# Statistical modelling of a terrorist network

June 12, 2016

Murray Aitkin and Duy Vu
School of Mathematics and Statistics
University of Melbourne, Victoria Australia
and Brian Francis
Department of Mathematics and Statistics
Lancaster University, Lancaster LA14YF United Kingdom

**Abstract**

This paper investigates the group structure in a terrorist network through the latent class model and a Bayesian model comparison method for the number of latent classes. The analysis of the terrorist network is sensitive to the model specification. Under one model it clearly identifies a group containing the leaders and organisers, and the group structure suggests a hierarchy of leaders, trainers and "footsoldiers" who carry out the attacks.

Keywords: terrorist groups, latent classes, Bayesian model comparison, Noordin Top terrorist network.

## 1 Introduction

This paper applies a statistical model for groups in a network, previously used for a small social network (Aitkin, Vu and Francis 2014) to the investigation of the group structure of the Noordin Top terrorist network. The approach addresses the identification of *actor* groups within the framework of a *two-mode* or *bipartite network* of actors attending events, through the latent class (*stochastic block*) model in which the groups of actors are represented by classes, which are not directly observable, but which can be probabilistically reconstructed from the event attendance patterns of the actors.

Our analysis is distinctly different from most applications of the stochastic block model, which have been to *one-mode* or *unipartite* networks. In the modelling process several questions are of critical importance: how many classes can be identified; the nature of the membership of the classes, for example whether individuals belong to one or to many classes; whether the *leadership* of the network can be identified, and if so the roles of those below the leaders; whether other models might give a better representation of the data.

We address these questions in a Bayesian framework, and use recent developments in Bayesian model comparisons (Aitkin 2010, Aitkin, Vu and Francis 2015) to illuminate the choice among possible models. Aitkin, Vu and Francis (2014) showed the application of the approach to a famous sociological data set from Davis, Gardner and Gardner (1941): the Bayesian analysis, unlike most other analyses in the literature, reproduced the conclusions of the original sociologists based on detailed interviews. In this paper we show its application to a recent data set (Everton 2012) on connections among members of the Noordin Top terrorist network.

## 2 The terrorist network

Information about Noordin Mohammad Top and his network was published regularly by the International Crisis Group (2009 for example). Details of his death can be found at

`http://www.nytimes.com/2009/09/18/world/asia/18indo.html`

We give a Wikipedia summary from 2014.

> Noordin Mohammad Top, a Malaysian citizen, was a Muslim extremist, also referred to as (Noordin) Din Moch Top, Muh Top, or Mat Top, and was Indonesia's most wanted Islamist militant. He is thought to have been a key bomb maker and/or financier for Jemaah Islamiyah (JI) and to have left JI and set up a more violent splinter group Tanzim Qaedat al-Jihad.
>
> Top and Azahari Husin were thought to have masterminded the 2003 Marriott hotel bombing in Jakarta, the 2004 Australian embassy bombing in Jakarta, the 2005 Bali bombings and the 2009 JW Marriott-Ritz-Carlton bombings, and Top may have assisted in the 2002 Bali bombings.
>
> Top, nicknamed "Moneyman", was an indoctrinator who specialized in recruiting militants into becoming suicide bombers and collecting funds for militant activities. Husin was killed in a police raid on his hideout in Batu, near Malang in East Java on 9 November 2005. Top was killed during a police raid in Solo, Central Java, on 17 September 2009 conducted by an Indonesian anti-terrorist team.

The data set in this paper is the terrorist network surrounding Noordin Top and Azahari Husin, documented in Everton (2012); the following quote is from Everton's Appendix 1:

> This subset of the Noordin Top terrorist network [was] drawn primarily from "Terrorism in Indonesia: Noordin's Networks," a 2006 publication of the International Crisis Group. It includes relational data on the 79 individuals listed in Appendix C of that publication. The data were initially coded by Naval Postgraduate School

students as part of the course "Tracking and Disrupting Dark Networks" under the direction of Professor Sean Everton, Co-Director of the CORE Lab, and Professor Nancy Roberts. CORE Lab Research Associate Dan Cunningham also reviewed and helped clean the data.

The network described in Everton (2012) and analysed there covers the period 2001-2010. The 79 individuals in this study – the actors – entered the network during this period, and either remained in it, or left through arrest and imprisonment or death. The presence of actors at events – meetings or joint actions – during this period is recorded, though the dates of the events are generally not given. Our analysis is restricted to 74 of these actors: five were eliminated as they were not present at any of the 45 events recorded. (These actors appear as "isolates" in several of Everton's analyses.) The presence of *links* – connections – between the actors and the events was inferred from their mention together in public reports in newspapers and elsewhere.

The presence of actor $i$ at event $j$ is denoted by an indicator variable $Y_{ij} = 1$; if actor $i$ was not present at event $j$ then $Y_{ij} = 0$. The complete set of indicator values forms a matrix, called the *adjacency matrix*, of general dimension $n \times r$, where $n$ is the number of actors and $r$ the number of events. In smaller networks this matrix can be printed and read. For large arrays like the terrorist network it is too large to be readable, and a "map" format is used instead for visualisation. Here the presence of a link (1) in cell $(i, j)$ is shown by a black square, and the absence of a tie (0) by a white square. The map for the terrorist network is shown in Figure 1. The full adjacency matrix is given in the supplementary materials.

The reported connections are classified by Everton into six categories of events, which we will call *scales*. These are the Organisations (8 events), Operations (5), Training (11), Finance (2), Logistics (7) and Meetings (12) scales.

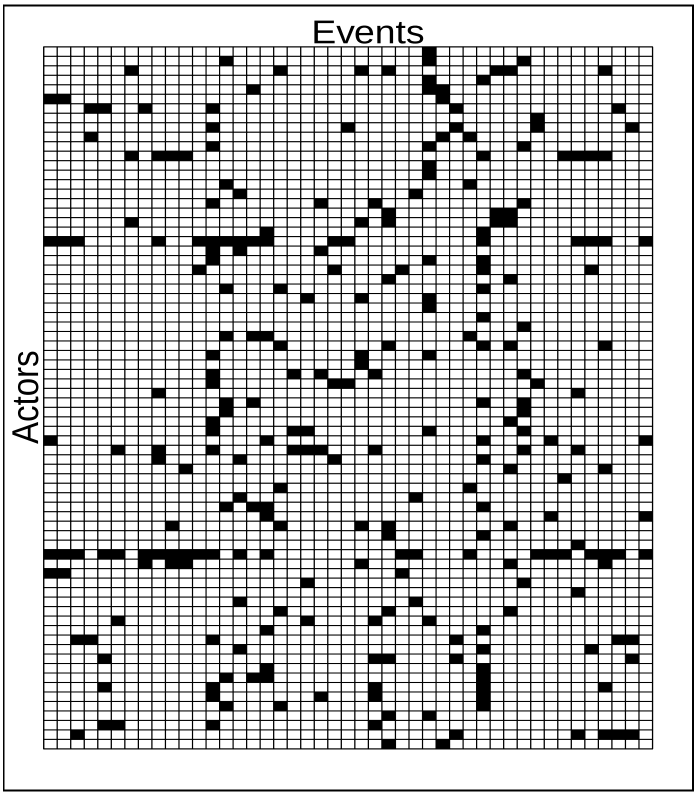


Figure 1: Noordin Top network matrix

The matrix is *very sparse*. There is no clear division into groups, though two actors (Top and Husin) are heavily involved in events (attending 23 and 17 events respectively). A conventional two-dimensional representation of the matrix is given in Figure 2, with actors shown as circles and events as squares (the online version is in colour). The two leaders are shown as large circles. It is difficult to identify group structure from this graph.

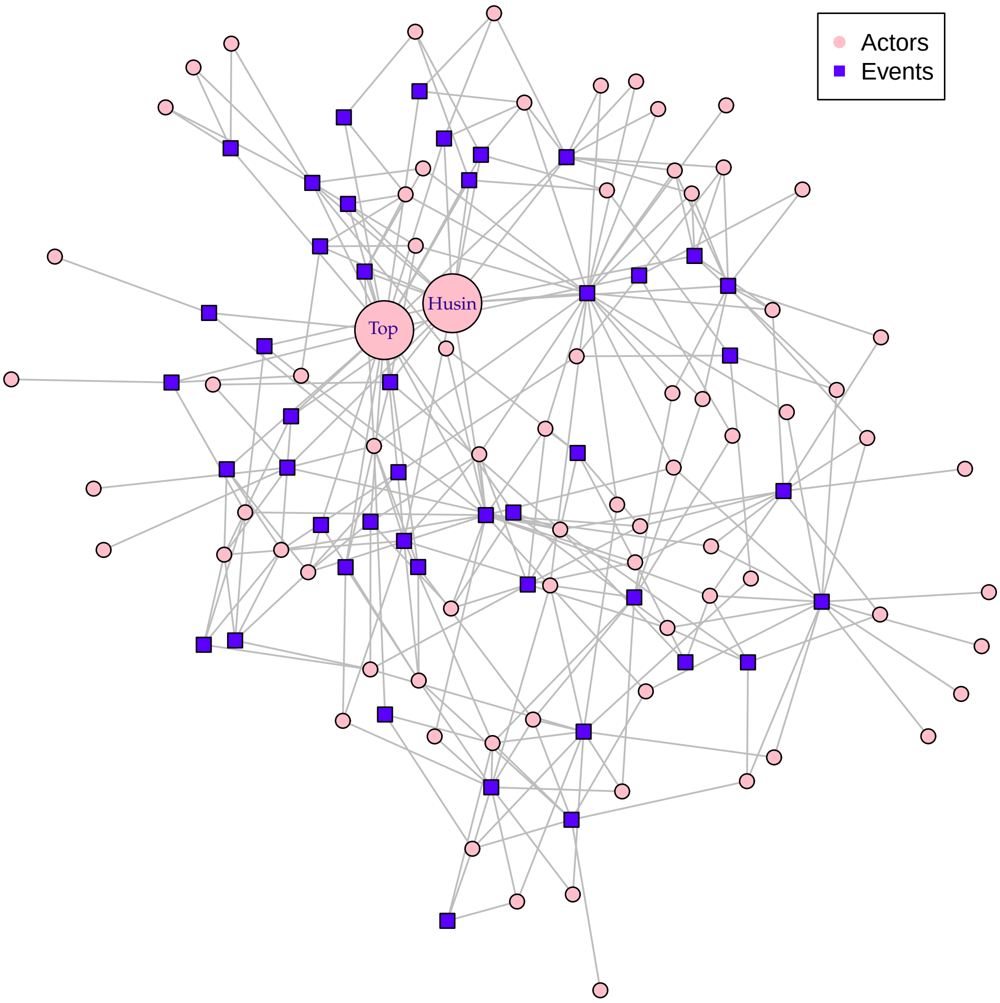


Figure 2: Noordin Top network map

# 3 Data strucures for network analysis

The standard methods for social network analysis have been developed for a type of network different from that given here. These methods assume that the data consist of binary "tie" variables, from *direct connections between actors* – not from the *indirect* connections between actors attending the same events. Networks of this type (for example friendship networks) – *one-mode* or *unipar-*

*tite* networks, have well-developed methods which however cannot be used for bipartite networks.

A common approach to the analysis of bipartite networks is to convert them to unipartite networks, by *projection* of the adjacency matrix $Y$ into its outer product $YY'$. For the terrorist network this involves summing over the events attended, to give a unipartite *valued network* whose $(i, j)$ entry is the number of events *jointly attended* by actors $i$ and $j$. The "value" would be a count, and not in general a 0 or 1 value. A common further approach is to convert the count to a binary, by dichotomizing the value using some horizon. We discuss in §10 the dangers of this approach, recently pointed out strongly by Neal (2014) and Gerdes (2014).

# 4 Analyses by Everton

The Noordin Top network is analysed extensively in the chapters of Everton's book, which is in addition a very detailed manual for network analysis using the major packages (UCINET, Pajek and ORA) developed for this purpose. The bipartite networks considered there are converted to unipartite networks in these analyses. The analyses provide a very detailed examination of the network properties of the standard kinds in social network analysis, and we do not detail them here. (Graphical representations of networks vary among packages.)

We follow a different form of analysis, based on explicit probability models for the bipartite network and its group structure, described in §6. We discuss the conclusions from our analysis, and relate them to those from Everton's analyses, in §10.

# 5 The meaning of a group or class

A fundamental question which has to be addressed first is what we mean by a *group*, in this social context. We should first note than even the *word* for this subset of actors is not consistent across research fields: it is also called *community*, *clique* and *class*.

We adopt, as in Aitkin, Vu and Francis (2014), the definition of a group or class as *an identifiable subset of actors who tend to attend the same events*. This definition will be made specific after we consider possible models for the event attendance variables.

