

Supplementary material 2. How to analyse the synthetic sentencing dataset

Referenced from “Developing a Complete Sentence Severity Scale using Extended Goodman RC models” J. Quant. Crim 2024 [<https://doi.org/10.1007/s10940-024-09591-6>]

Stephanie Wallace and Brian Francis

20 August, 2024

Reading the dataset and transforming it to frequency data

The synthetic dataset is stored in CSV format on the Lancaster University dataset repository <https://doi.org/10.17635/lancaster/researchdata/670> and is named *synthdata.csv*. In this RMarkdown document, we provide the code needed to fit Goodman RC models and extended Goodman RC models to this data so that sentences can be scored. Note that the results will differ slightly from the published paper as the dataset is a synthesized version of the real dataset, which is necessary for data confidentiality purposes. We first read the dataset and confirm that it has 61066 lines and five variables. We display the first six lines of the dataset.

```
sentcase=read.csv("synthdata.csv",header=T)
dim(sentcase)
```

```
## [1] 61066      6
```

```
head(sentcase,6)
```

```
##      X              SENTENCE1 OFFENSE PLEA_TYPE PREV_CONV NUM_OFF
## 1 1    Custody >6 Years <=7 Years      2    GUILTY  NO PRIOR  SINGLE
## 2 2    Custody >8 Years <=10 Years     2    GUILTY  NO PRIOR  SINGLE
## 3 3    Custody >4 Months <=6 Months     3    GUILTY  NO PRIOR  SINGLE
## 4 4    Custody >9 Months <=12 Months    3    GUILTY  NO PRIOR  SINGLE
## 5 5    Custody >9 Months <=12 Months    3    GUILTY  NO PRIOR  SINGLE
## 6 6    Custody >9 Months <=12 Months    3    GUILTY  NO PRIOR  SINGLE
```

Table 1: Variables extracted

Variable	Description
SENTENCE1	Categorised sentence (29 levels)
OFFENSE	Categorised offense categories (111 levels)
PLEA_TYPE	Plea made at court (guilty/not guilty)
PREV_CONV	Previous court appearances (yes/no)
NUM_OFF	Number of offenses at current court appearance (single/multiple)

We can see that the dataset starts with two records for offense code 2 (attempted murder) followed by four records of offense code 3 (threats to kill). These were all first offences, with no prior convictions and no other offenses being brought to court that the same time. Attempted murder offenses gained a more severe sentence than threats to kill.

Fitting the basic Goodman RC model

We start off simply and produce a two-way table of OFFENSE by SENTENCE1. We use *xtabs* to make the table and then convert the table to frequency form by using *as.data.frame*

```
basicsenttable=xtabs(~SENTENCE1+OFFENSE,data=sentcase)
basicsentfreq=as.data.frame(basicsenttable)
dim(basicsentfreq)
```

```
## [1] 3219    3
```

We can fit standard Poisson log-linear models to this dataset. First we try a main effects model:

```
basicmod1=glm(Freq~OFFENSE+SENTENCE1,family=poisson, data=basicsentfreq)
#summary(basicmod1) lots of output- we just want the deviance.
deviance(basicmod1);df.residual(basicmod1)
```

```
## [1] 63647.51
```

```
## [1] 3080
```

```
AIC(basicmod1)
```

```
## [1] 70018.49
```

This model has a deviance of 63648 on 3080 degrees of freedom. The model does not fit, and this is because of the strong relationship between SENTENCE1 and OFFENSE. WE need to fit a multiplicative term, where the categories of SENTENCE1 and of OFFENSE have estimated scores, which then interact with each other. we use the *gnm* package in R to do this. [*gnm* stands for generalised non-linear model]

```
library(gnm)
RC1model=gnm(Freq~SENTENCE1+OFFENSE+Mult(SENTENCE1,OFFENSE),family=poisson,
             tolerance=1e-3,iterStart=5,data=basicsentfreq)
```

```
## Initialising
## Running start-up iterations.....
## Running main iterations.....
## Done
```

```
deviance(RC1model); df.residual(RC1model)
```

```
## [1] 21073.86
```

```
## [1] 2943
```

