Package 'SMARTp'

January 20, 2025

Type Package

Title Sample Size for SMART Designs in Non-Surgical Periodontal Trials

Version 0.1.1

Author Jing Xu, Dipankar Bandyopadhyay, Douglas Azevedo, Bibhas Chakraborty

Maintainer Dipankar Bandyopadhyay <bandyopd@gmail.com>

Description Sample size calculation to detect dynamic treatment regime (DTR) effects based on change in clinical attachment level (CAL) outcomes from a non-surgical chronic periodontitis treatments study. The experiment is performed under a Sequential Multiple Assignment Randomized Trial (SMART) design. The clustered tooth (subunit) level CAL outcomes are skewed, spatially-referenced, and non-randomly missing. The implemented algorithm is available in Xu et al. (2019+) <arXiv:1902.09386>.

Depends R (>= 3.5)

Imports covr, sn (>= 1.5), mvtnorm (>= 1.0), stats, methods

URL https://github.com/bandyopd/SMARTp

License LGPL (>= 2)

Encoding UTF-8

LazyData true

RoxygenNote 6.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2019-05-17 07:10:04 UTC

Contents

CAR_cov_teeth *The within-mouth covariance matrix with conditional autoregressive structure*

Description

The covariance matrix of individual teeth measures for each subject follows a Conditional Autoregressive model (CAR) density

Usage

```
CAR_cov_teeth(m, rho, tau)
```

Arguments

m	Maximum number of units in each cluster, i.e., 28 teeth in each mouth (the 4 third-molars are usually ignored)
rho	Association parameter of the CAR model
tau	Variation parameter of the CAR model

Details

CAR_cov_teeth gives the covariance matrix among the teeth within each mouth based on the CAR structure (Besag *et al.*, 1991), given the maximum number of teeth for each subject (m), the variance (τ), and the association (ρ) parameters.

The CAR covariance matrix can be expressed as $\Sigma_{28\times 28} = \tau^2(W - \rho D)^{-1}$, where $\tau^2 > 0$, and $\rho \in [0, 1]$ are the parameters that control the magnitude of variation and the degree of spatial association, respectively. For matrix D, the element $D_{tt'}$ is 1 if locations t and t' are adjacent and 0 otherwise. The matrix W is diagonal with diagonal elements $W_{tt} = \sum_{t'} D_{tt'}$. Note, the argument τ in *CAR_cov_teeth* is the variance, and not the standard deviation.

Value

The covariance matrix among the teeth in each mouth (assuming full dentition, i.e., 28 teeth) based on a CAR model.

Author(s)

Jing Xu, Dipankar Bandyopadhyay, Douglas Azevedo, Bibhas Chakraborty

References

Besag, J., York, J. & Mollie, A. (1991), *"Bayesian image restoration, with two applications in spatial statistics (With Discussion)"*, Annals of the Institute of Statistical Mathematics 43, 159.

Reich, B. & Bandyopadhyay, D. (2010), *"A latent factor model for spatial data with informative missingness"*, The Annals of Applied Statistics 4, 439–459.

See Also

MC_var_yibar_mis, SampleSize_SMARTp

Examples

```
m <- 28
rho <- 0.975
tau <- 0.85
Sigma <- CAR_cov_teeth(m = m, rho = rho, tau = tau)
```

`MC_var_yibar_mis` *Estimated mean and variance of the average change in CAL for each subject*

Description

The estimated Monte Carlo mean and variance of the average change in clinical attachment level (CAL) for each subject

Usage

```
MC_var_yibar_mis(mu, Sigma, sigma1, lambda, nu, sigma0, Num, a0, b0, cutoff)
```

Arguments

`mu`	Mean matrix, where row represents each treatment path, and column represents each cluster unit
`Sigma`	Within-mouth teeth covariance matrix
`sigma1`	Standard deviation of the residual for the continuous outcome Y_{it}
`lambda`	The skewness parameter of the residual for the continuous outcome Y_{it}
`nu`	The degree freedom, or kurtosis parameter of the residual for the continuous outcome Y_{it}
`sigma0`	Standard deviation of the residual for the binary outcome M_{it}
`Num`	Number of samples to estimate mean or variance of $\bar{Y}_i$
`a0`	Intercept parameter in the probit model for the binary outcome M_{it}
`b0`	Slope parameter corresponding to the spatial random effect in the probit model for the binary outcome M_{it}
`cutoff`	Cut-off value in the binary outcome regression

Details

MC_var_yibar_mis computes the Monte-Carlo estimates of expectation and variance of the sample mean among the teeth within each mouth, i.e $\bar{Y}_i = \sum Y_{it}(1 - M_{it}) / \sum(1 - M_{it})$, where Y_{it} is the change in CAL (measured in mm) for patient i and tooth t, and M_{it} is the misingness indicator, i.e., $M_{it} = 1$ implies tooth t in subject i is mising. The joint regression models for Y_{it} and M_{it} are available in Reich & Bandyopadhyay (2010, *Annals of Applied Statistics*).