# 6 Statistical models

## 6.1 Models for a random process

We consider the presence or absence of an actor at an event as a *random process* – attendance was determined by a possibly large number of factors unknown to us, so we represent the process outcome as a *Bernoulli random variable*, taking the value $Y_{ij} = 1$ with probability $p_{ij}$, and $Y_{ij} = 0$ with probability $1 - p_{ij}$. The probability of the pattern of "responses" $\{y_{ij}\}$ given the set of probabilities $\{p_{ij}\}$ over all actors and all events is

$$\Pr[\{y_{ij}\} \mid \{p_{ij}\}] = \prod_i \prod_j p_{ij}^{y_{ij}} (1 - p_{ij})^{1-y_{ij}},$$

assuming (a strong assumption!) the independence of event attendance both within and among actors. We can bring the actor and event structures into the model in several ways. Aitkin, Vu and Francis (2104) gave an extensive list. Here we report on a subset of models used for the terrorist network.

## 6.2 The Rasch model

The Rasch model is widely used in *item response theory* (IRT) in psychology. Applied to a network, it is expressed through row and column parameters. Each actor $i = 1, ..., n$ has a *propensity* $\theta_i$ to attend any event. Each event $j = 1, ..., r$ has an *attractiveness* $\phi_j$ to any actor. Actors attend events *independently*, and independently of each other. The Rasch model is a *main effect* or *additive* model, in actors and events, on the logit scale:

$$\operatorname{logit} p_{ij} = \log\left(\frac{p_{ij}}{1 - p_{ij}}\right) = \theta_i + \phi_j.$$ [1]

For the Noordin Top network, the "events" attended by network members include meetings of various kinds (detailed above), and participation in attacks. The actor's "propensity" to attend events reflects his special or general abilities and level of responsibility.

The Rasch model has *no group structure* for actors, and so plays the role of a *baseline* model for comparison with models with group structure. Other "link" functions could be used instead of the logistic. Caron (2012) considered the complementary log-log link.

## 6.3 The latent class model

We now switch nomenclature and refer to groups as *classes*, as this is the tradition in latent class modelling. The use of this model in social networks dates from Holland, Laskey and Leinhardt (1983), though it has been used then and since only for one-mode networks, apart from Aitkin, Vu and Francis (2014). However it is very well-established in sociology for contingency table analysis, from Lazarsfeld and Henry (1968) and Goodman (1974) onwards, and more recently for binary incidence matrices at the individual level. The model specifies a *$K$-class latent structure* for actors; the classes are distinguished by different sets of event attractiveness parameters *among* classes but identical attractiveness parameters *within* classes.

[1] One of the parameters is unidentifiable in this parametrization: it is usually assumed that one of the $\theta_i$, or one of the $\phi_j$, is equal to zero.

Within each class $k$ the model fitted is a variation of the Rasch model with class-specific event attractiveness parameters $\phi_{jk}$, and class instead of actor "intercepts" $\psi_k$. The model assumes independence between the event attended *within classes*; this weakens the assumption of full independence. Marginally (summing over the unobserved classes), the events attended by actors are correlated, due to the omitted class variable. This *unconditional dependence* but *conditional independence* is characteristic of latent variable models, including the classical factor model. Assessing the validity of the conditional independence assumption is difficult, as for the factor model, since there is in general no *exact partition* of the actors into the latent classes which would allow this assessment. For sparse data, as in the Noordin Top network, the assessment of conditional independence would have further (lack of data) difficulties.

Replacing $p_{ij}$ by $q_{ijk}$ to incorporate the latent structure, the formal model is:

$$
\begin{aligned}
\Pr[\{Y_{ij}\} \mid k, i, \{q_{ijk}\}] &= \prod_{j=1}^{r} q_{ijk}^{y_{ij}} (1 - q_{ijk})^{1-y_{ij}} \\
\Pr[\{Y_{ij}\} \mid i, \{q_{ijk}\}] &= \sum_{k=1}^{K} \left[ \pi_k \prod_{j=1}^{r} q_{ijk}^{y_{ij}} (1 - q_{ijk})^{1-y_{ij}} \right] \\
\Pr[\{Y_{ij}\} \mid \{q_{ijk}\}] &= \prod_{1=1}^{n} \left\{ \sum_{k=1}^{K} \left[ \pi_k \prod_{j=1}^{r} q_{ijk}^{y_{ij}} (1 - q_{ijk})^{1-y_{ij}} \right] \right\} \\
\operatorname{logit} q_{ijk} &= \psi_k + \phi_{jk}.
\end{aligned}
$$

The class intercepts $\psi_k$ in this model are not identifiable separately from the event attractiveness parameters $\phi_{jk}$ without some form of constraints. The logit model can be represented alternatively as $\lambda_{jk}$, absorbing the $\psi_k$ into the $\phi_{jk}$. We make use also of the extended latent class model, retaining individual actor propensities: $\operatorname{logit} q_{ijk} = \theta_i + \phi_{jk}$.

The probability that actor $i$ is in class $k$ may depend on actor covariates $\mathbf{x}_i$ but is independent of the tie variables $y_{ij}$. The model parameters $\pi_k, \theta_i$ or $\psi_k$ and $\phi_{jk}$ do not have to be known or specified; they can be estimated by now-standard methods, discussed in §§7,8.

An important question is how to determine the number of classes $K$; this is discussed at length in §8.5.

# 7 The model likelihood

Statistical models are estimated, assessed and compared through the *model likelihood*. We have a model $f(\{y_{ij}\} \mid \boldsymbol{\lambda})$ for data $\{y_{ij}\}$, depending on model parameters $\boldsymbol{\lambda}$. The likelihood is

$$L(\boldsymbol{\lambda}) = L(\boldsymbol{\lambda} \mid \{y_{ij}\}) = f(\{y_{ij}\} \mid \boldsymbol{\lambda}).$$

For the extended latent group model with $K$ groups, the likelihood follows immediately from the mixture model specification above:

$$\begin{aligned} L(\boldsymbol{\lambda}) &= \prod_{1=1}^{n} \left\{ \sum_{k=1}^{K} \left[ \pi_k \prod_{j=1}^{r} q_{ijk}^{y_{ij}} (1 - q_{ijk})^{1-y_{ij}} \right] \right\}, \\ \operatorname{logit} q_{ijk} &= \theta_i + \phi_{jk}, \end{aligned}$$

with $\boldsymbol{\lambda} = (\{\pi_k\}, \{\theta_i\}, \{\phi_{jk}\}, K)$.

We present in the remainder of the paper the Bayesian analyis of the model, following Aitkin, Vu and Francis (2014). For reasons discussed in detail there (the greater precision achievable with the Bayesian analysis), we do not present the maximum likelihood analysis.

We examine the following sequence of latent class models of decreasing complexity for the logit of the probability $q_{ijk}$ that actor $i$ attends event $j$ (itself a part of scale $s$) in class $k$.

A: $\theta_i + \phi_{kj}$ – an extension of the latent class model, with individual actor propensities and class-specific event attendance parameters;

B: $\psi_k + \phi_{kj} \equiv \lambda_{kj}$ – the "standard" or "classic" latent class model: the class intercepts are not separately identifiable from the class-specific event parameters;

BS: $\psi_k + \phi_{ks} \equiv \lambda_{ks}$ – the "scale" version of B;

C: $\psi_k + \phi_j$ – constrained model B with common event attendance parameters across classes;

CS: $\psi_k + \phi_s$ – constrained model BS with common scale parameters across classes.

# 8 Bayesian analysis of two-mode networks

## 8.1 Priors and posteriors

We augment the model likelihood $L(\lambda)$ by a *prior distribution* $\pi(\lambda \mid \gamma)$ for the model parameters $\lambda$ depending in general on prior parameters $\gamma$, and use Bayes's theorem to update the prior distribution to the *posterior distribution* $\pi(\lambda \mid \mathbf{y}, \gamma)$:

$$\pi(\lambda \mid \mathbf{y}, \gamma) = \frac{L(\lambda) \cdot \pi(\lambda \mid \gamma)}{\int L(\lambda) \cdot \pi(\lambda \mid \gamma) \, \mathrm{d}\lambda}.$$

The denominator is a scaling term, depending on the data $\mathbf{y}$ and $\gamma$ but not $\lambda$.

If the prior is *flat* – constant – then the posterior distribution is a simple scaled version of the likelihood. Throughout this paper, following Aitkin (2010) we use flat, *reference* or *non-informative* priors, to allow as far as possible the *data* to determine the posterior distribution through the likelihood.

Inference about these parameters is through their posterior distributions, commonly through the posterior mean, posterior median and other percentiles. Credible intervals for the parameters follow from the posterior percentiles, in a way familiar from frequentist inference but without any assumption of normality. The posterior standard deviation is often quoted, though this is useful only for normal posterior distributions.

## 8.2 Priors for the latent class models

We implemented the Markov chain Monte Carlo (MCMC) procedure in OpenBugs. We followed the model structures A–CS set out in §7, with minimally informative priors: a Dirichlet (1,1,...,1) prior for the class proportions $\pi_k$; independent normal priors with mean zero and variance 10 for the propensity parameters $\theta_i$, class intercepts $\psi_k$ and event attendance parameters $\phi_j$ or scale parameters $\phi_s$.

After convergence we drew 10,000 random values from their joint posteriors and used every 10th value to reduce serial dependence.

## 8.3 Posteriors of functions of data and parameters

One of the powerful features of Bayesian analysis is its ability to provide posterior inference about complicated functions of the data and parameters. In frequentist theory we have to rely on the *delta method* – Taylor series expansions – to obtain the asymptotic sampling distributions of non-linear functions of the model parameters, especially *ratios* of parameters.

In Bayesian theory this is unnecessary; for *posterior sampling* inference about a non-linear function $g(\lambda)$ of the model parameters, we simply make $M$ random draws $\lambda^{[m]}$ of $\lambda$ from $\pi(\lambda \mid \mathbf{y})$, and substitute them into the function $g$, to give $M$ random draws $g^{[m]} = g(\lambda^{[m]})$ from the *full posterior distribution* of $g(\lambda)$.

## 8.4 Posterior distribution of class membership probabilities

A general problem with the use of posterior class membership probabilities from Bayes's theorem following the EM algorithm is familiar from the frequentist analysis of other complex models. This is the problem of *overstated precision* resulting from the substitution of ML estimates for true parameter values, without any allowance for the imprecision of the ML estimates. In frequentist analysis this is forced on us by the complexity of the exact sampling distributions, especially for non-linear functions of the parameters.

As described above, the posterior distributions of these quantities can be obtained in theory from the random draws of the parameters. Recall that the posterior probability of membership of case $i$ in class $k$, in the general $K$-component mixture with component $k$ density $f_k(y \mid \lambda_k)$ with component-specfic parameter $\lambda_k$, is

$$\pi_{ki} = \frac{\pi_k f_k(y_i \mid \lambda_k)}{\sum_{\ell=1}^{K} \pi_\ell f_\ell(y_i \mid \lambda_\ell)}.$$

From the posterior distributions of $\lambda_k$ and $\pi_k$, we make $M$ independent draws $\lambda_k^{[m]}$ and $\pi_k^{[m]}$ and substitute them into $\pi_{ki}$ to give $M$ draws

$$\pi_{ki}^{[m]} = \frac{\pi_k^{[m]} f_k(y_i \mid \lambda_k^{[m]})}{\sum_{\ell=1}^{K} \pi_\ell^{[m]} f_\ell(y_i \mid \lambda_\ell^{[m]})}.$$

However, a major issue in computing posterior distributions for *any* class-specific parameter is *label-switching*. In the ML estimation of the model parameters, the class labelling of the $K$ class parameter estimates is arbitrary, and causes no confusion. However in the MCMC iterations, the class labelling can vary during iterations and switch the class labels around, leading to class-specific posteriors that are mixtures of the posteriors from each true class. This can lead in the worst case to identical class-specific distributions for all the class parameters, despite convergence.

We follow the approach of Sperrin et al (2010), in which the labels to be attached to the $M$ sets of posterior parameter draws are treated as *missing data* and analysed with an EM algorithm. For the two-class model, there is almost no uncertainty about the draw labels, but there is considerably more uncertainty with the three- or more class models. We discuss this below.

## 8.5 Posterior distribution of the model deviance

### 8.5.1 General models

A particularly useful application of the posterior sampling approach is to the deviance. In Bayesian terms the deviance is $D(\lambda) = -2\log L(\lambda)$. Since this is a function of both $\lambda$ and the data $\mathbf{y}$, it also has a posterior distribution obtainable in this way: given $M$ random draws $\lambda^{[m]}$, we substitute them into the deviance to give $M$ random draws $D^{[m]} = D(\lambda^{[m]})$.

A full discussion of this approach, and many applications of it, are given in Aitkin (2010). Aitkin, Vu and Francis (2015) carried out a simulation study to evaluate this approach, for both normal mixtures and Bernoulli latent class models. For latent class models, correct identification of many classes required substantial sample sizes of actors, in the simulations based on binary symptoms in psychiatric patients. Our application here is to Bayesian model comparisons of the number of latent classes. A normal mixture example can be found in Chapter 9 of Aitkin (2010).

