Voting Patterns in the 32nd Dáil Éireann

During the 5th session of the 32nd Dáil, there were a number of issues on which the elected TD's voted on. Of these issues, six of them were votes on motions of no confidence in the Housing Minister, to provide rent freeze for all new and existing tenancies, amend the Social Welfare Acts, amend the environmental policy, amend the Gaming and Lotteries Bill and an amendment to the Planning and Development Bill affecting first-time buyers. In a study carried out, the voting behaviour of each of the 156 TD's present for these six votes were examined and analysis was carried out to try and identify potential groups/clusters of TD's with similar attendance and/or voting patterns.

The data used in the study contains 156 observations. The votes of each observation for each of the six motions were recorded as one of two values: 1 if the observation voted 'No' and 2 if the observation voted 'Yes' for the proposed motion. Therefore, the data used was binary in nature with 156 rows and 6 columns of voting data. The name of each observation was recorded but their party affiliation was not included in the data to reduce any bias in further analysis. As you will see later on in the report, the party membership of each observation is taken into account after model costruction is carried out.

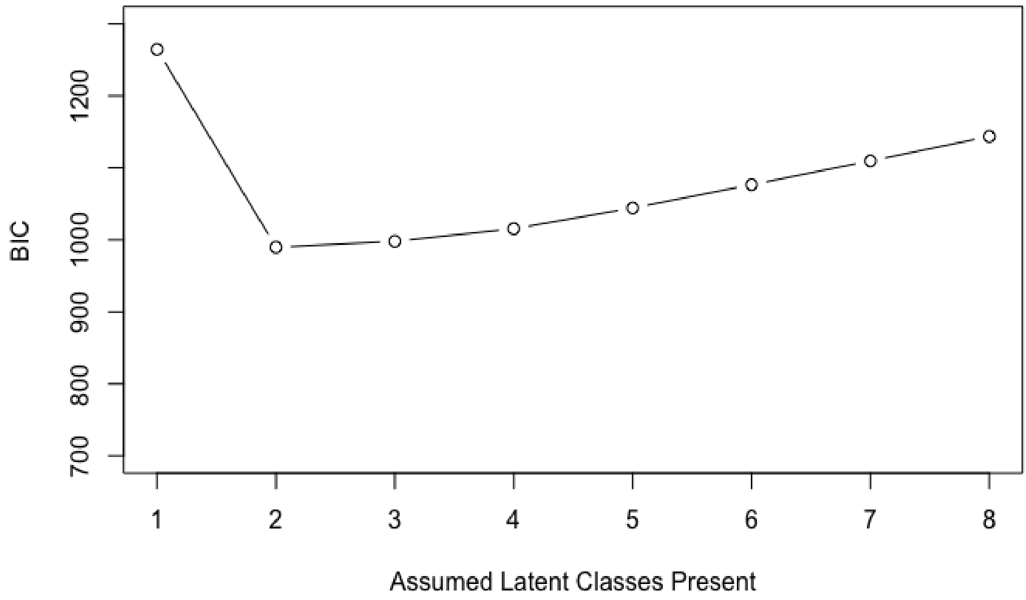
The aim of the study was to identify clusters of observations with similar voting characteristics. The technique used for this analysis is called model-based clustering. Model-based clustering assumes that the data were generated by a model and it tries to recover this model using the given data. It assumes that there is a given number of groups/clusters in the data present and each cluster has its own Gaussian/Normal multivariate distribution. The characteristics of each of these distributions is also derived from model based clustering. These characterisitcs include the shape, volume/size and orientation (with respect to the axes) of the clusters. Each of these characteristics are found easily by decomposing the covariance matrix assigned to each of the distributions. We will not get into the details here as this process was automated in the study. Finally, it is important to note that model-based clustering uses a maximum likelihood approach via an EM algorithm for fitting the model. We won't go into the details of the EM algorithm here as it was also automated in the process.

The study was carried out on the statistical computing environment, R, and a particular software package called **poLCA** which can be implemented in R was used for model-based clustering. A function within this package also called poLCA uses Latent Class Analysis as the model-based clustering approach which is useful when the underlying data is binary in nature. In short, latent class models probabilistically groups each observation into a "latent class", which in turn produces expectations about how each observation will respond on each manifest variable, which in this case, is how each observation will vote on each issue. In this case, the function does not automatically determine the number of latent classes (groups/clusters) in the data set. However, it does produce a number of goodness-of-fit statistics which were used to make a theoretically sound estimate of the number of clusters present. The latent class model also permits the use of covariates which in this case were not used because the only known covariate between observations was their party membership. As mentioned previously, this was preferred to be unknown in the data as these party memberships could be seen as clustrs themselves and may distort the results. We are only interested in clusters formed from voting patterns, independent of party membership. The party membership of observations in each latent class was analysed after the number of latent classes was found.

