

Package ‘phylosamp’

November 22, 2021

Type Package

Title The Probability of Transmission Given Phylogenetic Linkage

Version 0.1.6

Date 2021-11-20

Maintainer Justin Lessler <jlessler@unc.edu>

Description

Implements novel tools that estimate the probability of true transmission between two cases given phylogenetic linkage and the expected number of true transmission links in a sample. Methods described in Wohl, Giles, and Lessler (2021) <[doi:10.1371/journal.pcbi.1009182](https://doi.org/10.1371/journal.pcbi.1009182)>.

License GPL-2

URL <https://github.com/HopkinsIDD/phylosamp>

BugReports <https://github.com/HopkinsIDD/phylosamp/issues>

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Depends R (>= 2.10), stats

Imports reshape2, ggplot2, cowplot, RColorBrewer

Suggests knitr, rmarkdown, testthat

VignetteBuilder knitr

NeedsCompilation no

Author John R Giles [aut, ctb],
Shirlee Wohl [aut, ctb],
Justin Lessler [aut, cre]

Repository CRAN

Date/Publication 2021-11-22 09:40:02 UTC

R topics documented:

exp_links	2
falsediscoveryrate	3
gen_dists	5
gen_dist_sim	6
get_optim_roc	7
obs_pairs_mtml	8
obs_pairs_mtsl	9
obs_pairs_stsl	10
prob_trans_mtml	11
prob_trans_mtsl	12
prob_trans_stsl	13
samplesize	14
sens_spec_calc	15
sens_spec_roc	16
truediscoveryrate	17
true_pairs	19
true_pairs_mtml	20
true_pairs_mtsl	21
true_pairs_stsl	22
Index	23

exp_links	<i>Calculate expected number of links in a sample</i>
-----------	---

Description

This function calculates the expected number of observed pairs in the sample that are linked by the linkage criteria. The function requires the sensitivity η and specificity χ of the linkage criteria, and sample size M . Assumptions about transmission and linkage (single or multiple) can be specified.

Usage

```
exp_links(eta, chi, rho, M, R = NULL, assumption = "mtml")
```

Arguments

eta	scalar or vector giving the sensitivity of the linkage criteria
chi	scalar or vector giving the specificity of the linkage criteria
rho	scalar or vector giving the proportion of the final outbreak size that is sampled
M	scalar or vector giving the number of cases sampled
R	scalar or vector giving the effective reproductive number of the pathogen (default=NULL)
assumption	a character vector indicating which assumptions about transmission and linkage criteria. Default = 'mtml'. Accepted arguments are:

1. 'stsl' for the single-transmission single-linkage assumption ([prob_trans_stsl](#)).
2. 'mtsl' for the multiple-transmission single-linkage assumption ([prob_trans_mtsl](#)).
3. 'mtml' for the multiple-transmission multiple-linkage assumption ([prob_trans_mtml](#)).

Value

scalar or vector giving the expected number of observed links in the sample

Author(s)

John Giles, Shirlee Wohl, and Justin Lessler

See Also

Other obs_pairs: [obs_pairs_mtml\(\)](#), [obs_pairs_mtsl\(\)](#), [obs_pairs_stsl\(\)](#)

Examples

```
# The simplest case: single-transmission, single-linkage, and perfect sensitivity
exp_links(eta=1, chi=0.9, rho=0.5, M=100, assumption='stsl')

# Multiple-transmission and imperfect sensitivity
exp_links(eta=0.99, chi=0.9, rho=1, M=50, R=1, assumption='mtsl')

# Small outbreak, larger sampling proportion
exp_links(eta=0.99, chi=0.95, rho=1, M=50, R=1, assumption='mtml')

# Large outbreak, small sampling proportion
exp_links(eta=0.99, chi=0.95, rho=0.05, M=1000, R=1, assumption='mtml')
```

falsediscoveryrate *Calculate false discovery rate of a sample*

Description

This function calculates the false discovery rate (proportion of linked pairs that are false positives) in a sample given the sensitivity η and specificity χ of the linkage criteria, and sample size M . Assumptions about transmission and linkage (single or multiple) can be specified.