Within a model type (A, B, BS, C, CS) we have $K$ possible models, consisting of latent class models with $k = 1, ..., K$ classes, specified by $K$ sets of model parameters $\lambda_k$ containing the class proportion parameters $\pi_k$, the class propensity parameters $\theta_k$ and the class-conditional event attendance parameters $\phi_{kj}$, or scale parameters $\phi_{ks}$. For each $k$ we obtain the posterior distribution of $\lambda_k$, and the consequent posterior distribution of the deviance $D_k$. The deviance is unaffected by label-switching, since it is a *symmetric function* of the class labels, and invariant under their permutation.

The comparison of the class models is equivalent to the comparison of the *distributions* of the class deviances. In large samples from regular models with

MLEs internal to the parameter space, the second-order Taylor expansion of the deviance $D(\lambda)$ about the MLE $\hat{\lambda}$ gives the posterior distribution of the model deviance, by a Bayesian version of the derivation of the asymptotic $\chi^2$ distribution for the likelihood ratio test statistic:

$$D_k(\lambda_k) \sim D_k(\widehat{\lambda}_k) + \chi^2_{p_k},$$

where $p_k$ is the dimension of $\lambda_k$ (Aitkin 2010 p. 53). It is notable that the posterior distribution *starts from the frequentist deviance*, since this is minimised over all possible parameter values.

The asymptotic result, if it applies at all, is restricted to the simplest models: null and scale, for which the parameter dimension is small compared to the number of ties. For the mixture models, the actual posterior distributions depart from their asymptotic forms in two respects, as described in Aitkin (2010 pp. 216-220): the distributions *start from larger values than the frequentist deviances*, because with increasing parameter dimension it is increasingly difficult to sample by chance the MLE; and the distributions are *more diffuse* than the $\chi^2$ distributions because the parameter posteriors are skewed and/or heavy-tailed.

With increasing $K$ the frequentist deviances for the models (not shown) are increasingly far below the smallest values in the 1,000 draws from the posterior distributions of the model deviances. This phenomenon is discussed in Aitkin, Vu and Francis (2014) – it is the increasing difficulty of randomly sampling parameter values near the MLE as $K$ increases. It shows that the maximized likelihood is not a satisfactory summary of the likelihood evidence for the model, and explains why model comparison procedures based on the maximized log-likelihood – the likelihood ratio test, AIC and BIC – are not reliable indicators of model support for heavily parametrized models, especially with sparse data.

However, the deviance distributions can in many cases be *stochastically ordered*. There are several possibilities, which are illustrated in the next section with the Noordin Top network.

1. **Complete stochastic ordering**: the deviance cdfs are *ordered* and *do not cross* (except merge at 0 and 1); then the leftmost distribution gives the best-preferred model.

2. **Partial stochastic ordering**: the leftmost distribution does not cross the others, which may cross each other; then the leftmost distribution gives the best-preferred model.

3. **No stochastic order**: some or all of the distributions cross.

We discuss these further, and give a procedure for discriminating the models, with the analysis of the Noordin Top data.

# 9 Bayesian analysis of the Noordin Top network

We fitted one-, two-, three-, four- and five-class models A, B, BS, C and CS in a Bayesian analysis. For each model and set of classes we computed the posterior distribution of the class deviances. These are shown for each model in Figures 3–7.

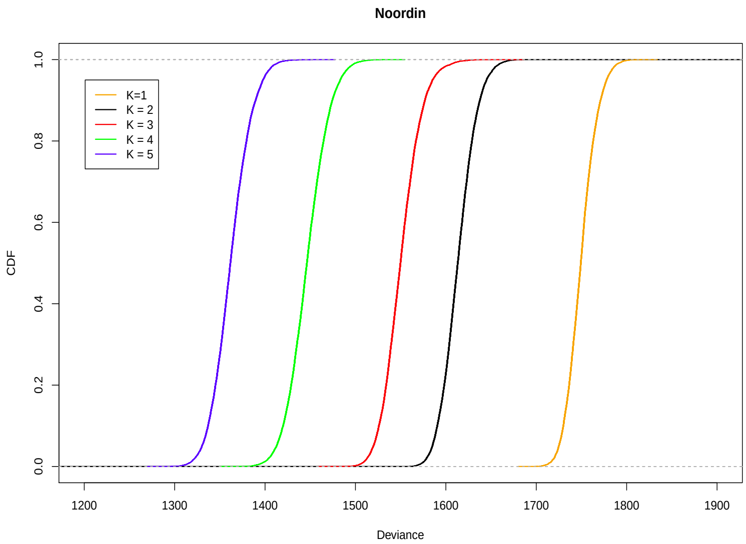


Figure 3: Noordin Top model A deviance distributions

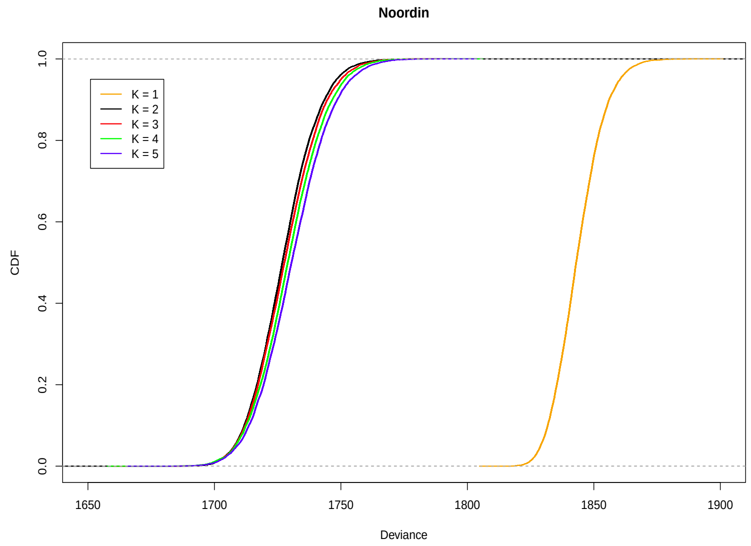


Figure 4: Noordin Top model B deviance distributions

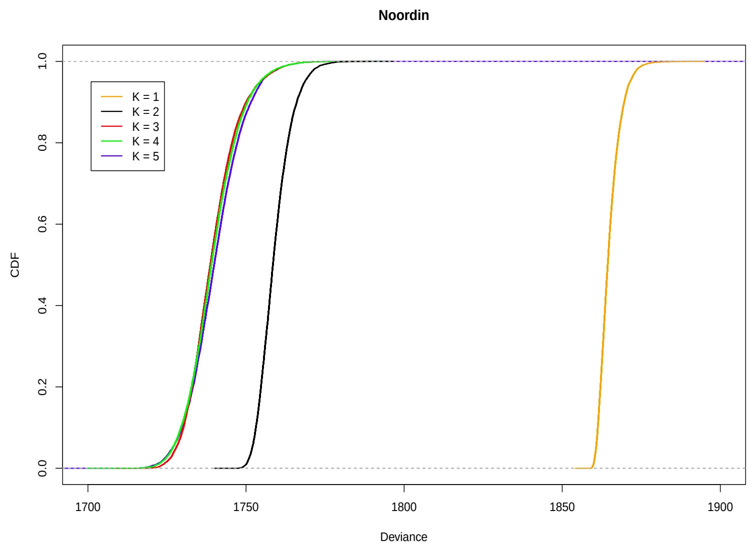


Figure 5: Noordin Top model BS deviance distributions

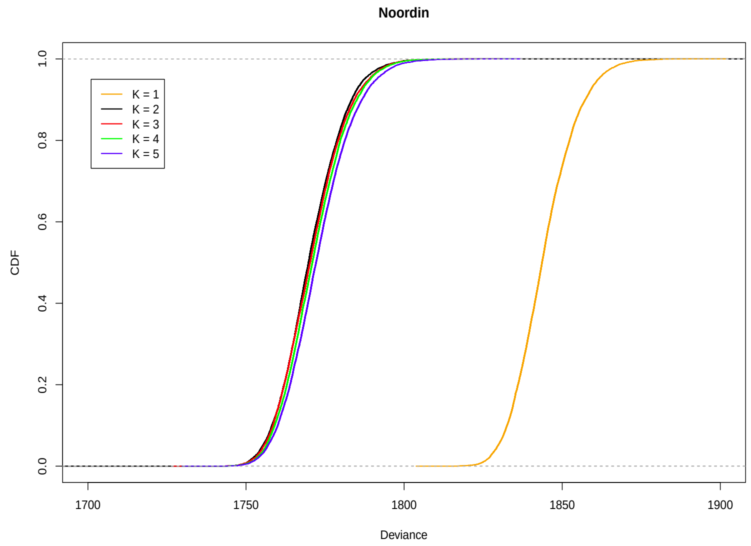


Figure 6: Noordin Top model C deviance distributions

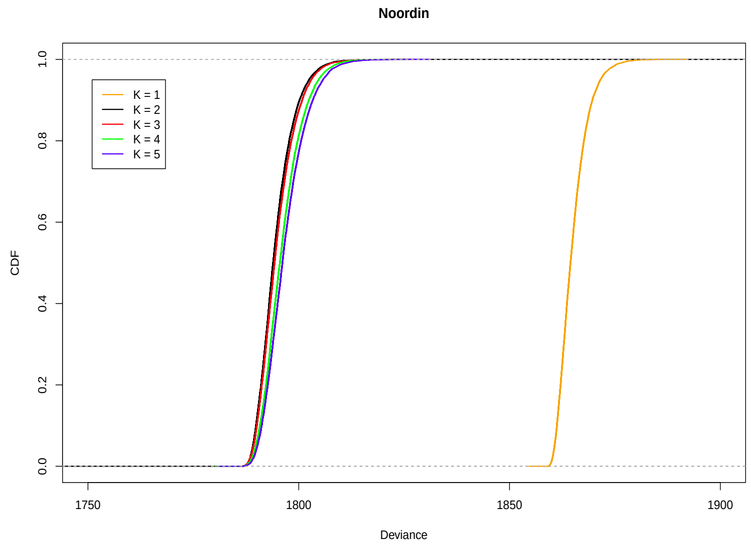


Figure 7: Noordin Top model CS deviance distributions

Apart from Figure 3, the deviance graphs show a consistent pattern: the deviance distributions decrease stochastically from $K = 1$ to $K = 3$, and then increase very slowly. These graphs suggest that three classes are established, but not more. Figure 3 shows a different pattern: the deviance distributions decrease stochastically from 1 all the way to 5 classes.

A careful comparison of the deviance scales in the five figures shows that the deviance distributions for 3, 4 and 5 classes with model A are stochastically smaller than the three-class distributions for models B, BS, C and CS (and for any other number of classes).

To understand the 3-class and 5-class A models, we compute the membership probabilities for each actor in each class, from the means of the membership indicator variables across the 1,000 random draws. Table 2 in the supplementary materials gives these probabilities for the 3-class model. We give in Figure 8 a ternary plot of these probabilities. The centrepoint of the triangle corresponds to probability 1/3 of membership in each class. Actors at a vertex have probability 1 of belonging to the vertex class. Actors spread along a side of the triangle have non-zero probabilities of belonging to both the vertex classes. Actors in the interior have non-zero probabilities of belonging to all three classes.

The labels (for the number of events attended) are jittered perpendicular to the axes. Labels are placed only for actors for whom the number of events attended is five or more. The membership pattern is unclear. Husin (attending 17 events) defines Class 1, but Top (attending 23 events) is some distance towards Class 2. Most of the members attending five or more events are spread out along the Class 1 - Class 2 axis. Class 2 has 22 members attending less than five events, and Class 3 a small group of seven actors clearly identified, with a small number spread along the Class 2 - Class 3 axis, and two members on the Class 3 - Class 1 axis. Several actors are in the interior of the triangle, with very diffuse membership probabilities. So the nature of the three classes (in terms of membership) is unclear.

The five-class model is even more unclear. Table 3 in the supplementary materials shows that the small Class 5 has only four clearly identified (probability 1.00) members ($i$ = 7, 9, 63 and 73), all with degree (number of events attended) 5 or 6. Class 1 is also sparse, with nine clearly identified members ($i$ = 3, 18, 19, 25, 32, 45, 51, 55, 60) with degrees 2, 3 or 5. Class 3 which contains the leaders ($i$ = 21 and 54) also has eight other clearly defined members ($i$ = 12, 16, 44, 48, 54, 56, 59 and 64) with degrees 2, 3, 4, 5 and 9. Class 4 has only eight clearly identified members ($i$ = 11, 17, 27, 33, 35, 41, 43, 61) with degrees 3, 4 or 5, and one with degree 9. So only 31 of the 74 actors are clearly identified with a class: the other 43 have appreciable probabilities of belonging to more than one class.