Obtaining the basic Goodman RC model scores for SENTENCE1

The row scores for SENTENCE1 can be normalised using the method described in section 7.1.1 of the *gnm* vignette . See <https://cran.r-project.org/web/packages/gnm/vignettes/gnmOverview.pdf>.

```
rowProbs<-with(basicsentfreq, tapply(Freq, SENTENCE1, sum)/sum(Freq))
mu<-getContrasts(RC1model, pickCoef(RC1model,"[.]SENTENCE1"),
  ref=rowProbs, scaleWeights=rowProbs)
mu
```

```
##
## Mult(., OFFENSE).SENTENCE1Community Order <=12 Months      0.43317968
## Mult(., OFFENSE).SENTENCE1Community Order >12 Months <=24 Months 0.43819141
## Mult(., OFFENSE).SENTENCE1Community Order >2 Years <=3 Years -0.98457019
## Mult(., OFFENSE).SENTENCE1Conditional Discharge            0.49308474
## Mult(., OFFENSE).SENTENCE1Custody <= 1 Month              0.51224175
## Mult(., OFFENSE).SENTENCE1Custody >1 Month <=2 Months     0.44364506
## Mult(., OFFENSE).SENTENCE1Custody >10 Years               -4.14923174
## Mult(., OFFENSE).SENTENCE1Custody >12 Months <=18 Months -0.88371230
## Mult(., OFFENSE).SENTENCE1Custody >18 Months <=24 Months -1.17551498
## Mult(., OFFENSE).SENTENCE1Custody >2 Months <=3 Months    0.37262608
## Mult(., OFFENSE).SENTENCE1Custody >2 Years <=3 Years      -1.46662011
## Mult(., OFFENSE).SENTENCE1Custody >3 Months <=4 Months    0.34345875
## Mult(., OFFENSE).SENTENCE1Custody >3 Years <=4 Years      -1.78868699
## Mult(., OFFENSE).SENTENCE1Custody >4 Months <=6 Months    -0.03387172
## Mult(., OFFENSE).SENTENCE1Custody >4 Years <=5 Years      -2.24020755
## Mult(., OFFENSE).SENTENCE1Custody >5 Years <=6 Years      -2.37385327
## Mult(., OFFENSE).SENTENCE1Custody >6 Months <=9 Months    -0.43028918
## Mult(., OFFENSE).SENTENCE1Custody >6 Years <=7 Years      -2.60846106
## Mult(., OFFENSE).SENTENCE1Custody >7 Years <=8 Years      -2.56455429
## Mult(., OFFENSE).SENTENCE1Custody >8 Years <=10 Years     -2.86516795
## Mult(., OFFENSE).SENTENCE1Custody >9 Months <=12 Months   -0.65381391
## Mult(., OFFENSE).SENTENCE1Fine                             0.57560533
## Mult(., OFFENSE).SENTENCE1Other                             0.59512532
## Mult(., OFFENSE).SENTENCE1Suspended <= 1 Month           0.49680209
## Mult(., OFFENSE).SENTENCE1Suspended >1 Month <=2 Months   0.40427762
## Mult(., OFFENSE).SENTENCE1Suspended >2 Months <=3 Months 0.38695255
## Mult(., OFFENSE).SENTENCE1Suspended >3 Months <=4 Months 0.35109618
## Mult(., OFFENSE).SENTENCE1Suspended >4 Months <=6 Months 0.02391151
## Mult(., OFFENSE).SENTENCE1Suspended >6 Months <=12 Months -0.44028465
##
## Std. Error
## Mult(., OFFENSE).SENTENCE1Community Order <=12 Months    0.006037932
## Mult(., OFFENSE).SENTENCE1Community Order >12 Months <=24 Months 0.007541200
## Mult(., OFFENSE).SENTENCE1Community Order >2 Years <=3 Years 0.038176141
## Mult(., OFFENSE).SENTENCE1Conditional Discharge           0.009378056
## Mult(., OFFENSE).SENTENCE1Custody <= 1 Month              0.008354503
## Mult(., OFFENSE).SENTENCE1Custody >1 Month <=2 Months     0.009599924
## Mult(., OFFENSE).SENTENCE1Custody >10 Years               0.164840303
## Mult(., OFFENSE).SENTENCE1Custody >12 Months <=18 Months 0.029762341
## Mult(., OFFENSE).SENTENCE1Custody >18 Months <=24 Months 0.031155721
## Mult(., OFFENSE).SENTENCE1Custody >2 Months <=3 Months    0.010688239
## Mult(., OFFENSE).SENTENCE1Custody >2 Years <=3 Years      0.031538248
## Mult(., OFFENSE).SENTENCE1Custody >3 Months <=4 Months    0.011615435
```

```
## Mult(., OFFENSE).SENTENCE1Custody >3 Years <=4 Years 0.037187070
## Mult(., OFFENSE).SENTENCE1Custody >4 Months <=6 Months 0.016955944
## Mult(., OFFENSE).SENTENCE1Custody >4 Years <=5 Years 0.046814785
## Mult(., OFFENSE).SENTENCE1Custody >5 Years <=6 Years 0.056145558
## Mult(., OFFENSE).SENTENCE1Custody >6 Months <=9 Months 0.027604853
## Mult(., OFFENSE).SENTENCE1Custody >6 Years <=7 Years 0.079404404
## Mult(., OFFENSE).SENTENCE1Custody >7 Years <=8 Years 0.086042668
## Mult(., OFFENSE).SENTENCE1Custody >8 Years <=10 Years 0.104302892
## Mult(., OFFENSE).SENTENCE1Custody >9 Months <=12 Months 0.028086474
## Mult(., OFFENSE).SENTENCE1Fine 0.010279160
## Mult(., OFFENSE).SENTENCE1Other 0.010910541
## Mult(., OFFENSE).SENTENCE1Suspended <= 1 Month 0.009812626
## Mult(., OFFENSE).SENTENCE1Suspended >1 Month <=2 Months 0.011368841
## Mult(., OFFENSE).SENTENCE1Suspended >2 Months <=3 Months 0.009916151
## Mult(., OFFENSE).SENTENCE1Suspended >3 Months <=4 Months 0.011102578
## Mult(., OFFENSE).SENTENCE1Suspended >4 Months <=6 Months 0.014640167
## Mult(., OFFENSE).SENTENCE1Suspended >6 Months <=12 Months 0.021377433
```

