

All_Final_Project_ERGM_Models

```
library(rio)
library(igraph)
library(dplyr)
library(readr)
library(ergm)
```

```
#Loading all the CA House and Senate data
```

```
#Each csv file in Edge_Lists represents 1 network. The first two columns are the sender/receiver
```

```
CA_2011_2012_Senate_edges <- read_csv("Edgelist/CA-2011-2012-Senate.csv")
CA_2011_2012_Senate <- read_csv("Metadata/CA-2011-2012-Senate.csv")
```

```
CA_2013_2014_Senate_edges <- read_csv("Edgelist/CA-2013-2014-Senate.csv")
CA_2013_2014_Senate <- read_csv("Metadata/CA-2013-2014-Senate.csv")
```

```
CA_2015_2016_Senate_edges <- read_csv("Edgelist/CA-2015-2016-Senate.csv")
CA_2015_2016_Senate <- read_csv("Metadata/CA-2015-2016-Senate.csv")
```

```
CA_2011_2012_House <- read_csv("Metadata/CA-2011-2012-House.csv")
CA_2011_2012_House_edges <- read_csv("Edgelist/CA-2011-2012-House.csv")
```

```
CA_2013_2014_House <- read_csv("Metadata/CA-2013-2014-House.csv")
CA_2013_2014_House_edges <- read_csv("Edgelist/CA-2013-2014-House.csv")
```

```
CA_2015_2016_House <- read_csv("Metadata/CA-2015-2016-House.csv")
CA_2015_2016_House_edges <- read_csv("Edgelist/CA-2015-2016-House.csv")
```

```
ERGM_data <- function(nodes, edges) {
  nodes_2 <- nodes %>%
    filter((PerDem >= 50 | PerRep >= 50), (CatCodeGroup == "Health" | CatCodeGroup == "Labor" | C
    mutate(party_node_2 = EID) %>%
    mutate(party_group_2 = CatCodeGroup) %>%
    mutate(Total_2 = Total) %>%
    select(party_node_2, party_group_2, Total_2)

  nodes_1 <- nodes %>%
    filter((PerDem >= 50 | PerRep >= 50), (CatCodeGroup == "Health" | CatCodeGroup == "Labor" | C
    mutate(party_node_1 = EID) %>%
    mutate(party_group_1 = CatCodeGroup) %>%
    mutate(Total_1 = Total) %>%
    select(EID, party_node_1, party_group_1, Total_1) %>%
    inner_join(edges, by = c("EID" = "node_1")) %>%
    inner_join(nodes_2, by = c("node_2" = "party_node_2")) %>%
```

```

  select(-EID, -edge) %>%
  mutate(party_node_2 = node_2) %>%
  mutate(party_1_group = party_group_1) %>%
  mutate(party_2_group = party_group_2) %>%
  mutate(party_1_total = Total_1) %>%
  mutate(party_2_total = Total_2) %>%
  select(-party_group_1, -party_group_2, -node_2, -Total_1, -Total_2) %>%
#https://www.statology.org/r-convert-true-false-to-1-0/
  mutate(same = as.integer(as.logical(party_1_group == party_2_group))) %>%
  mutate(total_diff = abs(party_1_total - party_2_total))
return(nodes_1)
}

```

#First step for ERGM with all House and Senate years prepared

```

CA_11_12_House <- graph.data.frame(ERGM_data(CA_2011_2012_House, CA_2011_2012_House_edges))
CA_11_12_House_sim <- get.adjacency(CA_11_12_House, sparse = FALSE, attr = 'same')
CA_11_12_House_total_diff <- get.adjacency(CA_11_12_House, sparse = FALSE, attr = 'total_diff')

CA_13_14_House <- graph.data.frame(ERGM_data(CA_2013_2014_House, CA_2013_2014_House_edges))
CA_13_14_House_sim <- get.adjacency(CA_13_14_House, sparse = FALSE, attr = 'same')
CA_13_14_House_total_diff <- get.adjacency(CA_13_14_House, sparse = FALSE, attr = 'total_diff')

CA_15_16_House <- graph.data.frame(ERGM_data(CA_2015_2016_House, CA_2015_2016_House_edges))
CA_15_16_House_sim <- get.adjacency(CA_15_16_House, sparse = FALSE, attr = 'same')
CA_15_16_House_total_diff <- get.adjacency(CA_15_16_House, sparse = FALSE, attr = 'total_diff')

CA_11_12_Senate <- graph.data.frame(ERGM_data(CA_2011_2012_Senate, CA_2011_2012_Senate_edges))
CA_11_12_Senate_sim <- get.adjacency(CA_11_12_Senate, sparse = FALSE, attr = 'same')
CA_11_12_Senate_total_diff <- get.adjacency(CA_11_12_Senate, sparse = FALSE, attr = 'total_diff')

CA_13_14_Senate <- graph.data.frame(ERGM_data(CA_2013_2014_Senate, CA_2013_2014_Senate_edges))
CA_13_14_Senate_sim <- get.adjacency(CA_13_14_Senate, sparse = FALSE, attr = 'same')
CA_13_14_Senate_total_diff <- get.adjacency(CA_13_14_Senate, sparse = FALSE, attr = 'total_diff')

CA_15_16_Senate <- graph.data.frame(ERGM_data(CA_2015_2016_Senate, CA_2015_2016_Senate_edges))
CA_15_16_Senate_sim <- get.adjacency(CA_15_16_Senate, sparse = FALSE, attr = 'same')
CA_15_16_Senate_total_diff <- get.adjacency(CA_15_16_Senate, sparse = FALSE, attr = 'total_diff')

```

#Detaching igraph to prepare for final step of ERGM
 detach(package:igraph)
 library(network)

```

CA_11_12_House_net <- as.network(ERGM_data(CA_2011_2012_House, CA_2011_2012_House_edges))

CA_11_12_House_ergm <- ergm(CA_11_12_House_net~edges + edgecov(CA_11_12_House_sim) + edgecov(CA_1

```

```
summary(CA_11_12_House_ergm)
```

Call:

```
ergm(formula = CA_11_12_House_net ~ edges + edgecov(CA_11_12_House_sim) +  
edgecov(CA_11_12_House_total_diff))
```

Maximum Likelihood Results:

	Estimate	Std. Error	MCMC %	z value	Pr(> z)
edges	-8.1738	0.1180	0	-69.251	<1e-04
edgecov.CA_11_12_House_sim		Inf	0	Inf	<1e-04
edgecov.CA_11_12_House_total_diff	2.6564	0.3492	0	7.608	<1e-04

edges ***
edgecov.CA_11_12_House_sim ***
edgecov.CA_11_12_House_total_diff ***

Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

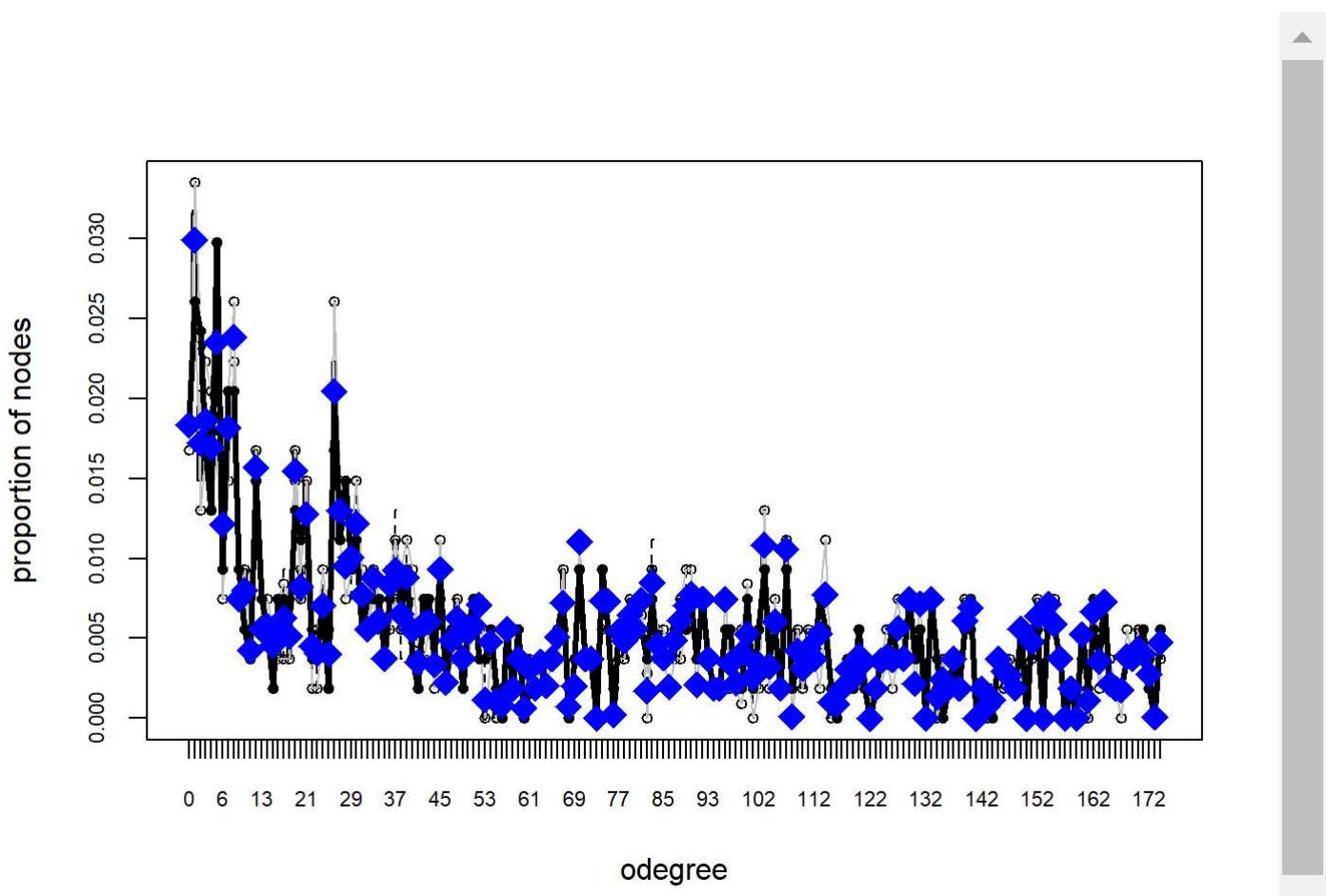
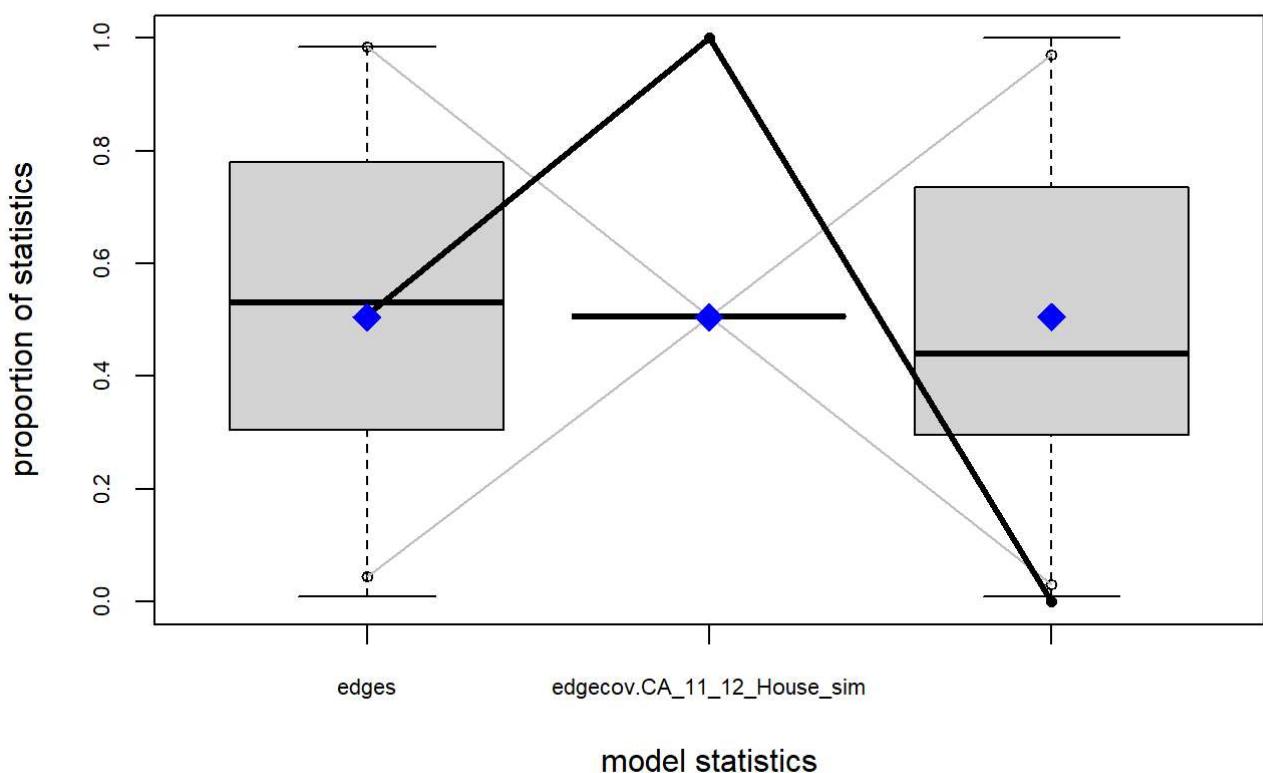
Null Deviance: 379326 on 273626 degrees of freedom
Residual Deviance: 1297 on 273623 degrees of freedom

