Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 51 (2007) 5205–5210
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

Editorial

Advances in Mixture Models

The importance of *mixture distributions* is not only remarked by a number of recent books on mixtures including Lindsay (1995), Böhning (2000), McLachlan and Peel (2000) and Frühwirth-Schnatter (2006) which update previous books by Everitt and Hand (1981), Titterington et al. (1985) and McLachlan and Basford (1988). Also, a diversity of publications on mixtures appeared in this journal since 2003 (which we take here as a milestone with the appearance of the first special issue on mixtures) including Hazan et al. (2003), Benton and Krishnamoorthy (2003), Woodward and Sain (2003), Besbeas and Morgan (2004), Jamshidian (2004), Hürlimann (2004), Bohacek and Rozovskii (2004), Tao et al. (2004), Vaz de Melo Mendes and Lopes (2006), Agresti et al. (2006), Bartolucci and Scaccia (2006), D'Elia and Piccolo (2005), Neerchal and Morel (2005), Klar and Meintanis (2005), Bocci et al. (2006), Hu and Sung (2006), Seidel et al. (2006), Nadarajah (2006), Almhana et al. (2006), Congdon (2006), Priebe et al. (2006), and Li and Zha (2006). In the following we give a brief introduction to the papers contributing novels aspects in this Special Issue. These come from a diversity of areas as different as capture–recapture modelling, likelihood based cluster analysis, semiparametric mixture modelling in microarray data, latent class analysis or integer lifetime data analysis—just to mention a few.

Mixture models are frequently used in capture–recapture studies for estimating population size (Chao, 1987; Link, 2003; Böhning and Schön, 2005; Böhning et al., 2005; Böhning and Kuhnert, 2006). In this issue, Mao (2007) highlights a variety of sources of difficulties in statistical inference using mixture models and uses a binomial mixture model as an illustration.

Random intercept models for binary data—as useful tools for addressing between-subject heterogeneity—are discussed by Caffo et al. (2007). The nonlinearity of link functions for binary data is blurred in probit models with a normally distributed random intercept because the resulting model implies a probit marginal link as well. Caffo et al. (2007) explore another family of random intercept models where the distribution associated with the marginal and conditional link function as well as the random effect distribution are all of the same family.

Formann (2007) extends the latent class approach (as a specific discrete multivariate mixture model) for situations where the discrete outcome variables (such as longitudinal binary data) experience nonignorable associations and, in addition and most importantly, have missing entries as it is rather typical for repeated observations in longitudinal studies. The modelling also incorporates potential covariates. This is illustrated using data from the Muscatine Coronary Risk Factor Study.

The contribution by Grün and Leisch (2007) introduces the `R`-package `flexmix` which provides flexible modelling of finite mixtures of regression models using the EM algorithm.

Alfò et al. (2007) consider a semiparametric mixture model for detecting differentially expressed genes in microarray experiments. An important goal of microarray studies is the detection of genes that show significant changes in observed expressions when two or more classes of biological samples (e.g. treatment and control) are compared. With the c-fold rule a gene is declared to be differentially expressed if its average expression level varies by more than a constant (typically 2). Instead, Alfò et al. (2007) introduce a gene-specific random term to control for both dependence among genes and variability with respect to the probability of yielding a fold change over *a threshold* c. Likelihood based inference is accomplished with a two-level finite mixture model while nonparametric Bayesian estimation is performed through the counting distribution of exceedances.

Mixtures-of-experts models (Jacobs et al., 1991) and their generalization, hierarchical mixtures-of-expert models (Jordan and Jacobs, 1994) have been introduced to account for nonlinearities and other complexities in the data.

0167-9473/$ - see front matter
doi:10.1016/j.csda.2006.10.025

Carvalho and Tanner (2007) study a class of hierarchical mixtures of Poisson experts to model nonlinear count time series. Identifiability and maximum likelihood estimation via the EM are discussed. Extending previous results for independent observations, the authors prove asymptotic normality of the maximum likelihood estimator under stationarity and nonstationarity of the covariates vector which may include lags of transformations of a response and lags of external predictors. The performance of standard criteria for selecting the number of experts is examined through Monte Carlo simulations. An application of the proposed model to real data is also given.