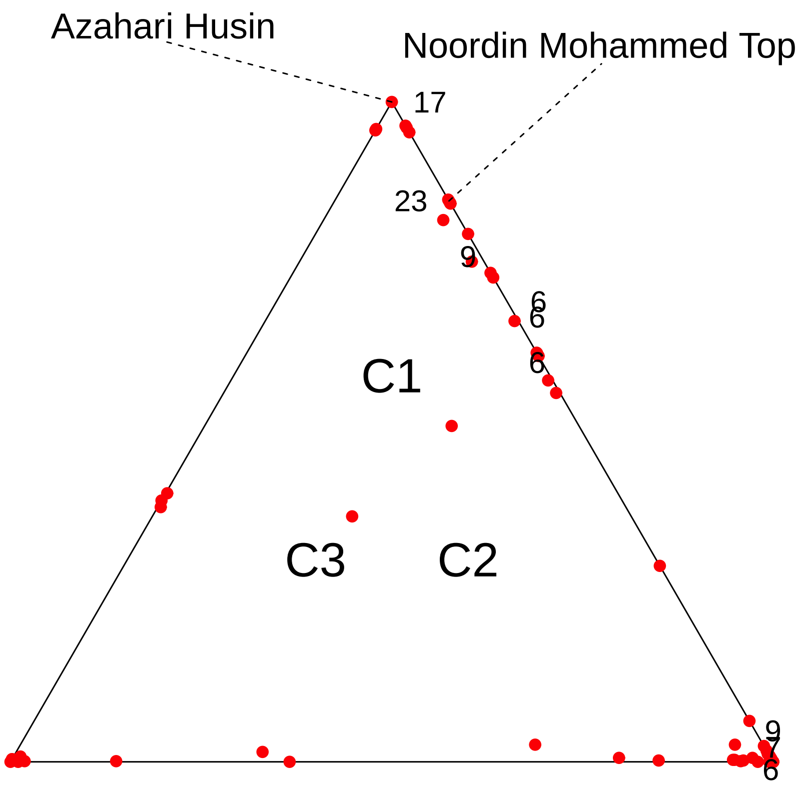


Figure 8: Three-class membership probabilities, Model A

### 9.1 Modelling the prior

Most of the latent class models considered here are heavily parametrised, and the flat priors on the parameters do not change this. The sparse data from the adjacency matrix provide little information about these unrelated parameters. A general Bayesian approach to heavily parametrised models is to use priors which replace the unrelated parameters by a *distributional model* with a small number of parameters. This approach is widely used in all fields of *random effect modelling*, including Rasch modelling (de Boeck 2008), though it is not so widely used in latent class analysis.

We evaluate this approach, and its effect on the interpretation of the latent classes, by *modelling the class and event parameters*. We write $s$ as before for scale, and change slightly the model notation from §7 to: $y_{ijsk} = 1$ if actor $i$ in class $k$ attended event $j$ in scale $s$, with probability $q_{ijsk}$. We now write $\psi_k$ for classes and $\phi_{sj}$ for events within scales, and use the model

$$\begin{aligned} \text{logit } q_{ijsk} &= \psi_k + \phi_{sj} \\ \psi_k &\sim N(\mu_\psi, \sigma^2_\psi) \\ \phi_{sj} &\sim N(\mu_\phi, \sigma^2_\phi) \end{aligned}$$

and priors

$$\begin{aligned} \mu_\psi &\sim N(0, 4) \\ \sigma^2_\psi &\sim U(0, 2) \\ \mu_\phi &\sim N(0, 100) \\ \sigma^2_\phi &\sim U(0, 10). \end{aligned}$$

For the membership indicator variables we use the same diffuse multinomial/Dirichlet model and prior as in the previous analysis:

$$\begin{aligned} Z_{ik} &\sim M(1, \pi) \\ \pi &\sim D(1, 1, ..., 1). \end{aligned}$$

We call this model the *random Rasch latent class model*.

The likelihood now contains the additional unobserved "random effect" terms $\psi_k$ and $\phi_{sj}$, which have to be integrated out of the likelihood with respect to their model distributions, leaving the likelihood as a function of the prior parameters $\mu_\psi, \sigma^2_\psi, \mu_\phi, \sigma^2_\phi$. As with the frequentist "fixed effect" and "random effect" models, this likelihood is not comparable with the "fixed effect" likelihood with flat priors on all the parameters, so the deviance scales for these two different models will not be comparable.

The random latent class model is related to all the previous models A – CS, but is more general in allowing for wide variations in the importance of the item and class parameters. The deviance distributions based on 10,000 draws of the model parameters and 10% thinning are shown for the Rasch model and the random Rasch latent class model with 2–5 classes in Figure 9.

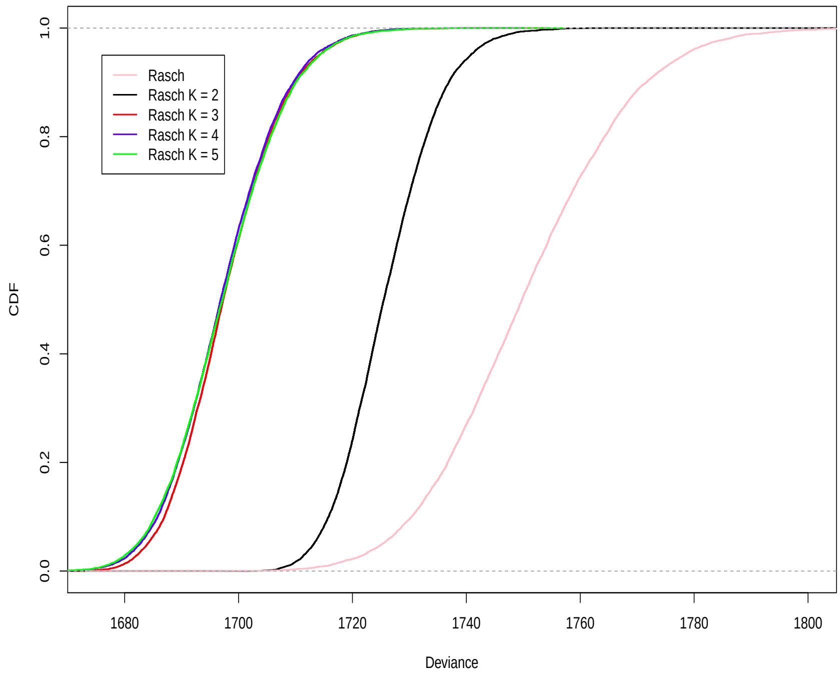


Figure 9: Noordin Top network deviance distributions, random Rasch latent class model

The deviance distributions improve (move left) substantially from the Rasch through the 2-class to the 3-class model. The distributions for 3, 4 and 5 classes overlap almost completely. Three classes are clearly established, while the fourth and fifth give unnecessary complexity.

Figure 10 shows a ternary plot of membership probabilities in the 3-class model, with labels (for the number of events attended) jittered perpendicular to the axes. Labels are placed only for actors for whom the number of events attended is five or more.

This plot is quite different from the corresponding plot in the "fixed effect" analysis with model A. Class 1 now contains only the two leaders and planners: almost no other actor has non-zero probability of belonging to Class 1. Class 2 contains most of the actors participating in 5 or more events, and class 3 contains mostly actors participating in 4 or fewer events. There is some uncertainty for some actors who have appreciable probabilities of being in both Classes 2 and 3.

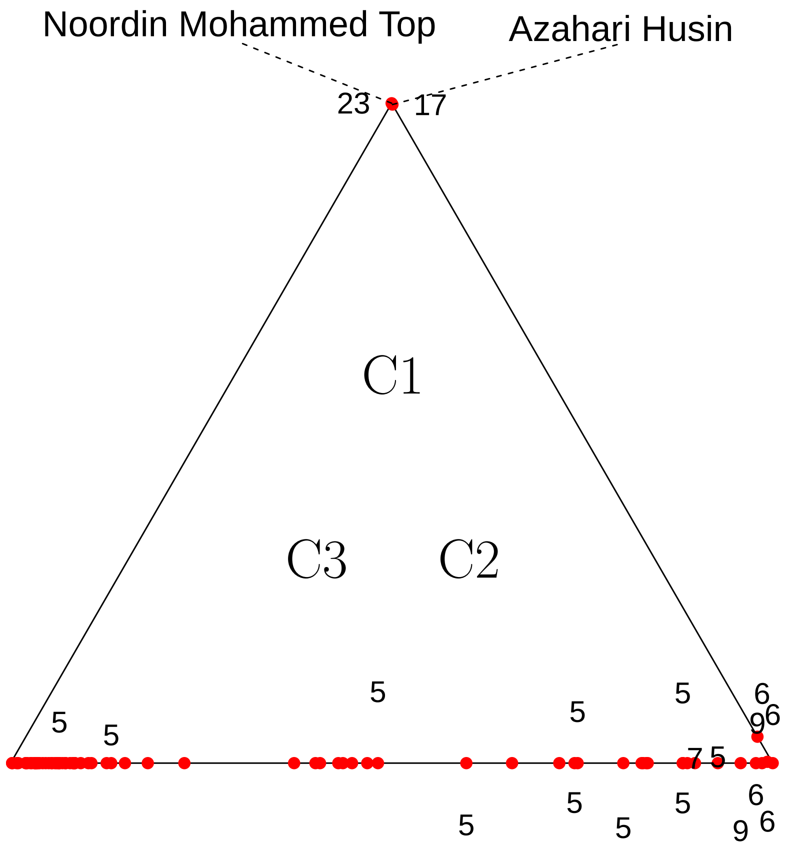


Figure 10: Three-class membership probabilities, random Rasch model

A natural question, given the infrequency of appearance of most actors at *any* event, is whether there is a real difference between classes 2 and 3 in the frequency profiles of their event participation. To address this this we give in the supplementary materials graphs of the posterior distributions of the event attendance probabilities for all three classes, broken down by the event scales.

It is immediately clear that the Class 1 leaders are heavily involved in Finance (one of them is present in every Finance meeting), Logistics, Operations and Meetings, while Class 2 are involved *more frequently* than Class 3 in Finance and Logistics, but *less frequently* than Class 3 in Operations. It is notable that both Class 2 (particularly) and Class 3 are involved rarely in more than one operation: this may be due to the high chance of death or capture in such operations, which were apparently planned, but rarely carried out, by the leaders.

We can compare the roles of the actors as given by Everton with their latent class memberships. Table 1 gives the actor roles defined by Everton. Role 9 (propagandist) is omitted: none of the 74 actors was so identified by Everton.

| Code | Role |
|---|---|
| 0 | No information/unclear |
| 1 | Strategist: high-level planner of a terrorist/insurgent network |
| 2 | Bomb maker: individual who constructs bombs |
| 3 | Bomber/fighter: individual who participates in bombing attacks or who is described as a fighter |
| 4 | Trainer/instructor: individual who trains or instructs new members of a terror network |
| 5 | Suicide bomber: individual who plans to perform, or already has performed a suicide attack |
| 6 | Recon and surveillance: individual who engages in the surveillance and recon of targets |
| 7 | Recruiter: individual who engages in identifying and recruiting new members (to include bombers) |
| 8 | Courier/go-between: individual used in communications between members |
| 10 | Facilitator: individual who assists in the operation of the network (especially with material and finance) |
| 11 | Religious leader: individual who provides religious training and support |
| 12 | Commander/tactical leader: individual in charge of operations at the local/tactical level |

Table 1: Actor roles from Everton

Figure 11 shows the 74 actors identified by the 13 role labels given by Everton (2012 p. 396) in Table 1. The labels are jittered away from the axes.

Notable differences between the 21 members of Class 2 and the 51 members of Class 3 (as assessed by the largest posterior probability) are that all seven suicide bombers, and 11 of the 14 bomber/fighters, are in Class 3, while 5 of the 7 couriers are in Class 2. The other role categories do not differ substantially between classes. Our interpretation of this analysis is that, allowing for possible progression in actor roles following actions, Class 2 are the "Trainers", intermediaries between the planners and operations directors in Class 1, and the "footsoldiers" in Class 3 carrying out the operations.