The estimates and standard errors are produced. Note that the scores are reversed in sign, with the highest severity having a large negative value. Switching the sign is straightforward. The algorithm does not guarantee the sign of the score - both the score for the offense and the score for the sentence may be reversed in sign and result in the same fitted cell counts.

Fitting the extended Goodman RC model using a five-way table

We now fit the extended Goodman RC model to control for the aggravating variables of PLEA_TYPE, PREV_CONV and NUM_OFF. We instead convert the original case by case dataframe into a **five-way** table of counts using *xtabs* and then convert the table to frequency form by using *as.data.frame*.

```
senttable=xtabs(~SENTENCE1+OFFENSE+PLEA_TYPE+PREV_CONV+NUM_OFF,data=sentcase)
sentfreq=as.data.frame(senttable)
dim(sentfreq)
```

```
## [1] 25752      6
```

```
knitr::kable(
  head(sentfreq,6), booktabs = TRUE,
  caption = 'The first six lines of sentfreq'
)
```

Table 2: The first six lines of sentfreq

SENTENCE1	OFFENSE	PLEA_TYPE	PREV_CONV	NUM_OFF	Freq
Community Order <=12 Months	1	GUILTY	NO PRIOR	MULTIPLE	0
Community Order >12 Months <=24 Months	1	GUILTY	NO PRIOR	MULTIPLE	0
Community Order >2 Years <=3 Years	1	GUILTY	NO PRIOR	MULTIPLE	0

SENTENCE1	OFFENSE	PLEA_TYPE	PREV_CONV	NUM_OFF	Freq
Conditional Discharge	1	GUILTY	NO PRIOR	MULTIPLE	0
Custody <= 1 Month	1	GUILTY	NO PRIOR	MULTIPLE	0
Custody >1 Month <=2 Months	1	GUILTY	NO PRIOR	MULTIPLE	0

There are 25752 lines in the frequency data. (29 x 111 x 2 x 2 x 2). Freq contains the frequencies for each combination of the five cross-classifying factors. Table 2 shows the first six lines of the frequency data. As OFFENSE 1 is “murder”, it is not surprising that the frequency counts for the low custodial categories of nine months or less are zero.

