Social Network Lab 4b: Emergence of Network Dynamics (Relational Event Modeling, REM)

Student: Yijian Li   Net ID: YLS9426

Part I: Constructing Hypotheses to Test

Hypothesis 1*: The likelihood of individual i sending a message to individual j is greater if (a) i has sent a message to j recently and (b) if j has sent a message to i recently.*
Hypotheses 2*: The likelihood of individual i sending a message to individual j is greater (a) if j has been more active in the past than others and (b) if j has been less active than others in the past.*
Hypothesis 3a*: The likelihood of individual i sending a message to individual j is greater if (a) i and j are in the same team and (b) if i and j are not in the same team.*
Hypothesis 3b*: The likelihood of individual i sending a message to individual j is greater if (a) i and j have the same role and (b) if i and j have different role.*
Hypothesis 4*: The likelihood of individual i sending a message to individual j is greater if the message immediately preceding that event is j sending a message to i.*
Hypothesis 5*: When an individual i sends a message to another person j, that person has a greater likelihood of communicating with another individual k, thereby creating an information-sharing chain.*

Part II: Testing your hypotheses

1. Open the file Lab4b-REM.R

2. Let's begin by fitting a very simple covariate model (model1) using the function FitEventNetworkCore that fits separate effects for repetition (H1a; RSndSnd) and reciprocity (H1b; RRecSnd) to see if prior events affect send/receipt rates differently.

3. The summary of the model will produce a typical regression style output. Please include a table of the model results. Interpret the sign of the coefficient as well as the significance. What can you conclude with regards to repetition and reciprocity effects?

Answer:

```
> summary(model1)
Relational Event Model (Interval Likelihood)

                    MLE     Std.Err  Z value  Pr(>|z|)
[Intercept] -5.7124e+00  3.9606e-02 -144.231 < 2.2e-16 ***
RRecSnd      5.2761e-04  7.6233e-05    6.921 4.483e-12 ***
RSndSnd      1.1998e-03  8.6674e-05   13.842 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual deviance: 17011.68 on 1129 degrees of freedom
AIC: 19275.68 BIC: 24971.61
```

```
> 5.2761e-04
[1] 0.00052761
> 1.1998e-03
[1] 0.0011998
```

```
> exp(0.00052761)
[1] 1.000528
> exp(0.0011998)
[1] 1.001201
```

In chart, All the parameters are significant here.

As we can see from above, both RRecSnd and RSndSnd are **positive**, which means there are **reciprocity** and **repetition** exist. The **hypothesis 1a** is **supported**:

The likelihood of individual i sending a message to individual j is greater if i has sent a message to j recently.

There is a positive value indicates **reciprocity** being observed at a higher than random frequency.

There is also a positive value indicates **repetition** being observed at a higher than random frequency.

Also, with the exponential values we can know that the case of **hypothesis 1a is more likely/often to happen than the hypothesis 1b.**

4. We interpret the coefficients in terms of hazard rates. In other words, when we take the exponential of the coefficient ($e^{coefficient}$), we get the relative likelihood of a particular event. How likely is it that an individual will send a message to their prior communication partner? What about someone responding to another?

Answer:

The hazard rate is indeed different from odd rate.

For an individual will send a message to their prior communication partner, we should pay attention to the **repetition** value which is **RSndSnd** in the model 1.

With a positive value 1.001201 above, we know the likelihood of this case happen will be **1.001201 times higher** than those people without previous messages.

For an individual will send a message to respond other people, we should pay attention to the **reciprocity** value which is **RRecSnd** in the model 1.

With a positive value 1.000528 above, we know the likelihood of this case happen will be **1.000528 times higher** than the case that j does not respond to i.

5. Now, fit a second model (model2) that accounts for a small number of individuals that were highly active in terms of communication. Please include a table of the model results. We want to test for the notion of "preferential attachment," or the likelihood to contact popular targets (H2). To do so, we fit another model (model1) add the effect "Normalized Total Degree Received" (NTDegRec). Thus, this effect measures how often an individual was previously contacted. Interpret the model in context, as with the previous models. What is the effect of preferential attachment? What are the implications?