An important task in mixture modelling is the determination of the number of components. The generalized likelihood ratio test is often suggested to be used in this context. The limiting distribution of the corresponding test statistic has long remained a mystery. It is well-known that the classical result giving a chi-squared asymptotic distribution is no longer valid here. This field has enjoyed an explosion in activity in the last decade. Garel (2007) gives an overview of very recent advances in the problem of testing homogeneity against a two-component mixture and adds some new results. The questions he addresses range from the determination of the asymptotic distribution to issues on the calculation of the percentage points, and to the assessment of asymptotic power, both for bounded and unbounded parameter sets.

In statistical survey data, usually a large number of values is missing. To deal with this problem, imputation is widely used. This means that missing values are replaced with plausible ones. For modelling incomplete data, parametric and nonparametric methods have been applied. To overcome the drawbacks of both approaches, the authors propose to use a finite mixture of Gaussian distributions as a semiparametric compromise. For estimating the parameters in presence of missing data, a version of the EM developed by Hunt and Jorgensen (2003) is applied by Di Zio et al. (2007). Two imputation strategies based on mixture models are suggested. Their performance in terms of preserving the mean and the covariance structure of the data is evaluated and compared with the commonly used nearest neighbour donor method.

In survival analysis, heterogeneity between individuals is often described by frailty models, where random effects act multiplicatively on the hazard function. This corresponds to a mixture model for survival times. Correlated frailty models (Yashin et al., 1995) are appropriate for clustered survival data, where one often observes correlation among survival times in each group as well as individual unknown risk factors. In Kheiri et al. (2007) the frailty of each individual is a sum of two independent components, a shared within group risk factor common to the individuals in the same group, and an individual frailty component. Kheiri et al. (2007) propose to use an inverse Gaussian distribution for the components of frailty in a correlated frailty model. In a Bayesian analysis framework, a MCMC approach is used to achieve parameter estimation and the deviance information criterion is used to compare models. The method is applied to a set of corneal graft survival time data for illustration.

In the normal mixture model-based approach to density estimation and clustering, the density of a p-dimensional random variable is modelled as a mixture of multivariate normal densities. If component covariance matrices are not restricted, this is a highly parameterized model. If the dimension p is large relative to the sample size, mixtures of factor analyzers have been proposed to reduce the number of free parameters (see for example McLachlan et al., 2003). The complexity of each component covariance matrix lies between that of the isotropic and full covariance structure model. However, such a model is sensitive to outliers since it uses normal errors and factors. McLachlan et al. (2007) propose to use mixtures of t-factor analyzers instead. They derive an EM-type algorithm for fitting the mixtures. The approach is applied to a real data set involving the clustering of gene expression levels on the basis of tissue samples from some microarray experiments.

Likelihood maximization in the case of normal mixtures with covariance heterogeneity suffers from the problem of singularities of the likelihood function at the boundary of the parameter space and of spurious local maxima. Although the usual unconstrained EM does a surprisingly good job if applied with some care, it is of interest to consider constrained versions in order to obtain a parameter space with no singularities and a smaller number of spurious local maxima of the likelihood function. In the spirit of Hathaway (1985), Ingrassia and Rocci (2007) impose constraints on the eigenvalues of the covariance matrices. They give conditions assuring that the EM based on such constraints maintains the monotonicity properties of the usual EM. Different approaches are presented and their performances are compared in a simulation study.

In a Bayesian context, the Markov chain Monte Carlo (MCMC) approach for simulation based approximations of intractable posterior distributions and predictive densities is popular. However, in very complex models it may result in unacceptably long computation times. Variational methods are a fast, deterministic alternative to MCMC methods which have gained recent popularity especially in the machine learning literature (see e.g. Jordan et al., 1999; Corduneanu and Bishop, 2001; Ueda and Ghahramani, 2002). These aim to construct a tight lower bound on the data marginal likelihood and then optimize this bound using an iterative scheme. McGrory and Titterington (2007) apply the variational method in the context of mixtures of Gaussian (*MoG*) models. They estimate simultaneously the number of mixture components as well as component-specific parameter vectors. They also use the variational algorithm to build a version of the *Deviance Information Criterion* (*DIC*, Spiegelhalter et al., 2002) which is able to handle missing—data models and is therefore applicable to mixtures (see also Celeux et al., 2006).

