# Package 'cities'

July 22, 2025

**Type** Package

**Title** Clinical Trials with Intercurrent Events Simulator

**Version** 0.1.3

**Maintainer** Ahmad Hakeem Abdul Wahab <`hakeemwahab00@gmail.com`>

**Description** Simulates clinical trials and summarizes causal effects and treatment policy estimands in the presence of intercurrent events in a transparent and intuitive manner.

**URL** `https://hakeemwahabapp.shinyapps.io/cities/`

**License** GPL (>= 3)

**Encoding** UTF-8

**RoxygenNote** 7.2.3

**VignetteBuilder** knitr

**Imports** dplyr, ggplot2, plotly, tidyr, ggthemes

**Suggests** testthat (>= 3.0.0), knitr, rmarkdown

**Config/testthat/edition** 3

**NeedsCompilation** no

**Author** Ahmad Hakeem Abdul Wahab [aut, cre] (ORCID:
<`https://orcid.org/0000-0002-7376-0006`>),
Yongming Qu [aut],
Hege Michiels [aut],
Junxiang Luo [aut],
Run Zhuang [aut],
Dominique McDaniel [aut],
Dong Xi [aut],
Elena Polverejan [aut],
Steven Gilbert [aut],
Stephen Ruberg [aut],
Arman Sabbagh [aut]

**Repository** CRAN

**Date/Publication** 2023-08-08 07:50:10 UTC

# Contents

---

`colSD` *colSD*

---

### Description

Helper function to calculate standard deviation of matrix by columns

### Usage

```
colSD(data_in)
```

### Arguments

| | |
|---|---|
| `data_in` | matrix of numeric values |

### Value

Vector of standard deivations of columns of data_in.

### Examples

```
set.seed(1)
colSD(matrix(rnorm(100), ncol=5))
```

---

`data_generator` *data_generator*

---

### Description

Helper function to simulate single clinical trial

### Usage

```
data_generator(
  n_patient_vector,
  p_loe_max,
  z_l_loe,
  z_u_loe,
  p_ee_max,
  z_l_ee,
  z_u_ee,
  timepoints,
  pacf_list,
  sigma_ar_vec,
  mean_list,
  beta_list,
  p_admin,
  rate_dc_ae,
  prob_ae,
  seed_val,
  reference_id,
  plot_po = FALSE,
  up_good = "Up",
  threshold,
  delta_adjustment_in,
  covariate_df
)
```

### Arguments

| | |
|---|---|
| `n_patient_vector` | Vector of number of patients |
| `p_loe_max` | The maximum probability of discontinuing due to LoE |
| `z_l_loe` | The lower (or left) threshold of the LoE curve |
| `z_u_loe` | The upper (or right) threshold of the LoE curve |
| `p_ee_max` | The maximum probability of discontinuing due to EE |
| `z_l_ee` | The lower (or left) threshold of the EE curve |
| `z_u_ee` | The upper (or right) threshold of the EE curve |
| `timepoints` | Vector of timepoints (e.g. weeks, days, time indices) |

| | |
|---|---|
| `pacf_list` | List of pacf vectors |
| `sigma_ar_vec` | Vector of variances per arm associated with list of pacf vectors |
| `mean_list` | List of vectors of means per arm |
| `beta_list` | List of vectors of beta coefficients per arm. All vectors must have the same length and must be the same as the number of columns for the covariate_df. |
| `p_admin` | Vector of probabilities of discontinuing due to admin reasons |
| `rate_dc_ae` | Vector of probabilities of observing at least one adverse event |
| `prob_ae` | Vector of proportions of discontinuing due to adverse event |
| `seed_val` | Starting seed value |
| `reference_id` | ID for pairwise comparisons, e.g. for three arms, if reference_id=1, then arms 2 and 3 will be compared only to arm 1 |
| `plot_po` | TRUE, if plotting data only. Otherwise, set to FALSE |
| `up_good` | "Up" if higher outcome values indicate better responses |
| `threshold` | Value to dichotomize continuous outcomes on |
| `delta_adjustment_in` | Vector of delta adjustment values or NA if none. E.g. (2,3,1) when reference_id = 1 means no delta adjustment on arm 1 (even though 2 was supplied, but since arm 1 is the reference arm, this will be defaulted to 0 regardless), 3 on arm 2 and 1 on arm 3. |
| `covariate_df` | Matrix or dataframe of covariates. Set NA if using default covariates, which comprises one continuous (standard normal) and binary (bernoulli with prob 0.5) covariates. Rows correspond to the total number of subjects. Order matters. For instance, if you want to simulate a trial with 3 arms, each of size 30,50 and 80, then covariate_df would have 30+50+80 rows such that the first 30 rows are covariates for arm 1, the next 50 rows are covariates for arm 2 and the last 80 rows are covariates for arm 3. |