Usage

```
falsediscoveryrate(eta, chi, rho, M, R = NULL, assumption = "mtml")
```

Arguments

<code>eta</code>	scalar or vector giving the sensitivity of the linkage criteria
<code>chi</code>	scalar or vector giving the specificity of the linkage criteria
<code>rho</code>	scalar or vector giving the proportion of the final outbreak size that is sampled
<code>M</code>	scalar or vector giving the number of cases sampled
<code>R</code>	scalar or vector giving the effective reproductive number of the pathogen (default=NULL)
<code>assumption</code>	a character vector indicating which assumptions about transmission and linkage criteria. Default = 'mtml'. Accepted arguments are: <ol style="list-style-type: none"> 1. 'stsl' for the single-transmission single-linkage assumption (prob_trans_stsl). 2. 'mtsl' for the multiple-transmission single-linkage assumption (prob_trans_mtsl). 3. 'mtml' for the multiple-transmission multiple-linkage assumption (prob_trans_mtml).

Value

scalar or vector giving the true discovery rate

Author(s)

John Giles, Shirlee Wohl, and Justin Lessler

See Also

Other discovery_rate: [truediscoveryrate\(\)](#)

Examples

```
# The simplest case: single-transmission, single-linkage, and perfect sensitivity
falsediscoveryrate(eta=1, chi=0.9, rho=0.5, M=100, assumption='stsl')

# Multiple-transmission and imperfect sensitivity
falsediscoveryrate(eta=0.99, chi=0.9, rho=1, M=50, R=1, assumption='mtsl')

# Small outbreak, larger sampling proportion
falsediscoveryrate(eta=0.99, chi=0.95, rho=1, M=50, R=1, assumption='mtml')

# Large outbreak, small sampling proportion
falsediscoveryrate(eta=0.99, chi=0.95, rho=0.5, M=1000, R=1, assumption='mtml')
```

gen_dists	<i>Calculate genetic distance distribution</i>
-----------	--

Description

Function calculates the distribution of genetic distances in a population of viruses with the given parameters

Usage

```
gen_dists(  
  mut_rate,  
  mean_gens_pdf,  
  max_link_gens = 1,  
  max_gens = NULL,  
  max_dist = NULL  
)
```

Arguments

mut_rate	mean number of mutations per generation, assumed to be poisson distributed
mean_gens_pdf	the density distribution of the mean number of generations between cases; the index of this vector is assumed to be the discrete distance between cases
max_link_gens	the maximum generations of separation for linked pairs
max_gens	the maximum number of generations to consider, if NULL (default) value is set to the highest number of generations in mean_gens_pdf with a non-zero probability
max_dist	the maximum distance to calculate, if NULL (default) value is set to max_gens * 99.9th percentile of mut_rate poisson distribution

Value

a data frame with distances and probabilities

Author(s)

Shirlee Wohl and Justin Lessler

See Also

Other mutrate_functions: [get_optim_roc\(\)](#), [sens_spec_calc\(\)](#), [sens_spec_roc\(\)](#)

Examples

```
# ebola-like pathogen
R <- 1.5
mut_rate <- 1

# use simulated generation distributions from the provided 'gen_dist_sim' data object
data("gen_dist_sim")
mean_gens_pdf <- as.numeric(gen_dist_sim[gen_dist_sim$R == R, -(1:2)])

# get theoretical genetic distance dist based on mutation rate and generation parameters
gen_dists(mut_rate = mut_rate,
           mean_gens_pdf = mean_gens_pdf,
           max_link_gens = 1)
```

gen_dist_sim

Simulations of the genetic distance distribution

Description

This data object contains the genetic distance distributions for 168 values of R between 1.3 and 18. The distributions represent the the average of 1000 simulations for each value, which can be used as a reasonable proxy for the generation distribution for large outbreaks.