Value

The simulated dataset of CAL change "Y_{it}", missingness "M_{it}" and function inside the indicator of "$M_{it}I_{it}$" for each tooth of each patient, with the corresponding estimated mean "mY_i", variance "$VarY_i$" and missing proportion "PM" for each patient

Author(s)

Jing Xu, Dipankar Bandyopadhyay, Douglas Azevedo, Bibhas Chakraborty

References

Besag, J., York, J. & Mollie, A. (1991), *"Bayesian image restoration, with two applications in spatial statistics (With Discussion)"*, Annals of the Institute of Statistical Mathematics 43, 159.

Reich, B. & Bandyopadhyay, D. (2010), *"A latent factor model for spatial data with informative missingness"*, The Annals of Applied Statistics 4, 439–459.

See Also

CAR_cov_teeth, SampleSize_SMARTp

Examples

```
m <- 28
Num <- 1000
cutoff <- 0
sigma1 <- 0.95
sigma0 <- 1
lambda <- 0
nu <- Inf
b0 <- 0.5
a0 <- -1.0
rho <- 0.975
tau <- 0.85
del1 <- 0.5
del2 <- 2

Sigma <- CAR_cov_teeth(m, rho, tau)
Sigma_comp <- array(Sigma, c(m, m, 4))
Sigma_sim <- array(Sigma, c(m, m, 10))

mu_comp <- array(0, c(2, m, 2))
mu_comp[, , 1] <- rbind(rep(0, m), rep(del1, m))
mu_comp[, , 2] <- rbind(rep(0, m), rep(del2, m))

VarYitd1R = MC_var_yibar_mis(mu = mu_comp[1, , 1], Sigma = Sigma,
                             sigma1 = sigma1,
                             lambda = lambda, nu = nu,
                             sigma0 = sigma0, Num = Num, a0 = a0, b0 = b0,
                             cutoff = cutoff)
PM <- VarYitd1R$PM
VarYid1R <- VarYitd1R$VarYi
```

```
mYid1R <- VarYitd1R$mYi
VarYitd1NR <- MC_var_yibar_mis(mu = mu_comp[2, , 1], Sigma = Sigma,
                               sigma1 = sigma1,
                               lambda = lambda, nu = nu,
sigma0 = sigma0, Num = Num, a0 = a0, b0 = b0, cutoff = cutoff)

PM <- VarYitd1NR$PM
VarYid1NR <- VarYitd1NR$VarYi
mYid1NR <- VarYitd1NR$mYi
VarYitd3R <- MC_var_yibar_mis(mu = mu_comp[1, , 2], Sigma = Sigma,
                              sigma1 = sigma1,
                              lambda = lambda, nu = nu,
                              sigma0 = sigma0, Num = Num, a0 = a0, b0 = b0,
                              cutoff = cutoff)

PM <- VarYitd3R$PM
VarYid3R <- VarYitd3R$VarYi
mYid3R <- VarYitd3R$mYi
VarYitd3NR <- MC_var_yibar_mis(mu = mu_comp[2,,2], Sigma = Sigma,
                               sigma1 = sigma1,
                               lambda = lambda, nu = nu,
sigma0 = sigma0, Num = Num, a0 = a0, b0 = b0, cutoff = cutoff)

PM <- VarYitd3NR$PM
VarYid3NR <- VarYitd3NR$VarYi
mYid3NR <- VarYitd3NR$mYi
```

`print.SMARTp` *Print for SMARTp class*

Description

Print for SMARTp class

Usage

```
## S3 method for class 'SMARTp'
print(x, ...)
```

Arguments

`x`	SMARTp object to print
`...`	Other parameters for print

`SampleSize_SMARTp` *Sample size calculation under a clustered SMART design for non-surgical treatment of chronic periodontitis*

Description

Sample size calculations to detect desired DTR effects, which includes (i) a single regime, (ii) difference between two regimes, and (iii) a specific regime is the best, based on CAL changes under the proposed clustered, two-stage, SMART trial given type I and type II error rates