## Value

List of dataframes of estimands and simulated data, including delta adjusted ones if requested:

| | |
|---|---|
| `estimand_mean` | List of means of the FULL, S_++, S_*+ and PP estimands |
| `estimand_sd` | List of standard deviations of the FULL, S_++, S_*+ and PP estimands |
| `dc_mean_list` | List of proportions of discontinuations |
| `observed_df` | Dataframe of the observed outcomes |
| `po_df` | Dataframe of the potential outcomes |
| `ir_df` | Dataframe of the outcomes that have been adjusted via immediate reference (IR) or delta adjustment (Delta) for treatment policy estimands. The IR outcomes are labelled as ir_data while the delta adjusted outcomes are labelled as delta_data. The delta adjusted outcomes will only be available if the correct inputs for delta_adjustment_in are provided |

## Examples

```
total_data = 3
reference_id = 1
threshold = NA
timepoints = c(0,24,48,72,96,120,144)
IR_display = TRUE
delta_adjustment_in = c(0,1)

n_patient_ctrl = 120
n_patient_expt = 150
n_patient_vector = c(n_patient_ctrl, n_patient_expt)
n_total = sum(n_patient_vector)

mean_control = c(0,0,0,0,0,0,0)
mean_treatment = c(0,0.1,0.2,0.4,0.6,0.8,1)
mean_list = list(mean_control, mean_treatment)

sigma_ar_vec = c(1, 1)
pacf_list = list(c(-0.2, 0.4),
                 c(-0.2, 0.4))

beta_list = list(c(1.25, 1.25),
                 c(1.25, 1.25))
covariate_df = NA

# LoE & EE
up_good = "Up"
p_loe_max = 0.75
z_l_loe = -7
z_u_loe = -1
p_ee_max = 0.1
z_l_ee = 4
z_u_ee = 10

# Admin & AE

p_admin_ctrl = 0.02
p_admin_expt = 0.02
p_admin = c(p_admin_ctrl, p_admin_expt)

prob_ae_ctrl = 0.7
prob_ae_expt = 0.9
prob_ae = c(prob_ae_ctrl, prob_ae_expt)

rate_dc_ae_ctrl = 0.1
rate_dc_ae_expt = 0.1
rate_dc_ae = c(rate_dc_ae_ctrl, rate_dc_ae_expt)

starting_seed_val = 1
static_output = TRUE

mean_out = plot_means(n_patient_vector = n_patient_vector, timepoints = timepoints,
```

```
pacf_list = pacf_list, sigma_ar_vec = sigma_ar_vec, mean_list = mean_list,
beta_list = beta_list, reference_id = reference_id, seed_val = starting_seed_val,
total_data = total_data, threshold = threshold, covariate_df = covariate_df,
static_output = static_output)

plot_loe_ee (mean_list = mean_list, ref_grp = reference_id,
stdev_vec = sigma_ar_vec, p_loe_max = p_loe_max, z_l_loe = z_l_loe,
z_u_loe = z_u_loe, p_ee_max = p_ee_max, z_l_ee = z_l_ee, z_u_ee = z_u_ee,
up_good = up_good, greyscale = FALSE, static_output = static_output)

data_out = data_generator(n_patient_vector = n_patient_vector,
p_loe_max = p_loe_max, z_l_loe = z_l_loe, z_u_loe = z_u_loe,
p_ee_max = p_ee_max, z_l_ee = z_l_ee, z_u_ee = z_u_ee, timepoints = timepoints,
pacf_list = pacf_list, sigma_ar_vec = sigma_ar_vec, mean_list = mean_list,
beta_list = beta_list, p_admin = p_admin, rate_dc_ae = rate_dc_ae,
prob_ae = prob_ae, seed_val = starting_seed_val, reference_id = reference_id,
plot_po = FALSE, up_good = up_good, threshold = threshold,
delta_adjustment_in = delta_adjustment_in,
covariate_df = covariate_df)
```

---

## `data_generator_loop` *data_generator_loop*

---

### Description

Simulate multiple or single clinical trial

### Usage

```
data_generator_loop(
  n_patient_vector,
  p_loe_max,
  z_l_loe,
  z_u_loe,
  p_ee_max,
  z_l_ee,
  z_u_ee,
  timepoints,
  pacf_list,
  sigma_ar_vec,
  mean_list,
  beta_list,
  p_admin,
  rate_dc_ae,
  prob_ae,
  seed_val,
  reference_id,
  plot_po = FALSE,
```

```
  up_good,
  threshold,
  total_data,
  delta_adjustment_in,
  covariate_df
)
```

