

GenePair: Statistical Methods for Modeling Spatially-Referenced Paired Genetic Relatedness Data

Clustered_Example

[1] Simulate data from the proposed model:

- Setting the reproducibility seed and initializing packages for data simulation:

```
set.seed(7783)
```

```
library(GenePair)
```

```
library(mnormt) #Multivariate normal distribution
```

- Setting the global data values:

```
n<-75 #Number of individuals
```

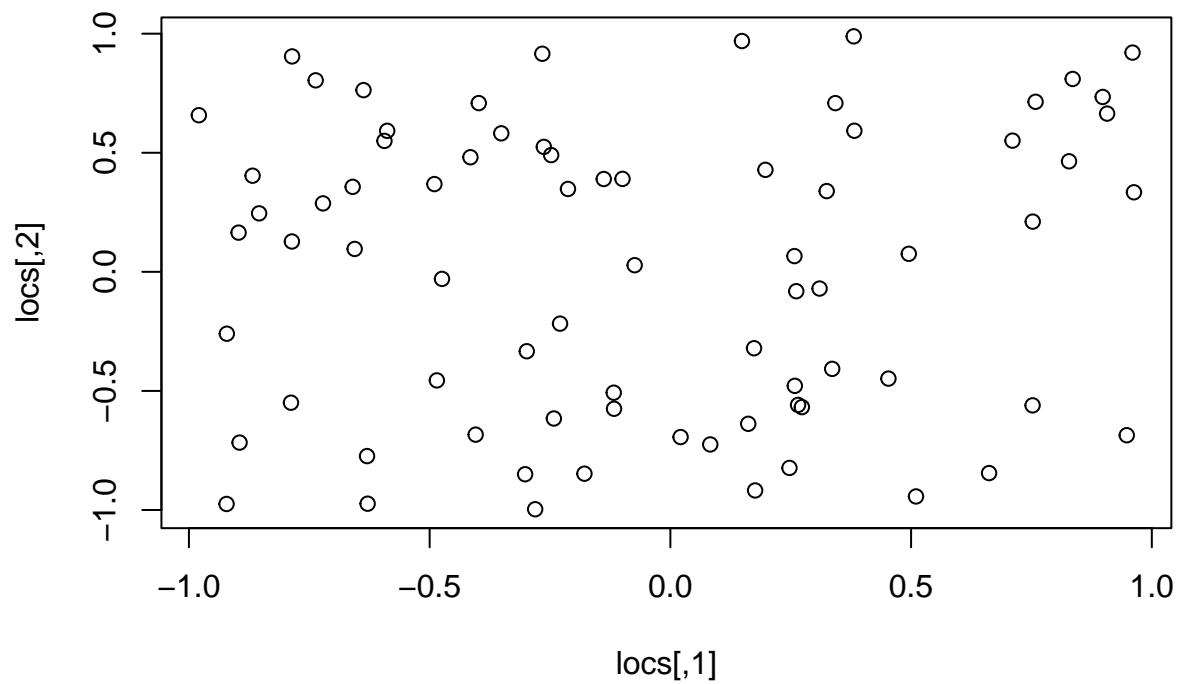
```
n_star<-0 #Number of paired responses
```

```
for(j in 1:(n - 1)){  
  for(k in (j + 1):n){  
    n_star<-n_star +  
      1  
  }  
}
```

```
#Unique locations (one for each individual)
```

```
locs<-matrix(runif((2*n),  
                  min = -1.00,  
                  max = 1.00),  
             nrow = n,  
             ncol = 2)
```

```
plot(locs)
```



```
v<-diag(n)

spatial_dists<-as.matrix(dist(locs,
                              diag = TRUE,
                              upper = TRUE))

#x_pair
x_pair<-matrix(1,
               nrow = n_star,
               ncol = 2)
x1<-rnorm(n = n)
x2<-rnorm(n = n)
counter<-1
for(j in 1:(n - 1)){
  for(k in (j + 1):n){

    x_pair[counter, 2]<-abs(x1[j] - x2[k])
    counter<-counter +
      1

  }
}
x_pair[,2]<-scale(x_pair[,2])

#x_ind
x1<-rnorm(n = n)
```

```

x2<-rnorm(n = n)
x_ind<-matrix(0.00,
              nrow = n_star,
              ncol = 2)
counter<-1
for(j in 1:(n - 1)){
  for(k in (j + 1):n){

    x_ind[counter, 1]<-(x1[k] + x1[j])
    x_ind[counter, 2]<-(x2[k] + x2[j])
    counter<-counter +
      1

  }
}
x_ind<-scale(x_ind)

#z
z<-matrix(0,
          nrow = n_star,
          ncol = n)
counter<-1
for(j in 1:(n - 1)){
  for(k in (j + 1):n){

    z[counter, j]<-1
    z[counter, k]<-1
    counter<-counter +
      1

  }
}

