Package 'SMARTp'

January 2019

Title Sample size calculation under a sequential multiple assignment randomized trial (SMART) design for clustered clinical attachment level (CAL) outcomes observed from a chronic periodontitis study.

Version 0.1.0

Date 2019-01-06

Author Jing Xu, Bibhas Chakraborty, Dipankar Bandyopadhyay.

Maintainer Jing Xu <kennyxu1983@hotmail.com>

Description Sample size calculation to detect dynamic treatment regime (DTR) effect based on change in CAL outcomes measured in chronic periodontitis studies. The experiment is performed under a SMART design. The clustered sub-unit level CAL outcomes are skewed, spatially-referenced, and non-randomly missing.

Depends pwr

Needs Compilation yes

License LGPL (>= 2)

Repository CRAN

Date/Publication 2019-

R topics documented:

CAR.Cov.Teeth *The within-mouth covariance matrix with conditional autoregressive structure*

Description
The covariance matrix of individual teeth measure (clinical attachment level) for each patient is obtained by conditional autoregressive model (CAR)

Usage
CAR.Cov.Teeth(m, rho, tau)

Arguments

m	maximum number of individuals in each cluster, i.e. 28 teeth in each mouth
tau	variation parameter of the CAR model
rho	correlation parameter of the CAR model

Details
CAR.Cov.Teeth gives a covariance matrix among the teeth within each mouth based on the CAR structure (Besag et al. ,1991), given the maximum number of teeth of each patient 'm', the variation 'tau' and correlation 'rho' parameters of the CAR model.

For the covariance matrix with CAR model, i.e. $\mathbf{\Sigma}_{28\times28}=\tau^2(\mathbf{W}_{28\times28}-\rho\mathbf{D}_{28\times28})^{-1}$, 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 $\mathbf{D}_{tt'}$ is one if locations t and t' are adjacent and zero otherwise. The matrix **W** is diagonal with diagonal elements $\mathbf{W}_{tt}=\sum_{t'}\mathbf{D}_{tt'}$.

Value
The covariance matrix among the teeth in each mouth based on conditional autoregressive model is obtained.

Authors
Jing Xu, Bibhas Chakraborty, Dipankar Bandyopadhyay

References
Besag, J., York, J. & Molli´e, 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.SMART.2Stage.Clust.CAR

Examples

```
m=28
rho=0.975
tau=0.85
Sigma = CAR.Cov.Teeth(m, rho, tau)
```

MC.Var.Yibar.Mis *Estimated variance of the mean CAL for each patient by Monte Carlo method*

Description
The variance of the mean clinical attachment level for each patient is estimated using Monte Carlo method

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, column represents each individual within a cluster
Sigma	covariance matrix of CAL among the teeth within each mouth
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 the residual for the continuous outcome Y_{it}
sigma0	standard deviation of the residual for the binary outcome M_{it}
Num	number of sampling to estimate variance of \bar{Y}_i
a0	intercept parameter of the probit model for the binary outcome M_{it}
b0	slope parameter of spatial random effect of the probit model for the binary outcome M_{it}
cutoff	cut off value of the binary outcome regression

Details
MC.Var.Yibar.Mis computes the expectation and variance of the sample mean among the teeth within each mouth, i.e. $\bar{Y}_i = \frac{\sum_{i=1}^{28} Y_{it}(1-M_{it})}{\sum_{i=1}^{28}(1-M_{it})}$, where Y_{it} states the clinical attachment level (mm) for patient 'i' and tooth 't', and the corresponding teeth can be missing, i.e. $M_{it} = 1$. It uses Monte-Carlo method. The joint regression models for Y_{it} and M_{it} are defined in Reich & Bandyopadhyay (2010, *Annals of Applied Statistics*).

Value
The simulated dataset of CAL change 'Yit', missingness 'Mit' and function inside the indicator of 'Mit' 'lit' for each tooth of each patient, with the corresponding estimated mean 'mYi', variance 'VarYi' and missing proportion 'PM' for each patient

Authors
Jing Xu, Bibhas Chakraborty, Dipankar Bandyopadhyay

References
Besag, J., York, J. & Molli´e, 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.SMART.2Stage.Clust.CAR

Examples

```
m=28;Num=1000000
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$PM #0.8
VarYid1R=VarYitd1R$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$PM #0.8
VarYid1NR=VarYitd1NR$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$PM #0.8
VarYid3R=VarYitd3R$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$PM; #0.8
VarYid3NR=VarYitd3NR$VarYi
mYid3NR=VarYitd3NR$mYi
```

SampleSize.SMARTp *Sample size calculation under a clustered SMART design for chronic periodontitis*

Description

Sample size calculation to detect desired DTR effects based on change in CAL outcomes under the proposed clustered, two-stage, SMART trial given type I and type II error rates