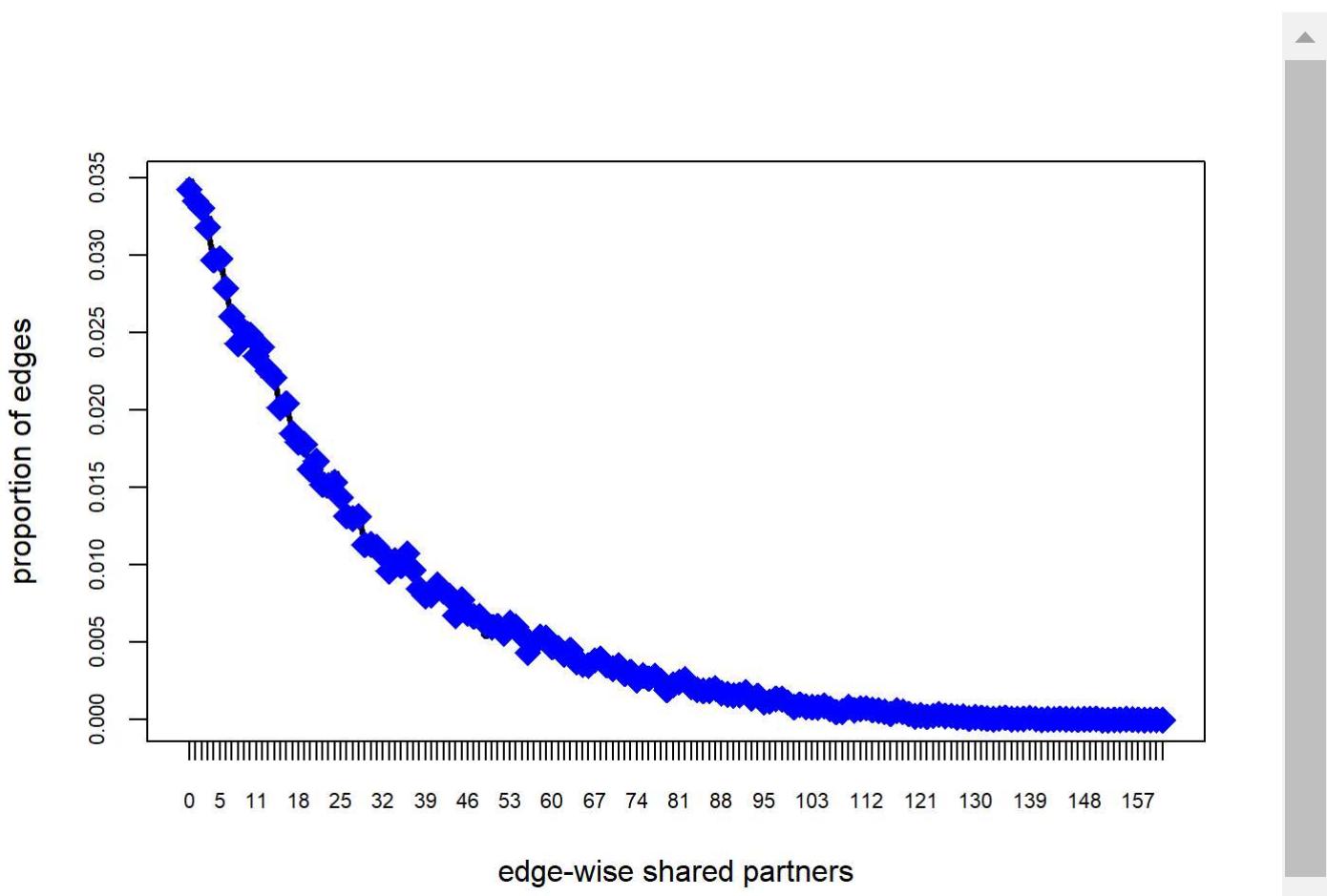
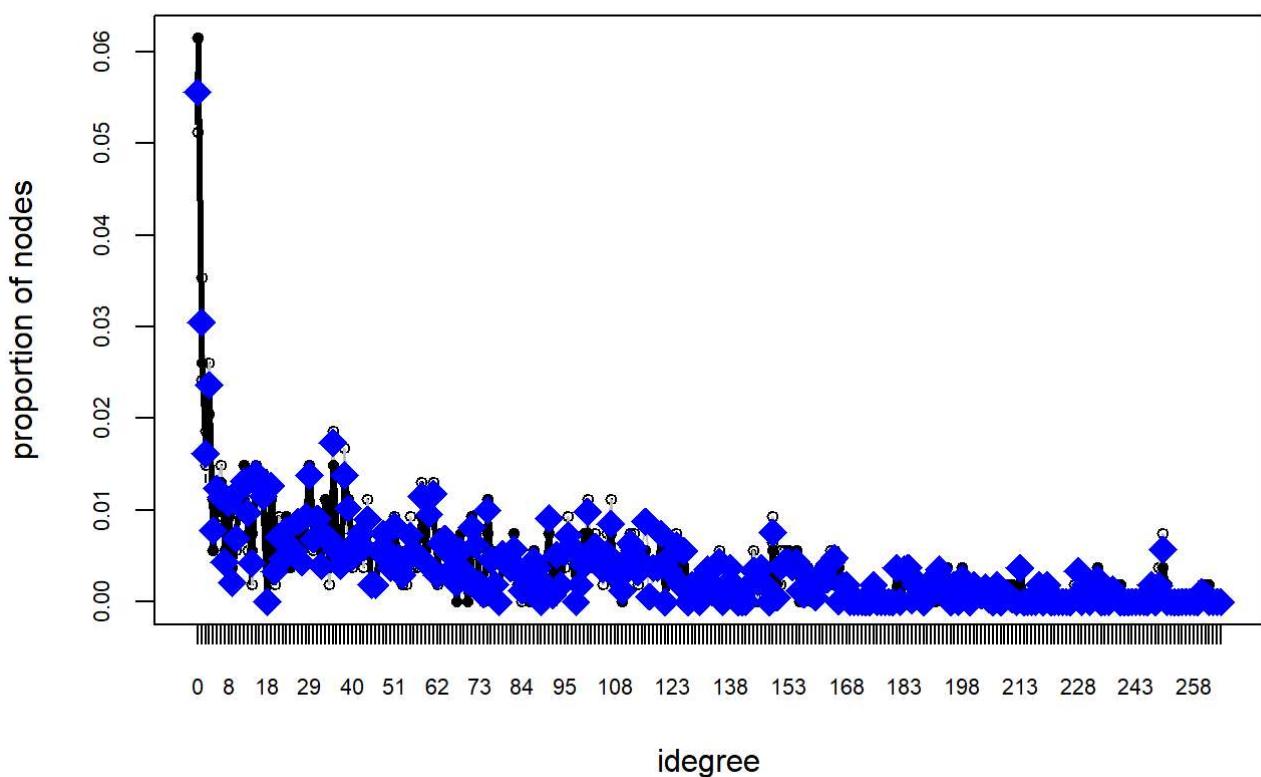
AIC: 1301 BIC: 1322 (Smaller is better. MC Std. Err. = 0)

Warning: The following terms have infinite coefficient estimates:
edgecov.CA_11_12_House_sim

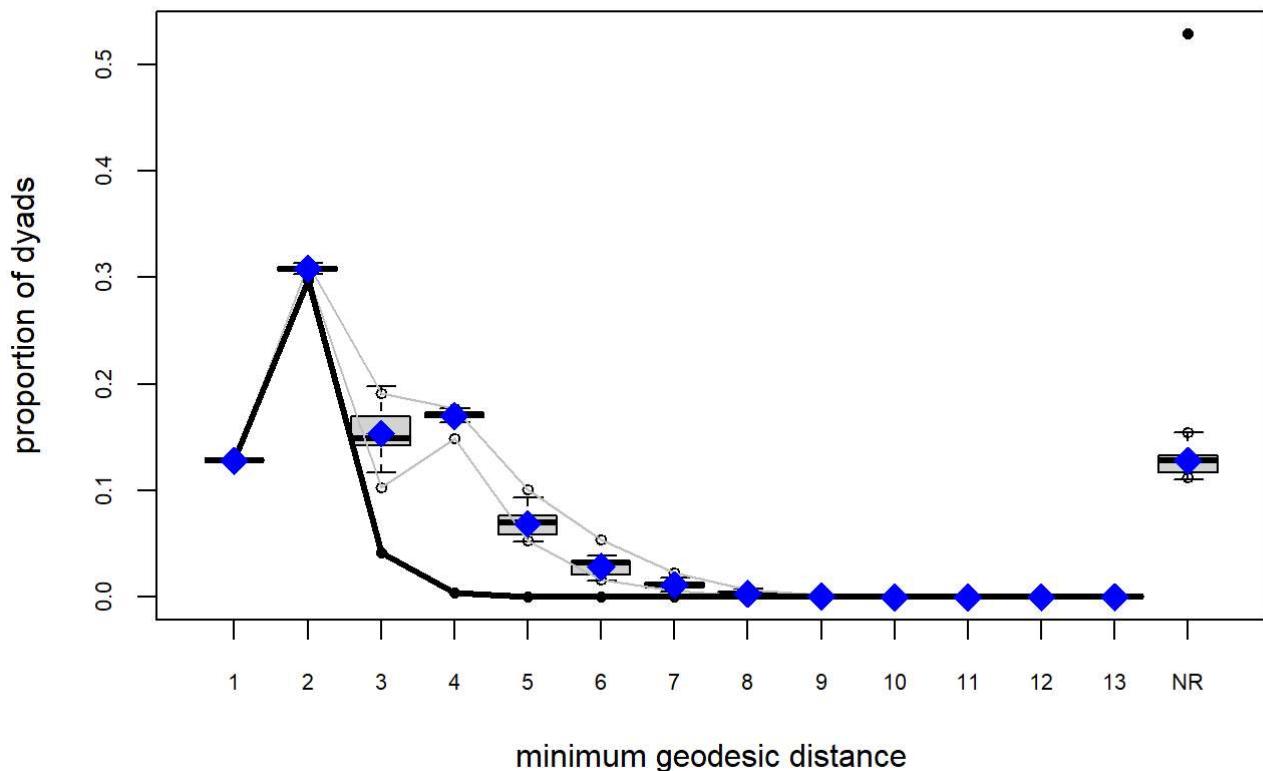
```
gof_CA_11_12_House_ergm <- gof(CA_11_12_House_ergm)
```

```
plot(gof_CA_11_12_House_ergm)
```





Goodness-of-fit diagnostics



```
CA_13_14_House_net <- as.network(ERGM_data(CA_2013_2014_House, CA_2013_2014_House_edges))

CA_13_14_House_ergm <- ergm(CA_13_14_House_net ~ edges + edgecov(CA_13_14_House_sim) + edgecov(CA_1

summary(CA_13_14_House_ergm)
```

Call:

```
ergm(formula = CA_13_14_House_net ~ edges + edgecov(CA_13_14_House_sim) +
  edgecov(CA_13_14_House_total_diff))
```

Maximum Likelihood Results:

	Estimate	Std. Error	MCMC %	z value	Pr(> z)
edges	-8.4375	0.1443	0	-58.460	< 1e-04
edgecov.CA_13_14_House_sim	Inf	0.0000	0	Inf	< 1e-04
edgecov.CA_13_14_House_total_diff	0.6676	0.2194	0	3.042	0.00235

edges ***

edgecov.CA_13_14_House_sim ***

edgecov.CA_13_14_House_total_diff **

Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

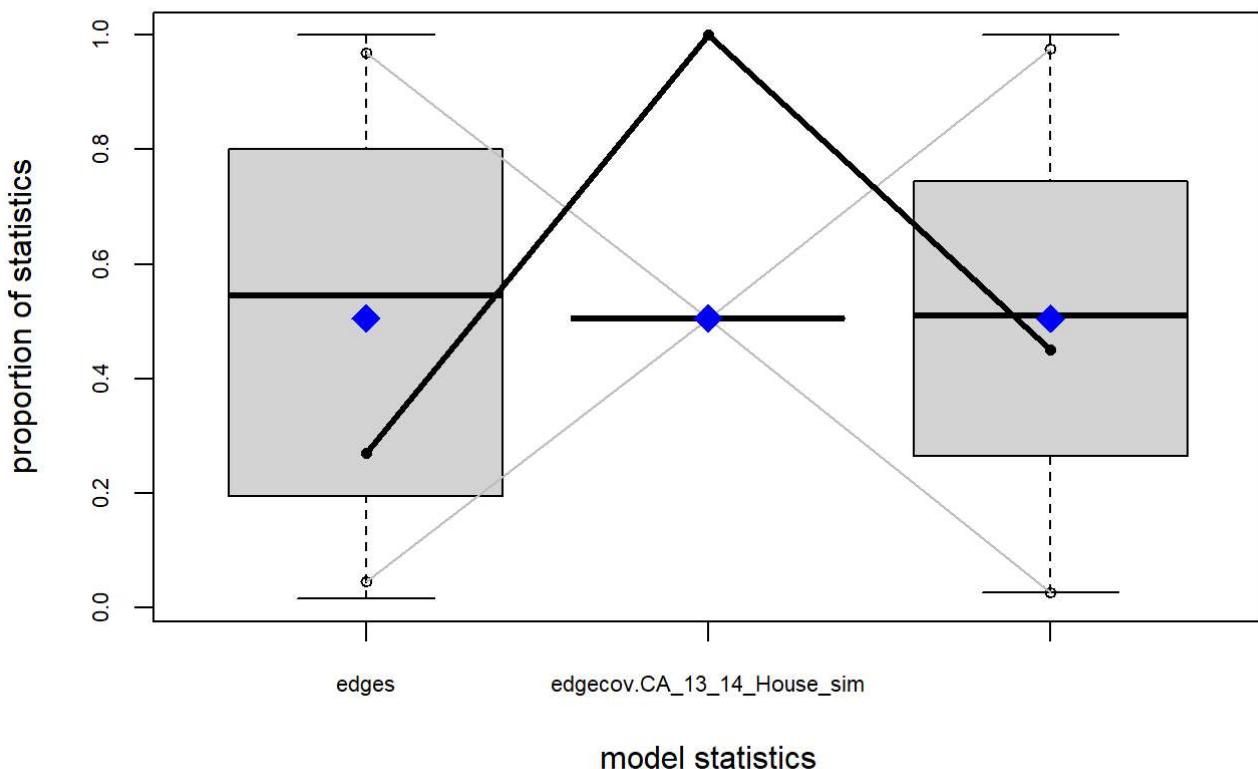
```
Null Deviance: 347119.8 on 250394 degrees of freedom
Residual Deviance: 906.1 on 250391 degrees of freedom
```