## Arguments

| | |
|---|---|
| `n_patient_vector` | Vector of number of patients |
| `p_loe_max` | The maximum probability of discontinuing due to LoE |
| `z_l_loe` | The lower (or left) threshold of the LoE curve |
| `z_u_loe` | The upper (or right) threshold of the LoE curve |
| `p_ee_max` | The maximum probability of discontinuing due to EE |
| `z_l_ee` | The lower (or left) threshold of the EE curve |
| `z_u_ee` | The upper (or right) threshold of the EE curve |
| `timepoints` | Vector of timepoints (e.g. weeks, days, time indices) |
| `pacf_list` | List of pacf vectors |
| `sigma_ar_vec` | Vector of variances per arm associated with list of pacf vectors |
| `mean_list` | List of vectors of means per arm |
| `beta_list` | List of vectors of beta coefficients per arm. All vectors must have the same length and must be the same as the number of columns for the covariate_df |
| `p_admin` | Vector of probabilities of discontinuing due to admin reasons |
| `rate_dc_ae` | Vector of probabilities of observing at least one adverse event |
| `prob_ae` | Vector of proportions of discontinuing due to adverse event |
| `seed_val` | Starting seed value |
| `reference_id` | ID for pairwise comparisons, e.g. for three arms, if reference_id=1, then arms 2 and 3 will be compared only to arm 1 |
| `plot_po` | TRUE, if plotting data only. Otherwise, set to FALSE |
| `up_good` | "Up" if higher outcome values indicate better responses |
| `threshold` | Value to dichotomize continuous outcomes on |
| `total_data` | Total number of clinical trials to simulate |
| `delta_adjustment_in` | Vector of delta adjustment values or NA if none. E.g. (2,3,1) when reference_id = 1 means no delta adjustment on arm 1 (even though 2 was supplied, but since arm 1 is the reference arm, this will be defaulted to 0 regardless), 3 on arm 2 and 1 on arm 3. |
| `covariate_df` | Matrix or dataframe of covariates. Set NA if using default covariates, which comprises one continuous (standard normal) and binary (bernoulli with prob 0.5) covariates. Rows correspond to the total number of subjects. Order matters. For instance, if you want to simulate a trial with 3 arms, each of size 30,50 and 80, then covariate_df would have 30+50+80 rows such that the first 30 rows are covariates for arm 1, the next 50 rows are covariates for arm 2 and the last 80 rows are covariates for arm 3. |

## Value

List of dataframes of estimands and simulated data, including delta adjusted ones if requested:

| | |
|---|---|
| `estimand_mean` | List of means of the FULL, S_++, S_*+ and PP estimands |
| `estimand_sd` | List of standard deviations of the FULL, S_++, S_*+ and PP estimands |
| `dc_mean_list` | List of proportions of discontinuations |
| `observed_df` | Dataframe of the observed outcomes |
| `po_df` | Dataframe of the potential outcomes |
| `ir_df` | Dataframe of the outcomes that have been adjusted via immediate reference (IR) or delta adjustment (Delta) for treatment policy estimands. The IR outcomes are labelled as ir_data while the delta adjusted outcomes are labelled as delta_data.The delta adjusted outcomes will only be available if the correct inputs for delta_adjustment_in are provided. |