Answer:

```
> summary(model2)
Relational Event Model (Interval Likelihood)

                 MLE     Std.Err   Z value  Pr(>|z|)
         -6.2311e+00  4.8513e-02 -128.4428 < 2.2e-16 ***
RRecSnd   3.7346e-04  7.5755e-05    4.9298 8.231e-07 ***
RSndSnd   1.0267e-03  8.4918e-05   12.0903 < 2.2e-16 ***
NTDegRec  2.9458e-01  9.6379e-03   30.5650 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual deviance: 16361.53 on 1128 degrees of freedom
AIC: 18625.53 BIC: 24321.46
```

Social Network Lab 4b: Emergence of Network Dynamics (Relational Event Modeling, REM)

Student: Yijian Li Net ID: YLS9426

```
> 3.7346e-04
[1] 0.00037346
> 1.0267e-03
[1] 0.0010267
> 2.9458e-01
[1] 0.29458
```

```
> exp(0.00037346)
[1] 1.000374
> exp(0.0010267)
[1] 1.001027
> exp(0.29458)
[1] 1.342562
```

In chart, All the parameters are significant here.
From all the data above, we can easily see that with the positive value of **NTDegRec in 0.29458**.
Combined with the positive values of **RRecSnd** and **RSndSnd**, we can conclude that the person who have been more active in the past are more likely to receive messages than expected by chance.

The case of **hypothesis 2(a) is more likely/often to happen than the hypothesis 2(b).**
The likelihood of 2(a) happens will be **1.342562 times higher** than the case 2(b).

6. Suppose we want to determine which model is a better fit for the data. We can use the BIC criterion to compare models; remember, given the same data, lower BIC is better. Which model do we prefer?

Answer:

BIC for Model 1: **24971.61**
BIC for Model 2: **24321.46**
24971.61 – 24321.46 = 650.15

Therefore, the model 2 is better and we prefer to use model 2.

7. These groups have different teams and roles. Fit a model (model3) using the function SameConstGroup. This term captures tendencies to speak to other members of your team (H3a; SameTeam) and with your same role (H3b; SameRole). Please include a table of the model results. Interpret the results.

Answer:

```
> summary(model3)
Relational Event Model (Interval Likelihood)

                    MLE     Std.Err  Z value  Pr(>|z|)
[Intercept] -6.8623e+00  7.0890e-02 -96.8031 < 2.2e-16 ***
RRecSnd      3.3598e-04  7.6651e-05   4.3832 1.170e-05 ***
RSndSnd      9.9219e-04  8.5181e-05  11.6480 < 2.2e-16 ***
NTDegRec     2.2949e-01  1.1204e-02  20.4826 < 2.2e-16 ***
SameTeam     1.0511e+00  7.4215e-02  14.1631 < 2.2e-16 ***
SameRole     5.7107e-01  9.5832e-02   5.9591 2.537e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual deviance: 16139.69 on 1126 degrees of freedom
AIC: 18403.69 BIC: 24099.62
```

```
> 1.0511e+00
[1] 1.0511
> 5.7107e-01
[1] 0.57107
> exp(1.0511)
[1] 2.860796
> exp(0.57107)
[1] 1.77016
```

First of all, we can see all these parameters are significant.

As we can see from above, value of SameTeam is **positive**, which means people who are in the **same team** being observed at a higher than random frequency.
The **hypothesis 3a(a)** is **supported**:
The likelihood of individual i sending a message to individual j is **greater** if i and j are in the **same team**.

With a positive value 2.860796 above, we know the likelihood of this case happen will be **2.860795 times higher** than the case that i and j are not in the same team.

As we can see from above, value of SameRole is **positive**, which means people who are in the **same role** being observed at a higher than random frequency.
The **hypothesis 3b(a)** is **supported**:
The likelihood of individual i sending a message to individual j is **greater** if i and j have the **same role**.

With a positive value 1.77016 above, we know the likelihood of this case happen will be **1.77016 times higher** than the case that i and j do not have the same role.

8. Compare this model to (model2). Which model is better?

Answer:

BIC for Model 2: **24321.46**
BIC for Model 3: **24099.62**

24321.46 - 24099.62 = 221.84

Therefore, the model 3 is better and we prefer to use model 3.

9. Communication over electronic channels is governed by strong conversational norms, which among other things mandate systematic turn-taking reciprocity. Test for this in a new model (model4), via the use of participation shifts (P-shift), particularly the AB-BA shift (a tendency for B to call A, given that A has just called B). Interpret the model. Is there a tendency to immediately reciprocate communication (H4; PSAB.BA(data))? Is model4 preferable to model3? To answer this question, investigate BIC scores for both models. What does this suggest about immediate response rates? Using the hazard rates, how likely is it that any given event is driven by a response?