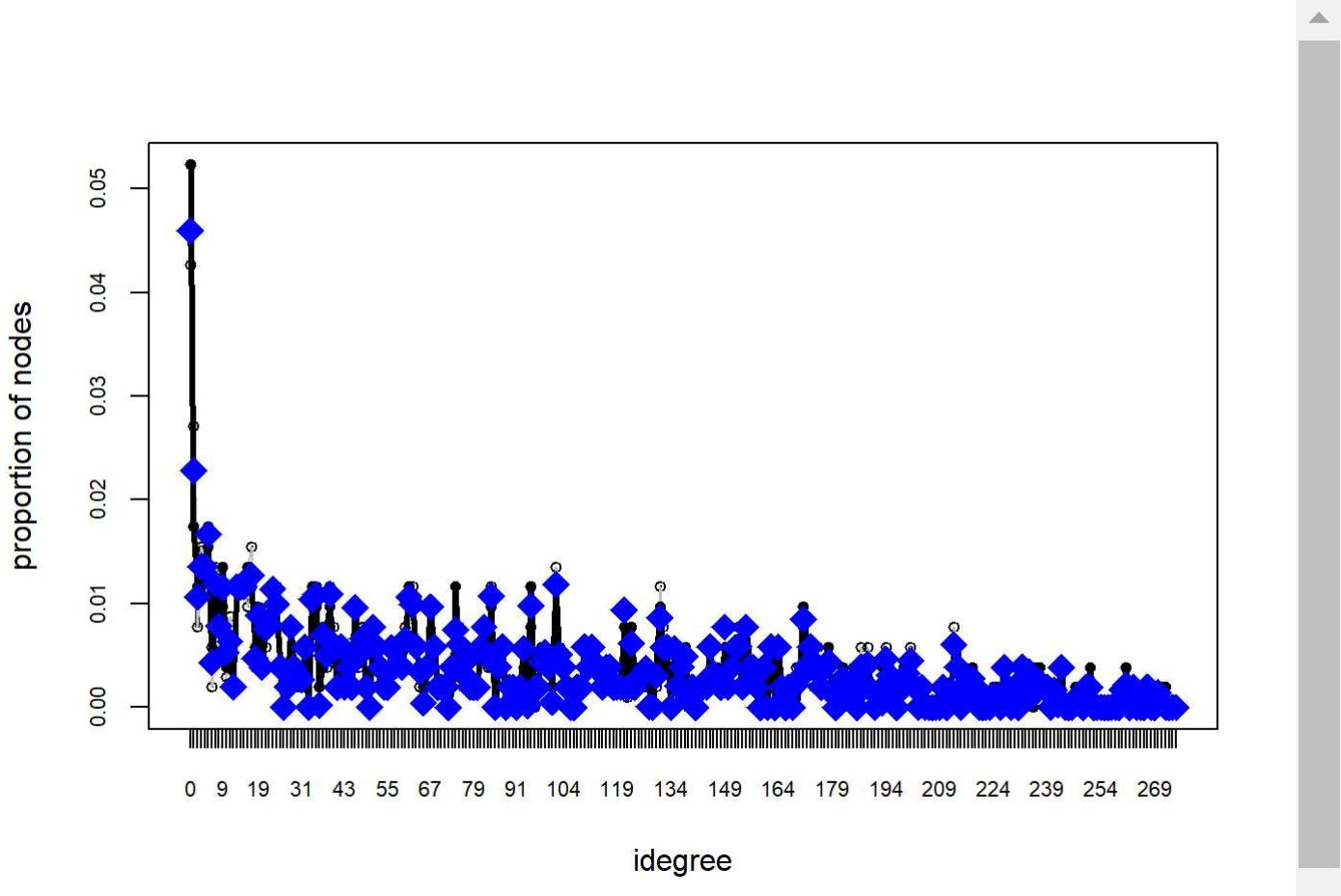
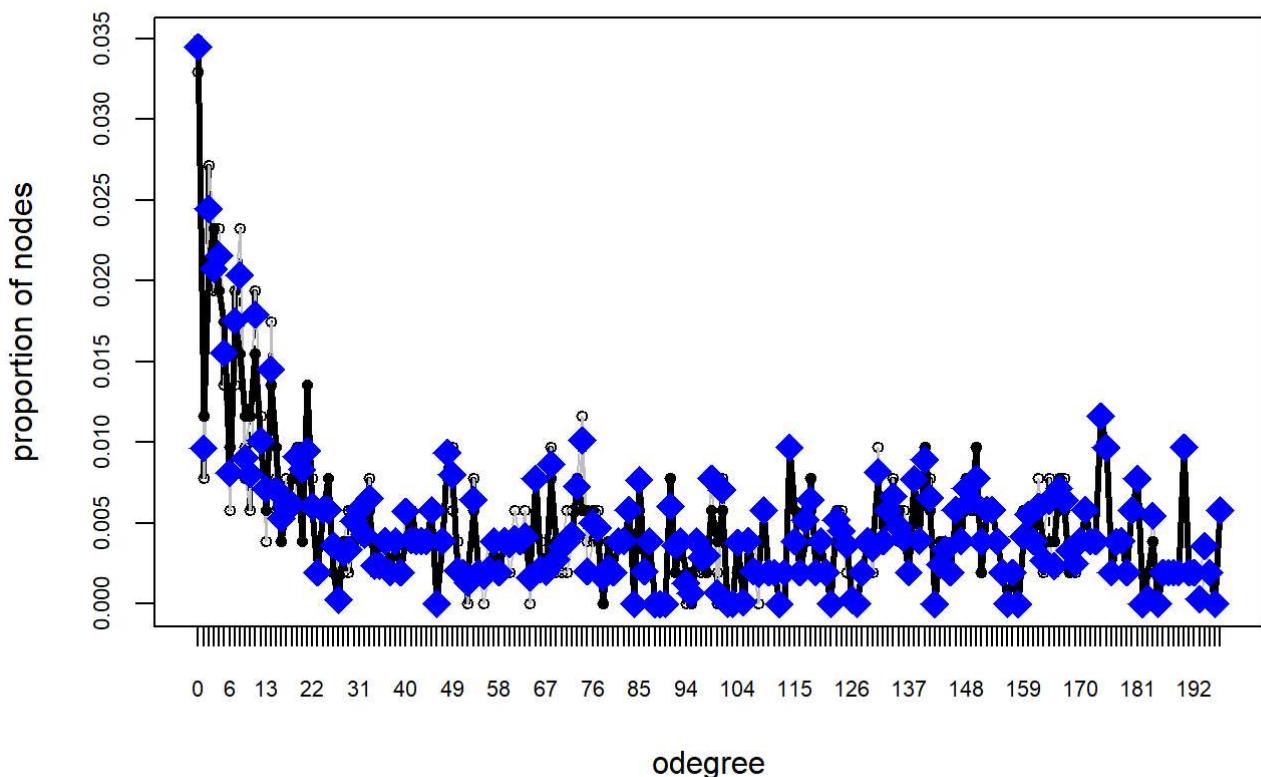
```
AIC: 910.1 BIC: 930.9 (Smaller is better. MC Std. Err. = 0)
```

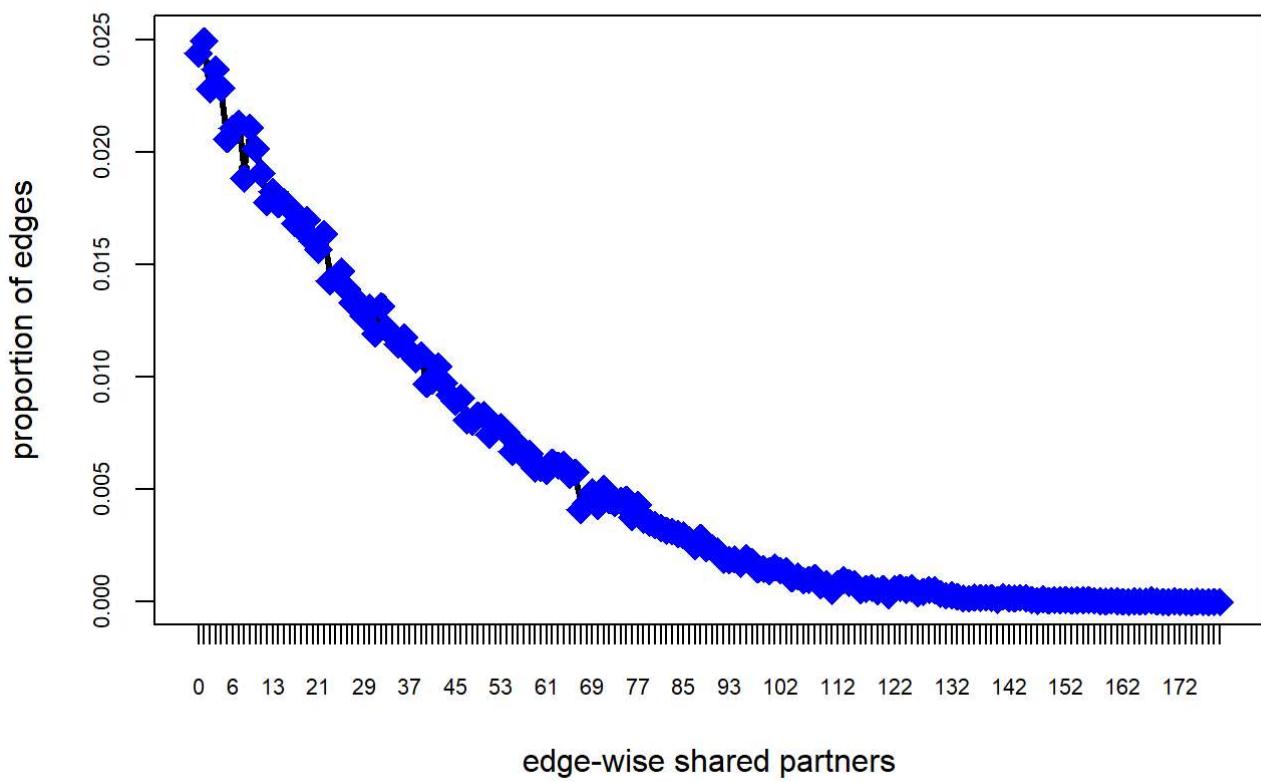
```
Warning: The following terms have infinite coefficient estimates:
edgecov.CA_13_14_House_sim
```

```
gof_CA_13_14_House_ergm <- gof(CA_13_14_House_ergm)
```

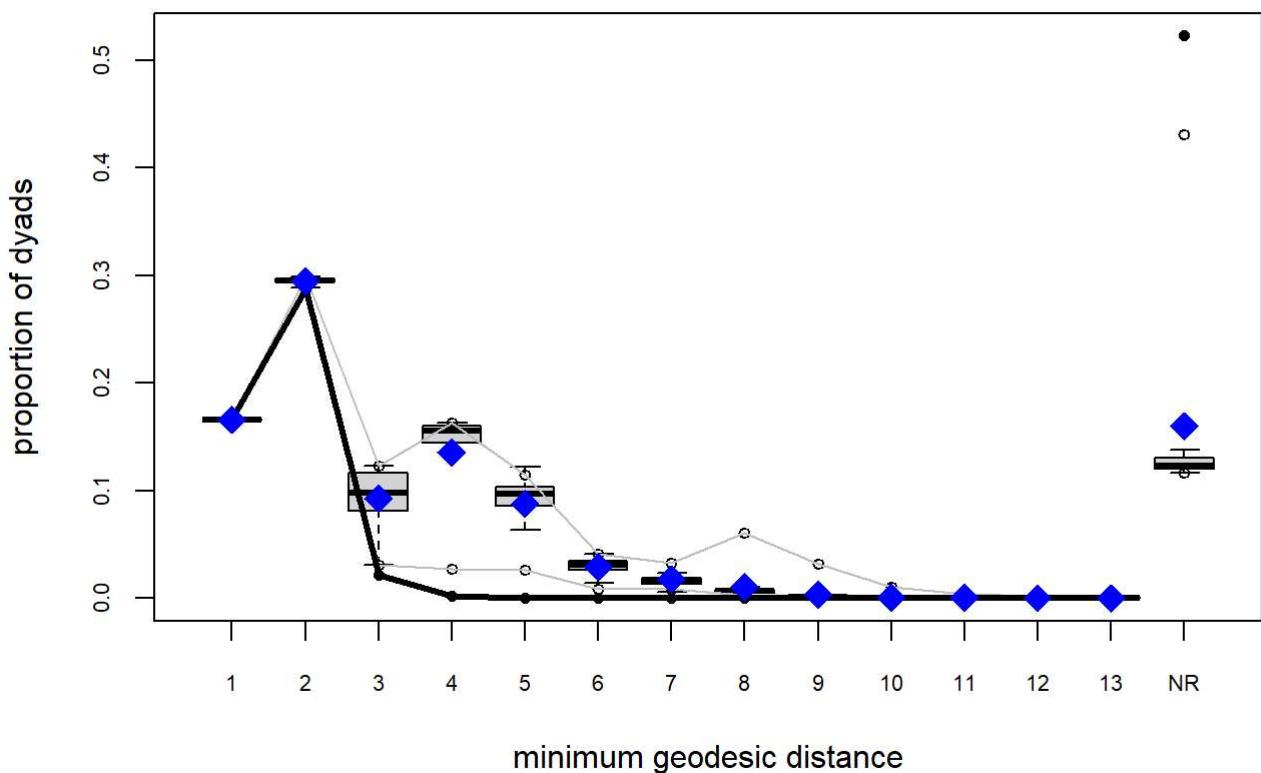
```
plot(gof_CA_13_14_House_ergm)
```







Goodness-of-fit diagnostics



```
CA_15_16_House_net <- as.network(ERGM_data(CA_2015_2016_House, CA_2015_2016_House_edges))

CA_15_16_House_ergm <- ergm(CA_15_16_House_net~edges + edgecov(CA_15_16_House_sim) + edgecov(CA_1

summary(CA_15_16_House_ergm)
```