## Examples

```
total_data = 3
reference_id = 1
threshold = NA
timepoints = c(0,24,48,72,96,120,144)
IR_display = TRUE
delta_adjustment_in = c(0,1)

n_patient_ctrl = 120
n_patient_expt = 150
n_patient_vector = c(n_patient_ctrl, n_patient_expt)
n_total = sum(n_patient_vector)

mean_control = c(0,0,0,0,0,0,0)
mean_treatment = c(0,0.1,0.2,0.4,0.6,0.8,1)
mean_list = list(mean_control, mean_treatment)

sigma_ar_vec = c(1, 1)
pacf_list = list(c(-0.2, 0.4),
                 c(-0.2, 0.4))

beta_list = list(c(1.25, 1.25),
                 c(1.25, 1.25))
covariate_df = NA

# LoE & EE
up_good = "Up"
p_loe_max = 0.75
z_l_loe = -7
z_u_loe = -1
p_ee_max = 0.1
z_l_ee = 4
z_u_ee = 10

# Admin & AE
```

```
p_admin_ctrl = 0.02
p_admin_expt = 0.02
p_admin = c(p_admin_ctrl, p_admin_expt)

prob_ae_ctrl = 0.7
prob_ae_expt = 0.9
prob_ae = c(prob_ae_ctrl, prob_ae_expt)

rate_dc_ae_ctrl = 0.1
rate_dc_ae_expt = 0.1
rate_dc_ae = c(rate_dc_ae_ctrl, rate_dc_ae_expt)

starting_seed_val = 1
static_output = TRUE

mean_out = plot_means(n_patient_vector = n_patient_vector, timepoints = timepoints,
pacf_list = pacf_list, sigma_ar_vec = sigma_ar_vec, mean_list = mean_list,
beta_list = beta_list, reference_id = reference_id, seed_val = starting_seed_val,
total_data = total_data, threshold = threshold, covariate_df = covariate_df,
static_output = static_output)

plot_loe_ee (mean_list = mean_list, ref_grp = reference_id,
stdev_vec = sigma_ar_vec, p_loe_max = p_loe_max, z_l_loe = z_l_loe,
z_u_loe = z_u_loe, p_ee_max = p_ee_max, z_l_ee = z_l_ee, z_u_ee = z_u_ee,
up_good = up_good, greyscale = FALSE, static_output = static_output)

data_out = data_generator_loop(n_patient_vector = n_patient_vector,
p_loe_max = p_loe_max, z_l_loe = z_l_loe, z_u_loe = z_u_loe,
p_ee_max = p_ee_max, z_l_ee = z_l_ee, z_u_ee = z_u_ee, timepoints = timepoints,
pacf_list = pacf_list, sigma_ar_vec = sigma_ar_vec, mean_list = mean_list,
beta_list = beta_list, p_admin = p_admin, rate_dc_ae = rate_dc_ae,
prob_ae = prob_ae, seed_val = starting_seed_val, reference_id = reference_id,
plot_po = FALSE, up_good = up_good, threshold = threshold,
total_data = total_data, delta_adjustment_in = delta_adjustment_in,
covariate_df = covariate_df)
```

---

## `line_parameters` *line_parameters*

---

### Description

Helper function that returns slope and intercept for line equation using two points in the cartesian plot: (x1, x2) and (y1, y2)

### Usage

```
line_parameters(x1, y1, x2, y2)
```

## Arguments

| | |
|---|---|
| `x1` | first value of the point (x1, x2) in the cartesian plot |
| `y1` | first value of the point (y1, y2) in the cartesian plot |
| `x2` | second value of the point (x1, x2) in the cartesian plot |
| `y2` | second value of the point (y1, y2) in the cartesian plot |

## Value

Vector of slope and intercept for equation of line.

## Examples

```
line_parameters(1,2,4,2)
```

---

# `pacf_vec_to_acf` *pacf_vec_to_acf*

## Description

Generate correlation matrix from partial autocorrelations

## Usage

```
pacf_vec_to_acf(pacf_vec, n_repeat)
```

## Arguments

| | |
|---|---|
| `pacf_vec` | Vector of partial autocorrelations |
| `n_repeat` | number of repeat measures (must be longer than length of pacf_vec) |

## Value

Correlation matrix from partial autocorrelations.

## Examples

```
pacf_vec_to_acf(c(0.5, -0.1), 5)
```

---

`plot_dc` *plot_dc*

---

## Description

Plots the discontinuation rates by timepoints

## Usage

```
plot_dc(
  data_out,
  total_data,
  timepoints,
  normal_output = TRUE,
  static_output = FALSE,
  greyscale = FALSE
)
```

## Arguments

| | |
|---|---|
| `data_out` | The output from data_generator_loop() |
| `total_data` | Total number of clinical trials to simulate |
| `timepoints` | Vector of timepoints (e.g. weeks, days, time indices) |
| `normal_output` | TRUE if both plots and numeric values of estimands are requested. FALSE if only plots are requested |
| `static_output` | TRUE if static mode requested and FALSE if dynamic plot is requested |
| `greyscale` | TRUE if greyscale requested and FALSE for color |

## Value

Plot and dataframe of proportion of discontinuations.