Usage

```
gen_dist_sim
```

Format

```
dataframe
```

Author(s)

Shirlee Wohl, John Giles, and Justin Lessler

Examples

```
data(gen_dist_sim)
```

get_optim_roc	<i>Find optimal ROC threshold</i>
---------------	-----------------------------------

Description

This function takes the dataframe output of the 'sens_spec_roc()' function and finds the optimal threshold of sensitivity and specificity by minimizing the distance to the top left corner of the Receiver Operating Characteristic (ROC) curve

Usage

```
get_optim_roc(roc)
```

Arguments

roc a dataframe produced by the 'sens_spec_roc()' function containing the Receiver Operating Characteristic (ROC) curve

Value

vector containing optimal thresholds of sensitivity and specificity

Author(s)

Shirlee Wohl, John Giles, and Justin Lessler

See Also

Other mutrate_functions: [gen_dists\(\)](#), [sens_spec_calc\(\)](#), [sens_spec_roc\(\)](#)

Examples

```
# ebola-like pathogen
R <- 1.5
mut_rate <- 1

# use simulated generation distributions
data("gen_dist_sim")
mean_gens_pdf <- as.numeric(gen_dist_sim[gen_dist_sim$R == R, -(1:2)])

# get theoretical genetic distance dist based on mutation rate and generation parameters
dists <- as.data.frame(gen_dists(mut_rate = mut_rate,
                                mean_gens_pdf = mean_gens_pdf,
                                max_link_gens = 1))

# reshape dataframe for plotting
dists <- reshape2::melt(dists,
                       id.vars = "dist",
                       variable.name = "status",
```

```

value.name = "prob")

# get sensitivity and specificity using the same paramters
roc_calc <- sens_spec_roc(cutoff = 1:(max(dists$dist)-1),
                        mut_rate = mut_rate,
                        mean_gens_pdf = mean_gens_pdf)

# get the optimal value for the ROC plot
optim_point <- get_optim_roc(roc_calc)

```

obs_pairs_mtml	<i>Expected number of observed pairs assuming multiple-transmission and multiple-linkage</i>
----------------	--

Description

This function calculates the expected number of pairs observed in a sample of size M . The multiple-transmission and multiple-linkage method assumes the following:

1. Each case i is, on average, the infector of R cases in the population (N)
2. Each case i is allowed to be linked by the linkage criteria to multiple cases j in the sampled population (M).
3. Linkage events are independent of one another (i.e, linkage of case i to case j has no bearing on linkage of case i to any other sample).

Usage

```
obs_pairs_mtml(chi, eta, rho, M, R)
```

Arguments

chi	scalar or vector giving the specificity of the linkage criteria
eta	scalar or vector giving the sensitivity of the linkage criteria
rho	scalar or vector giving the proportion of the final outbreak size that is sampled
M	scalar or vector giving the number of cases sampled
R	scalar or vector giving the effective reproductive number of the pathogen

Value

scalar or vector giving the expected number of linked pairs observed in the sample

Author(s)

John Giles, Shirlee Wohl and Justin Lessler

See Also

Other obs_pairs: [exp_links\(\)](#), [obs_pairs_mtsl\(\)](#), [obs_pairs_stsl\(\)](#)

Examples

```
obs_pairs_mtml(eta=1, chi=1, rho=0.5, M=100, R=1) # Perfect sensitivity and specificity
obs_pairs_mtml(eta=0.99, chi=0.9, rho=1, M=50, R=1)
obs_pairs_mtml(eta=0.99, chi=0.9, rho=0.5, M=100, R=1)
```

obs_pairs_mtsl	<i>Expected number of observed pairs assuming multiple-transmission and single-linkage</i>
----------------	--

Description

This function calculates the expected number of pairs observed in a sample of size M . The multiple-transmission and single-linkage method assumes the following:

1. Each case i is, on average, the infector of R cases in the population (N)
2. Each case i is allowed to be linked by the linkage criteria to only one other case j in the sampled population (M).

