</code>	numeric value for winsorization constant. If set to a finite value, estimates will be winsorized at the constant multiple of the inter-quartile range below the 25th percentile or above the 75th percentile of the distribution. For instance, setting $\text{winz} = 3$ will truncate estimates that fall below $P25 - 3 * \text{IQR}$ or above $P75 + 3 * \text{IQR}$ .
<code>nested</code>	logical value controlling the format of the output. If FALSE (the default), then the results will be returned as a data frame with rows for each distinct number of bootstraps. If TRUE, then the results will be returned as a data frame with a single row, with each performance criterion containing a nested data frame.
<code>format</code>	character string controlling the format of the output when <code>CI_subsamples</code> has results for more than one type of confidence interval. If "wide" (the default), then each performance criterion will have a separate column for each CI type. If "long", then each performance criterion will be a single variable, with separate rows for each CI type.
<code>width_trim</code>	numeric value specifying the trimming percentage to use when summarizing CI widths across replications from a single set of bootstraps, with a default of 0.0 (i.e., use the regular arithmetic mean).
<code>cover_na_val</code>	numeric value to use for calculating coverage if bootstrap CI end-points are missing. Default is NA.
<code>width_na_val</code>	numeric value to use for calculating width if bootstrap CI end-points are missing. Default is NA.



**Value**

A tibble containing the number of simulation iterations, performance criteria estimate(s) and the associated MCSE.

**References**

Boos DD, Zhang J (2000). “Monte Carlo evaluation of resampling-based hypothesis tests.” *Journal of the American Statistical Association*, **95**(450), 486–492. doi:10.1080/01621459.2000.10474226.

**Examples**

```
dgp <- function(N, mu, nu) {
  mu + rt(N, df = nu)
}

estimator <- function(
  dat,
  B_vals = c(49,59,89,99),
  m = 4,
  trim = 0.1
) {

  # compute estimate and standard error
  N <- length(dat)
  est <- mean(dat, trim = trim)
  se <- sd(dat) / sqrt(N)

  # compute booties
  booties <- replicate(max(B_vals), {
    x <- sample(dat, size = N, replace = TRUE)
    data.frame(
      M = mean(x, trim = trim),
      SE = sd(x) / sqrt(N)
    )
  }, simplify = FALSE) |>
  dplyr::bind_rows()

  # confidence intervals for each B_vals
  CIs <- bootstrap_CIs(
    boot_est = booties$M,
    boot_se = booties$SE,
    est = est,
    se = se,
    CI_type = c("normal", "basic", "student", "percentile"),
    B_vals = B_vals,
    reps = m,
    format = "wide-list"
  )

  res <- data.frame(
    est = est,
```

```

      se = se
    )
    res$CIs <- CIs

    res
  }

#' build a simulation driver function
simulate_bootCIs <- bundle_sim(
  f_generate = dgp,
  f_analyze = estimator
)

boot_results <- simulate_bootCIs(
  reps = 50, N = 20, mu = 2, nu = 3,
  B_vals = seq(49, 199, 50),
)

extrapolate_coverage(
  data = boot_results,
  CI_subsamples = CIs,
  true_param = 2
)

extrapolate_coverage(
  data = boot_results,
  CI_subsamples = CIs,
  true_param = 2,
  B_target = 999,
  format = "long"
)

```

---

extrapolate\_rejection *Extrapolate coverage and width using sub-sampled bootstrap confidence intervals.*

---

### Description

Given a set of bootstrap confidence intervals calculated across sub-samples with different numbers of replications, extrapolates confidence interval coverage and width of bootstrap confidence intervals to a specified (larger) number of bootstraps. The function also calculates the associated Monte Carlo standard errors. The confidence interval percentage is based on how you calculated the lower and upper bounds.

### Usage

```

extrapolate_rejection(
  data,
  pvalue_subsamples,

```

```

  B_target = Inf,
  alpha = 0.05,
  nested = FALSE,
  format = "wide"
)

```

### Arguments

data	data frame or tibble containing the simulation results.
pvalue_subsamples	list or name of column from data containing list of confidence intervals calculated based on sub-samples with different numbers of replications.
B_target	number of bootstrap replications to which the criteria should be extrapolated, with a default of B = Inf.
alpha	scalar or vector indicating the nominal alpha level(s). Default value is set to the conventional .05.
nested	logical value controlling the format of the output. If FALSE (the default), then the results will be returned as a data frame with rows for each distinct number of bootstraps. If TRUE, then the results will be returned as a data frame with a single row, with each performance criterion containing a nested data frame.
format	character string controlling the format of the output when CI_subsamples has results for more than one type of confidence interval. If "wide" (the default), then each performance criterion will have a separate column for each CI type. If "long", then each performance criterion will be a single variable, with separate rows for each CI type.

### Value

A tibble containing the number of simulation iterations, performance criteria estimate(s) and the associated MCSE.

### References

Boos DD, Zhang J (2000). "Monte Carlo evaluation of resampling-based hypothesis tests." *Journal of the American Statistical Association*, **95**(450), 486–492. doi:10.1080/01621459.2000.10474226.