Usage

SampleSize.SMARTp(mu, st1, dtr, regime, pow, b, 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, column represents each individual (i.e. tooth) within a cluster (i.e. mouth)
st1	stage-1 treatment matrix, where rows represent the corresponding stage-1 treatments, the first column includes the numbers of treatment options for responder, the second column includes the numbers of treatment options for non-responder, the third column are the responding rates, the fourth column includes the row numbers
dtr	matrix of dimension (# of DTRs X 4), the first column represents the DTR numbers, the second column represents the treatment path number of responders for the corresponding DTRs in the first column, the third column represents the corresponding treatment path number of the non-responders for the corresponding DTRs in the first column, while the fourth column represents the corresponding initial treatment
regime	a single or a paired treatment regime, i.e. length should be 1 or 2
pow	power or 1- type II error rate, default is pow=0.8
b	beta, type-II error rate, default is b=0.2
a	alpha, type- I error rate, default is a=0.05
tau	variance parameter of the CAR model, default is tau=0.85
rho	association parameter of the CAR model, default is rho=0.975
sigma1	standard deviation of the residual for the continuous outcome Y_{it}, default is sigma1=0.95
lambda	skewness parameter of the residual for the continuous outcome Y_{it}, default is lambda=0
nu	the degrees of freedom parameter of the residual for Y_{it}, default is nu=Inf
sigma0	standard deviation of the residual for the binary outcome M_{it}, default is sigma0=1
Num	number of sampling to estimate variance of \bar{Y}_i, default is Num=1000000
p_i	the expected proportion of available teeth for patient 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 a0=-1
b0	slope parameter corresponding to the spatial random effect in the probit model for binary M_{it}, default is b0=0.5, note that a0 and b0 can be determined given p_i and c_i
cutoff	cut-off value of the binary outcome regression, default is cutoff=0

Details

SampleSize.SMARTp computes the sample size required to detect the dynamic treatment regime (DTR) (Murphy (2005)) effects for periodontitis, performed by the experimental design of sequential multiple assignment randomized trial with two-stage.

Outcome measures (i.e. change in CAL) are continuous and clustered (i.e. tooth within a patient's mouth, where each patient/mouth is a cluster) with non-random missingness captured via a shared parameter setting

specified in Reich and Bandyopadhyay, 2010. 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 treatment. 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. In the paper of Xu *et al* (2019+), 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').

Value

N	the estimated sample size
Del	effect size
Del_std	standardized effect size
ybard1	the regime mean, corresponds to the first element of input `regime'
ybard2	the regime mean, corresponds to the second element of input `regime', ybard2=0 if the number of elements of regime is 1
Sigma	the CAR covariance matrix of Q_{it}
sig.d1.sq	N*the variance of the estimated mean of regime, which corresponds to the first element of input `regime'
sig.d2.sq	N*the variance of the estimated mean of regime, which corresponds to the second element of input `regime', sig.d2.sq=0 if the number of element of `regime' is one
sig.d1d2	N*the covariance of the estimated means between the regimes within `regime', sig.d1d2=0, if the number of element of `regime' is one
sig.e.sq	N*the variance of the difference between the estimated mean of the regimes correspond to `regime', sig.e.sq=sig.d1.sq if the number of element 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 responders, or non-responders corresponding to a treatment path
ga	the response rates of initial treatments corresponding to each treatment path
initr	one column matrix with dimension of number of treatment paths, the elements are the corresponding row number of st1

Authors

Jing Xu, Bibhas Chakraborty, Dipankar Bandyopadhyay

References

Besag, J., York, J. & Molli´e, 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., Bibhas Chakraborty, B., Mirzaei, S. and Michalowicz, B. (2019+), 'SMARTp: A SMART design for non-surgical treatments of chronic periodontitis with spatially-referenced and non-randomly missing skewed outcomes', *Submitted*

See also CAR.Cov.Teeth , MC.Var.Yibar.Mis

Examples

```
m=28;
del1=0.5;
del2=2
del3=5

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)) )   #DTR information
regime=c(1, 5)

pow=0.8;
b=1-pow;
a=0.05
Num=1000000
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, rho, tau)
p_i=pifun(cutoff, a0, b0, Sigma, 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)

SampleSize=SampleSize.SMARTp(mu=mu_sim, st1=st1, dtr=dtr, regime=regime, pow=pow, b=b, 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$N);N
Del=SampleSize$Del;Del
sig.e.sq=SampleSize$sig.e.sq;sig.e.sq;sqrt(sig.e.sq/N)
Del_std=Del/sqrt(sig.e.sq/2);Del_std
```

```
sig.d1.sq=SampleSize$sig.d1.sq; sig.d1.sq; sqrt(sig.d1.sq /N)
sig.d2.sq=SampleSize$sig.d2.sq; sig.d2.sq; sqrt(sig.d2.sq /N)
sig.d1d2=SampleSize$sig.d1d2; sig.d1d2/N

SampleSize=SampleSize.SMARTp(mu=mu_sim, st1=st1, dtr=dtr, regime=regime, pow=pow, b=b, a=a, rho=rho,
tau=tau, sigma1=sigma1, lambda=0, nu=Inf, sigma0=sigma0, Num=Num, a0=a0, b0=b0, cutoff=cutoff);
SampleSize=SampleSize.SMARTp(mu=mu_sim, st1=st1, dtr=dtr, regime=regime, p_i=p_i, c_i=c_i);
```