# 10 Comparisons with Everton's analyses

Everton's analyses are much more extensive than ours, as they are used to illustrate some of the features of the packages used for social network analysis. The philosophy of the packages is also different: descriptive statistics of the network are computed, and assessed for importance by comparison with the distributions of the same statistics from randomly generated networks. Explicit probability models for the ties are mentioned only once, in Chapter 11 (p. 360) giving the exponential random graph model (ERGM), as the analysis of these

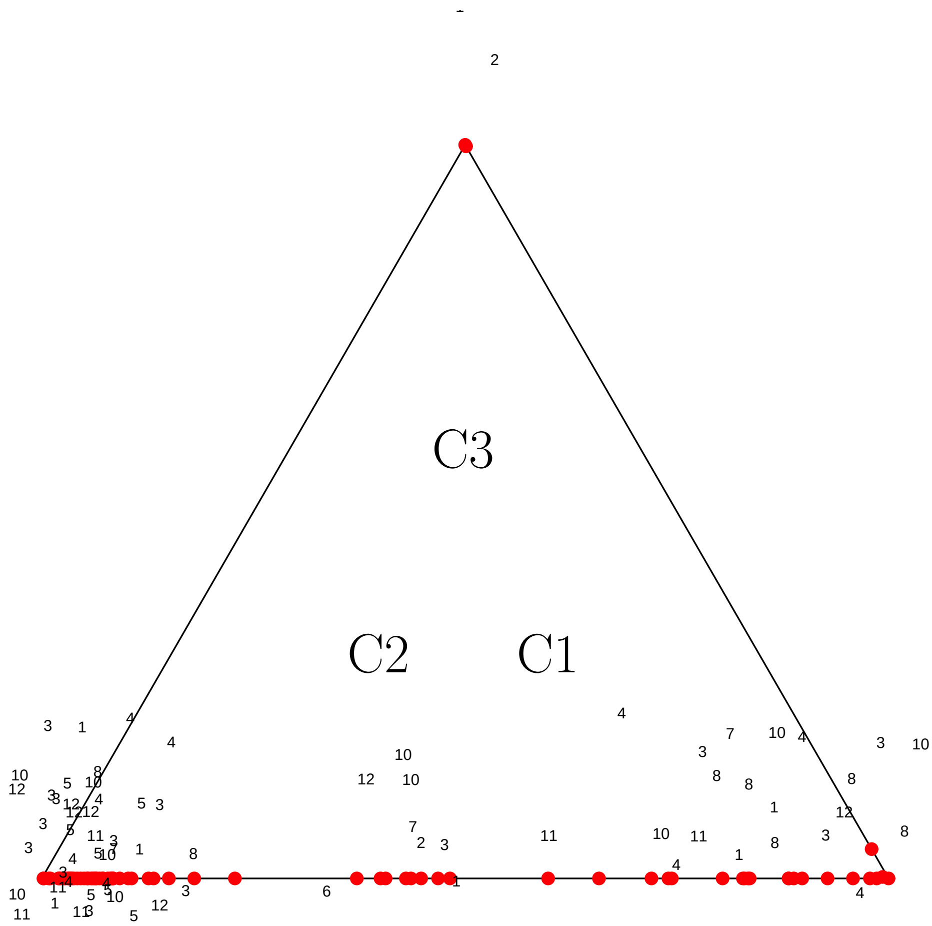


Figure 11: Three-class membership probabilities by roles, random Rasch model

models is not implemented in the packages used for his analyses.

The closest comparison is in Chapter 6 "Cohesion and Clustering" which considers the identification of cohesive subgroups. Everton evaluates four sets of algorithms for detecting subgroups: components, cores, factions, and Newman groups. An immediate difficulty with these comparisons is that some of the analyses are applied to different sub-networks of actors: the "alive" network has 69 actors, but the "alive and free" network has only 24: 45 actors are imprisoned.

A further difficulty is that the ties themselves are of several kinds, and all the analyses are of *one-mode* networks, by collapsing the adjacency matrix across events (formally, by transforming the $n \times p$ adjacency matrix $Y$ to the $n \times n$ symmetric matrix $YY'$). The resulting counts of the numbers of events attended in common by pairs of actors were further dichotomized for procedures which required binary data. This differs fundamentally from the latent class analysis, which requires the full information in the adjacency matrix.

The difficulty with this approach is clear. Summing over the event classification gives an entry $n_{i,i'}$ in cell $(i, i')$ of the resulting projected table, given by $n_{i,i'} = \sum_{j=1}^{r} y_{ij} y_{i'j}$, a sum of products of Bernoulli variables which has no simple form. Dichotomising this sum will make the distribution even more complex. Gerdes (2014) and Neal (2014) discuss this problem at length and suggest some alternatives, though the direct modelling of the bipartite table is not one of them.

A summary of the results reported by Everton follows; we do not give details of the algorithms used.

The components analysis of the "aggregated trust network" in Pajek placed all actors in the one component.

The "alive" network gave the same single component.

The $k$-core analysis of the "alive trust" network in UCINET showed a well-connected central subgroup surrounded by four slightly less connected subgroups.

The $k$-core analysis of the "alive operational" network found a set of clusters centering on Noordin Top.

The faction analysis in UCINET of the "alive and free" combined network gave 8, 9 or 3 factions depending on the "measure-of-fit" option chosen; other considerations suggested 3 or 4 factions.

The Newman group analysis gave similar results to the faction analysis for the "alive and free" combined network.

This does not exhaust the list of possible algorithms for subgroup structure; as Everton wrote (p. 204),

> There are several more we have not considered. What should be clear by now is that we may have to use multiple algorithms before we succeed in detecting cohesive subgroups. ..

# 11 Conclusions

It may be frustrating to readers that there is little connection between our analyses and those of Everton. This lack of connection follows from the different latent-class model-based approach we are following. The multiple algorithms for subgroup structure examined by Everton do not lead to any clear subgroup structure. In our view this is a consequence of analysing the one-mode data structure, with its loss of information.

The approach to sub-grouping of actors in the full two-mode network through latent class modelling and analysis of the adjacency matrix is successful in identifying the leadership structure of the Noordin Top network. The identification of the number of classes through the posterior distributions of the competing model deviances also appears to be successful. While the sparsity of this network, and (in our interpretation) the short operational life of many of the actors, obscure the connections among them, the identification of the leaders is clear, and we can give a plausible interpretation of the “command structure” of the network from the three-class model.

A possibility not considered in our paper is the modelling of change over time in the terrorist network structure. This comes about through the time-stamping of membership of the actors: over the period 2000-2010 we know the first time at which an actor is mentioned, and so is known to have entered the network. We know also when actors die or are arrested and therefore are withdrawn from the network. However the events at which attendance is recorded are not time-stamped, so we cannot model the successive structures of the network, and changes to it, as each new event occurs. If the events were also time-stamped, much richer modelling would be possible.

An extension not considered in this paper is *biclustering*, or double latent-class modelling of both actors and events. In the Noordin Top network, the reason is clear: we have Everton’s manifest categorization of events by scales. In other bipartite networks, biclustering may be of interest. The observed data likelihood in this model is particularly complex, though MCMC analysis is fairly straightforward. Large networks will require very long computation times for large models with many latent classes; variational methods are increasingly used for such networks. Most variational methods are developed for the symmetric unipartite network model (for example Bickel et al 2013). Vu, Hunter and Schweinberger (2013) and Vu and Aitkin (2015) give examples for bipartite networks.

# 12 Acknowledgements

In preparing this paper we have benefitted greatly from interactions with the Social Network group in the Department of Psychology, School of Psychological Sciences at the University of Melbourne. Professors Pip Pattison (a former Deputy Vice-Chancellor of the University) and Garry Robins have been particularly helpful in comparing and interpreting the differences between this approach and others used in social network modelling, especially the exponential random graph model approach.

We are also grateful for assistance from Sean Everton, and research support from the Australian Research Council under Grant DP120102902 for Pip Pattison’s participation in the project, the support of Duy Vu for the period of this research (2012-15), and visits to Melbourne from the University of Lancaster by Brian Francis.

The preparation of this paper has benefitted from very helpful comments by two referees and an Editor. Supplementary material, including the adjacency matrix (the data source we have used) and graphs of posterior distributions of event attendance probabilities in the three latent classes, can be found at ***.

## 13 References


Aitkin, M. (2010). Statistical Inference: an Integrated Bayesian/Likelihood Approach. Boca Raton FL: Chapman and Hall/CRC Press.

Aitkin, M., Vu, D. and Francis, B.J. (2014). Statistical modelling of the group structure of social networks. Social Networks 38, 74–87.

Aitkin, M., Vu, D. and Francis, B.J. (2015). A new Bayesian approach for determining the number of components in a finite mixture. Metron 73, 155–176.

Bickel, P., Choi, D., Chang, X. and Zhang, H. (2013). Asymptotic normality of maximum likelihood and its variational approximation for stochastic blockmodels. Annals of Statistics 41, 1922-1943.

Davis, A., Gardner, B.B. and Gardner, M.R. (1941). Deep South: A Social Anthropological Study of Caste and Class. Chicago: University Press.

De Boeck, P. (2008). Random item IRT models. Psychometrika 73, 533-559.

Everton, S.F. (2012). Disrupting Dark Networks. Cambridge: University Press.

Gerdes, L.M. (2014) MAPPing dark networks: a data transformation method to study clandestine organizations. in Network Science, Cambridge Univesity Press, pp. 1–41.

Goodman, L.A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika 61, 215–231.

International Crisis Group (2009). Indonesia: Noordin Top's Support Base. Asia briefing 95. Online at

`http://www.crisisgroup.org/~/media/Files/asia/south-east-asia /b95_indonesia_noordin_tops_support_base.pdf`

Lazarsfeld, P.F. and Henry, N.W. (1968). Latent Structure Analysis. Boston: Houghton Mifflin.

Neal, Z. (2014) The backbone of bipartite projections: inferring relationships from co-authorship, co-sponsorship, co-attendance and other co-behaviors. Social Networks 39, 84–97.

Sperrin, M., Jaki, T. and Wit, E. (2010). Probabilistic relabelling strategies for the label switching problem in Bayesian mixture models. Statistics and Computing 20, 357–366.

Vu, D. and Aitkin, M. (2015). Variational algorithms for biclustering models. Computational Statistics and Data Analysis 89, 12–24.

Vu, D., Hunter, D. and Schweinberger, M, (2013). Model-based clustering of large networks. Annals of Applied Statistics 7, 1010–1039.

# Statistical modelling of a terrorist network – supplementary materials

June 12, 2016

Murray Aitkin and Duy Vu
School of Mathematics and Statistics
University of Melbourne, Victoria Australia
and Brian Francis
Department of Mathematics and Statistics
Lancaster University, Lancaster LA14YF United Kingdom
%sectionAppendix

## 1 Adjacency matrix 74 rows, 45 columns

1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
1,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,1,1,0,0,0,0,0,0,1,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0
1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
1,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0
0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,1,0,0,1,0,0,0,0
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0
1,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1,1,0,0,0,0,0,0,0,1,0,1,1,1,0
1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0
0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,1,1,1,1,1,0,0,0,0,1,1,0,0,0,0,0,1,0,0,0,0,1,1,1,0,1,1,1,0,0,0,0,0,1,0,0,1
0,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
1,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1
0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,1,0,0,0,0,0,1,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,1,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0
1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0
0,0,0,0,1,0,0,1,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
1,0,0,0,0,0,0,1,1,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,1,1,0,0,0,0,0,1,1,1,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,0,1,0,0,0
0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0
0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0

0,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,0,1,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,1,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0
0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,1,0,0,0,0,1,0,1,0,1,0,0,0,0,0,0,0,0,0,1,1,0,1,1,1,1,0,1,1,1,1,1,1,0,1,1,0,1,1,1,1,1
0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,1,0,1,1,0
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,1,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0
0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,1,1,0,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,0,1,0,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,1,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,1,0,0,0,0,1,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0,0,0,0,0,0
0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,1,0,1,1,0,0,1,0,0,0,0,0,0,0,0,0
0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0