Usage

```
obs_pairs_mtsl(chi, eta, rho, M, R)
```

Arguments

chi	scalar or vector giving the specificity of the linkage criteria
eta	scalar or vector giving the sensitivity of the linkage criteria
rho	scalar or vector giving the proportion of the final outbreak size that is sampled
M	scalar or vector giving the number of cases sampled
R	scalar or vector giving the effective reproductive number of the pathogen

Value

scalar or vector giving the expected number of linked pairs observed in the sample

Author(s)

John Giles, Shirlee Wohl and Justin Lessler

See Also

Other obs_pairs: [exp_links\(\)](#), [obs_pairs_mtml\(\)](#), [obs_pairs_stsl\(\)](#)

Examples

```
obs_pairs_mtsl(eta=1, chi=1, rho=0.5, M=100, R=1) # Perfect sensitivity and specificity
obs_pairs_mtsl(eta=0.99, chi=0.9, rho=1, M=50, R=1)
obs_pairs_mtsl(eta=0.99, chi=0.9, rho=0.5, M=100, R=1)
```

obs_pairs_stsl	<i>Expected number of observed pairs assuming single-transmission and single-linkage</i>
----------------	--

Description

This function calculates the expected number of link pairs observed in a sample of size M . The single-transmission and single-linkage method assumes the following:

1. Each case i is linked by transmission to only one other case j in the population (N).
2. Each case i is linked by the linkage criteria to only one other case j in the sampled population (M).

Usage

```
obs_pairs_stsl(eta, chi, rho, M)
```

Arguments

eta	scalar or vector giving the sensitivity of the linkage criteria
chi	scalar or vector giving the specificity of the linkage criteria
rho	scalar or vector giving the proportion of the final outbreak size that is sampled
M	scalar or vector giving the number of cases sampled

Value

scalar or vector giving the expected number of linked pairs observed in the sample

Author(s)

John Giles, Shirlee Wohl, and Justin Lessler

See Also

Other obs_pairs: [exp_links\(\)](#), [obs_pairs_mtml\(\)](#), [obs_pairs_mtsl\(\)](#)

Examples

```
obs_pairs_stsl(eta=1, chi=1, rho=0.5, M=100) # perfect sensitivity and specificity
obs_pairs_stsl(eta=0.99, chi=0.9, rho=1, M=50)
obs_pairs_stsl(eta=0.99, chi=0.9, rho=0.5, M=100)
```

prob_trans_mtml	<i>Probability of transmission assuming multiple-transmission and multiple-linkage</i>
-----------------	--

Description

This function calculates the probability that two cases are linked by direct transmission given that they have been linked by phylogenetic criteria. The multiple-transmission and multiple-linkage method assumes the following:

1. Each case i is, on average, the infector of R cases in the population (N)
2. Each case i is allowed to be linked by the linkage criteria to multiple cases j in the sampled population (M).
3. Linkage events are independent of one another (i.e, linkage of case i to case j has no bearing on linkage of case i to any other sample).

Usage

```
prob_trans_mtml(eta, chi, rho, M, R)
```

Arguments

eta	scalar or vector giving the sensitivity of the linkage criteria
chi	scalar or vector giving the specificity of the linkage criteria
rho	scalar or vector giving the proportion of the final outbreak size that is sampled
M	scalar or vector giving the number of cases sampled
R	scalar or vector giving the effective reproductive number of the pathogen

Value

scalar or vector giving the probability of transmission between two cases given linkage by phylogenetic criteria

Author(s)

John Giles, Shirlee Wohl, and Justin Lessler

See Also

Other prob_trans: [prob_trans_mtsl\(\)](#), [prob_trans_stsl\(\)](#)

Examples

```

prob_trans_mtml(eta=1, chi=1, rho=0.5, M=100, R=1) # Perfect sensitivity and specificity
prob_trans_mtml(eta=0.99, chi=0.9, rho=1, M=50, R=1)
prob_trans_mtml(eta=0.99, chi=0.9, rho=0.5, M=100, R=1)

```

prob_trans_mtsl	<i>Probability of transmission assuming multiple-transmission and single-linkage</i>
-----------------	--

Description

This function calculates the probability that two cases are linked by direct transmission given that they have been linked by phylogenetic criteria. The multiple-transmission and single-linkage method assumes the following:

1. Each case i is, on average, the infector of R cases in the population (N)
2. Each case i is allowed to be linked by the linkage criteria to only one other case j in the sampled population (M).