The poLCA function works by fitting a model to the data with the assumed number of latent classes present used as an input. As mentioned before, the function does not determine the optimal number of latent classes so in order to find the optimal model, it was required to fit the model using the poLCA function for several assumed numbers of latent classes. The optimal model was then chosen based on the goodness-of-fit statistics output by each model. The goodness-of-fit statistics of interest were the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC). Both are measures of prediction of out of sample error. Preferred models are those that give a lower BIC and/or AIC. These were calculated as follows:

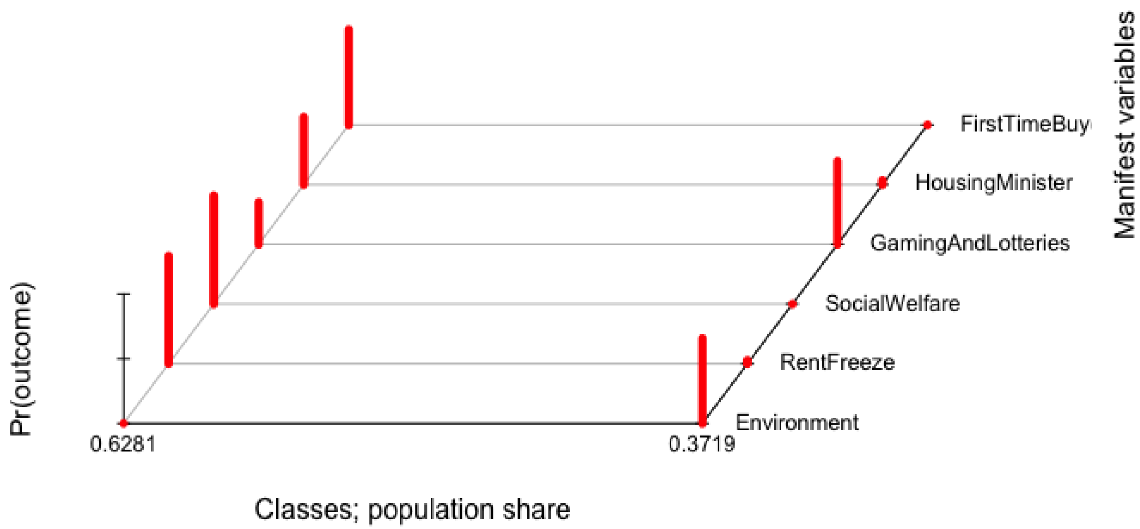
- $BIC = -2\Lambda + \Phi \ln N$
- $AIC = -2\Lambda + 2\Phi$

where Λ is the maximum likelihood of the model, Φ is the number of estimated paramters of the model and N is the number of observations in the data, in this case is equal to 156. It was found that the addition of extra clusters (more complexity) to the model only contributed minimally to improving the AIC of the model, making it hard to distinguish if the tradeoff between a more complex model and a marginally lower AIC is worth it. Therefore, the model with the lowest BIC was selected as this gave a clearer, more confident result. The function was run 8 times, producing 8 models each with assumed number latent classes equal 1 to 8 repectively. The BIC of each model were plotted and are shown below:



From the above plot, it is clear that the model with the lowest BIC is that which assumes that there are 2 latent classes present in the data. It produces a relatively low BIC of 989.6813 This result shows that this is the optimal model for selection and we can conclude that based on model based clustering on the voting patterns of the 156 TD's on the 6 aforementioned motions, there are 2 clusters present within the group of TD's.

Now that we have a good estimate of how many clusters of TD's are present, we can examine the properties of each cluster. It is of concern what kind voting patterns each cluster has and what each member of the given cluster was likely to have voted on each issue. This may give us an idea of which issues the government and the opposition had conflicts of interest. For the purpose of simplicity, we will call the first cluster Group 1 and the second cluster Group 2. In the below plot, you can see the issues on which the TD's voted labelled along the right. The red bars on the left represent the probability of voting 'Yes' on each motion given that you belong to Group 1 and the bars on the right represent the probability of voting 'Yes' on each motion given you belong to Group 2.



Based on this graph, the following table summarizes the voting patterns of each group:

Group	Vote					
	Environment	Rent Freeze	Social Welfare	Gaming and Lotteries	Housing Minister	First-Time-Buyers
1	No	Yes	Yes	No	Yes	Yes
2	Yes	No	No	Yes	No	No

As you can see from the above results, Group 1 is a more radical group who desire more change in current policies especially those regarding housing issues and social welfare. However, this group voted against changes in environmental policy. On the other hand, Group 2 is a more conservative group and voted against changes regarding housing and social welfare issues but voted for change in environmental policy. There was not an obvious divide between groups regarding the gaming and lotteries motion suggesting that no particular group of TD's had one particular view on this issue.