```

- Setting the values for the statistical model parameters:

```

beta_true<- c(-1.00,
              -0.30)
gamma_true<-c(0.10,
              -0.25)

tau2_true<-0.25
sigma2_zeta_true<-0.10

phi_true<-5.00
Sigma_true<-tau2_true*exp(-phi_true*spatial_dists)
eta_true<-rmnorm(n = 1,
                mean = rep(0.00,
                          times = n),
                varcov = Sigma_true)

theta_true<-eta_true +
  rnorm(n = n,
        mean = 0.00,
        sd = sqrt(sigma2_zeta_true))

```

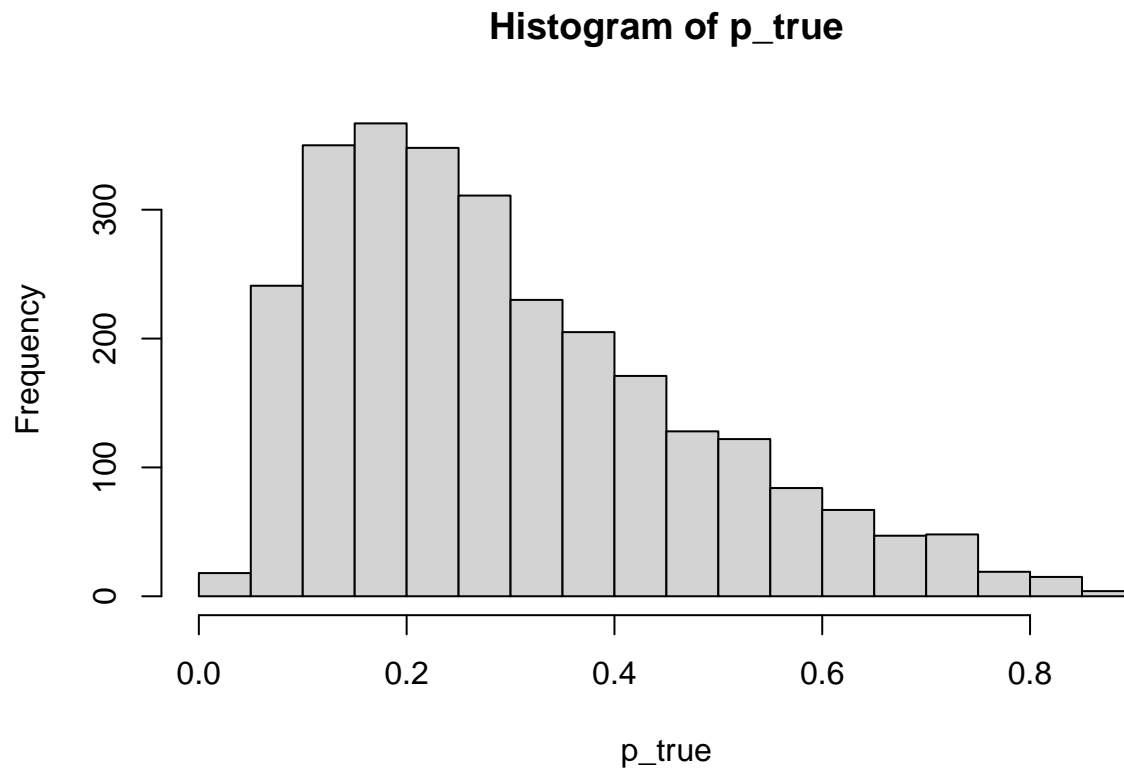
```

theta_true<-theta_true -
  mean(theta_true)

mu_true<-x_pair%%beta_true +
  x_ind%%gamma_true +
  z%%theta_true

p_true<-1.00/(1.00 + exp(-mu_true))
hist(p_true)