## 2 3-class posterior membership probabilities

| $i$\C | 3 | 2 | 1 | d | $i$\C | 3 | 2 | 1 | d |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.05 | 0.95 | 0.00 | 1 | 41 | 0 | 1 | 0 | 5 |
| 2 | 0.63 | 0.37 | 0 | 3 | 42 | 0.04 | 0 | 0.96 | 5 |
| 3 | 0 | 1 | 0 | 7 | 43 | 0 | 1 | 0 | 9 |
| 4 | 0.19 | 0.80 | 0.01 | 2 | 44 | 0 | 0.05 | 0.95 | 4 |
| 5 | 0.30 | 0.67 | 0.03 | 3 | 45 | 0 | 0.99 | 0.01 | 3 |
| 6 | 0.02 | 0.16 | 0.82 | 3 | 46 | 0.02 | 0.23 | 0.75 | 1 |
| 7 | 0 | 0.38 | 0.62 | 6 | 47 | 0.66 | 0.32 | 0.02 | 2 |
| 8 | 0.01 | 0.32 | 0.67 | 1 | 48 | 0 | 0.15 | 0.85 | 2 |
| 9 | 0 | 0.38 | 0.62 | 5 | 49 | 1.00 | 0 | 0.00 | 4 |
| 10 | 0.17 | 0.32 | 0.51 | 3 | 50 | 0.04 | 0 | 0.96 | 3 |
| 11 | 0.00 | 1.00 | 0 | 3 | 51 | 0 | 1 | 0 | 5 |
| 12 | 0 | 0.20 | 0.80 | 9 | 52 | 0.02 | 0.97 | 0.01 | 2 |
| 13 | 0.04 | 0.96 | 0.00 | 1 | 53 | 0.01 | 0.41 | 0.58 | 1 |
| 14 | 0.04 | 0.96 | 0.00 | 1 | 54 | 0 | 0.15 | 0.85 | 23 |
| 15 | 0.98 | 0.01 | 0.01 | 2 | 55 | 0 | 0.99 | 0.01 | 6 |
| 16 | 0 | 0.15 | 0.85 | 2 | 56 | 0 | 0.15 | 0.85 | 3 |
| 17 | 0 | 1 | 0 | 4 | 57 | 0.02 | 0.98 | 0 | 2 |
| 18 | 0 | 1 | 0 | 3 | 58 | 0.01 | 0.43 | 0.56 | 1 |
| 19 | 0 | 1 | 0 | 5 | 59 | 0 | 0.15 | 0.85 | 2 |
| 20 | 0.60 | 0 | 0.40 | 2 | 60 | 0 | 1 | 0 | 3 |
| 21 | 0 | 0 | 1 | 17 | 61 | 0 | 1 | 0 | 4 |
| 22 | 0 | 0.70 | 0.30 | 3 | 62 | 0.61 | 0.00 | 0.39 | 2 |
| 23 | 0.01 | 0.99 | 0 | 3 | 63 | 0 | 0.38 | 0.62 | 6 |
| 24 | 0 | 0.04 | 0.96 | 5 | 64 | 0 | 0.04 | 0.96 | 3 |
| 25 | 0 | 1 | 0 | 2 | 65 | 0 | 0.94 | 0.06 | 5 |
| 26 | 0.98 | 0.02 | 0.00 | 3 | 66 | 0.59 | 0.00 | 0.41 | 2 |
| 27 | 0 | 1 | 0 | 3 | 67 | 0.99 | 0 | 0.01 | 4 |
| 28 | 0.05 | 0.95 | 0.00 | 1 | 68 | 0 | 0.98 | 0.02 | 5 |
| 29 | 0.37 | 0.26 | 0.37 | 1 | 69 | 0 | 0.99 | 0.01 | 4 |
| 30 | 0.15 | 0.85 | 0.00 | 1 | 70 | 0.99 | 0.01 | 0.00 | 3 |
| 31 | 1 | 0 | 0 | 4 | 71 | 0.00 | 1.00 | 0 | 2 |
| 32 | 0 | 1 | 0 | 5 | 72 | 0 | 0.98 | 0.02 | 4 |
| 33 | 0 | 1 | 0 | 3 | 73 | 0 | 0.38 | 0.62 | 6 |
| 34 | 0.00 | 0.99 | 0.01 | 1 | 74 | 0.04 | 0.94 | 0.02 | 2 |
| 35 | 0 | 1 | 0 | 5 | | | | | |
| 36 | 0 | 0.26 | 0.74 | 4 | | | | | |
| 37 | 0 | 0.27 | 0.73 | 2 | | | | | |
| 38 | 0.99 | 0.01 | 0 | 4 | | | | | |
| 39 | 0.86 | 0.14 | 0.00 | 2 | | | | | |
| 40 | 0 | 1 | 0 | 2 | | | | | |
| T | | | | | | 12.84 | 38.90 | 22.26 | |

Table 1: Posterior class probabilities for actors, 3-class model A

## 3 5-class posterior membership probabilities

| $i$\C | 1 | 2 | 3 | 4 | 5 | d | $i$\C | 1 | 2 | 3 | 4 | 5 | d |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.00 | 0.32 | 0.00 | 0.68 | 0.00 | 1 | 41 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 5 |
| 2 | 0.00 | 0.93 | 0.00 | 0.07 | 0.00 | 3 | 42 | 0.00 | 0.04 | 0.96 | 0.00 | 0.00 | 5 |
| 3 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 7 | 43 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 9 |
| 4 | 0.00 | 0.69 | 0.00 | 0.31 | 0.00 | 2 | 44 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 4 |
| 5 | 0.00 | 0.98 | 0.00 | 0.02 | 0.00 | 3 | 45 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 3 |
| 6 | 0.00 | 0.10 | 0.90 | 0.00 | 0.00 | 3 | 46 | 0.00 | 0.04 | 0.92 | 0.03 | 0.01 | 1 |
| 7 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 6 | 47 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 2 |
| 8 | 0.00 | 0.04 | 0.74 | 0.03 | 0.19 | 1 | 48 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 2 |
| 9 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 5 | 49 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 4 |
| 10 | 0.00 | 0.98 | 0.01 | 0.00 | 0.00 | 3 | 50 | 0.00 | 0.04 | 0.96 | 0.00 | 0.00 | 3 |
| 11 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 3 | 51 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 5 |
| 12 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 9 | 52 | 0.06 | 0.90 | 0.01 | 0.03 | 0.00 | 2 |
| 13 | 0.00 | 0.32 | 0.00 | 0.68 | 0.00 | 1 | 53 | 0.00 | 0.02 | 0.73 | 0.20 | 0.05 | 1 |
| 14 | 0.00 | 0.33 | 0.00 | 0.67 | 0.00 | 1 | 54 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 23 |
| 15 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 2 | 55 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 6 |
| 16 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 2 | 56 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 3 |
| 17 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 4 | 57 | 0.00 | 0.01 | 0.00 | 0.99 | 0.00 | 2 |
| 18 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 3 | 58 | 0.00 | 0.02 | 0.75 | 0.18 | 0.05 | 1 |
| 19 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 5 | 59 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 2 |
| 20 | 0.00 | 0.70 | 0.30 | 0.00 | 0.00 | 2 | 60 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 3 |
| 21 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 17 | 61 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 4 |
| 22 | 0.00 | 0.00 | 0.12 | 0.88 | 0.00 | 3 | 62 | 0.00 | 0.67 | 0.33 | 0.00 | 0.00 | 2 |
| 23 | 0.00 | 0.02 | 0.00 | 0.98 | 0.00 | 3 | 63 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 6 |
| 24 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 5 | 64 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 3 |
| 25 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 2 | 65 | 0.00 | 0.00 | 0.00 | 0.01 | 0.99 | 5 |
| 26 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 3 | 66 | 0.00 | 0.68 | 0.32 | 0.00 | 0.00 | 2 |
| 27 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 3 | 67 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 4 |
| 28 | 0.00 | 0.33 | 0.00 | 0.67 | 0.00 | 1 | 68 | 0.00 | 0.00 | 0.04 | 0.91 | 0.05 | 5 |
| 29 | 0.00 | 0.56 | 0.35 | 0.09 | 0.00 | 1 | 69 | 0.00 | 0.00 | 0.01 | 0.99 | 0.00 | 4 |
| 30 | 0.00 | 0.34 | 0.00 | 0.66 | 0.00 | 1 | 70 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 3 |
| 31 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 4 | 71 | 0.02 | 0.85 | 0.00 | 0.13 | 0.00 | 2 |
| 32 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 5 | 72 | 0.00 | 0.00 | 0.00 | 0.98 | 0.02 | 4 |
| 33 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 3 | 73 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 6 |
| 34 | 0.20 | 0.04 | 0.02 | 0.73 | 0.00 | 1 | 74 | 0.06 | 0.92 | 0.00 | 0.02 | 0.00 | 2 |
| 35 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 5 | | | | | | | |
| 36 | 0.00 | 0.00 | 0.77 | 0.01 | 0.22 | 4 | | | | | | | |
| 37 | 0.00 | 0.00 | 0.93 | 0.07 | 0.00 | 2 | | | | | | | |
| 38 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 4 | | | | | | | |
| 39 | 0.00 | 0.97 | 0.00 | 0.03 | 0.00 | 2 | | | | | | | |
| 40 | 0.44 | 0.00 | 0.02 | 0.49 | 0.05 | 2 | | | | | | | |
| T | | | | | | | | 9.78 | 19.84 | 19.19 | 19.54 | 5.63 | 73.98 |