## Examples

```
total_data = 3
reference_id = 1
threshold = NA
timepoints = c(0,24,48,72,96,120,144)
IR_display = TRUE
delta_adjustment_in = c(0,1)
n_patient_ctrl = 120
n_patient_expt = 150
n_patient_vector = c(n_patient_ctrl, n_patient_expt)
n_total = sum(n_patient_vector)
mean_control = c(0,0,0,0,0,0,0)
mean_treatment = c(0,0.1,0.2,0.4,0.6,0.8,1)
mean_list = list(mean_control, mean_treatment)
```

```
sigma_ar_vec = c(1, 1)
pacf_list = list(c(-0.2, 0.4),
                 c(-0.2, 0.4))
beta_list = list(c(1.25, 1.25),
                 c(1.25, 1.25))
covariate_df = NA
up_good = "Up"
p_loe_max = 0.75
z_l_loe = -7
z_u_loe = -1
p_ee_max = 0.1
z_l_ee = 4
z_u_ee = 10
p_admin_ctrl = 0.02
p_admin_expt = 0.02
p_admin = c(p_admin_ctrl, p_admin_expt)
prob_ae_ctrl = 0.7
prob_ae_expt = 0.9
prob_ae = c(prob_ae_ctrl, prob_ae_expt)
rate_dc_ae_ctrl = 0.1
rate_dc_ae_expt = 0.1
rate_dc_ae = c(rate_dc_ae_ctrl, rate_dc_ae_expt)
starting_seed_val = 1
static_output = TRUE
data_out = data_generator_loop(n_patient_vector = n_patient_vector,
p_loe_max = p_loe_max, z_l_loe = z_l_loe, z_u_loe = z_u_loe,
p_ee_max = p_ee_max, z_l_ee = z_l_ee, z_u_ee = z_u_ee, timepoints = timepoints,
pacf_list = pacf_list, sigma_ar_vec = sigma_ar_vec, mean_list = mean_list,
beta_list = beta_list, p_admin = p_admin, rate_dc_ae = rate_dc_ae,
prob_ae = prob_ae, seed_val = starting_seed_val, reference_id = reference_id,
plot_po = FALSE, up_good = up_good, threshold = threshold,
total_data = total_data, delta_adjustment_in = delta_adjustment_in,
covariate_df = covariate_df)

estimates_out = plot_estimates(data_out = data_out, total_data = total_data,
timepoints = timepoints, reference_id = reference_id, IR_display = IR_display,
normal_output = TRUE, static_output = static_output)

dc_out = plot_dc(data_out = data_out, total_data = total_data,
timepoints = timepoints, static_output = static_output)
```

---

## `plot_estimates` *plot_estimates*

### Description

Plots the estimates of the estimands

### Usage

```
plot_estimates(
  data_out,
  total_data,
  timepoints,
  reference_id,
  IR_display = TRUE,
  delta_display = TRUE,
  normal_output = TRUE,
  static_output = FALSE,
  greyscale = FALSE
)
```

### Arguments

| | |
|---|---|
| `data_out` | The output from data_generator_loop() |
| `total_data` | Total number of clinical trials to simulate |
| `timepoints` | Vector of timepoints (e.g. weeks, days, time indices) |
| `reference_id` | ID for pairwise comparisons, e.g. for three arms, if reference_id=1, then arms 2 and 3 will be compared only to arm 1 |
| `IR_display` | TRUE if requested to display Immediate Reference estimand. FALSE otherwise |
| `delta_display` | TRUE if requested to display Delta estimand. FALSE otherwise |
| `normal_output` | TRUE if both plots and numeric values of estimands are requested. FALSE if only plots are requested |
| `static_output` | TRUE if static mode requested and FALSE if dynamic plot is requested |
| `greyscale` | TRUE if greyscale requested and FALSE for color |

### Value

Plot and dataframe of estimands.