Usage

```
SampleSize_SMARTp(mu, st1, dtr, regime, pow, a, rho, tau, sigma1,
                  lambda, nu, sigma0, Num, p_i, c_i, a0, b0, cutoff)
```

Arguments

`mu`	Mean matrix, where row represents each treatment path from the SMART design diagram (see Xu *et al.*, 2019), and column represents each unit (i.e. tooth) within a cluster (i.e. mouth)
`st1`	Stage-1 treatment matrix, where rows represent the corresponding stage-1 treatments, the 1st column includes the number of treatment options for the responder, the 2nd column include the numbers of treatment options for the non-responder, the 3rd column are the response rates, and the 4th column includes the row numbers
`dtr`	Matrix of dimension (# of DTRs X 4), the 1st column represents the DTR numbers, the 2nd column represents the treatment path number of responders for the corresponding DTRs in the 1st column, the 3rd column represents the corresponding treatment path number of the non-responders for the corresponding DTRs in the 1st column, while the 4th column represents the corresponding initial treatment
`regime`	Treatment regime vector. For detecting regime 1 as the best, use c(1, 2, 3, 4, 5, 6, 7, 8). Similarly, if regime 2 is the best, use c(2, 1, 3, 4, 5, 6, 7, 8), and so on
`pow`	Power or 1 - Type II error rate, default is 0.8
`a`	Type I error rate, default is 0.05
`rho`	Association parameter of the CAR model, default is 0.975
`tau`	Variance parameter of the CAR model, default is 0.85
`sigma1`	Standard deviation of the residual for the continuous outcome Y_{it}, default is 0.95
`lambda`	Skewness parameter of the residual for the continuous outcome Y_{it}, default is 0
`nu`	The degrees of freedom parameter of the residual for Y_{it}, default is Inf
`sigma0`	Standard deviation of the residual for the binary outcome M_{it}, default is 1
`Num`	Iteration size to estimate variance of $\bar{Y}_i$, default is 100000
`p_i`	The expected proportion of available teeth for subject i

c_i	The average Pearson correlation coefficient between Y_{it} and M_{it} over the 28 teeth
a0	Intercept parameter in the probit model for the binary M_{it}, default is -1
b0	Slope parameter corresponding to the spatial random effect in the probit model for binary M_{it}, default is 0.5; note that a_0 and b_0 can be determined given p_i and c_i
cutoff	Cut-off value of the binary outcome regression, default is 0

Details

SampleSize_SMARTp computes the sample size required to detect the dynamic treatment regime (DTR) (Murphy, 2005, *Statistics in Medicine*) effects in a study comparing non-surgical treatments of chronic periodontitis, via the sequential multiple assignment randomized trial (SMART) design, with two-stages.

Outcome measures (i.e. change in CAL) are continuous and clustered (i.e. tooth within a subject's mouth, where each subject/mouth is a cluster) with non-random missingness captured via a shared parameter setting, specified in Reich and Bandyopadhyay (2010, *Annals of Applied Statistics*). Each cluster sub-unit has a binary missingness indicator, which is associated to its corresponding change of CAL through a joint model. The covariance structure within a cluster is captured by the conditionally autoregressive (CAR) structure (Besag et al, 1991).

The DTR effect can be detected based on either a single treatment regime, or the difference between two treatment regimes (with or without sharing initial treatments), or when one regime is considered the best among others. The mean and variance of the CAL change for each DTR can be estimated by the inverse probability weighting method via method of moments.

Note that the first three inputs "mu", "st1" and "dtr" define the SMART design in term of matrices. From Xu *et al.* (2019+, Under Review), stage-1 includes two treatments, e.g., treatments "3" and "8". Participants who respond to the stage-1 treatment will receive same treatment at stage-2, while non-responders will be randomly allocated to other treatments, i.e. non-responders who received treatment "3" at stage-1 will be randomly allocated to treatments "4"-"7" at stage-2, while non-responders receiving treatment "8" at stage-1 will be randomly allocated to treatments "4"-"7" at stage-2.