### Examples

```

# function to generate data from two distinct populations
dgp <- function(N_A, N_B, shape_A, scale_A, shape_B, scale_B) {
  data.frame(
    group = rep(c("A", "B"), c(N_A, N_B)),
    y = c(
      rgamma(N_A, shape = shape_A, scale = scale_A),
      rgamma(N_B, shape = shape_B, scale = scale_B)
    )
  )
}

```

```

# function to do a bootstrap t-test
estimator <- function(
  dat,
  B_vals = c(49,59,89,99), # number of booties to evaluate
  pval_reps = 4L
) {
  stat <- t.test(y ~ group, data = dat)$statistic

  # create bootstrap replications under the null of no difference
  boot_dat <- dat
  booties <- replicate(max(B_vals), {
    boot_dat$group <- sample(dat$group)
    t.test(y ~ group, data = boot_dat)$statistic
  })

  # calculate multiple bootstrap p-values using sub-sampling of replicates
  res <- data.frame(stat = stat)

  res$pvalue_subsamples <- bootstrap_pvals(
    boot_stat = booties,
    stat = stat,
    B_vals = B_vals,
    reps = pval_reps,
    enlist = TRUE
  )

  res
}

# create simulation driver
simulate_boot_pvals <- bundle_sim(
  f_generate = dgp,
  f_analyze = estimator
)

# replicate the bootstrap process
x <- simulate_boot_pvals(
  reps = 50L,
  N_A = 20, N_B = 25,
  shape_A = 7, scale_A = 2,
  shape_B = 4, scale_B = 3,
  B_vals = c(49, 99, 149, 199),
  pval_reps = 2L
)

extrapolate_rejection(
  data = x,
  pvalue_subsamples = pvalue_subsamples,
  B_target = 1999,
  alpha = c(.01, .05, .10)
)

extrapolate_rejection(

```

```
data = x,  
pvalue_subsamples = pvalue_subsamples,  
B_target = Inf,  
alpha = c(.01, .05, .10),  
nested = TRUE  
)
```

---

repeat_and_stack	<i>Repeat an expression multiple times and (optionally) stack the results.</i>
------------------	--

---

### Description

Repeat an expression (usually involving random number generation) multiple times. Optionally, organize the results into a `data.frame` that stacks the output from all replications of the expression.

### Usage

```
repeat_and_stack(n, expr, stack = TRUE)
```

### Arguments

<code>n</code>	Number of times to repeat the expression
<code>expr</code>	An expression to be evaluated.
<code>stack</code>	Logical value indicating whether to organize the results into a <code>data.frame</code> .

### Value

If `stack = TRUE` (the default), the results of each evaluation of `expr` will be stacked together using `rbind`. If `stack = FALSE`, a list of length `n` with entries corresponding to the output of each replication of `expr`.

### Examples

```
repeat_and_stack(n = 3, data.frame(x = rexp(2)))  
  
repeat_and_stack(n = 3, data.frame(x = rexp(2)), stack = FALSE)
```

---

Tipton\_Pusto

*Results for Figure 2 of Tipton & Pustejovsky (2015)*


---

### Description

A dataset containing simulation results comparing small sample correction methods for cluster robust variance estimation in meta-analysis.

### Usage

Tipton\_Pusto

### Format

A tibble with 15,300 rows and 8 variables:

**num\_studies** the number of studies included in the meta-analysis.

**r** correlation between outcomes.

**Isq** measure of heterogeneity of true effects.

**contrast** type of contrast that was tested.

**test** small sample method used.

**q** the number of parameters in the hypothesis test.

**rej\_rate** the Type 1 error rate.

**mcse** the Monte Carlo standard error for the estimate of the Type 1 error rate.

### Source

Tipton E, Pustejovsky JE (2015). “Small-Sample Adjustments for Tests of Moderators and Model Fit Using Robust Variance Estimation in Meta-Regression.” *Journal of Educational and Behavioral Statistics*, **40**(6), 604–634. ISSN 1076-9986, 1935-1054, doi:10.3102/1076998615606099, <https://journals.sagepub.com/doi/10.3102/1076998615606099>.

---

t\_res

*t-test simulation results*


---

### Description

A dataset containing simulation results from a study that just runs a t-test.

### Usage

t\_res

**Format**

A tibble with 1,000 rows and 5 variables:

**est** estimate of the mean difference.

**p\_val** p-value from the t-test.

**lower\_bound** lower bound of the confidence interval.

**upper\_bound** upper bound of the confidence interval.

**true\_param** true mean difference used to generate the data.

---

welch\_res

*Welch t-test simulation results*

---

**Description**

A dataset containing simulation results from a study comparing Welch t-test to the conventional t-test.

**Usage**

welch\_res

**Format**

A tibble with 16,000 rows and 11 variables:

**n1** sample size for Group 1.

**n2** sample size for Group 2.

**mean\_diff** true difference in means of two groups used to generate the data.

**iterations** number of iterations.

**seed** seed used to generate data.

**method** indicates whether Welch or conventional t-test was used.

**est** estimate of the mean difference.

**var** variance of the estimate.

**p\_val** p-value from the t-test.

**lower\_bound** lower bound of the confidence interval.

**upper\_bound** upper bound of the confidence interval.

# Index

## \* datasets

alpha\_res, [2](#)

t\_res, [22](#)

Tipton\_Pusto, [22](#)

welch\_res, [23](#)

alpha\_res, [2](#)

boot.ci, [4](#)

bootstrap\_CIs, [3](#)

bootstrap\_pvals, [5](#)

bundle\_sim, [7](#)

calc\_absolute, [9](#)

calc\_coverage, [10](#)

calc\_rejection, [11](#)

calc\_relative, [12](#)

calc\_relative\_var, [13](#)

create\_skeleton, [14](#)

evaluate\_by\_row, [14](#)

extrapolate\_coverage, [15](#)

extrapolate\_rejection, [18](#)

furrr\_options(), [15](#)

repeat\_and\_stack, [21](#)

t\_res, [22](#)

Tipton\_Pusto, [22](#)

welch\_res, [23](#)