Table 2: Posterior class probabilities for actors, 5-class model A

## 4 Event-and class-specific posterior distributions of probabilities of event attendance

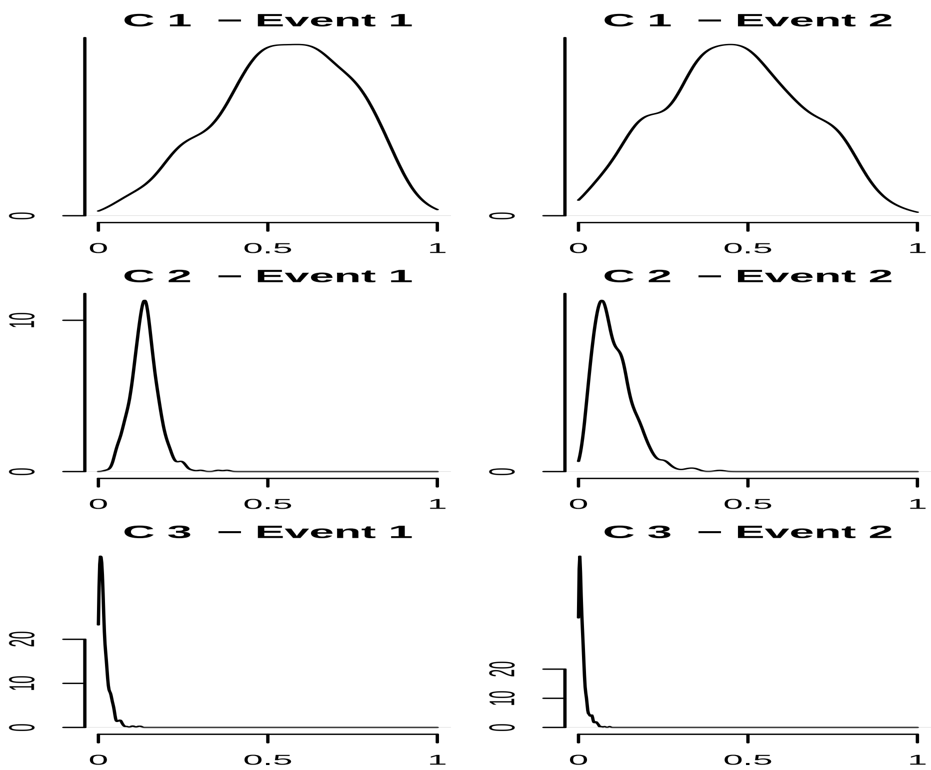


Figure 1: Finance, three classes

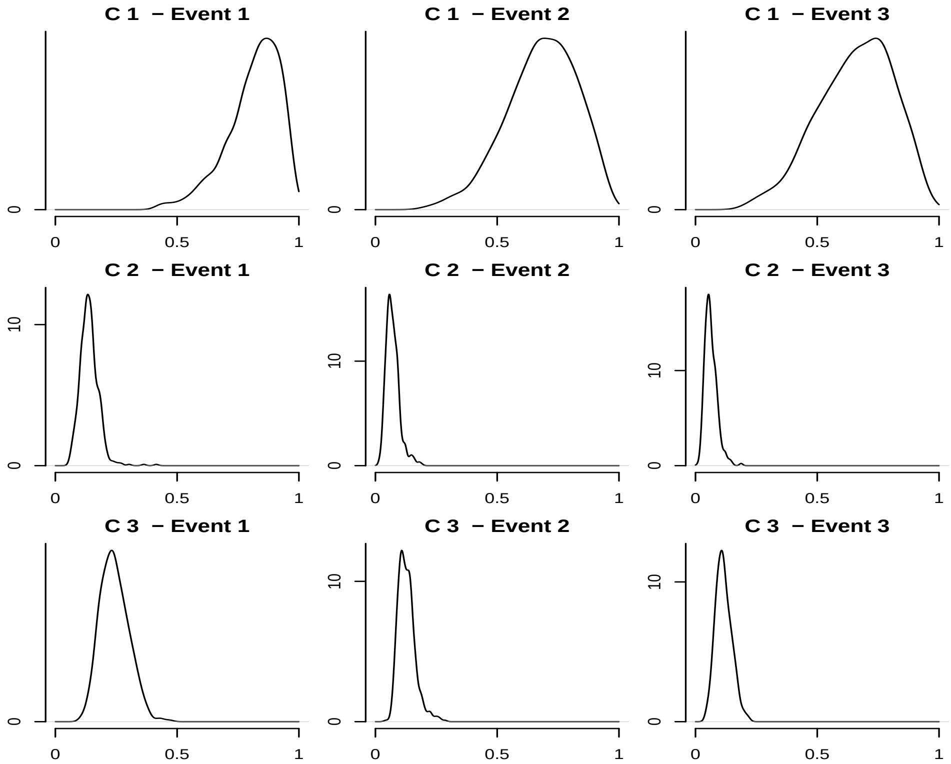


Figure 2: Operations, three classes

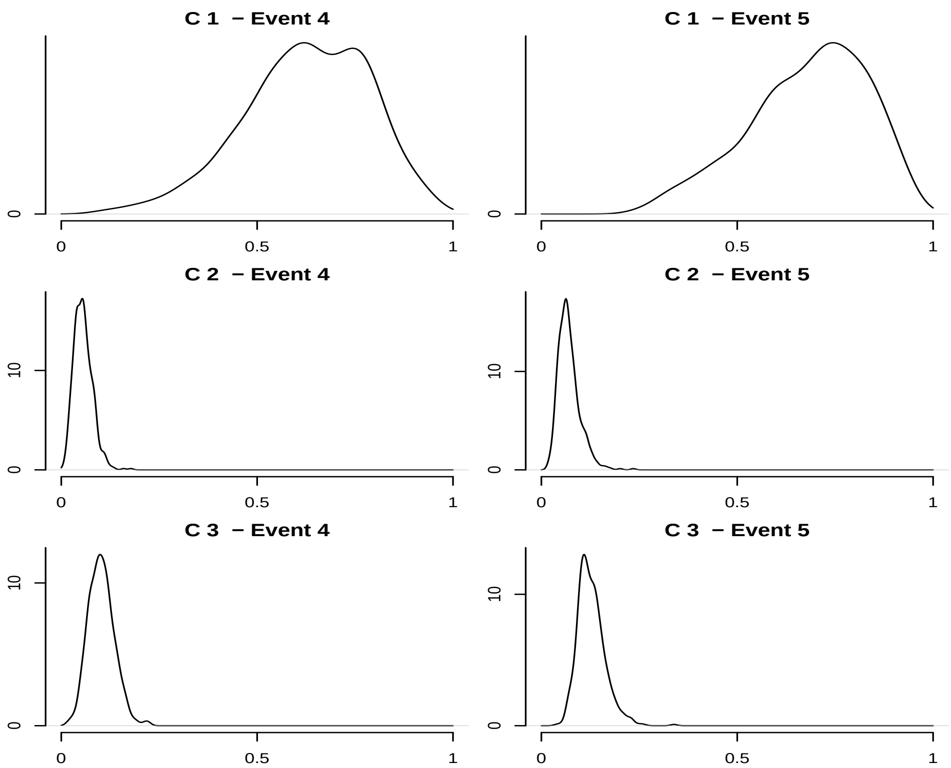


Figure 3: Operations, three classes

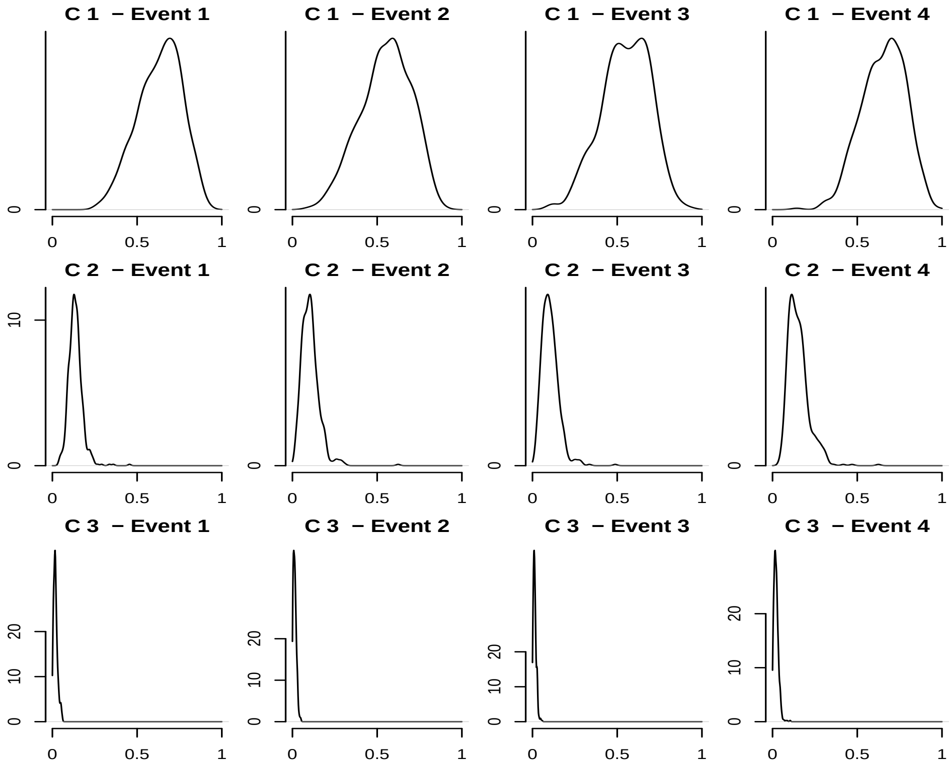


Figure 4: Logistics, three classes

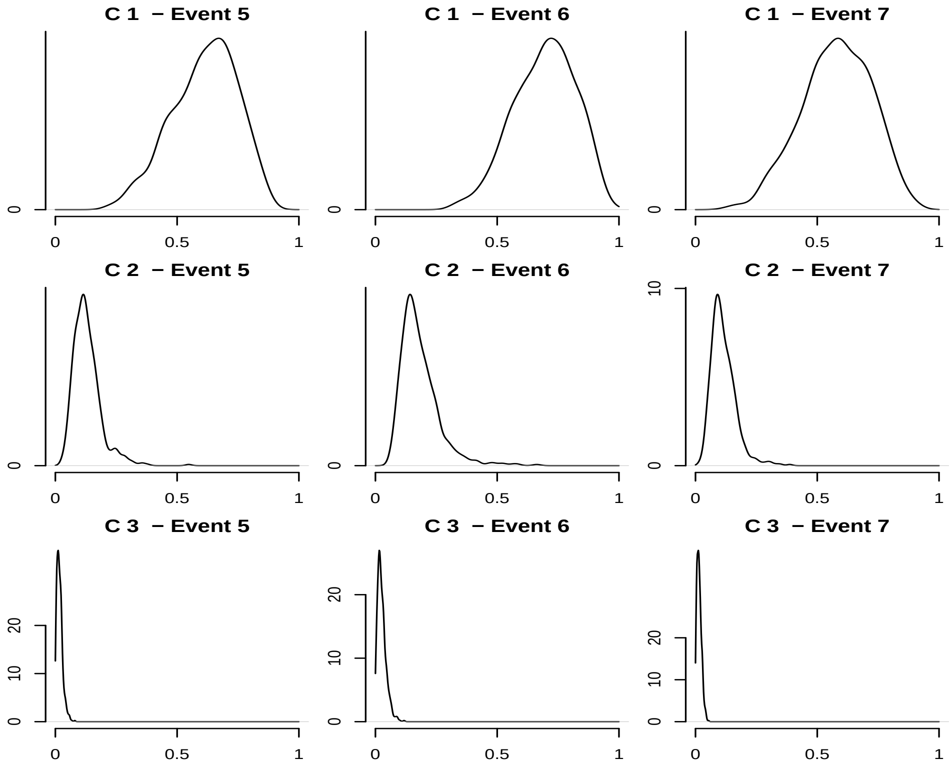


Figure 5: Logistics, three classes

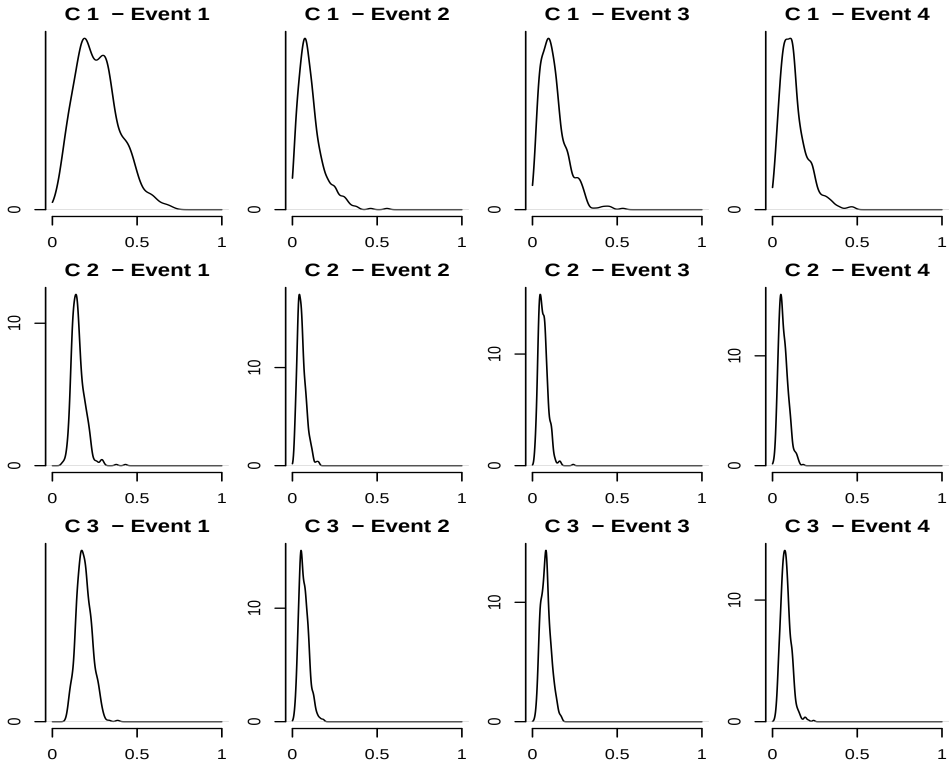