Answer:

Social Network Lab 4b: Emergence of Network Dynamics (Relational Event Modeling, REM)

Student: Yijian Li   Net ID: YLS9426

```
> summary(model4)
Relational Event Model (Interval Likelihood)

                   MLE     Std.Err  Z value  Pr(>|z|)
[Intercept] -6.8623e+00 7.2141e-02 -95.1236 < 2.2e-16 ***
RRecSnd      3.3597e-04 7.6697e-05   4.3805 1.184e-05 ***
RSndSnd      9.9219e-04 8.5223e-05  11.6422 < 2.2e-16 ***
NTDegRec     2.2950e-01 1.1231e-02  20.4333 < 2.2e-16 ***
SameTeam     1.0511e+00 7.4863e-02  14.0404 < 2.2e-16 ***
SameRole     5.7106e-01 9.5977e-02   5.9500 2.681e-09 ***
PSAB-BA     -4.9245e-03 1.0026e+00  -0.0049    0.9961
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual deviance: 16139.69 on 1125 degrees of freedom
AIC: 18403.69 BIC: 24099.62
```

```
> -4.9245e-03
[1] -0.0049245
> exp(-0.0049245)
[1] 0.9950876
```

a)

From the chart, we can see that the value of **PSAB-BA is not significant**, so we **cannot** say there is a tendency to immediately reciprocate communication.
The hypothesis 4 **cannot be determined** by this data.

b)

BIC for Model 3: **24099.62**
BIC for Model 4: **24099.62**
The AIC and BIC for these 2 models are the same, so we cannot say which one is better.

c)

If the PSAB-BA value is significant, we can say hypothesis 4 is not supported with a **negative** value:
The likelihood of 4 happens will be **0.9550876 times lower** than the case of message preceding in normal speed.

The likelihood of individual i sending a message to individual j is **less** if the message immediately preceding that event is j sending a message to i.

10. What about other conversational norms? We may expect that the current participants in an interaction may be likely to initiate the next call, a tendency that can also be captured with P-shift effects. The "AB-BY" effect captures spread of information in a chain (H5; PSAB.BY(data)). Fit another model (model5) to test for this effect, and interpret the results. What do the sign and significance tell you about each effect? Interpret the implications for team functioning. Check BIC to see if we improved our model over the reciprocity model (model4).

Answer:

Social Network Lab 4b: Emergence of Network Dynamics (Relational Event Modeling, REM)

Student: Yijian Li Net ID: YLS9426

```
> summary(model5)
Relational Event Model (Interval Likelihood)

                    MLE     Std.Err  Z value  Pr(>|z|)
[Intercept] -6.8630e+00  7.2114e-02 -95.1684 < 2.2e-16 ***
RRecSnd      3.3615e-04  7.6702e-05   4.3825 1.173e-05 ***
RSndSnd      9.9129e-04  8.5223e-05  11.6317 < 2.2e-16 ***
NTDegRec     2.2927e-01  1.1239e-02  20.4007 < 2.2e-16 ***
SameTeam     1.0511e+00  7.4843e-02  14.0446 < 2.2e-16 ***
SameRole     5.7094e-01  9.5962e-02   5.9496 2.688e-09 ***
PSAB-BA     -6.2468e-01  1.2176e+00  -0.5131    0.6079
PSAB-BY      6.2071e-01  7.1117e-01   0.8728    0.3828
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual deviance: 16139.04 on 1124 degrees of freedom
AIC: 18403.04 BIC: 24098.97
```

```
> 6.2071e-01
[1] 0.62071
> exp(0.62071)
[1] 1.860248
> -6.2468e-01
[1] -0.62468
> exp(-0.62468)
[1] 0.5354327
```

a)

From the chart, we can see that the value of **PSAB-BY is not significant**, so we cannot say there is a tendency to have a spread of information in a chain.

The hypothesis 5 **cannot be determined** by this data.

b)

BIC for Model 4: **24099.62**

BIC for Model 5: **24098.97**

24099.62 – 24098.97 = 0.65

Therefore, the model 5 is better and we prefer to upgrade model 4 to model 5.

c)

If the PSAB-BY value is significant, we can say hypothesis 5 is supported with a **positive** value:

The likelihood of 5 happens will be **1.860248 times higher** than the case that there is not a spread in the chain.

When an individual i sends a message to another person j, that person has a **greater** likelihood of communicating with another individual k, thereby creating an information-sharing chain.