Call:

```
ergm(formula = CA_15_16_House_net ~ edges + edgecov(CA_15_16_House_sim) +
  edgecov(CA_15_16_House_total_diff))
```

Maximum Likelihood Results:

	Estimate	Std. Error	MCMC %	z value	Pr(> z)
edges	-8.4461	0.1854	0	-45.555	<1e-04
edgecov.CA_15_16_House_sim	Inf	0.0000	0	Inf	<1e-04
edgecov.CA_15_16_House_total_diff	1.0873	0.2310	0	4.708	<1e-04

edges ***
edgecov.CA_15_16_House_sim ***
edgecov.CA_15_16_House_total_diff ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

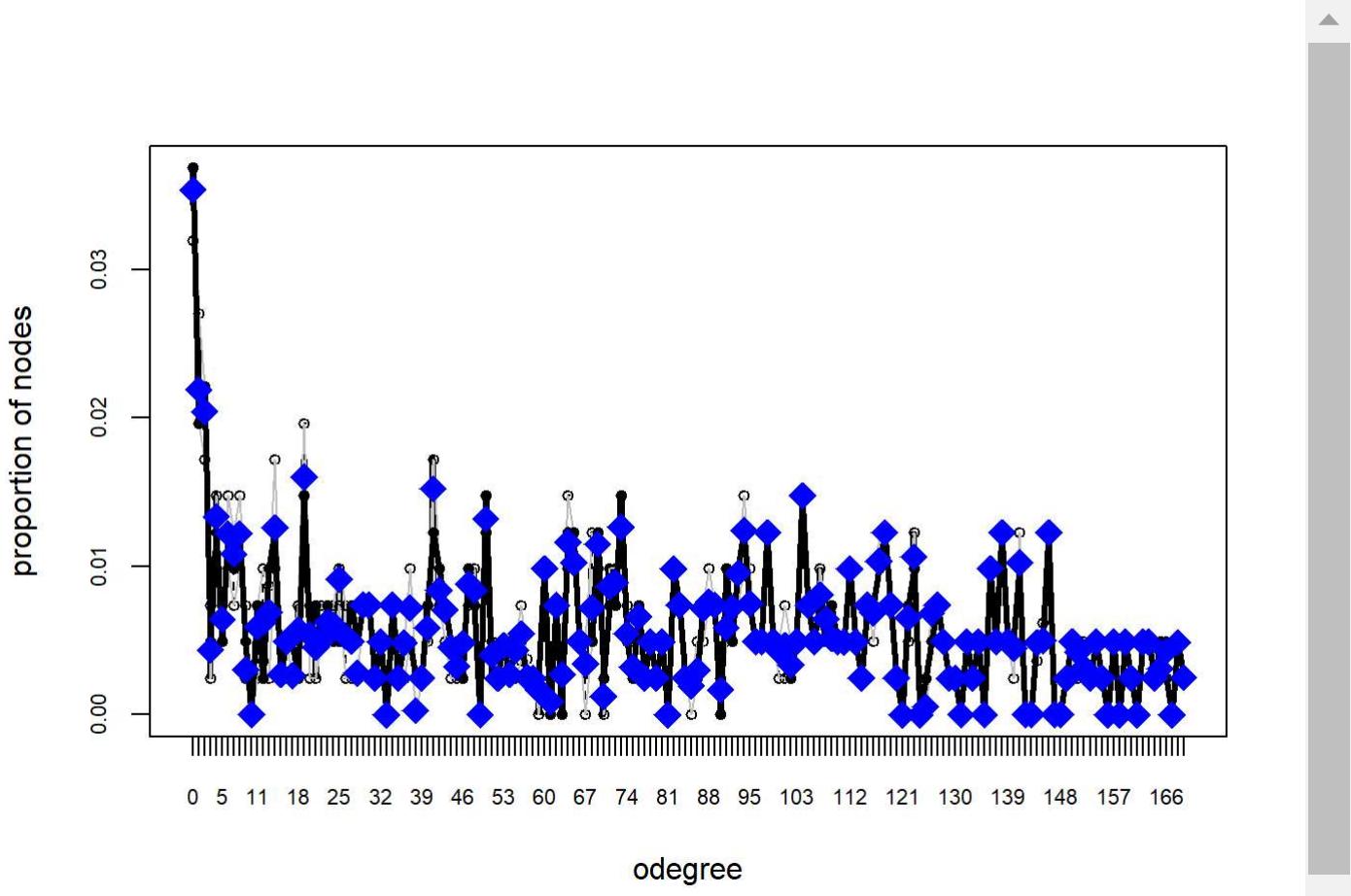
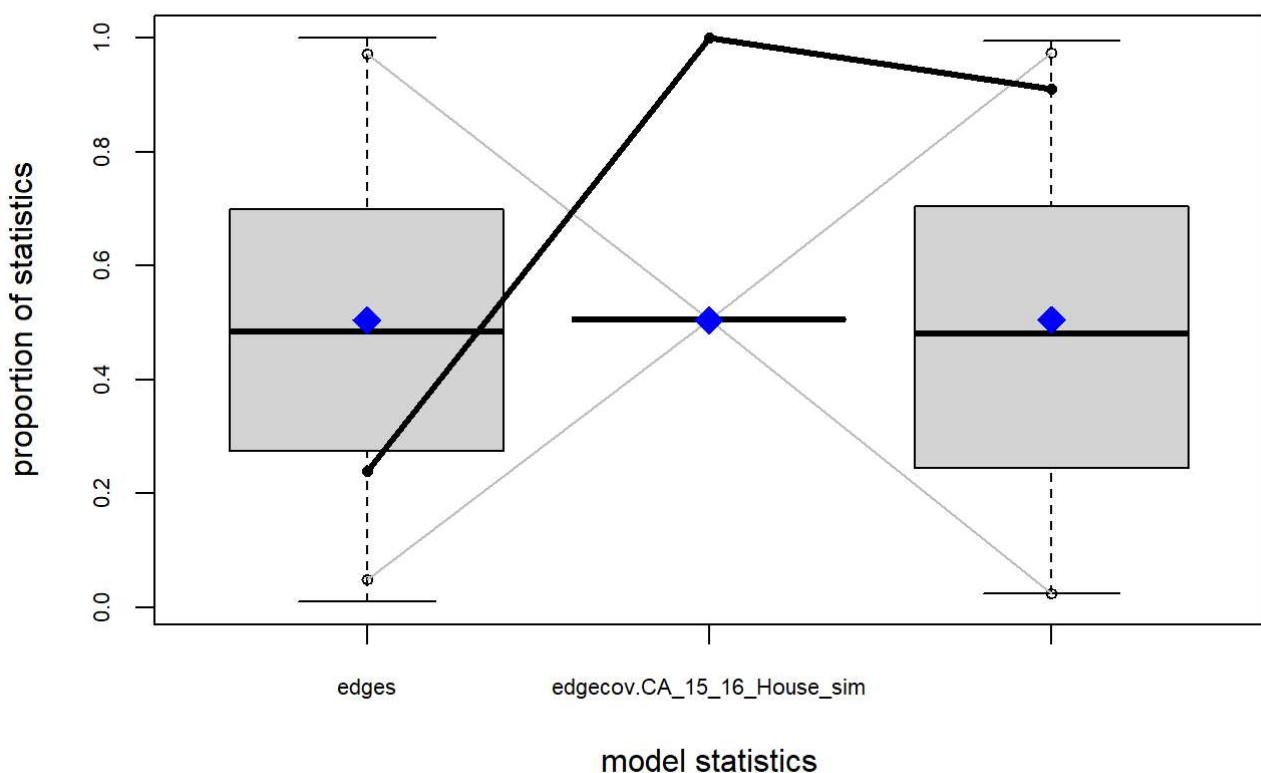
Null Deviance: 214460 on 154700 degrees of freedom
Residual Deviance: 548 on 154697 degrees of freedom

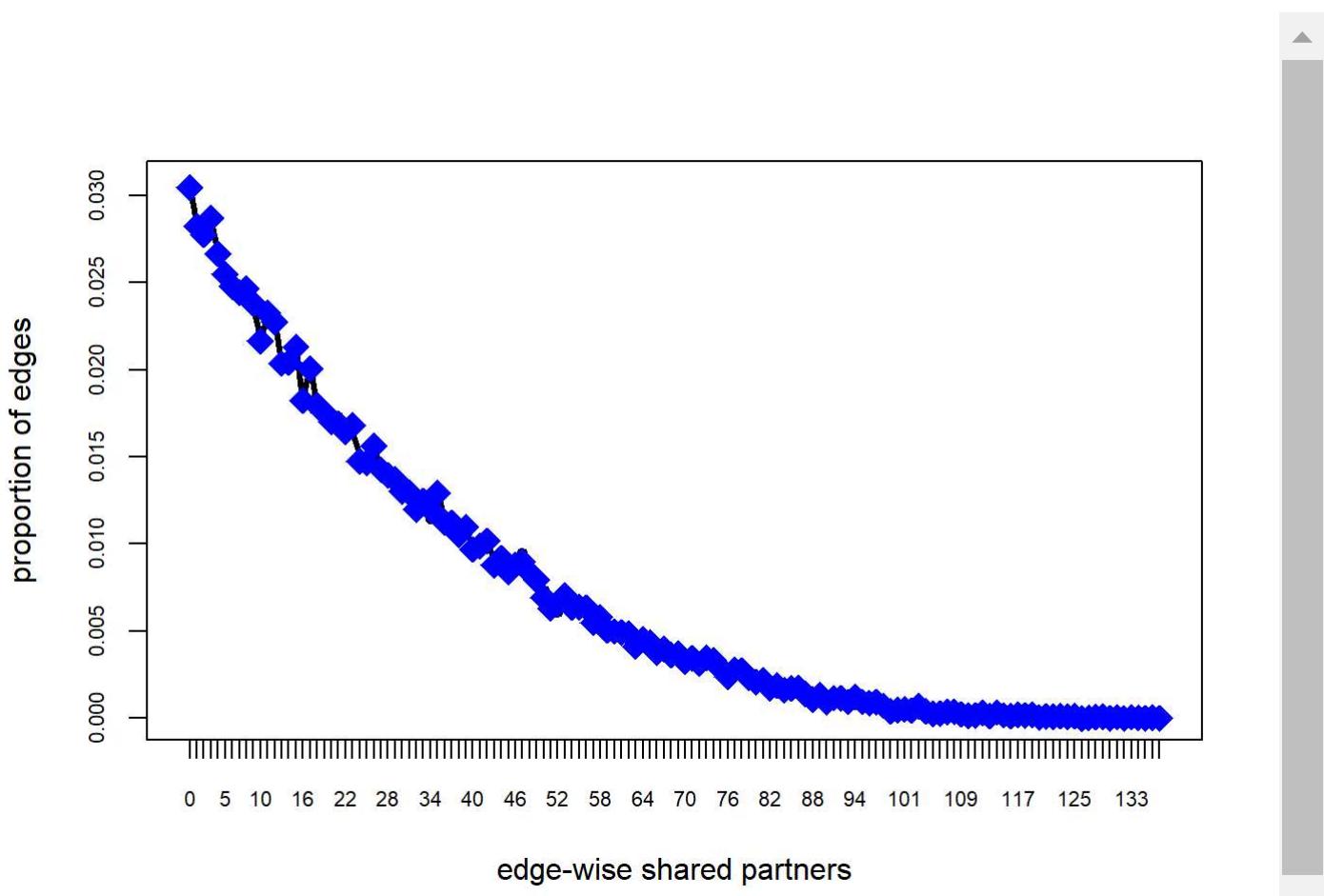
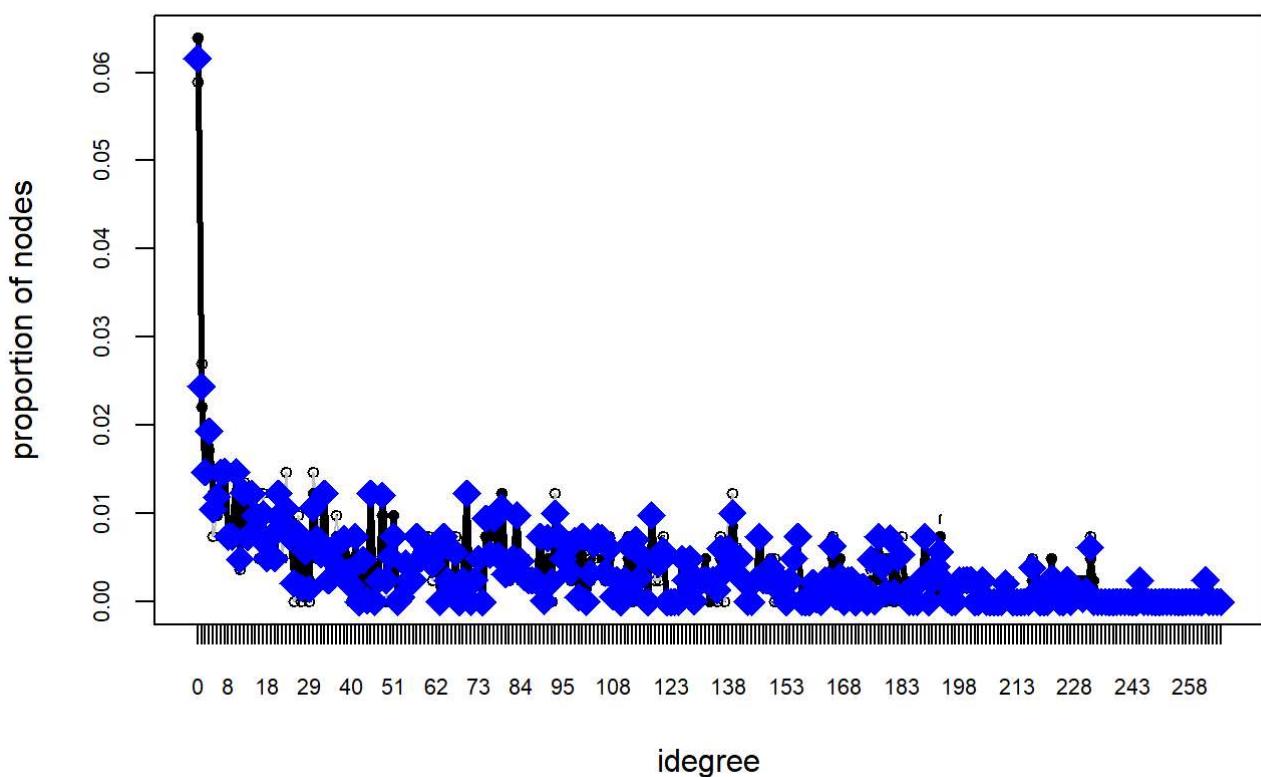
AIC: 552 BIC: 571.9 (Smaller is better. MC Std. Err. = 0)