Figure 6: Organization, three classes

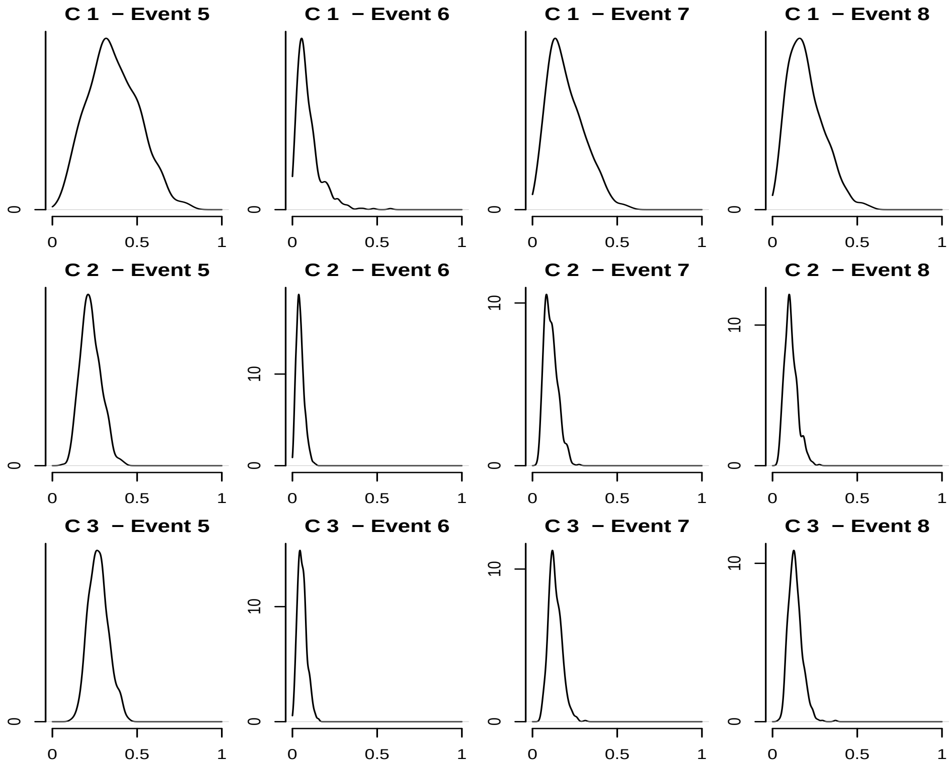


Figure 7: Organization, three classes

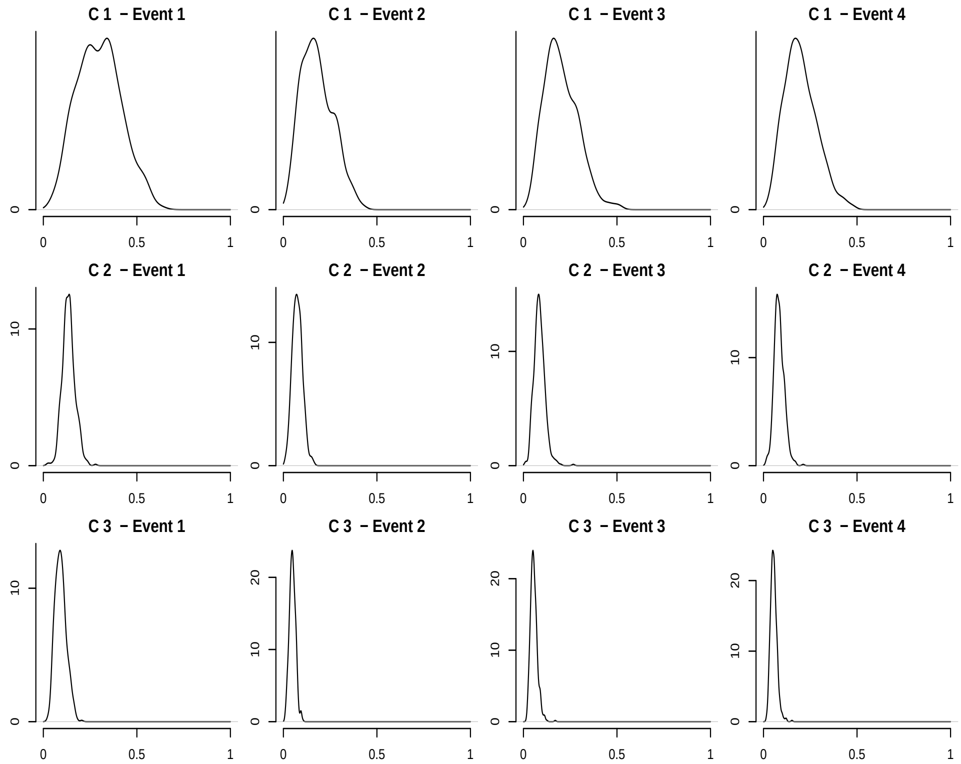


Figure 8: Training, three classes

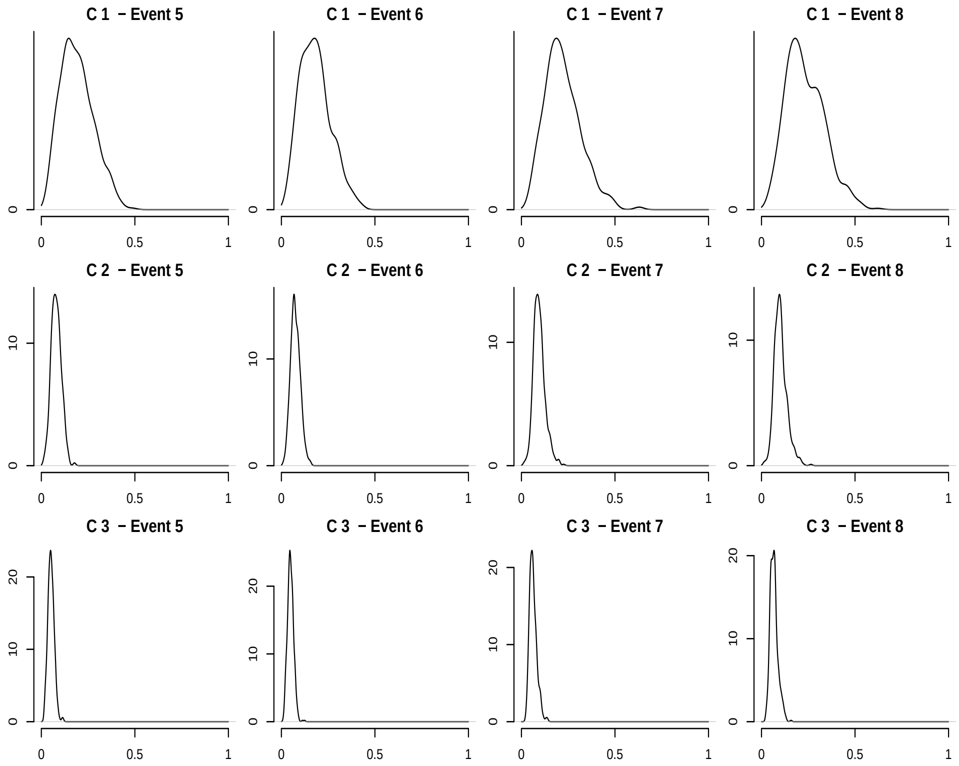


Figure 9: Training, three classes

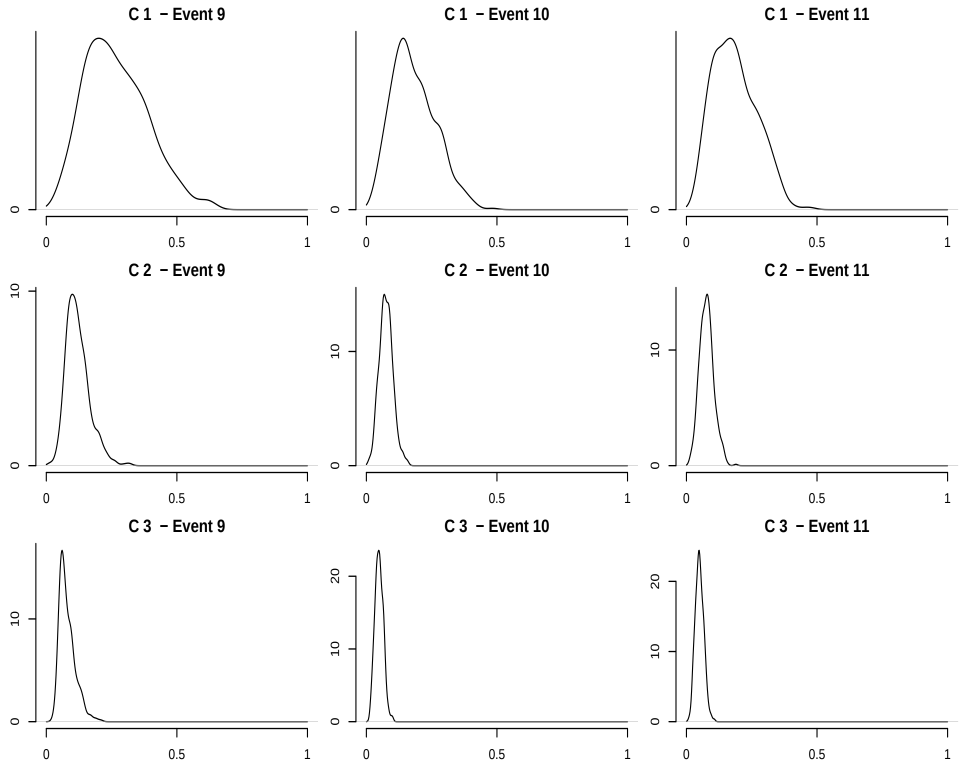


Figure 10: Training, three classes

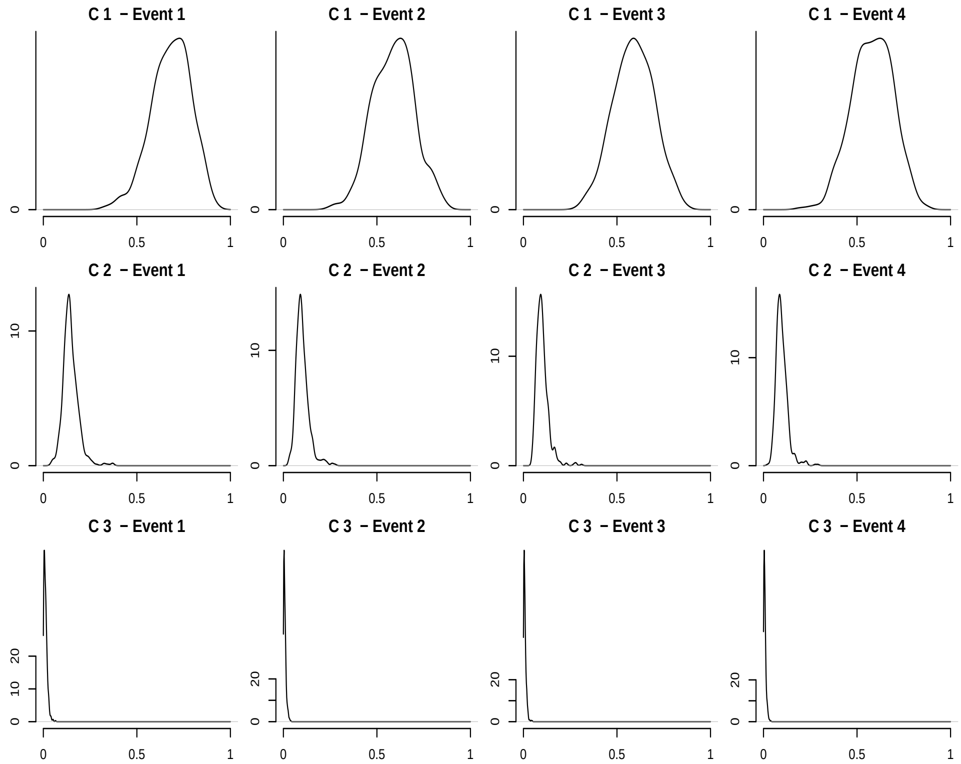


Figure 11: Meetings, three classes

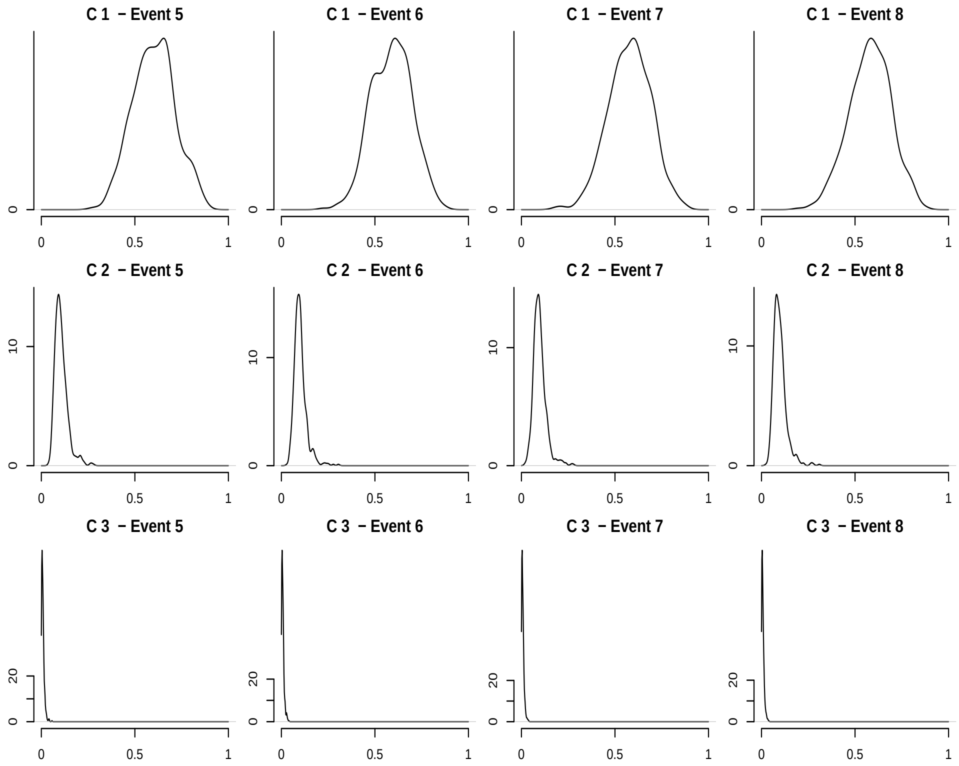


Figure 12: Meetings, three classes

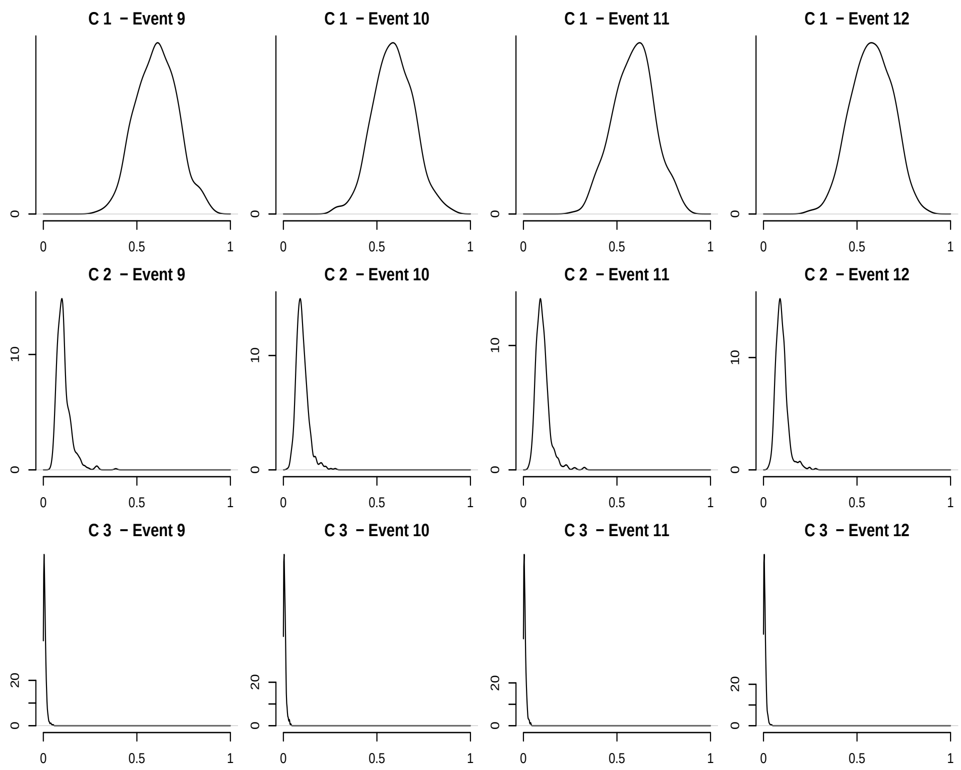


Figure 13: Meetings, three classes