Usage

```
prob_trans_mtsl(chi, eta, rho, M, R)
```

Arguments

chi	scalar or vector giving the specificity of the linkage criteria
eta	scalar or vector giving the sensitivity of the linkage criteria
rho	scalar or vector giving the proportion of the final outbreak size that is sampled
M	scalar or vector giving the number of cases sampled
R	scalar or vector giving the effective reproductive number of the pathogen

Value

scalar or vector giving the probability of transmission between two cases given linkage by phylogenetic criteria

Author(s)

John Giles, Shirlee Wohl, and Justin Lessler

See Also

Other prob_trans: [prob_trans_mtml\(\)](#), [prob_trans_stsl\(\)](#)

Examples

```
prob_trans_mtsl(eta=1, chi=1, rho=0.5, M=100, R=1) # Perfect sensitivity and specificity
prob_trans_mtsl(eta=0.99, chi=0.9, rho=1, M=50, R=1)
prob_trans_mtsl(eta=0.99, chi=0.9, rho=0.5, M=100, R=1)
```

prob_trans_stsl	<i>Probability of transmission assuming single-transmission and single-linkage</i>
-----------------	--

Description

This function calculates the probability that two cases are linked by direct transmission given that they have been linked by phylogenetic criteria. The single-transmission and single-linkage method assumes the following:

1. Each case i is linked by transmission to only one other case j in the population (N).
2. Each case i is linked by the linkage criteria to only one other case j in the sampled population (M).

Usage

```
prob_trans_stsl(eta, chi, rho, M)
```

Arguments

eta	scalar or vector giving the sensitivity of the linkage criteria
chi	scalar or vector giving the specificity of the linkage criteria
rho	scalar or vector giving the proportion of the final outbreak size that is sampled
M	scalar or vector giving the number of cases sampled

Details

For perfect sensitivity, set eta = 1.

Value

scalar or vector giving the probability of transmission between two cases given linkage by phylogenetic criteria

Author(s)

John Giles, Shirlee Wohl, and Justin Lessler

See Also

Other prob_trans: [prob_trans_mtml\(\)](#), [prob_trans_mtsl\(\)](#)

Examples

```

prob_trans_stsl(eta=1, chi=1, rho=0.2, M=100) # perfect sensitivity and specificity
prob_trans_stsl(eta=1, chi=0.95, rho=0.2, M=100) # perfect sensitivity only
prob_trans_stsl(eta=0.99, chi=0.95, rho=0.9, M=50)
prob_trans_stsl(eta=0.99, chi=0.95, rho=0.05, M=100)

```

samplesize

Calculate sample size

Description

This function calculates the sample size needed to obtain at least a defined false discovery rate given a final outbreak size N .

Usage

```
samplesize(eta, chi, N, R = NULL, phi, min_pairs = 1, assumption = "mtml")
```

Arguments

eta	scalar or vector giving the sensitivity of the linkage criteria
chi	scalar or vector giving the specificity of the linkage criteria
N	scalar or vector giving the final outbreak size
R	scalar or vector giving the effective reproductive number of the pathogen
phi	scalar or vector giving the desired true discovery rate (1-false discovery rate)
min_pairs	minimum number of linked pairs observed in the sample, defaults to 1 pair (2 samples); this is to ensure reasonable results are obtained
assumption	a character vector indicating which assumptions about transmission and linkage criteria. Default = 'mtml'. Accepted arguments are: <ol style="list-style-type: none"> 1. 'stsl' for the single-transmission single-linkage assumption (prob_trans_stsl). 2. 'mtsl' for the multiple-transmission single-linkage assumption (prob_trans_mtsl). 3. 'mtml' for the multiple-transmission multiple-linkage assumption (prob_trans_mtml).

Value

scalar or vector giving the sample size needed to meet the given conditions

Author(s)

John Giles, Shirlee Wohl, and Justin Lessler

Examples

```
samplesize(eta=0.99, chi=0.995, N=100, R=1, phi=0.75)
```

sens_spec_calc	<i>Calculate sensitivity and specificity</i>
----------------	--

Description

Function to calculate the sensitivity and specificity of a genetic distance cutoff given an underlying mutation rate and mean number of generations between cases

Usage

```
sens_spec_calc(  
  cutoff,  
  mut_rate,  
  mean_gens_pdf,  
  max_link_gens = 1,  
  max_gens = NULL,  
  max_dist = NULL  
)
```