Finite mixture models have been frequently used to cluster two-way data sets. Recently, three-way data sets have become popular, containing for example attributes (variables) measured on objects (statistical units) in several conditions (occasions, time points, environments, etc.). Basford and McLachlan (1985) have proposed a finite mixture model for the analysis of such data, where the aim is to cluster objects by explicitly taking simultaneously into account the information on variables and occasions. Hunt and Basford (1999, 2001) have extended the approach to deal with categorical variables in unbalanced panels, while Meulders et al. (2002) have proposed a constrained latent class model for the analysis of three-way binary data. All these models assume that cases belong to the same cluster in all investigated situations. Vermunt (2007) proposes an extension of this approach assuming that objects may be in a different latent class depending on the situation or, more specifically, objects are clustered with respect to the probability of being in a particular latent class at a given situation. Relevant work in this topic include, among others, Böhning et al. (2000) and Knorr-Held and Raßer (2000). Vermunt (2007) considers the three ways as hierarchically nested levels and models a mixture distribution at each of the two higher levels; i.e., one at the object and one at the object-in-occasion level. The proposed model is an adaptation of the multilevel latent class model by Vermunt (2003) to continuous responses and has the advantage that it may yield more parsimonious and insightful solutions than the Basford and McLachlan (1985) model.

Finite mixture models also play a vital role in life testing and reliability. An identifiable mixture model of two inverse Weibull distributions is introduced by Sultan et al. (2007). The corresponding hazard function is computed and the estimation of the parameters is solved via the EM algorithm. Some simulation studies are carried out to illustrate the performance of the estimation technique. Discrete lifetime data frequently arise in reliability, economics, finance, etc. A simple estimator of the hazard rate is usually based on empirical proportions of failure. However, tail probabilities may be badly estimated. Following Böhning and Patilea (2005), Karlis and Patilea (2007) consider semiparametric mixtures of power series distributions, a class which is flexible enough to model adequately a wide range of discrete lifetime data. By estimating the unknown mixing distribution through nonparametric maximum likelihood they obtain smoother individual probabilities estimates and therefore a smoother estimator for the hazard rate. Wald-type as well as bootstrap confidence intervals are constructed. Parametric bootstrapping from the estimated mixture is quite common in the mixture literature, see for instance Böhning (2000), McLachlan and Peel (2000) for theoretical motivations and Zheng and Frey (2004), Böhning et al. (2005) and van der Heijden et al. (2003) for discussions in specific contexts. Using a simulation study, Karlis and Patilea (2007) show that bootstrap confidence intervals have better properties than Wald-type CIs, like smaller length for fixed confidence levels especially for tail hazards.

Consider now a response vector of correlated binary variables. A common approach is to introduce some vector of latent variables and to model their joint distribution given the covariates, like for instance in the multivariate probit model. Jara et al. (2007) propose a more flexible modeling for the distribution of the latent vector using mixtures of multivariate normal distributions with respect to both location and covariance. Their approach is a nonparametric Bayesian one and relies on Dirichlet process priors. Jara et al. (2007) illustrate their proposal by an application to caries research.

Mixture models are quite popular for classifying data, but they become of little use with very large amounts of observations due to the limited memory and computational power. Steiner and Hudec (2007) look at huge data sets and propose a two-step strategy for estimating a mixture of multivariate normal laws. First, the data is pre-clustered into a much smaller number of groups using any partition method suitable for large data sets. The groups are represented by their sufficient statistics (mean vector, estimated covariance matrix, number of observations in the group). In the second step, the mixture is estimated using a variant of the EM algorithm that takes into account the structural information summarized by the covariance within groups. Steiner and Hudec (2007) illustrate their method with a real data set from a web-usage mining application.

Sometimes it is desirable to work with a mixture model that is more flexible than a parametric mixture. Bordes et al. (2007) consider the semiparametric location mixture $\sum_{j=1}^{m} \lambda_j f(x-\mu_j)$, where the λ_j, μ_j are unknown parameters and furthermore *f* is a symmetric but otherwise unspecified density. They propose a generalization of the EM algorithm for the analysis of such semiparametric mixtures and study its performance with Monte Carlo simulations as well as comparisons with existing methods. Eilers and Borgdorff (2007) pursue flexibility of the component density f in the context of the EM algorithm by estimating f in the M-step with a penalized smoother that moves the smoothed estimate towards a log-concave density. They use the resulting model for clustering univariate data and show how to extend the methodology to bivariate data with independent marginals.

Computational advances on approximation techniques such as MCMC methods have been a keystone to Bayesian analysis of mixture models. Mixtures of multinomial and negative-multinomial distributions are addressed by Rufo et al. (2007). In their paper, a classical framework addressing the main topics in a Bayesian analysis is developed for the case with a known number of components. In particular, a Stephens' type algorithm for solving the label-switching problem is adapted.