```



```

clustered_indicators<-rbinom(n = n_star,
  size = 1,
  prob = p_true)

mean(clustered_indicators)

## [1] 0.3084685

[2] Apply the Clustered Function within the GenePair Package:
results<-Clustered(mcmc_samples = 20000,
  clustered_indicators = clustered_indicators,
  x_pair = x_pair,
  x_ind = x_ind,
  z = z,
  spatial_dists = spatial_dists,
  v = v,

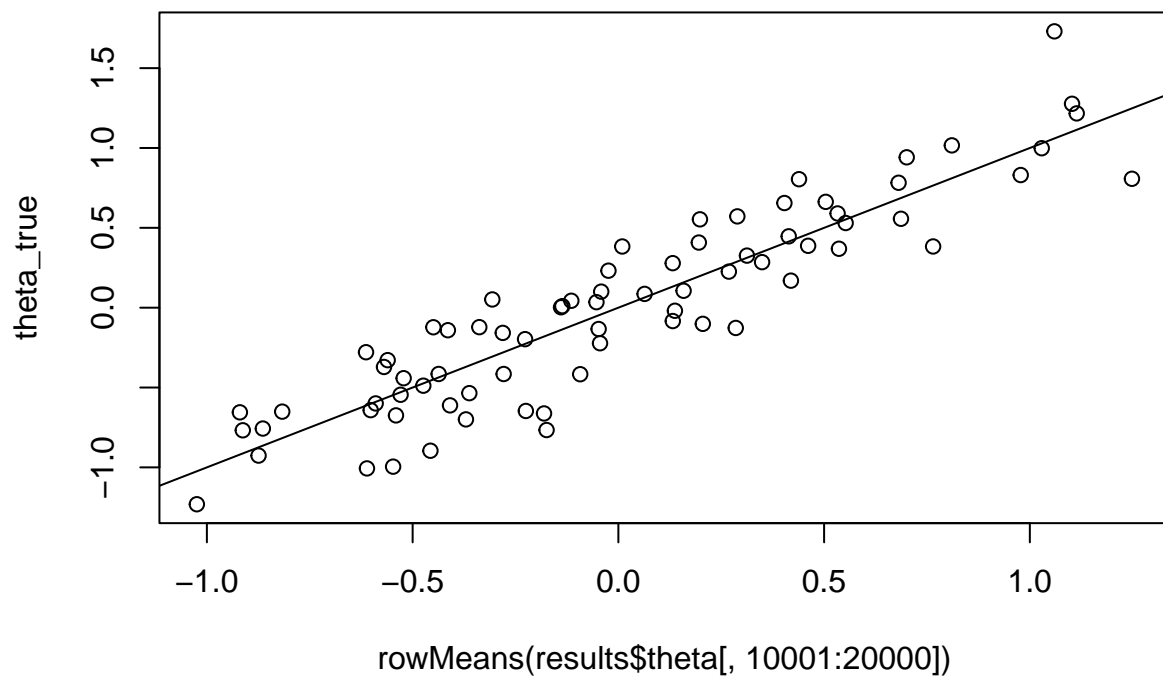
```

```
metrop_var_phi_trans = 0.75)
```

```
## Progress: 10%
## phi Acceptance: 25%
## *****
## Progress: 20%
## phi Acceptance: 26%
## *****
## Progress: 30%
## phi Acceptance: 25%
## *****
## Progress: 40%
## phi Acceptance: 25%
## *****
## Progress: 50%
## phi Acceptance: 25%
## *****
## Progress: 60%
## phi Acceptance: 25%
## *****
## Progress: 70%
## phi Acceptance: 25%
## *****
## Progress: 80%
## phi Acceptance: 25%
## *****
## Progress: 90%
## phi Acceptance: 25%
## *****
## Progress: 100%
## phi Acceptance: 25%
## *****
```

[3] Analyzing Output:

```
#Compare random effects
plot(rowMeans(results$theta[,10001:20000]),
      theta_true)
abline(0,1)
```



```
#Compare \beta
round(mean(results$beta[2, 10001:20000]), 2)

## [1] -0.27

round(beta_true[2], 2)

## [1] -0.3

#Compare \gamma
round(rowMeans(results$gamma[, 10001:20000]), 2)

## [1] 0.10 -0.24

round(gamma_true, 2)

## [1] 0.10 -0.25
```