When examining the two resulting clusters, is also of interest to see which TD's and their respective parties are in each. This is of interest because the voting patterns seen above might align with some parties' policies. When the chosen optimal model is run, it predicts the membership of each observation to either group and classifies them accordingly. This output splits the observations into their assigned class, either Group 1 or Group 2. From the model estimates, approximately 64.1% of observations were placed in Group 1 and the remaining observations were placed in Group 2. The party membership of each TD in both groups was then compared. The following table shows the party composition of Groups 1 and 2:

Party	Number of TD's	
	Group 1	Group 2
Solidarity - People Before Profit	6	0
Fianna Fáil	43	0
Fine Gael	1	48
Green Party	3	0
Independents 4 Change	2	1
Independent	11	7
Labour Party	7	0
Social Democrats	3	0
Sinn Féin	24	0

The table above shows a significant difference in party composition between Group 1 and 2. The clear and obvious difference is that all but one Fine Gael TD's have been placed into Group 2. Of the 56 TD's in Group 2, 48 of them are in Fine Gael while the remaining TD's in this group were independent. The table shows that there was one Independents 4 Change TD in Group 2, however, further research showed that he was in fact an independent TD. Only one Fine Gael TD was placed in Group 1. These results are not surprising as Fine Gael were in government during the 32nd Dáil. The voting patterns of Group 2 reflect this fact as members of Group 2 are predicted to mostly vote against change, more notably wanting to protect their housing minister, Eoghan Murphy. It is also worth noting that Group 2 voted against the proposed changes in housing policies. These were changes Fine Gael were opposed against during the housing crisis. Group 1 consists of members of all parties but includes only one member of Fine Gael. Of the larger parties, all Fianna Fáil and Sinn Féin members are assigned to this group. Again this is unsurprising as these parties were in opposition during the 32nd Dáil. Likewise, the voting patterns of this group reflect the parties' propositions for change.

To conclude, using Latent Class Analysis as a model-based clustering approach, we have exposed two clear clusters of TD's in the 32nd Dáil based on the voting patterns of the TD's for six different voting sessions. The first cluster was found to be a more radical group, desiring change in housing and social welfare policies. All members of Fianna Fáil and Sinn Féin were found to be members of this group along with all members of the smaller parties. The second group was found to be a more conservative group during these voting sessions, voting against changes in housing and social welfare policies but voting for change in environmental policies. All but one member of the party in government at the time of voting, Fine Gael, were found to be member sof this group alongside 8 indpendent TD's. There was an approximate 60:40 ratio with respect to the size of groups 1 and 2, suggesting that there was a larger desire for change in housing and social welfare policies than there was for no change in these policies across all TD's in the 32nd Dáil Éireann.

Appendix

A) BIC and AIC plots

```
load("/Users/timhowes/Downloads/32ndDail_FifthSession_BinaryVotes.Rdata") #load the data in
bin.votes = bin.votes+1 #add 1 for applicability of function
f1 <- cbind(Environment,RentFreeze,SocialWelfare,GamingAndLotteries,HousingMinister,FirstTimeBuyers)~1
AIC = c(rep(0,len=8))
BIC = c(rep(0,len=8))
for(i in 1:8){
  BIC[i] = poLCA(f1, bin.votes, nclass=i,verbose =FALSE,nrep=30)$bic
  AIC[i] = poLCA(f1, bin.votes, nclass=i,verbose =FALSE,nrep=30)$aic
}
plot(1:8, BIC, xlab="Assumed Latent Classes Present", ylab="BIC", ylim=c(700,1300),type = "b")
plot(1:8, AIC, xlab="Assumed Latent Classes Present", ylab="AIC",type = "b")
```

B) Classification Plot

```
mod = poLCA(f1, bin.votes, nclass=2)
plot(mod,what="classification")
```

C) Calculations

```
parties = read.csv("/Users/timhowes/Downloads/TDs_names_parties.csv"). #read in party names
Party = data.frame(parties$Party)
bin.votes = cbind(bin.votes,Party)  #add party name column to data
colnames(bin.votes)[7] = "Party"

class1 = bin.votes[c(which(mod$predclass == 1)),-c(1:6),drop=FALSE]   #create group 1
class2 = bin.votes[c(which(mod$predclass == 2)),-c(1:6),drop=FALSE]   #create group 2
group1 = matrix(rep(0,2*length(levels(class1$Party))),ncol=2)
group2 = matrix(rep(0,2*length(levels(class2$Party))),ncol=2)
for(i in 1:nrow(group1)){
  group1[i,1] = levels(class1$Party)[i]
  group1[i,2] = sum(class1$Party == levels(class1$Party)[i])     #find number of each party in group 1
}
for(i in 1:nrow(group2)){
  group2[i,1] = levels(class2$Party)[i]
  group2[i,2] = sum(class2$Party == levels(class2$Party)[i])   #find number of each party in group 2
}
numFGclass1 = sum(class1$Party=="FG")/nrow(class1)   #analyse Fine Gael proportions
numFGclass2 = sum(class2$Party=="FG")/nrow(class2)
```