This special issue on mixture models stems from a special track on mixture models at the recent IASC-CSDA-conference on *Computational Statistics and Data Analysis*, taken place in Limassol, Cyprus, October 28–31, 2005. A conference in Hamburg (Germany) in 2001 on *Recent Developments in Mixture Models* tried to capture contemporary issues of debate in this area and have them discussed in a broader audience. A special issue of CSDA collected some of these topics and put them together in a common context (Böhing and Seidel , 2003). Since the appearance of that special issue new developments took place which are the subject of the current special issue. The guest editors would also acknowledge the help of Geoff McLachlan (Brisbane, Australia) and Matthias Greiner (Berlin, Germany) in organizing the special track. Special thanks go to several referees. Without their continuous efforts the reviewing could not have been done. Finally, the guest editors would express their sincerest thanks to the CoEditor Erricos John Kontoghiorghes.

References

Agresti, A., Caffo, B., Ohman-Strickland, P., 2006. Examples in which misspecification of a random effects distribution reduces efficiency, and possible remedies. Comput. Statist. Data Anal. 47, 639–653.

Alfò, M., Farcomeni, A., Tardella, L., 2007. Robust semiparametric mixing for detecting differentially expressed genes in microarray experiments. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.08.009.

Bartolucci, F., Scaccia, L., 2006. The use of mixtures for dealing with non-normal regression errors. Comput. Statist. Data Anal. 48, 821–834.

Benton, D., Krishnamoorthy, D., 2003. Computing discrete mixtures of continuous distributions: noncentral chisquare, noncentral t and the distribution of the square of the sample multiple correlation coefficient. Comput. Statist. Data Anal. 43, 249–267.

Besbeas, P., Morgan, B.T.J., 2004. Integrated squared error estimation of normal mixtures. Comput. Statist. Data Anal. 44, 517–526.

Bocci, L., Vicari, D., Vichi, M., 2006. A mixture model for the classification of three-way proximity data. Comput. Statist. Data Anal. 50, 1625–1654.

Bohacek, S., Rozovskii, B., 2004. A diffusion model of roundtrip time. Comput. Statist. Data Anal. 45, 25–50.

Böhning, D., 2000. Computer-Assisted Analysis of Mixtures and Applications. Meta-Analysis, Disease Mapping and Others. Chapman & Hall/CRC, Boca Raton.

Böhning, D., Kuhnert, R., 2006. The equivalence of truncated count mixture distributions and mixtures of truncated count distributions. Biometrics. 62, 1207–1215.

Böhning, D., Patilea, V., 2005. Asymptotic normality in mixtures of discrete distributions. Scand. J. Statist. 32, 115–132.

Böhning, D., Schön, D., 2005. Nonparametric maximum likelihood estimation of the population size based upon the counting distribution. J. Roy. Statist. Soc., Ser. C, Appl. Statist. 54, 721–737.

Böhing, D., Seidel, W., 2003. Recent developments in mixture models. Comput. Statist. Data Anal. 41, 349–357.

Böhning, D., Dietz, E., Schlattmann, P., 2000. Space-time mixture modelling of public health data. Statist. Med. 9, 2333–2344.

Böhning, D., Dietz, E., Kuhnert, R., Schön, D., 2005. Mixture models for capture–recapture count data. Statist. Methods Appl. 4, 29–43.

Bordes, L., Chauveau, D., Vandekerkhove, P., 2007. A stochastic EM algorithm for a semiparametric mixture model. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.08.015.

Caffo, B., An, M.-W., Rohde, C., 2007. Flexible random intercept models for binary outcomes using mixtures of normals. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.09.031.

Carvalho, A.X., Tanner, M.A., 2007. Modelling nonlinear count time series with local mixtures of Poisson autoregressions. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.09.032.

Celeux, G., Forbes, F., Robert, C., Titterington, D.M., 2006. Deviance Information Criteria for missing data models with discussion. Bayesian Anal., in print, published online.

Chao, A., 1987. Estimating the population size for capture–recapture data with unequal capture probabilities. Biometrics 43, 783–791.

Congdon, P., 2006. A model for non-parametric spatially varying regression effects. Comput. Statist. Data Anal. 50, 422–445.

Corduneanu, A., Bishop, C.M., 2001. Variational Bayesian model selection for mixture distributions. In: Jaakkola, T., Richardson, T. (Eds.), Artificial Intelligence and Statistics. Morgan Kaufmann, Los Altos, CA, pp. 27–34.