Warning: The following terms have infinite coefficient estimates:
edgecov.CA_15_16_House_sim

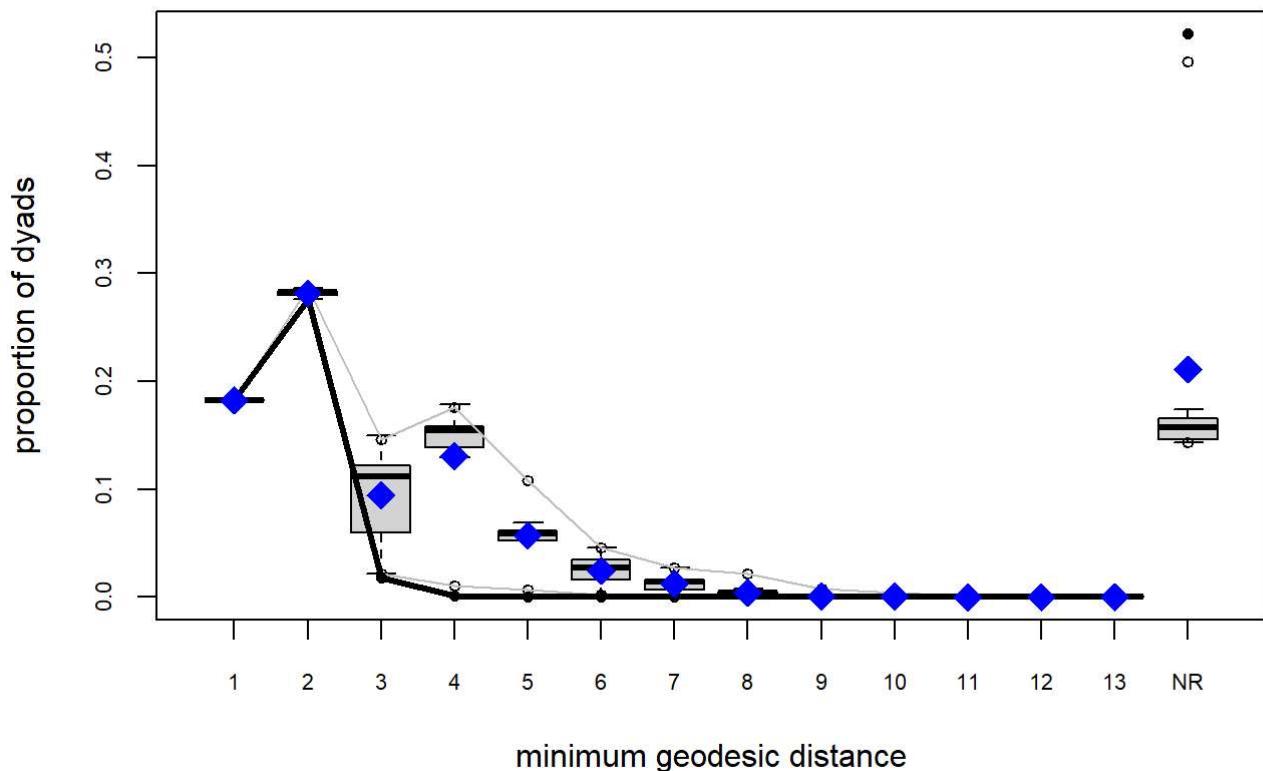
```
gof_CA_15_16_House_ergm <- gof(CA_15_16_House_ergm)
```

```
plot(gof_CA_15_16_House_ergm)
```





Goodness-of-fit diagnostics



```
CA_11_12_Senate_net <- as.network(ERGM_data(CA_2011_2012_Senate, CA_2011_2012_Senate_edges))

CA_11_12_Senate_ergm <- ergm(CA_11_12_Senate_net ~ edges + edgecov(CA_11_12_Senate_sim) + edgecov(CA_11_12_Senate_total_diff))

summary(CA_11_12_Senate_ergm)
```

Call:
`ergm(formula = CA_11_12_Senate_net ~ edges + edgecov(CA_11_12_Senate_sim) + edgecov(CA_11_12_Senate_total_diff))`

Maximum Likelihood Results:

	Estimate	Std. Error	MCMC %	z value	Pr(> z)
edges	-6.67874	0.09817	0	-68.031	<1e-04
edgecov.CA_11_12_Senate_sim	Inf	0.00000	0	Inf	<1e-04
edgecov.CA_11_12_Senate_total_diff	27.89942	4.15477	0	6.715	<1e-04

edges ***
`edgecov.CA_11_12_Senate_sim ***`
`edgecov.CA_11_12_Senate_total_diff ***`

Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

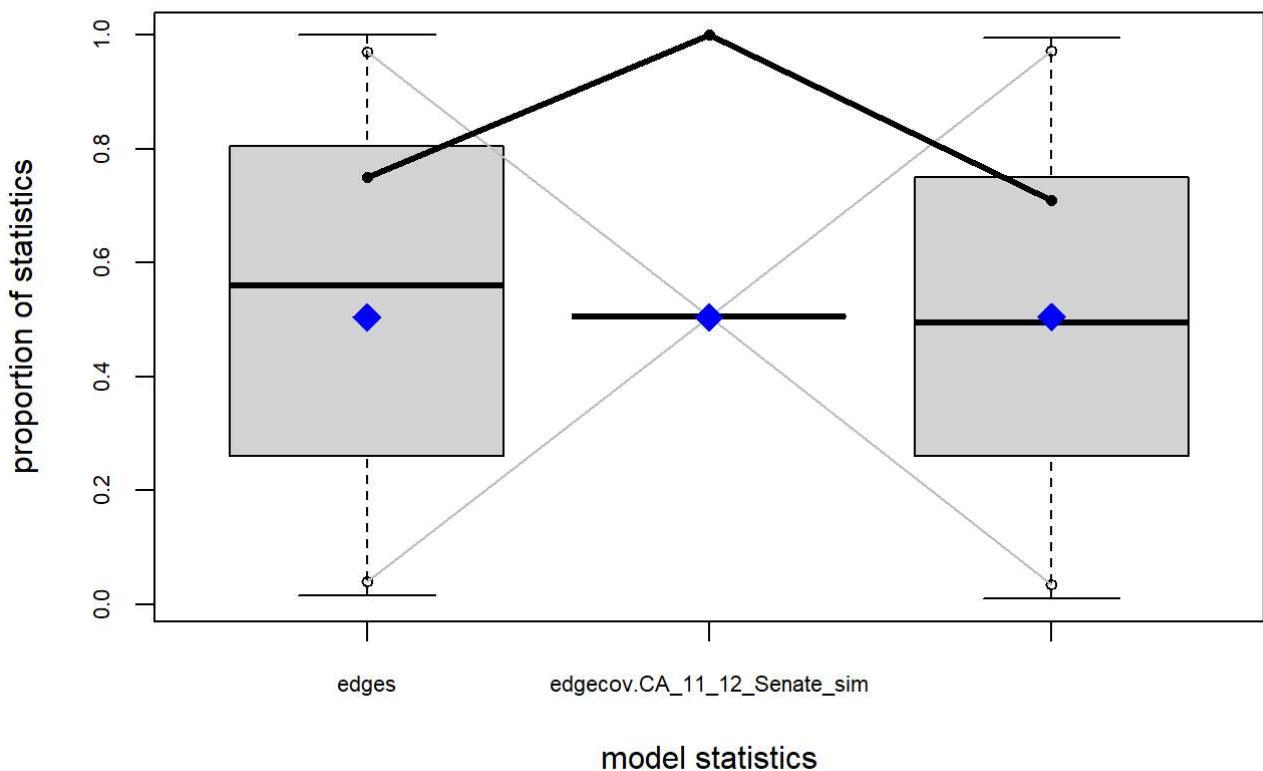
```
Null Deviance: 138401  on 99835  degrees of freedom
Residual Deviance:    1583  on 99832  degrees of freedom
```

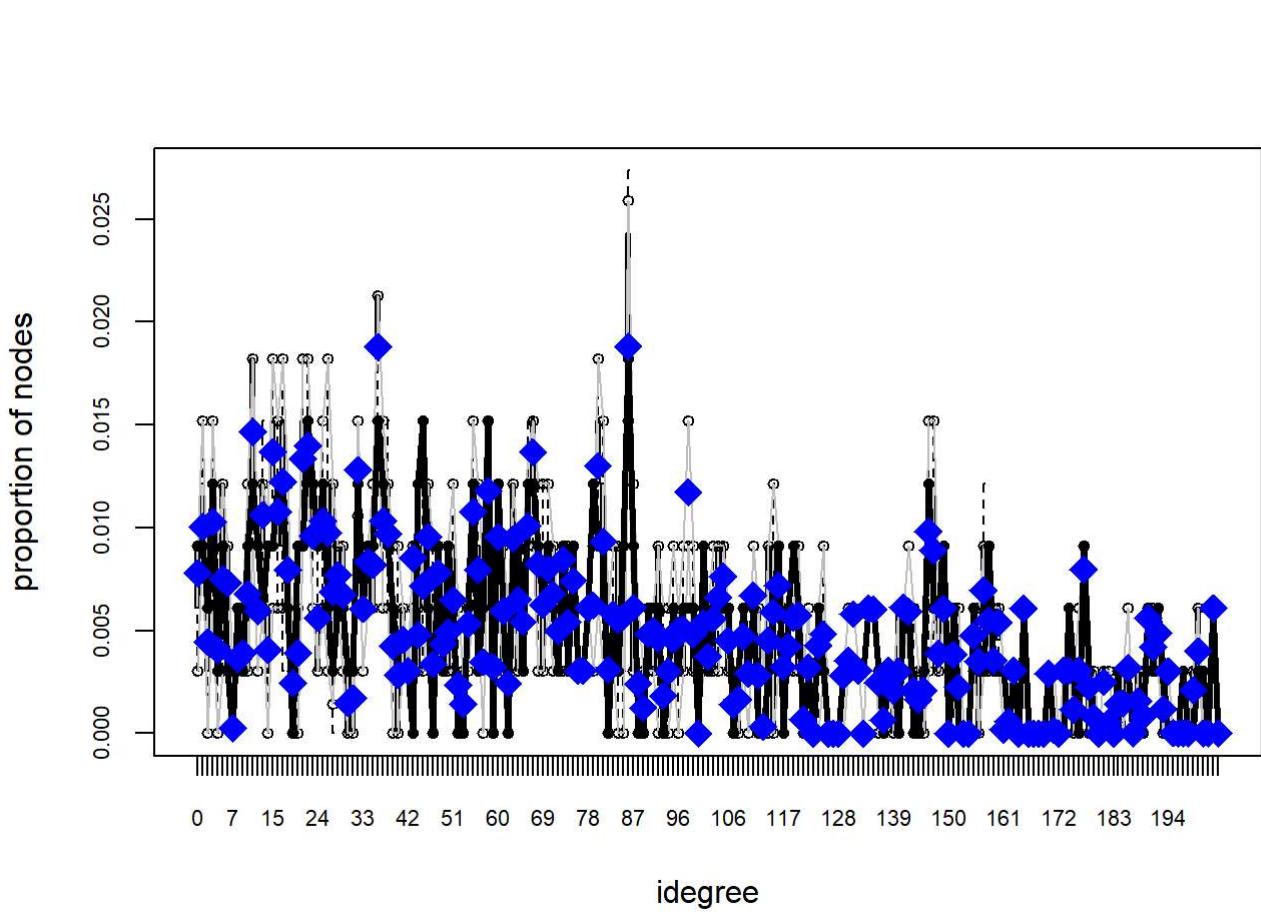
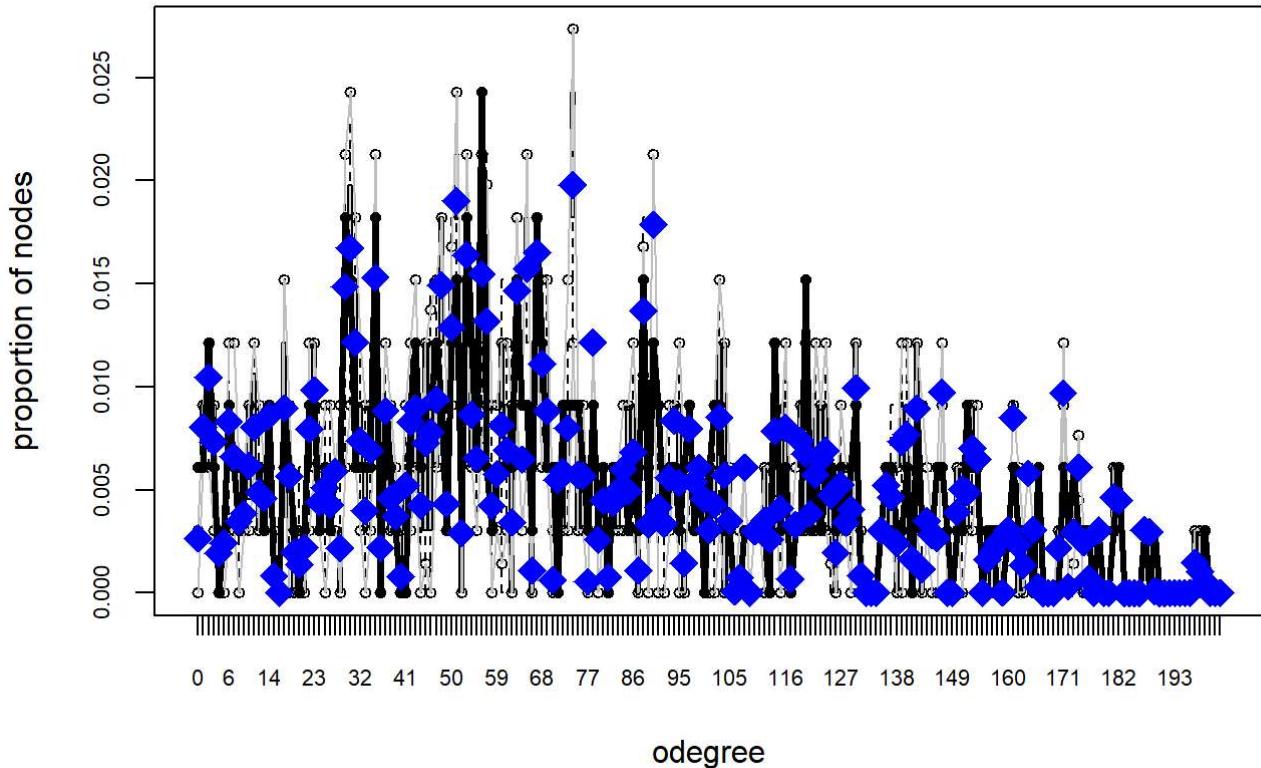
```
AIC: 1587  BIC: 1606  (Smaller is better. MC Std. Err. = 0)
```