Arguments

cutoff	the maximum genetic distance at which to consider cases linked
mut_rate	mean number of mutations per generation, assumed to be poisson distributed
mean_gens_pdf	the density distribution of the mean number of generations between cases; the index of this vector is assumed to be the discrete distance between cases
max_link_gens	the maximum generations of separation for linked pairs
max_gens	the maximum number of generations to consider, if NULL (default) value set to the highest number of generations in mean_gens_pdf with a non-zero probability
max_dist	the maximum distance to calculate, if NULL (default) value set to max_gens * 99.9th percentile of mut_rate poisson distribution

Value

a data frame with the sensitivity and specificity for a particular genetic distance cutoff

Author(s)

Shirlee Wohl and Justin Lessler

See Also

Other mutrate_functions: [gen_dists\(\)](#), [get_optim_roc\(\)](#), [sens_spec_roc\(\)](#)

Examples

```
# calculate the sensitivity and specificity for a specific genetic distance threshold of 2 mutations
sens_spec_calc(cutoff=2,
              mut_rate=1,
              mean_gens_pdf=c(0.02,0.08,0.15,0.75),
              max_link_gens=1)

# calculate the sensitivity and specificity for a a range of genetic distance thresholds
sens_spec_calc(cutoff=1:10,
              mut_rate=1,
              mean_gens_pdf=c(0.02,0.08,0.15,0.75),
              max_link_gens=1)
```

sens_spec_roc	<i>Make ROC from sensitivity and specificity</i>
---------------	--

Description

This is a wrapper function that takes output from the ‘sens_spec_calc()’ function and constructs values for the Receiver Operating Characteric (ROC) curve

Usage

```
sens_spec_roc(
  cutoff,
  mut_rate,
  mean_gens_pdf,
  max_link_gens = 1,
  max_gens = NULL,
  max_dist = NULL
)
```

Arguments

cutoff	the maximum genetic distance at which to consider cases linked
mut_rate	mean number of mutations per generation, assumed to be poisson distributed
mean_gens_pdf	the density distribution of the mean number of generations between cases; the index of this vector is assumed to be the discrete distance between cases
max_link_gens	the maximum generations of separation for linked pairs
max_gens	the maximum number of generations to consider, if NULL (default) value set to the highest number of generations in mean_gens_pdf with a non-zero probability
max_dist	the maximum distance to calculate, if NULL (default) value set to max_gens * 99.9th percentile of mut_rate poisson distribution

Value

data frame with cutoff, sensitivity, and 1-specificity

Author(s)

Shirlee Wohl and Justin Lessler

See Also

Other mutrate_functions: [gen_dists\(\)](#), [get_optim_roc\(\)](#), [sens_spec_calc\(\)](#)

Examples

```
# ebola-like pathogen
R <- 1.5
mut_rate <- 1

# use simulated generation distributions
data("gen_dist_sim")
mean_gens_pdf <- as.numeric(gen_dist_sim[gen_dist_sim$R == R, -(1:2)])

# get theoretical genetic distance dist based on mutation rate and generation parameters
dists <- as.data.frame(gen_dists(mut_rate = mut_rate,
                                mean_gens_pdf = mean_gens_pdf,
                                max_link_gens = 1))

dists <- reshape2::melt(dists,
                        id.vars = "dist",
                        variable.name = "status",
                        value.name = "prob")

# get sensitivity and specificity using the same paramters
roc_calc <- sens_spec_roc(cutoff = 1:(max(dists$dist)-1),
                          mut_rate = mut_rate,
                          mean_gens_pdf = mean_gens_pdf)
```

truediscoveryrate	<i>Calculate true discovery rate of a sample</i>
-------------------	--

Description

This function calculates the true discovery rate (proportion of true transmission pairs) in a sample given the sensitivity η and specificity χ of the linkage criteria, and sample size M . Assumptions about transmission and linkage (single or multiple) can be specified.