There are 8 treatment regimes for this design. They are 1 (treatment "3" at stage-1 and treatment "3" at stage- 2 if responder, otherwise treatment "4"), 2 (treatment "3" at stage-1 and treatment "3" at stage-2 if responder, otherwise treatment "5"), 3 (treatment "3" at stage-1 and treatment "3" at stage-2 if responder, otherwise treatment "6"), 4 (treatment "3" at stage-1 and treatment "3" at stage-2 if responder, otherwise treatment "7"), 5 (treatment "8" at stage-1 and treatment "8" at stage-2 if responder, otherwise treatment "4"), 6 (treatment "8" at stage-1 and treatment "8" at stage-2 if responder, otherwise treatment "5"), 7 (treatment "8" at stage-1 and treatment "8" at stage-2 if responder, otherwise treatment "6") and 8 (treatment "8" at stage-1 and treatment "8" at stage-2 if responder, otherwise treatment "7"). See Figure 2 in Xu *et al.* (2019+, Under Review)

Value

N	the estimated sample size
Del	effect size
Del_std	standardized effect size

ybar	the estimated regime means corresponding to "regime"
Sigma	the CAR covariance matrix corresponding to the latent Q_{it}; see Xu *et al.* (2019+, Under Review)
sig.dd	N*the variance or covariance matrix of the estimated regime means corresponding to "regime"
sig.e.sq	N*the variance or covariance matrix of the difference between first and rest of estimated regime means corresponding to "regime", sig.e.sq = sig.dd if the element number of "regime" is one
p_st1	the randomization probability of stage-1 for each treatment path
p_st2	the randomization probability of stage-2 for each treatment path
res	a vector with binary indicators represent responses or non-responses that corresponds to a treatment path
ga	the response rates of initial treatments corresponding to each treatment path
initr	column matrix with dimension = the number of treatment paths, the elements are the corresponding row number of st1

Author(s)

Jing Xu, Dipankar Bandyopadhyay, Douglas Azevedo, Bibhas Chakraborty

References

Besag, J., York, J. & Mollie, A. (1991) *"Bayesian image restoration, with two applications in spatial statistics (with discussion)"*, Annals of the Institute of Statistical Mathematics 43, 159.

Murphy, S. A. (2005), *"An experimental design for the development of adaptive treatment strategies"*, Statistics in Medicine 24, 1455–1481.

Reich, B. & Bandyopadhyay, D. (2010), *A latent factor model for spatial data with informative missingness*, The Annals of Applied Statistics 4, 439–459.

Xu, J., Bandyopadhyay, D., Mirzaei, S., Michalowicz, B and Bibhas Chakraborty. (2019+), *"SMARTp: A SMART design for non-surgical treatments of chronic periodontitis with spatially-referenced and non-randomly missing skewed outcomes"*, Under Review