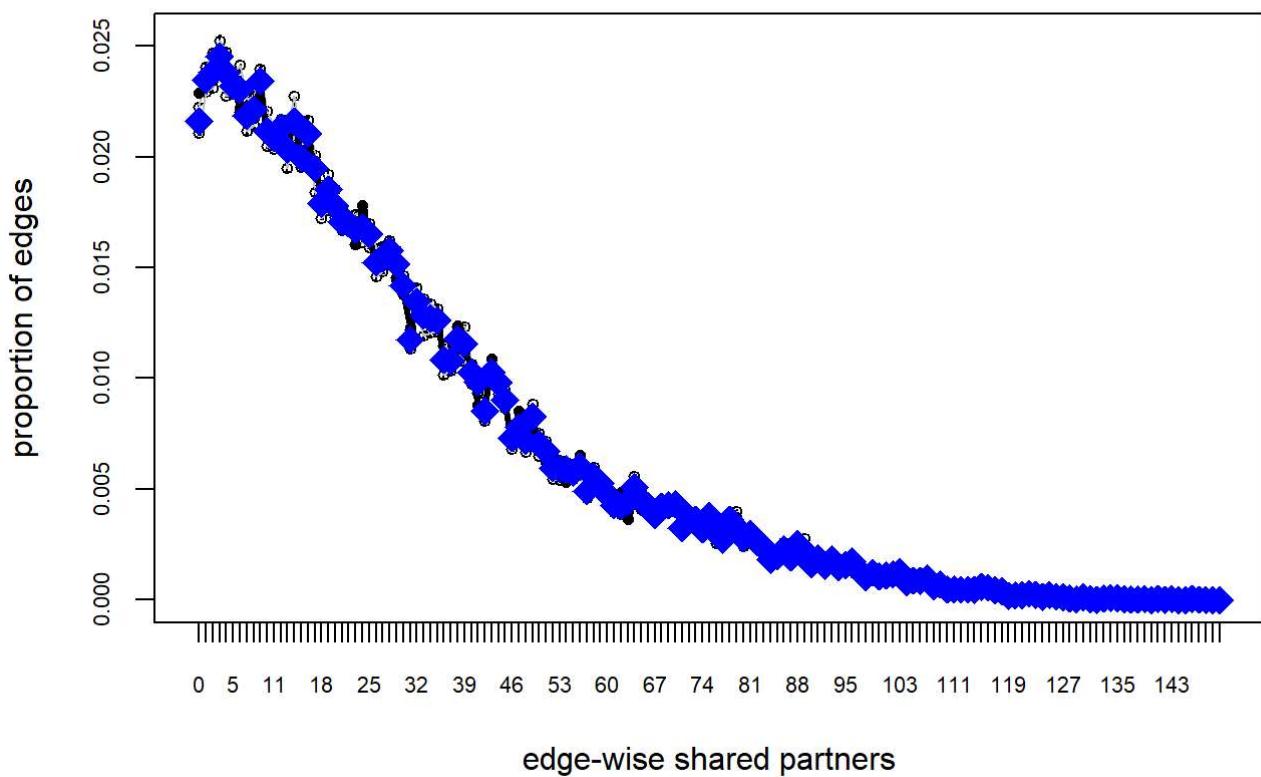
```
Warning: The following terms have infinite coefficient estimates:
edgecov.CA_11_12_Senate_sim
```

```
gof_CA_11_12_Senate_ergm <- gof(CA_11_12_Senate_ergm)
```

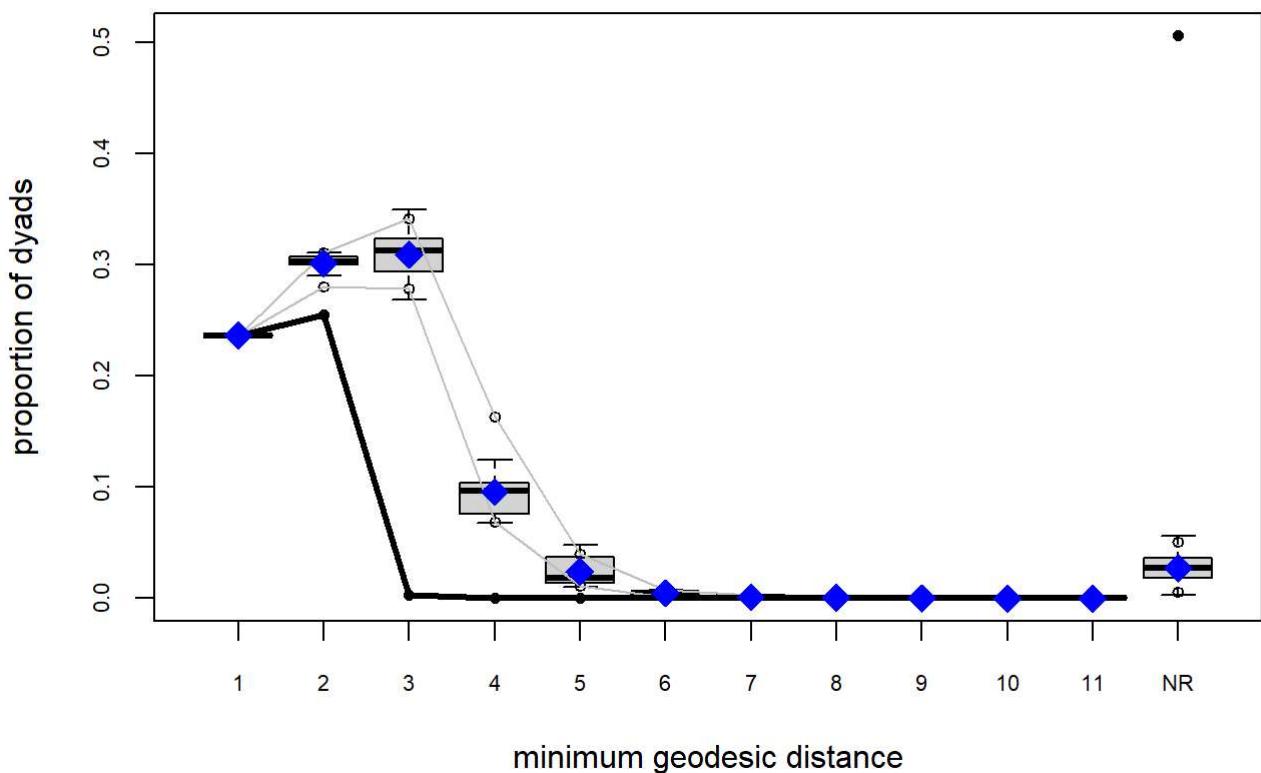
```
plot(gof_CA_11_12_Senate_ergm)
```







Goodness-of-fit diagnostics



```
CA_13_14_Senate_net <- as.network(ERGM_data(CA_2013_2014_Senate, CA_2013_2014_Senate_edges))

CA_13_14_Senate_ergm <- ergm(CA_13_14_Senate_net~edges + edgecov(CA_13_14_Senate_sim) + edgecov(CA_13_14_Senate_total_diff))

summary(CA_13_14_Senate_ergm)
```

Call:

```
ergm(formula = CA_13_14_Senate_net ~ edges + edgecov(CA_13_14_Senate_sim) +
  edgecov(CA_13_14_Senate_total_diff))
```

Maximum Likelihood Results:

	Estimate	Std. Error	MCMC %	z value	Pr(> z)	
edges	-6.46876	0.08549	0	-75.667	<1e-04	
edgecov.CA_13_14_Senate_sim		Inf	0.00000	0	Inf	<1e-04
edgecov.CA_13_14_Senate_total_diff	6.20811	9.90832	0	0.627	0.531	

edges ***
edgecov.CA_13_14_Senate_sim ***
edgecov.CA_13_14_Senate_total_diff

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

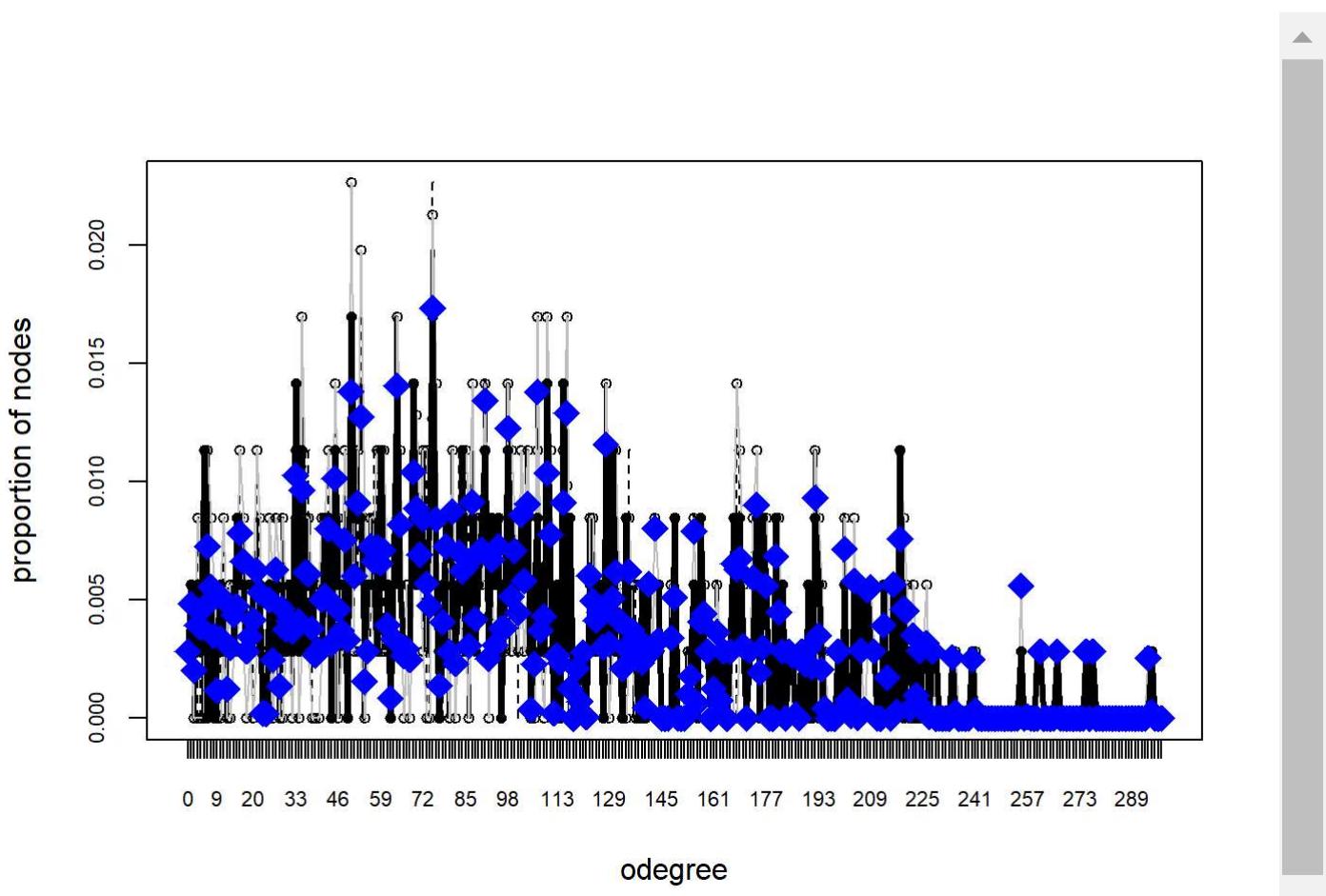
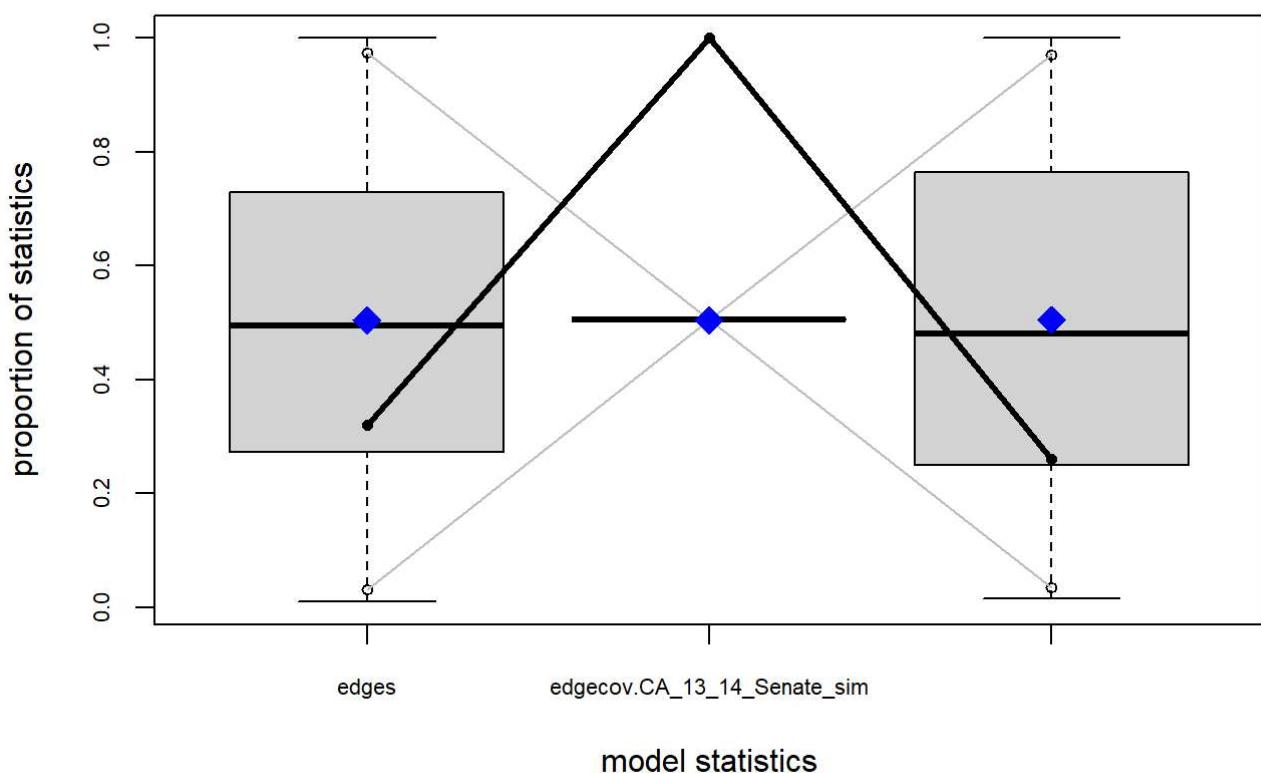
Null Deviance: 156585 on 112952 degrees of freedom
Residual Deviance: 2047 on 112949 degrees of freedom