D'Elia, A., Piccolo, D., 2005. A mixture model for preferences data analysis. Comput. Statist. Data Anal. 49, 917–934.

Di Zio, M., Guarnera, U., Luzi, O., 2007. Imputation through finite Gaussian mixture models. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.10.002.

Eilers, P.H.C., Borgdorff, M.W., 2007. Non-parametric log-concave mixtures. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.08.027.

Formann, A.K., 2007. Mixture analysis of multivariate categorical data with covariates and missing entries. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.08.020.

Frühwirth-Schnatter, S., 2006. Finite Mixture and Markov Switching Models. Springer, Heidelberg.

Garel, B., 2007. Recent asymptotic results in testing for mixtures. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.09.033.

Grün, B., Leisch, F., 2007. Fitting finite mixtures of generalized linear regressions in *R*. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.08.014.

Hathaway, R.J., 1985. A constrained formulation of maximum-likelihood estimation for normal mixture distributions. Ann. Statist. 13, 795–800.

Hazan, A., Landsmanm, Z., Makov, U.E., 2003. Robustness via a mixture of exponential power distributions. Comput. Statist. Data Anal. 42, 111–121.

Hu, T., Sung, S.Y., 2006. A hybrid EM approach to spatial clustering. Comput. Statist. Data Anal. 50, 1188–1205.

Hunt, L.A., Basford, K.E., 1999. Fitting a mixture model to three-mode three-way data with categorical and continuous missing information. J. Classification 18, 283–296.

Hunt, L.A., Basford, K.E., 2001. Fitting a mixture model to three-mode three-way data with missing information. J. Classification 18, 209–226.

Hunt, L., Jorgensen, M., 2003. Mixture model clustering for mixed data with missing information. Comput. Statist. Data Anal. 41, 429–440.

Hürlimann, W., 2004. Fitting bivariate cumulative returns with copulas. Comput. Statistist. Data Anal. 45, 355–372.

Ingrassia, S., Rocci, R., 2007. Constrained monotone EM algorithms for finite mixture of multivariate Gaussians. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.10.011.

Jacobs, R., Jordan, M., Nowlan, S., Hinton, G., 1991. Adaptive mixtures of local experts. Neural Comput. 3, 79–87.

Jamshidian, M., 2004. On algorithms for restricted maximum likelihood estimation. Comput. Statist. Data Anal. 45, 137–157.

Jara, A., Garcia-Zattera, M.-J., Lesaffre, E., 2007. A Dirichlet process mixture model for the analysis correlated binary responses. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.09.010.

Jordan, M., Jacobs, R., 1994. Hierarchical mixtures-of-experts and the EM algorithm. Neural Comput. 6, 181–214.

Jordan, M.I., Ghahramani, Z., Jaakkola, T.S., Saul, L.K., 1999. An introduction to variational methods for graphical models. Mach. Learning 37, 183–233.

Karlis, D., Patilea, V., 2007. Confidence intervals of the hazard rate function for discrete distributions using mixtures. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.07.037.

Kheiri, S., Kimber, A., Meshkani, M.R., 2007. Bayesian analysis of an inverse Gaussian correlated frailty model. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.07.026.

Klar, B., Meintanis, S.G., 2005. Tests for normal mixtures based on the empirical characteristic function. Comput. Statist. Data Anal. 49, 227–242.

Knorr-Held, L., Raßer, S., 2000. Bayesian detection of clusters and discontinuities in disease maps. Biometrics 56, 13–21.

Li, J., Zha, H., 2006. Two-way Poisson mixture models for simultaneous document classification and word clustering. Comput. Statist. Data Anal. 50, 163–180.

Lindsay, B.G., 1995. Mixture models: theory geometry and applications. NSF-CBMS Regional Conference Series in Probability and Statistics, vol. 5. Institute of Mathematical Statistics, Hayward.

Link, W.A., 2003. Nonidentifiability of population size from capture–recapture data with heterogenous detection populations. Biometrics 59, 1123–1130.

Mao, C.X., 2007. Estimating population sizes for capture-recapture sampling with binomial mixtures. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.09.025.

McGrory, C.A., Titterington, D.M., 2007. Variational approximations in Bayesian model selection for finite mixture distributions. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.07.020.

McLachlan, G.J., Basford, K.E., 1988. Mixture Models. Inference and Applications to Clustering. Marcel Dekker, New York.