### Examples

```
total_data = 3
reference_id = 1
threshold = NA
timepoints = c(0,24,48,72,96,120,144)
IR_display = TRUE
delta_adjustment_in = c(0,1)
delta_display = TRUE
n_patient_ctrl = 120
n_patient_expt = 150
n_patient_vector = c(n_patient_ctrl, n_patient_expt)
n_total = sum(n_patient_vector)
mean_control = c(0,0,0,0,0,0,0)
mean_treatment = c(0,0.1,0.2,0.4,0.6,0.8,1)
mean_list = list(mean_control, mean_treatment)
```

```
sigma_ar_vec = c(1, 1)
pacf_list = list(c(-0.2, 0.4),
                c(-0.2, 0.4))
beta_list = list(c(1.25, 1.25),
                c(1.25, 1.25))
covariate_df = NA
up_good = "Up"
p_loe_max = 0.75
z_l_loe = -7
z_u_loe = -1
p_ee_max = 0.1
z_l_ee = 4
z_u_ee = 10
p_admin_ctrl = 0.02
p_admin_expt = 0.02
p_admin = c(p_admin_ctrl, p_admin_expt)
prob_ae_ctrl = 0.7
prob_ae_expt = 0.9
prob_ae = c(prob_ae_ctrl, prob_ae_expt)
rate_dc_ae_ctrl = 0.1
rate_dc_ae_expt = 0.1
rate_dc_ae = c(rate_dc_ae_ctrl, rate_dc_ae_expt)
starting_seed_val = 1
static_output = TRUE
data_out = data_generator_loop(n_patient_vector = n_patient_vector,
p_loe_max = p_loe_max, z_l_loe = z_l_loe, z_u_loe = z_u_loe,
p_ee_max = p_ee_max, z_l_ee = z_l_ee, z_u_ee = z_u_ee, timepoints = timepoints,
pacf_list = pacf_list, sigma_ar_vec = sigma_ar_vec, mean_list = mean_list,
beta_list = beta_list, p_admin = p_admin, rate_dc_ae = rate_dc_ae,
prob_ae = prob_ae, seed_val = starting_seed_val, reference_id = reference_id,
plot_po = FALSE, up_good = up_good, threshold = threshold,
total_data = total_data, delta_adjustment_in = delta_adjustment_in,
covariate_df = covariate_df)

estimates_out = plot_estimates(data_out = data_out, total_data = total_data,
timepoints = timepoints, reference_id = reference_id, IR_display = IR_display,
delta_display = delta_display, normal_output = TRUE, static_output = static_output)
```

---

## `plot_loe_ee` *plot_loe_ee*

---

### Description

Plots the lack of efficacy (LoE) and excess efficacy (EE) graphs

### Usage

```
plot_loe_ee(
  mean_list,
```

```
  ref_grp,
  stdev_vec,
  p_loe_max,
  z_l_loe,
  z_u_loe,
  p_ee_max,
  z_l_ee,
  z_u_ee,
  up_good,
  greyscale,
  static_output = FALSE
)
```

## Arguments

| | |
|---|---|
| `mean_list` | List of vectors of means per arm |
| `ref_grp` | ID for pairwise comparisons, e.g. for three arms, if reference_id=1, then arms 2 and 3 will be compared only to arm 1 |
| `stdev_vec` | Vector of standard deviations per arm. This is used to adjust the x-axis for display |
| `p_loe_max` | The maximum probability of discontinuing due to LoE |
| `z_l_loe` | The lower (or left) threshold of the LoE curve |
| `z_u_loe` | The upper (or right) threshold of the LoE curve |
| `p_ee_max` | The maximum probability of discontinuing due to EE |
| `z_l_ee` | The lower (or left) threshold of the EE curve |
| `z_u_ee` | The upper (or right) threshold of the EE curve |
| `up_good` | "Up" if higher outcome values indicate better responses and "Down" otherwise |
| `greyscale` | TRUE for greyscale setting and FALSE for color setting |
| `static_output` | TRUE, if static and FALSE if dynamic plot is requested |

## Value

The plot for LoE and EE.

## Examples

```
total_data = 3
reference_id = 1
threshold = NA
timepoints = c(0,24,48,72,96,120,144)
IR_display = TRUE
delta_adjustment_in = c(0,1)

n_patient_ctrl = 120
n_patient_expt = 150
n_patient_vector = c(n_patient_ctrl, n_patient_expt)
n_total = sum(n_patient_vector)
```

```
mean_control = c(0,0,0,0,0,0,0)
mean_treatment = c(0,0.1,0.2,0.4,0.6,0.8,1)
mean_list = list(mean_control, mean_treatment)

sigma_ar_vec = c(1, 1)
pacf_list = list(c(-0.2, 0.4),
                 c(-0.2, 0.4))

beta_list = list(c(1.25, 1.25),
                 c(1.25, 1.25))
covariate_df = NA

# LoE & EE
up_good = "Up"
p_loe_max = 0.75
z_l_loe = -7
z_u_loe = -1
p_ee_max = 0.1
z_l_ee = 4
z_u_ee = 10

# Admin & AE

p_admin_ctrl = 0.02
p_admin_expt = 0.02
p_admin = c(p_admin_ctrl, p_admin_expt)

prob_ae_ctrl = 0.7
prob_ae_expt = 0.9
prob_ae = c(prob_ae_ctrl, prob_ae_expt)

rate_dc_ae_ctrl = 0.1
rate_dc_ae_expt = 0.1
rate_dc_ae = c(rate_dc_ae_ctrl, rate_dc_ae_expt)

starting_seed_val = 1
static_output = TRUE

mean_out = plot_means(n_patient_vector = n_patient_vector, timepoints = timepoints,
pacf_list = pacf_list, sigma_ar_vec = sigma_ar_vec, mean_list = mean_list,
beta_list = beta_list, reference_id = reference_id, seed_val = starting_seed_val,
total_data = total_data, threshold = threshold, covariate_df = covariate_df,
static_output = static_output)

plot_loe_ee (mean_list = mean_list, ref_grp = reference_id,
stdev_vec = sigma_ar_vec, p_loe_max = p_loe_max, z_l_loe = z_l_loe,
z_u_loe = z_u_loe, p_ee_max = p_ee_max, z_l_ee = z_l_ee, z_u_ee = z_u_ee,
up_good = up_good, greyscale = FALSE, static_output = static_output)
```