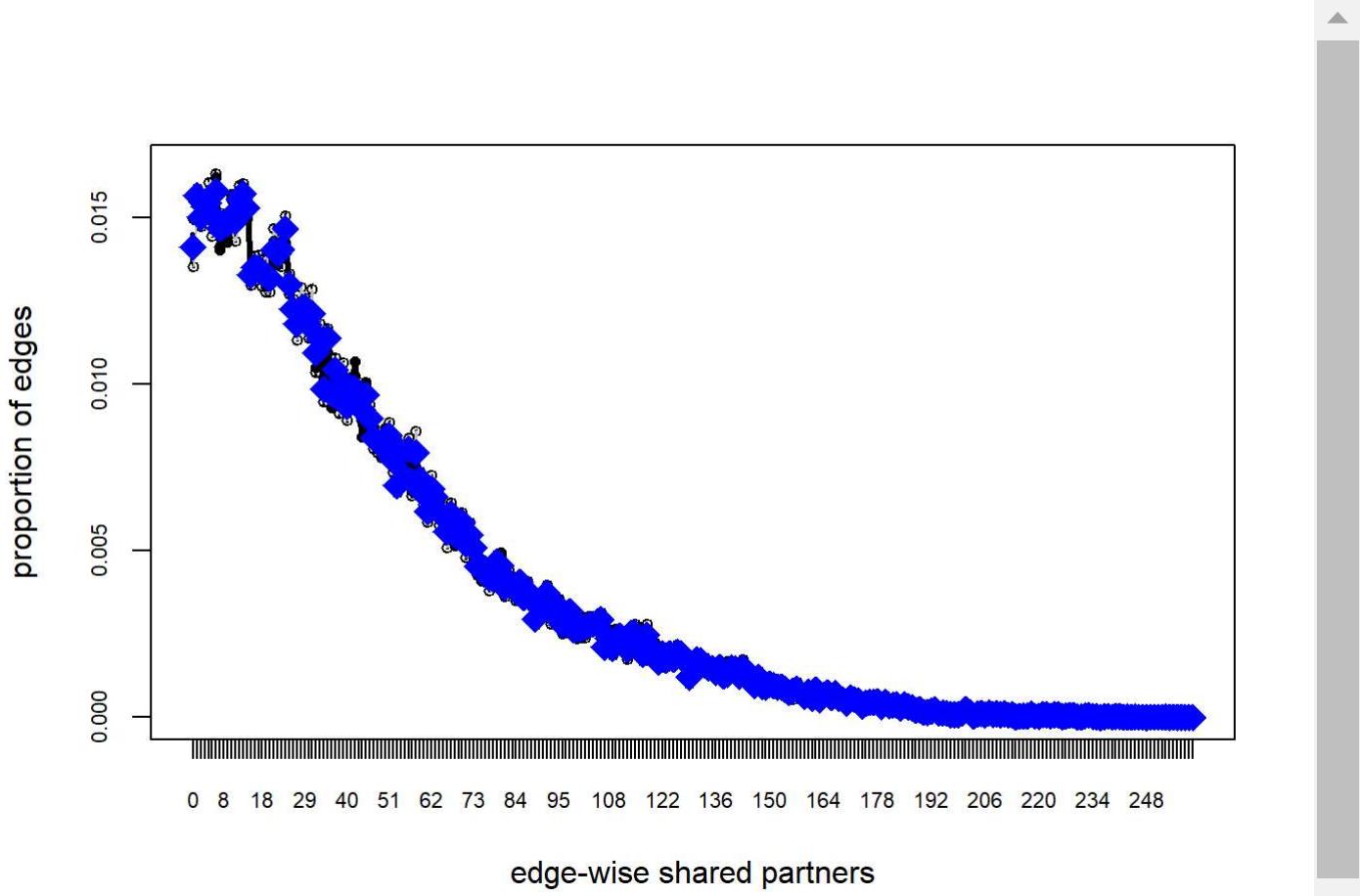
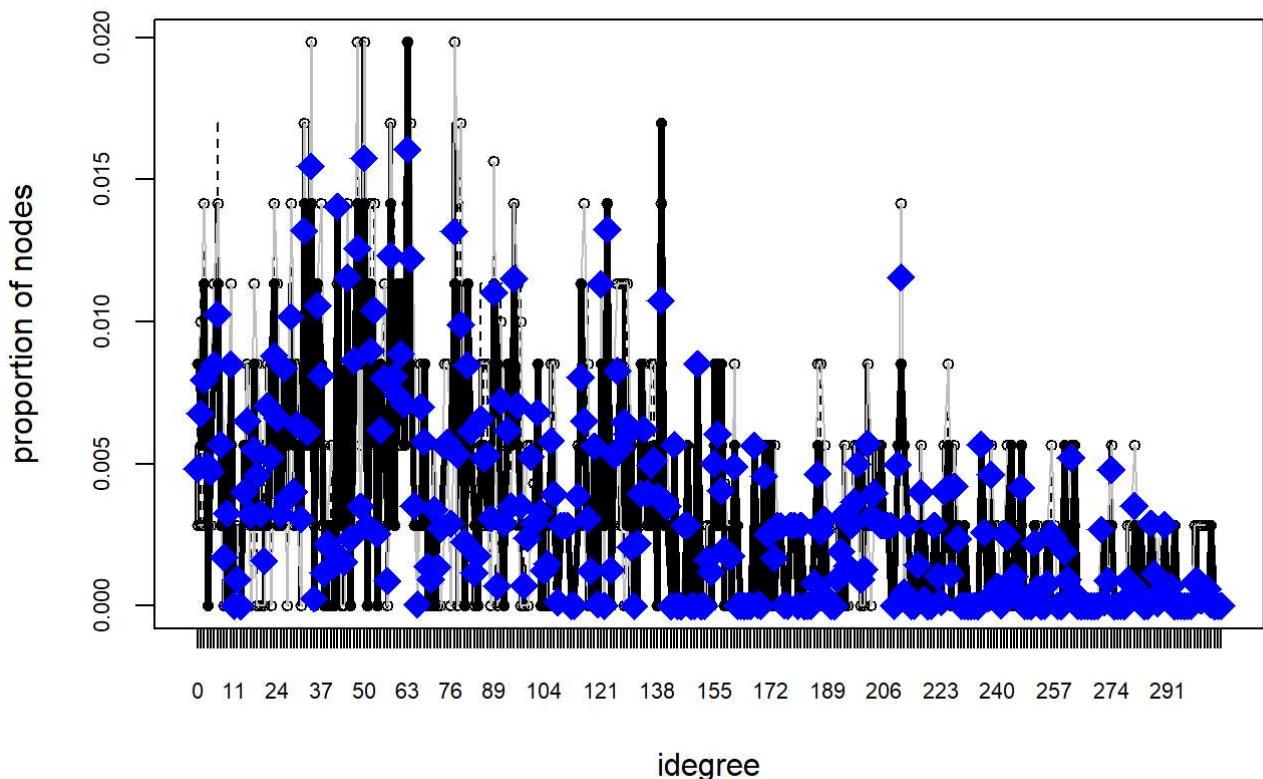
AIC: 2051 BIC: 2070 (Smaller is better. MC Std. Err. = 0)

Warning: The following terms have infinite coefficient estimates:
edgecov.CA_13_14_Senate_sim

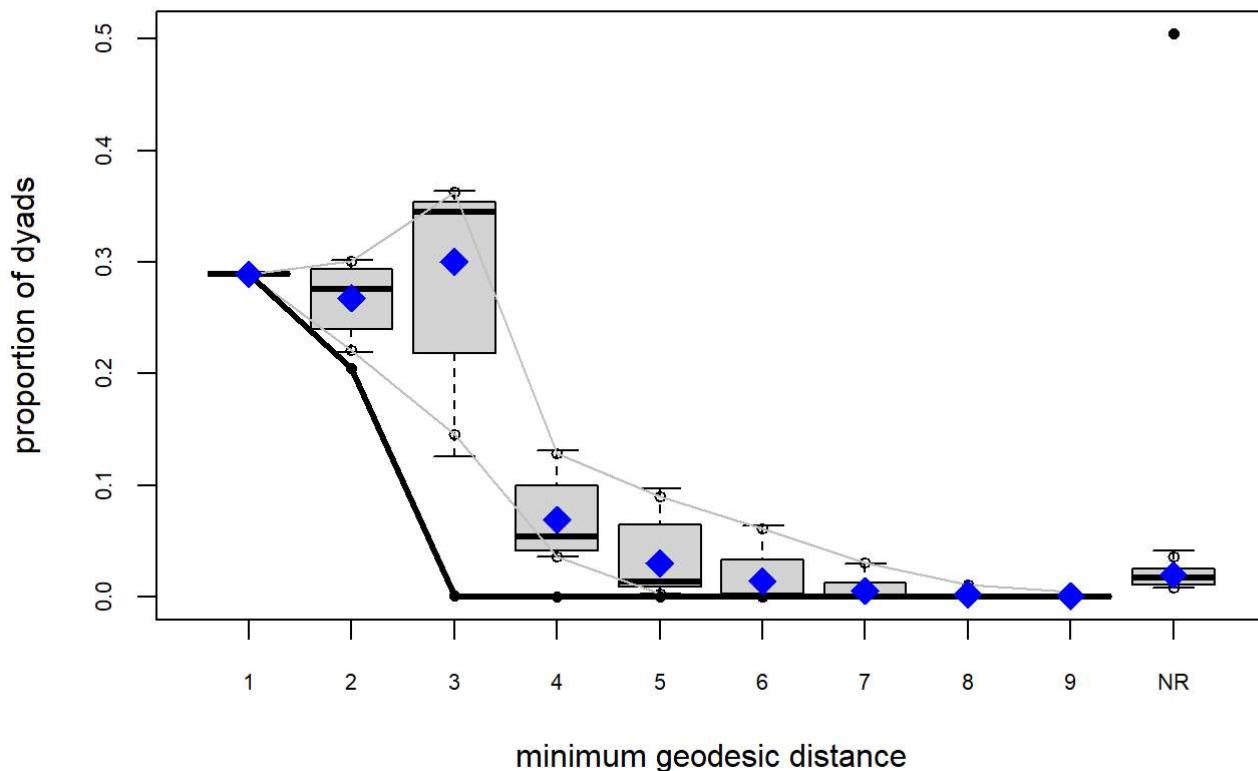
```
gof_CAI_13_14_Senate_ergm <- gof(CA_13_14_Senate_ergm)
```

```
plot(gof_CAI_13_14_Senate_ergm)
```





Goodness-of-fit diagnostics



```
CA_15_16_Senate_net <- as.network(ERGM_data(CA_2015_2016_Senate, CA_2015_2016_Senate_edges))

CA_15_16_Senate_ergm <- ergm(CA_15_16_Senate_net~edges + edgecov(CA_15_16_Senate_sim) + edgecov(CA_15_16_Senate_total_diff))

summary(CA_15_16_Senate_ergm)
```

Call:
`ergm(formula = CA_15_16_Senate_net ~ edges + edgecov(CA_15_16_Senate_sim) + edgecov(CA_15_16_Senate_total_diff))`

Maximum Likelihood Results:

	Estimate	Std. Error	MCMC %	z value	Pr(> z)
edges	-7.0087	0.1222	0	-57.349	< 1e-04
edgecov.CA_15_16_Senate_sim	Inf	0.0000	0	Inf	< 1e-04
edgecov.CA_15_16_Senate_total_diff	1.3889	0.3974	0	3.495	0.000474

edges ***
`edgecov.CA_15_16_Senate_sim ***`
`edgecov.CA_15_16_Senate_total_diff ***`
`---`
Signif. codes: 0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

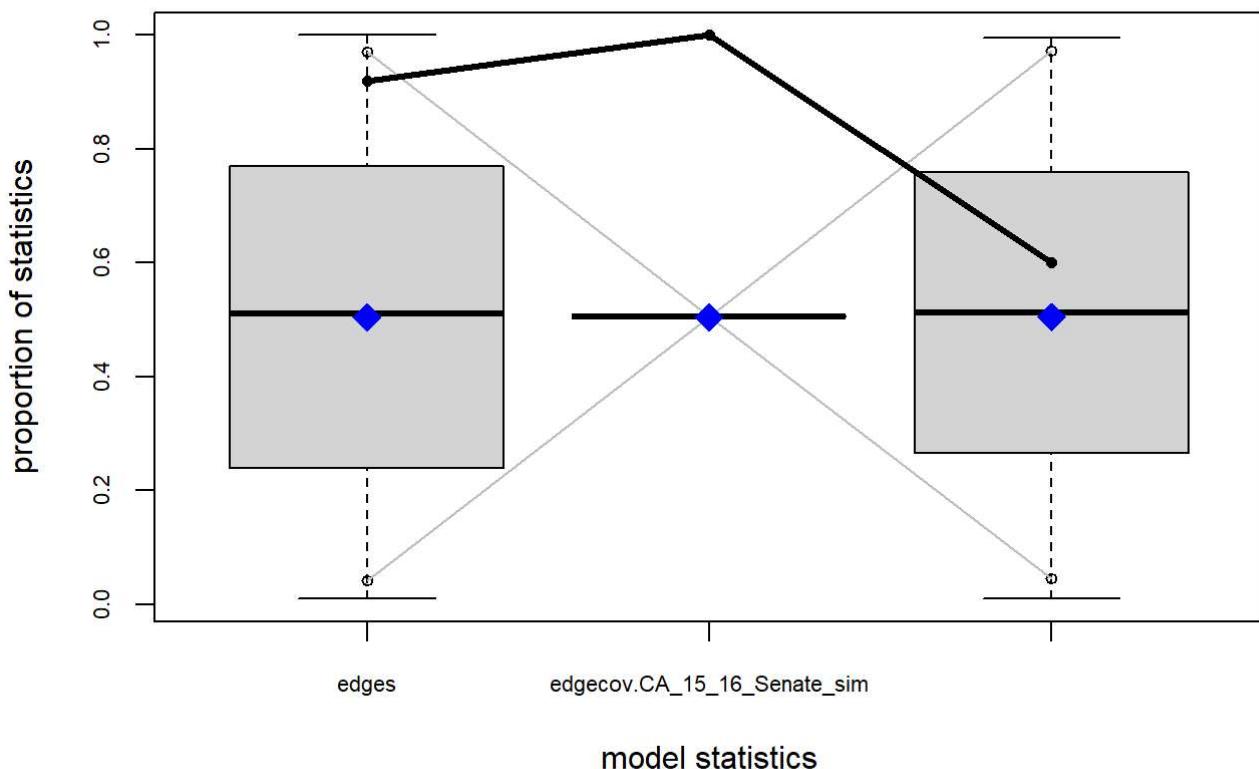
```
Null Deviance: 132714  on 95733  degrees of freedom
Residual Deviance:    1073  on 95730  degrees of freedom

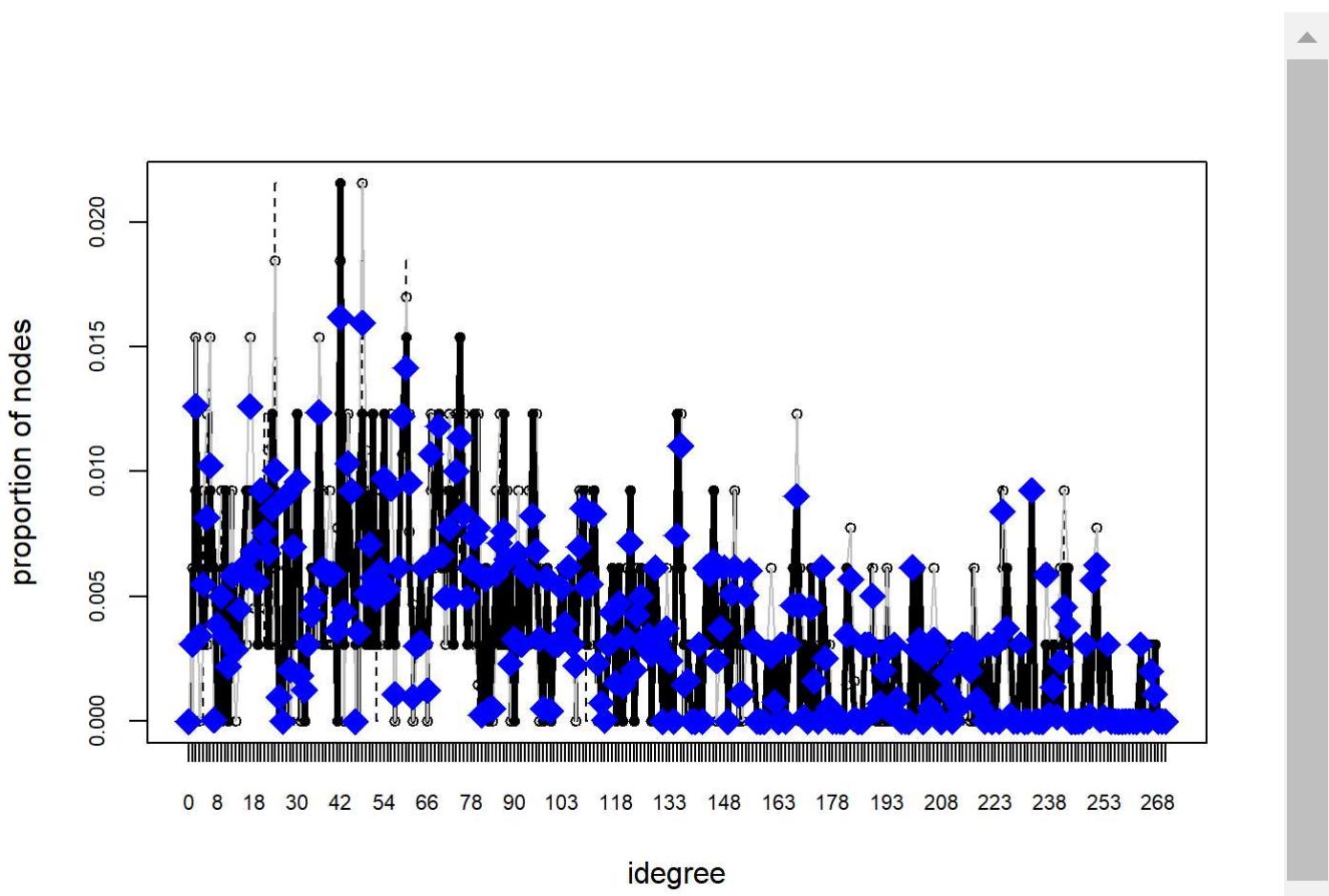
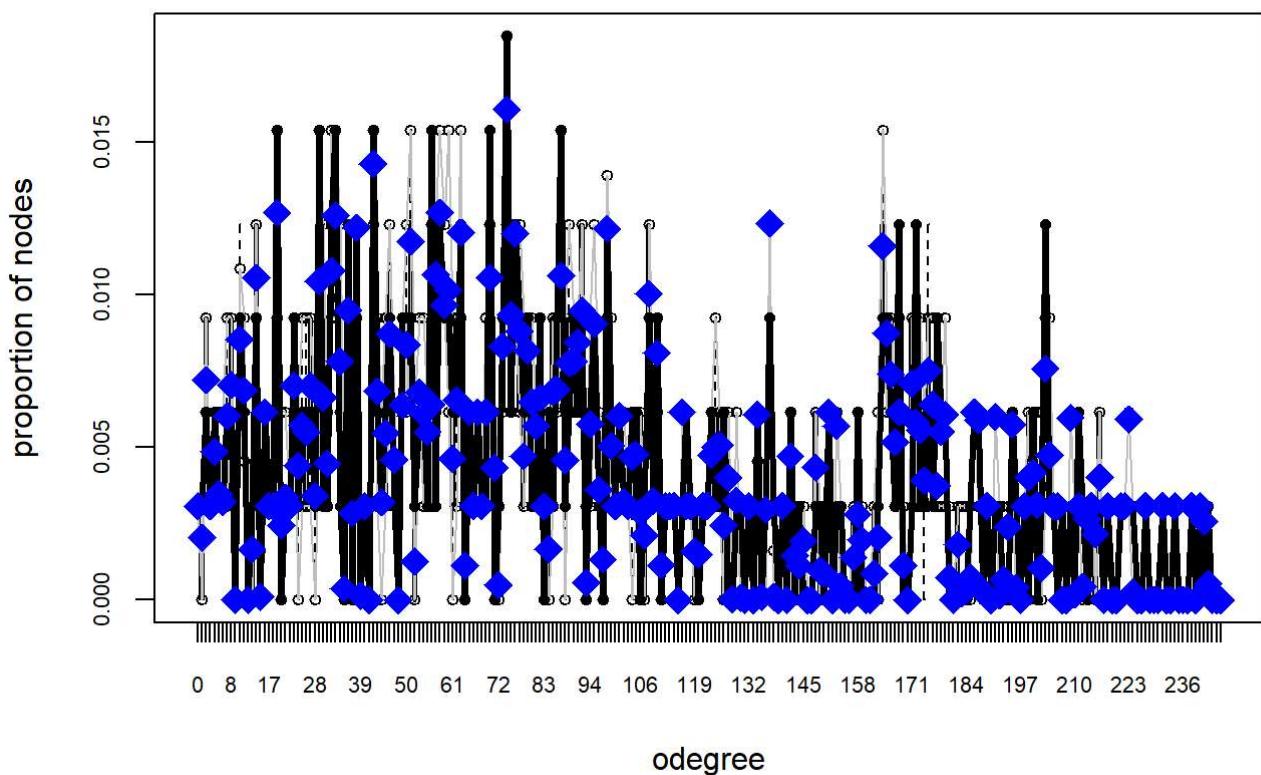
AIC: 1077  BIC: 1096  (Smaller is better. MC Std. Err. = 0)
```

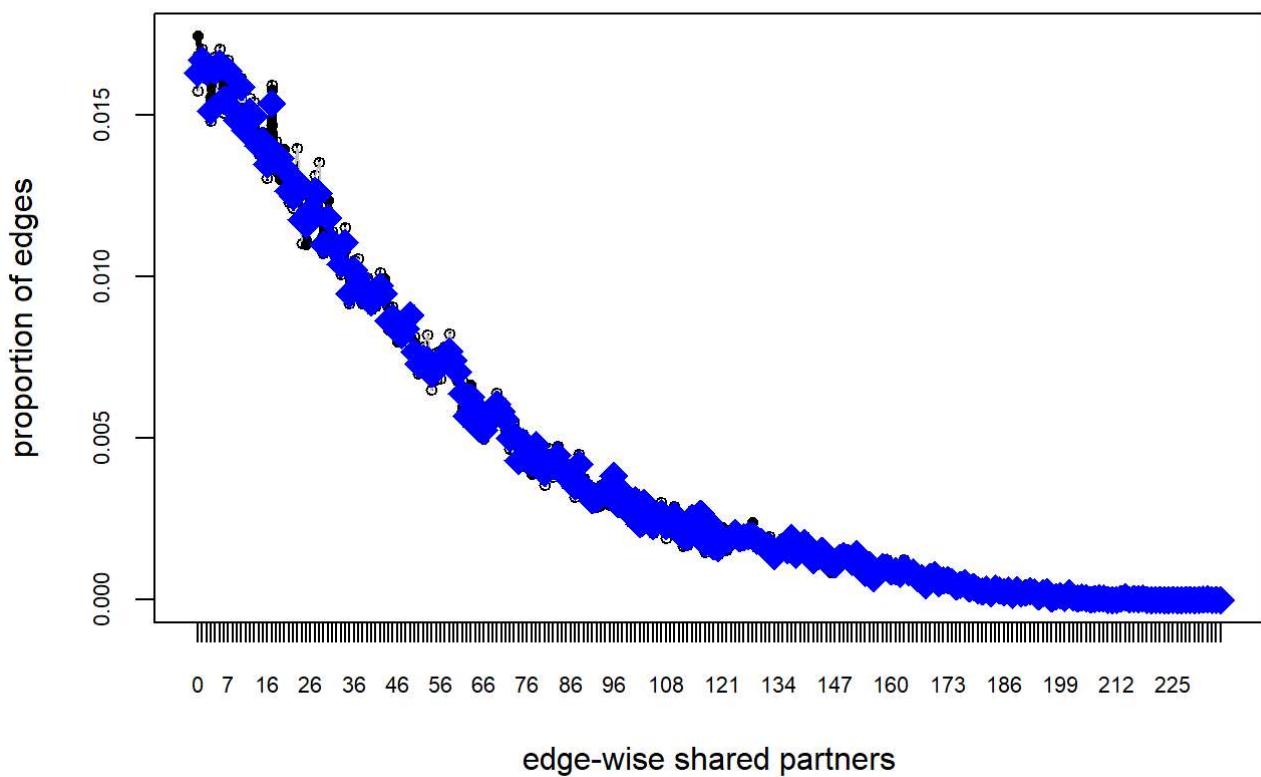
```
Warning: The following terms have infinite coefficient estimates:
edgecov.CA_15_16_Senate_sim
```

```
gof_CA_15_16_Senate_ergm <- gof(CA_15_16_Senate_ergm)
```

```
plot(gof_CA_15_16_Senate_ergm)
```







Goodness-of-fit diagnostics