We now fit the extended Goodman RC model, again using the package *gnm*.

```
library(gnm)
RC2model=gnm(Freq~SENTENCE1+OFFENSE+
              Mult((SENTENCE1+PLEA_TYPE+PREV_CONV+NUM_OFF),OFFENSE)
              + PLEA_TYPE*PREV_CONV*NUM_OFF,family=poisson,tolerance=1e-4,
              data=sentfreq)
```

```
## Initialising
## Running start-up iterations..
## Running main iterations.....
## Done
```

```
deviance(RC2model); df.residual(RC2model)
```

```
## [1] 61424.83
```

```
## [1] 25466
```

Note that deviances fitted to different sized tables are not comparable.

The model is constructed by including the extra factors as multiplicative interactions with OFFENSE. We also fit a main effect of the layer dimension, which is achieved by including the three-way interaction “PLEA_TYPE*PREV_CON*NUMOFF”

As with the simple model, we can then extract the sentence scores:

```
rowProbs<-with(sentfreq, tapply(Freq, SENTENCE1, sum)/sum(Freq))
mu<-getContrasts(RC2model, pickCoef(RC2model,"[.]SENTENCE1"),
  ref=rowProbs, scaleWeights=rowProbs)
mu
```

```
##
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Community Order <=12 Months -0.42857737
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Community Order >12 Months <=24 Months -0.41482413
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Community Order >2 Years <=3 Years 1.21128745
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Conditional Discharge -0.49564485
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody <= 1 Month -0.58019919
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >1 Month <=2 Months -0.47439377
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >10 Years 3.79120983
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >12 Months <=18 Months 0.96196466
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >18 Months <=24 Months 1.20827101
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >2 Months <=3 Months -0.38739721
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >2 Years <=3 Years 1.42366128
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >3 Months <=4 Months -0.36099854
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >3 Years <=4 Years 1.71754312
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >4 Months <=6 Months 0.08517937
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >4 Years <=5 Years 2.16178596
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >5 Years <=6 Years 2.28038607
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >6 Months <=9 Months 0.54770807
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >6 Years <=7 Years 2.50354795
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >7 Years <=8 Years 2.43922030
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >8 Years <=10 Years 2.75549083
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >9 Months <=12 Months 0.76051291
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Fine -0.66890872
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Other -0.70468992
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Suspended <= 1 Month -0.54963796
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Suspended >1 Month <=2 Months -0.40631456
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Suspended >2 Months <=3 Months -0.38448031
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Suspended >3 Months <=4 Months -0.34734232
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Suspended >4 Months <=6 Months 0.02874075
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Suspended >6 Months <=12 Months 0.52298680
## Std. Error
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Community Order <=12 Months 0.005311266
```

## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Community Order >12 Months <=24 Months	0.007942585
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Community Order >2 Years <=3 Years	0.034659205
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Conditional Discharge	0.011710571
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody <= 1 Month	0.008765346
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >1 Month <=2 Months	0.010453444
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >10 Years	0.135227772
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >12 Months <=18 Months	0.026966321
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >18 Months <=24 Months	0.027911550
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >2 Months <=3 Months	0.011348371
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >2 Years <=3 Years	0.028153716
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >3 Months <=4 Months	0.012328631
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >3 Years <=4 Years	0.033535152
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >4 Months <=6 Months	0.018653634
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >4 Years <=5 Years	0.042348116
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >5 Years <=6 Years	0.050878794
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >6 Months <=9 Months	0.027882426
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >6 Years <=7 Years	0.071508049
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >7 Years <=8 Years	0.077568653
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >8 Years <=10 Years	0.093166211
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Custody >9 Months <=12 Months	0.026274830
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Fine	0.010403603
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Other	0.010271014
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Suspended <= 1 Month	0.011635033
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Suspended >1 Month <=2 Months	0.012597484
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Suspended >2 Months <=3 Months	0.010750314
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Suspended >3 Months <=4 Months	0.012095082
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Suspended >4 Months <=6 Months	0.016332981
## Mult(. + PLEA_TYPE + PREV_CONV + NUM_OFF, OFFENSE).SENTENCE1Suspended >6 Months <=12 Months	0.020228565

These scores can be compared with those published in the article. 95% confidence intervals were calculated using the estimate with plus or minus two standard deviations.