---

## `plot_means` *plot_means*

---

## Description

Plots the means of simulation parameters.

## Usage

```
plot_means(
  n_patient_vector,
  timepoints,
  pacf_list,
  sigma_ar_vec,
  mean_list,
  beta_list,
  reference_id,
  seed_val,
  threshold,
  total_data,
  covariate_df,
  static_output = FALSE
)
```

## Arguments

| | |
|---|---|
| `n_patient_vector` | Vector of number of patients |
| `timepoints` | Vector of timepoints (e.g. weeks, days, time indices) |
| `pacf_list` | List of pacf vectors |
| `sigma_ar_vec` | Vector of variances per arm associated with list of pacf vectors |
| `mean_list` | List of vectors of means per arm |
| `beta_list` | List of vectors of beta coefficients per arm. All vectors must have the same length and must be the same as the number of columns for the covariate_df. |
| `reference_id` | ID for pairwise comparisons, e.g. for three arms, if reference_id=1, then arms 2 and 3 will be compared only to arm 1 |
| `seed_val` | Starting seed value |
| `threshold` | Value to dichotomize continuous outcomes on |
| `total_data` | Total number of clinical trials to simulate |
| `covariate_df` | Matrix or dataframe of covariates. Rows correspond to the total number of subjects. Order matters, For instance, if you want to simulate a trial with 3 arms, each of size 30,50 and 80, then covariate_df would have 30+50+80 rows such that the first 30 rows are covariates for arm 1, the next 50 rows are covariates for arm 2 and the last 80 rows are covariates for arm 3. |
| `static_output` | TRUE, if static and FALSE if dynamic plot is requested |

## Value

The plot of raw means.

## Examples

```
total_data = 3
reference_id = 1
threshold = NA
timepoints = c(0,24,48,72,96,120,144)
IR_display = TRUE
delta_adjustment_in = c(0,1)

n_patient_ctrl = 120
n_patient_expt = 150
n_patient_vector = c(n_patient_ctrl, n_patient_expt)
n_total = sum(n_patient_vector)

mean_control = c(0,0,0,0,0,0,0)
mean_treatment = c(0,0.1,0.2,0.4,0.6,0.8,1)
mean_list = list(mean_control, mean_treatment)

sigma_ar_vec = c(1, 1)
pacf_list = list(c(-0.2, 0.4),
                 c(-0.2, 0.4))

beta_list = list(c(1.25, 1.25),
                 c(1.25, 1.25))
covariate_df = NA

# LoE & EE
up_good = "Up"
p_loe_max = 0.75
z_l_loe = -7
z_u_loe = -1
p_ee_max = 0.1
z_l_ee = 4
z_u_ee = 10

# Admin & AE

p_admin_ctrl = 0.02
p_admin_expt = 0.02
p_admin = c(p_admin_ctrl, p_admin_expt)

prob_ae_ctrl = 0.7
prob_ae_expt = 0.9
prob_ae = c(prob_ae_ctrl, prob_ae_expt)

rate_dc_ae_ctrl = 0.1
rate_dc_ae_expt = 0.1
rate_dc_ae = c(rate_dc_ae_ctrl, rate_dc_ae_expt)

starting_seed_val = 1
static_output = TRUE

mean_out = plot_means(n_patient_vector = n_patient_vector, timepoints = timepoints,
```

```
pacf_list = pacf_list, sigma_ar_vec = sigma_ar_vec, mean_list = mean_list,
beta_list = beta_list, reference_id = reference_id, seed_val = starting_seed_val,
total_data = total_data, threshold = threshold, covariate_df = covariate_df,
static_output = static_output)
```

---

## `p_ae_poisson` *p_ae_poisson*

---

### Description

Helper function that returns probability of discontinuing due to adverse events (AE)

### Usage

```
p_ae_poisson(rate_dc_ae, prob_ae)
```

### Arguments

| | |
|---|---|
| `rate_dc_ae` | Probability of observing at least one AE |
| `prob_ae` | Proportion of discontinuation due to AE |

### Value

Probabilities of discontinuing due to AE.