Usage

```
truediscoveryrate(eta, chi, rho, M, R = NULL, assumption = "mtml")
```

Arguments

<code>eta</code>	scalar or vector giving the sensitivity of the linkage criteria
<code>chi</code>	scalar or vector giving the specificity of the linkage criteria
<code>rho</code>	scalar or vector giving the proportion of the final outbreak size that is sampled
<code>M</code>	scalar or vector giving the number of cases sampled
<code>R</code>	scalar or vector giving the effective reproductive number of the pathogen (default=NULL)
<code>assumption</code>	a character vector indicating which assumptions about transmission and linkage criteria. Default = 'mtml'. Accepted arguments are: <ol style="list-style-type: none"> 1. 'stsl' for the single-transmission single-linkage assumption (prob_trans_stsl). 2. 'mtsl' for the multiple-transmission single-linkage assumption (prob_trans_mtsl). 3. 'mtml' for the multiple-transmission multiple-linkage assumption (prob_trans_mtml).

Value

scalar or vector giving the true discovery rate

Author(s)

John Giles, Shirlee Wohl, and Justin Lessler

See Also

Other discovery_rate: [falsediscoveryrate\(\)](#)

Examples

```
# The simplest case: single-transmission, single-linkage, and perfect sensitivity
truediscoveryrate(eta=1, chi=0.9, rho=0.5, M=100, assumption='stsl')

# Multiple-transmission and imperfect sensitivity
truediscoveryrate(eta=0.99, chi=0.9, rho=1, M=50, R=1, assumption='mtsl')

# Small outbreak, larger sampling proportion
truediscoveryrate(eta=0.99, chi=0.95, rho=1, M=50, R=1, assumption='mtml')

# Large outbreak, small sampling proportion
truediscoveryrate(eta=0.99, chi=0.95, rho=0.5, M=1000, R=1, assumption='mtml')
```

true_pairs	<i>Calculate expected number of true transmission pairs</i>
------------	---

Description

This function calculates the expected number true transmission pairs in a sample of size M . Assumptions about transmission and linkage (single or multiple) can be specified.

Usage

```
true_pairs(eta, rho, M, R = NULL, assumption = "mtml")
```

Arguments

eta	scalar or vector giving the sensitivity of the linkage criteria
rho	scalar or vector giving the proportion of the final outbreak size that is sampled
M	scalar or vector giving the number of cases sampled
R	scalar or vector giving the effective reproductive number of the pathogen (default=NULL)
assumption	a character vector indicating which assumptions about transmission and linkage criteria. Default = 'mtml'. Accepted arguments are: <ol style="list-style-type: none">1. 'stsl' for the single-transmission single-linkage assumption (prob_trans_stsl).2. 'mtsl' for the multiple-transmission single-linkage assumption (prob_trans_mtsl).3. 'mtml' for the multiple-transmission multiple-linkage assumption (prob_trans_mtml).

Value

scalar or vector giving the expected number of true transmission pairs in the sample

Author(s)

John Giles, Shirlee Wohl, and Justin Lessler

See Also

Other true_pairs: [true_pairs_mtml\(\)](#), [true_pairs_mtsl\(\)](#), [true_pairs_stsl\(\)](#)

Examples

```
true_pairs(eta=0.99, rho=0.75, M=100, R=1)
```

true_pairs_mtml	<i>Expected number of true transmission pairs assuming multiple-transmission and multiple-linkage</i>
-----------------	---

Description

This function calculates the expected number of true transmission pairs in a sample of size M . The multiple-transmission and multiple-linkage method assumes the following:

1. Each case i is, on average, the infector of R cases in the population (N)
2. Each case i is allowed to be linked by the linkage criteria to multiple cases j in the sampled population (M).
3. Linkage events are independent of one another (i.e, linkage of case i to case j has no bearing on linkage of case i to any other sample).

Usage

```
true_pairs_mtml(eta, rho, M, R)
```

Arguments

eta	scalar or vector giving the sensitivity of the linkage criteria
rho	scalar or vector giving the proportion of the final outbreak size that is sampled
M	scalar or vector giving the number of cases sampled
R	scalar or vector giving the effective reproductive number of the pathogen

Value

scalar or vector giving the expected number of true transmission pairs in the sample

Author(s)

John Giles, Shirlee Wohl and Justin Lessler

See Also

Other true_pairs: [true_pairs_mtsl\(\)](#), [true_pairs_stsl\(\)](#), [true_pairs\(\)](#)

Examples

```
true_pairs_mtml(eta=0.95, rho=0.2, M=1000, R=1)
```

true_pairs_mtsl	<i>Expected number of true transmission pairs assuming multiple-transmission and single-linkage</i>
-----------------	---

Description

This function calculates the expected number true transmission pairs in a sample of size M . The multiple-transmission and single-linkage method assumes the following:

1. Each case i is, on average, the infector of R cases in the population (N)
2. Each case i is allowed to be linked by the linkage criteria to only one other case j in the sampled population (M).