See Also

CAR_cov_teeth, MC_var_yibar_mis

Examples

```
m <- 28
pow <- 0.8
a <- 0.05
Num <- 1000
cutoff <- 0
sigma1 <- 0.95
sigma0 <- 1
lambda <- 0
nu <- Inf
```

```
b0 <- 0.5
a0 <- -1.0
rho <- 0.975
tau <- 0.85

Sigma <- CAR_cov_teeth(m = m, rho = rho, tau = tau)
p_i <- SMARTp:::pifun(cutoff = cutoff, a0 = a0, b0 = b0,
                      Sigma = Sigma, sigma0 = sigma0)
cit4 <- b0*diag(Sigma)/sqrt((diag(Sigma) +
            (sigma1^2 - 2/pi*sigma1^2*(0^2/(1+0^2))))*(b0^2*diag(Sigma) +
            sigma0^2))
c_i <- mean(cit4)

del1 <- 5
del2 <- 0
del3 <- 0

mu_sim <- matrix(0, 10, m)
mu_sim[2, ] <- rep(del1, m)
mu_sim[4, ] <- rep(del2, m)
mu_sim[7, ] <- rep(del3, m)

st1 <- cbind(c(1, 1), c(4, 4), c(0.25, 0.5), 1:2)

##-- Stage-1 information
dtr <- cbind(1:8, c(rep(1, 4), rep(6, 4)),
             c(2, 3, 4, 5, 7, 8, 9, 10), c(rep(1, 4), rep(2, 4)))

##-- Detecting a single regime, e.g., Regime 1
regime <- 1
SampleSize <- SampleSize_SMARTp(mu = mu_sim, st1 = st1, dtr = dtr,
                                regime = regime,
                                pow = pow, a = a, rho = rho,
                                tau = tau, sigma1 = sigma1, lambda = 0,
                                nu = Inf, sigma0 = sigma0, Num = Num,
                                p_i = p_i, c_i = c_i,
                                cutoff = cutoff)

N <- ceiling(SampleSize$N)
sig.e.sq <- SampleSize$sig.e.sq

sqrt(diag(sig.e.sq)/N)
SampleSize$Del_std
SampleSize$Del
SampleSize$sig.dd
sqrt(diag(SampleSize$sig.dd)/N)
SampleSize$ybar

##-- Now using a0 and b0
SampleSize_SMARTp(mu = mu_sim, st1 = st1, dtr = dtr, regime = regime,
                  pow = pow, a = a, rho = rho,
                  tau = tau, sigma1 = sigma1, lambda = 0, nu = Inf,
                  sigma0 = sigma0, Num = Num, a0 = a0, b0 = b0,
```

```
                  cutoff = cutoff)
SampleSize_SMARTp(mu = mu_sim, st1 = st1, dtr = dtr, regime = regime,
                  p_i = p_i, c_i = c_i)

##-- Detecting the difference between two regimes that shares initial treatment,
##-- e.g., Regimes 1 vs 3
regime <- c(1, 3)
SampleSize = SampleSize_SMARTp(mu = mu_sim, st1 = st1, dtr = dtr, regime = regime,
                               pow = pow, a = a, rho = rho,
                               tau = tau, sigma1 = sigma1, lambda = 0, nu = Inf,
                               sigma0 = sigma0, Num = Num, a0 = a0, b0 = b0,
                               cutoff = cutoff)

N <- ceiling(SampleSize$N)
sig.e.sq <- SampleSize$sig.e.sq

sqrt(diag(sig.e.sq)/N)
SampleSize$Del_std
SampleSize$Del
SampleSize$sig.dd

##-- Detecting the difference between two regimes that do not share initial treatment,
##-- e.g., Regimes 1 vs 5
regime <- c(1, 5)

SampleSize <- SampleSize_SMARTp(mu = mu_sim, st1 = st1, dtr = dtr, regime = regime,
                                pow = pow, a = a, rho = rho,
                                tau = tau, sigma1 = sigma1, lambda = 0, nu = Inf,
                                sigma0 = sigma0, Num = Num, a0 = a0, b0 = b0,
                                cutoff = cutoff)

N <- ceiling(SampleSize$N)
sig.e.sq <- SampleSize$sig.e.sq

sqrt(diag(sig.e.sq)/N)
SampleSize$Del_std
SampleSize$Del
SampleSize$sig.dd

##-- Detecting when Regime 1 is the best, e.g., comparing Regimes 1 vs 2, 3, 4, 5, 6, 7 and 8, i.e.
##-- the alternative hypothesis is \mu_{d1}>\mu_{d2} & \mu_{d1}>\mu_{d3} ... & \mu_{d1}>\mu_{d8}
##-- Note that this is a one-side test with Type-1 error rate of 0.025.
regime <- c(1, 2, 3, 4, 5, 6, 7, 8)
##-- To detect Regime 2 is the best, just use regime = c(2, 1, 3, 4, 5, 6, 7, 8), and so on
SampleSize <- SampleSize_SMARTp(mu = mu_sim, st1 = st1, dtr = dtr, regime = regime,
                                pow = pow, a = a, rho = rho,
                                tau = tau, sigma1 = sigma1, lambda = 0, nu = Inf,
                                sigma0 = sigma0, Num = Num, a0 = a0, b0 = b0,
                                cutoff = cutoff)

N <- ceiling(SampleSize$N)
sig.e.sq <- SampleSize$sig.e.sq
```

```
sqrt(diag(sig.e.sq)/N)
SampleSize$Del_std
SampleSize$Del
SampleSize$sig.dd
```

`SMARTp-class` *An object of "SMARTp" class*

Description

An object of "SMARTp" class

Slots

`N` The estimated sample size

`sig.dd` N*the variance or covariance matrix of the estimated regime means correspond to "regime"

`sig.e.sq` N*the variance or covariance matrix of the difference between first and rest of estimated regime means correspond to *regime*, sig.e.sq = sig.dd if the element number of *regime* is one

`Del` Effect size

`Del_std` Standardized effect size

`ybar` The estimated regime means corresponding to "regime"

`initr` column matrix with dimension = the number of treatment paths, the elements are the corresponding row number of st1

`ga` The response rates of initial treatments corresponding to each treatment path

`res` A vector with binary indicators represent responders, or non-responders corresponding to a treatment path

`p_st1` The randomization probability of stage-1 for each treatment path

`p_st2` The randomization probability of stage-2 for each treatment path

`Sigma` The CAR covariance matrix of the latent Q_{it}

`summary.SMARTp` *Summary for SMARTp class*

Description

Summary for SMARTp class

Usage

```
## S3 method for class 'SMARTp'
summary(object, ...)
```

Arguments

`object`	SMARTp object to summarise
`...`	Other parameters for summary

Index