### Examples

```
p_ae_poisson(c(0.9, 0.8), c(0.1, 0.1))
```

---

## `p_loe_ee_function` *p_loe_ee_function*

---

### Description

Helper function that returns probability of discontinuing due to lack of efficacy (LoE) or excess efficacy (EE) via a piecewise linear function

### Usage

```
p_loe_ee_function(z, p_max, z_l, p_min = 0, z_u, up_good = TRUE)
```

### Arguments

| | |
|---|---|
| `z` | Vector of numeric values, i.e. change from baseline values |
| `p_max` | Maximum probability of discontinuing |
| `z_l` | The lower (or left) threshold of the piecewise linear function |
| `p_min` | Maximum probability of discontinuing (set to 0) |
| `z_u` | The upper (or right) threshold of the piecewise linear function |
| `up_good` | TRUE if higher outcome values indicate better responses |

### Value

Probabilities of discontinuing due to LoE or EE.

### Examples

```
line_parameters(1,2,4,2)
```

---

## `rep_col` *rep_rcol*

### Description

Helper function to repeat a matrix by column

### Usage

```
rep_col(x, n)
```

### Arguments

| | |
|---|---|
| `x` | vector to repeat |
| `n` | number of repetions |

### Value

matrix with vector x repeated n-times by columns.

### Examples

```
set.seed(1)
rep_col(rnorm(5), 5)
```

---

## `rep_row` *rep_row*

---

### Description

Helper function to repeat a matrix by row

### Usage

```
rep_row(x, n)
```

### Arguments

| | |
|---|---|
| `x` | vector to repeat |
| `n` | number of repetions |

### Value

Matrix with vector x repeated n-times by rows.

### Examples

```
set.seed(1)
rep_row(rnorm(5), 5)
```

---

## `simulated_data_output` *simulated_data_output*

---

### Description

Helper function to combine simulated data

### Usage

```
simulated_data_output(
  n_patient_cumsum,
  i,
  first_patient,
  data_in,
  covariate_df,
  timepoints,
  beta_list,
  seed_val,
  potential_outcomes = FALSE,
  observed_indicator = NA
)
```

## Arguments

| | |
|---|---|
| `n_patient_cumsum` | Vector of number of patients |
| `i` | Index for arm |
| `first_patient` | Index for first patient of arm |
| `data_in` | Simulated data from data_generator() |
| `covariate_df` | Matrix or dataframe of covariates. Rows correspond to the total number of subjects. Order matters, For instance, if you want to simulate a trial with 3 arms, each of size 30,50 and 80, then covariate_df would have 30+50+80 rows such that the first 30 rows are covariates for arm 1, the next 50 rows are covariates for arm 2 and the last 80 rows are covariates for arm 3. |
| `timepoints` | Vector of timepoints (e.g. weeks, days, time indices) |
| `beta_list` | List of vectors of beta coefficients per arm. All vectors must have the same length and must be the same as the number of columns for the covariate_df. |
| `seed_val` | Current seed value |
| `potential_outcomes` | TRUE if data to be combined is for potential outcomes, and FALSE otherwise |
| `observed_indicator` | Dataframe containing which subjects/arms/timepoints were observed (necessary for potential outcomes), else default to NA |

## Value

Dataframe of for either potential outcomes, observed outcomes, outcomes with immediate reference assumption or delta adjustment assumption

## Examples

```
n_patient_ctrl = 120
n_patient_expt = 150
n_patient_vector = c(n_patient_ctrl, n_patient_expt)
n_patient_cumsum = cumsum(n_patient_vector)
total_patients = sum(n_patient_vector)
timepoints = c(0,24,48,72,96,120,144)
data_in = matrix(rnorm(length(timepoints)*n_patient_ctrl), ncol = length(timepoints))
i = 1
first_patient = 1
covariate_df = data.frame(continuous  = rnorm(n = total_patients, mean = 0, sd = 1),
binary = rbinom(n = total_patients, size = 1, prob = 0.5))
beta_list = NA
seed_val = 1
potential_outcomes = FALSE
observed_indicator = NA
simulated_data_output(n_patient_cumsum = n_patient_cumsum, i = i,
first_patient = first_patient, data_in = data_in, covariate_df = covariate_df,
timepoints = timepoints, beta_list = beta_list, seed_val = seed_val,
potential_outcomes = FALSE, observed_indicator = NA)
```

# Index