Usage

```
true_pairs_mtsl(eta, rho, M, R)
```

Arguments

eta	scalar or vector giving the sensitivity of the linkage criteria
rho	scalar or vector giving the proportion of the final outbreak size that is sampled
M	scalar or vector giving the number of cases sampled
R	scalar or vector giving the effective reproductive number of the pathogen

Value

scalar or vector giving the expected number of true transmission pairs in the sample

Author(s)

John Giles, Shirlee Wohl and Justin Lessler

See Also

Other true_pairs: [true_pairs_mtml\(\)](#), [true_pairs_stsl\(\)](#), [true_pairs\(\)](#)

Examples

```
true_pairs_mtsl(eta=0.95, rho=0.2, M=200, R=1)
```

true_pairs_stsl	<i>Expected number of true transmission pairs assuming single-transmission and single-linkage</i>
-----------------	---

Description

This function calculates the expected number of true transmission pairs in a sample of size M . The single-transmission and single-linkage method assumes the following:

1. Each case i is linked by transmissino to only one other case j in the population (N).
2. Each case i is linked by the linkage criteria to only one other case j in the sampled population (M).

Usage

```
true_pairs_stsl(eta, rho, M)
```

Arguments

eta	scalar or vector giving the sensitivity of the linkage criteria
rho	scalar or vector giving the proportion of the final outbreak size that is sampled
M	scalar or vector giving the number of cases sampled

Value

scalar or vector giving the expected number of true transmission pairs in the sample

Author(s)

John Giles, Shirlee Wohl, and Justin Lessler

See Also

Other true_pairs: [true_pairs_mtml\(\)](#), [true_pairs_mtsl\(\)](#), [true_pairs\(\)](#)

Examples

```
true_pairs_stsl(eta=0.95, rho=0.2, M=200)
```

Index

- * **datasets**
 - gen_dist_sim, 6
 - * **discovery_rate**
 - falsediscoveryrate, 3
 - truediscoveryrate, 17
 - * **inverse_functions**
 - samplesize, 14
 - * **nutrate_functions**
 - gen_dists, 5
 - get_optim_roc, 7
 - sens_spec_calc, 15
 - sens_spec_roc, 16
 - * **obs_pairs**
 - exp_links, 2
 - obs_pairs_mtml, 8
 - obs_pairs_mtsl, 9
 - obs_pairs_stsl, 10
 - * **prob_trans**
 - prob_trans_mtml, 11
 - prob_trans_mtsl, 12
 - prob_trans_stsl, 13
 - * **true_pairs**
 - true_pairs, 19
 - true_pairs_mtml, 20
 - true_pairs_mtsl, 21
 - true_pairs_stsl, 22
- exp_links, 2, 8–10
- falsediscoveryrate, 3, 18
- gen_dist_sim, 6
- gen_dists, 5, 7, 15, 17
- get_optim_roc, 5, 7, 15, 17
- obs_pairs_mtml, 3, 8, 9, 10
- obs_pairs_mtsl, 3, 8, 9, 10
- obs_pairs_stsl, 3, 8, 9, 10
- prob_trans_mtml, 3, 4, 11, 12–14, 18, 19
- prob_trans_mtsl, 3, 4, 11, 12, 13, 14, 18, 19
- prob_trans_stsl, 3, 4, 11, 12, 13, 14, 18, 19
- samplesize, 14
- sens_spec_calc, 5, 7, 15, 17
- sens_spec_roc, 5, 7, 15, 16
- true_pairs, 19, 20–22
- true_pairs_mtml, 19, 20, 21, 22
- true_pairs_mtsl, 19, 20, 21, 22
- true_pairs_stsl, 19–21, 22
- truediscoveryrate, 4, 17