McLachlan, G.J., Peel, D., 2000. Finite Mixture Models. Wiley, New York.

McLachlan, G.J., Peel, D., Bean, R.W., 2003. Modelling high-dimensional data by mixtures of factor analyzers. Comput. Statist. Data Anal. 41, 379 –388.

McLachlan, G.J., Bean, R.W., Ben-Tovim Jones, L., 2007 Extension of the mixture of factor analyzers model to incorporate the multivariate *t*-Distribution. Comput. Statist. Data Anal. (this issue), doi:10.1016/j.csda.2006.09.015.

Meulders, M., DeBoeck, P., Kuppens, P., Van Mechelen, I., 2002. Constrained latent class analysis of three-way three-mode data. J. Classification 19, 277–302.

Nadarajah, S., 2006. Information matrices for Laplace and Pareto mixtures. Comput. Statist. Data Anal. 50, 950–966.

Neerchal, N.K., Morel, J.G., 2005. An improved method for the computation of maximum likelihood estimates for multinomial overdispersion models. Comput. Statist. Data Anal. 49, 33–43.

Priebe, C.E., Miller, M.I., Ratnanather, T., 2006. Segmenting magnetic resonance images via hierarchical mixture modelling. Comput. Statist. Data Anal. 50, 551–567.

Rufo, M.J., Pérez, C.J., Martín, J., 2007. Bayesian analysis of finite mixtures of multinomial and negative-multinomial distributions. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.09.026.

Seidel, W., Sever, K., Ševčiková, H., 2006. Efficient calculation of the NPMLE of a mixing distribution for mixtures of exponentials. Comput. Statist. Data Anal. 50, 1248–1271.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van der Linde, A., 2002. Bayesian measures of model complexity and fit (with discussion). J. Roy. Statist. Soc., Ser. B 64, 583–639.

Steiner, P.M., Hudec, M., 2007. Classification of large data sets with mixture models via sufficient EM. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.09.014.

Sultan, K.S., Ismail, M.A., Al-Moisheer, A.S., 2007. Mixture of two inverse Weibull distributions: properties and estimation. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.09.016.

Tao, J., Shi, N.-Z., Lee, S.Y., 2004. Drug risk assessment with determining the number of sub-populations under finite mixture normal models. Comput. Statist. Data Anal. 46, 661–676.

Titterington, D.M., Smith, A.F.M., Makov, U.E., 1985. Statistical Analysis of Finite Mixture Distributions. Wiley, New York.

Ueda, N., Ghahramani, Z., 2002. Bayesian model search for mixture models based on optimizing variational bounds. Neural Networks 15, 1223–1241.

van der Heijden, P.G.M., Bustami, R., Cruyff, M., Engbersen, G., van Houwelingen, H.C., 2003. Point and interval estimation of the population size using the truncated Poisson regression model. Statist. Model.—An Internat. J. 3, 305–322.

Vaz de Melo Mendes, B., Lopes, F.H., 2006. Data driven estimates for mixtures. Comput. Statist. Data Anal. 47, 583–598.

Vermunt, J.K., 2003. Multilevel latent class models. Sociological Methodology 33, 213–239.

Vermunt, J.K., 2007. A hierarchical mixture model for clustering three-way data sets. Comput. Statist. Data Anal. (this issue), 10.1016/j.csda.2006.08.005.

Woodward, W.A., Sain, S.R., 2003. Testing for outliers from a mixture distribution when some data are missing. Comput. Statist. Data Anal. 44, 193–210.

Yashin, A.I., Vaupel, J.W., Iachine, I., 1995. Correlated individual frailty: an advantageous approach to survival analysis of bivariate data. Math. Popul. Studies 5, 1–15.

Zheng, J., Frey, H.C., 2004. Quantification of variability and uncertainty using mixture distributions: evaluation of sample size, mixing weights, and separation between components. Risk Anal. 24, 553–571.

Dankmar Böhning
Applied Statistics, School of Biological Sciences, University of Reading,
Whiteknights, Early Gate, Harry Pitt Building, Reading, RG6 5SA, UK
E-mail address: d.a.w.bohning@reading.ac.uk

Wilfried Seidel
Department of Economic and Organizational Sciences, Helmut Schmidt University,
University of the Federal Armed Forces Hamburg, D-22039 Hamburg, Germany
E-mail address: Wilfried.Seidel@hsu-hh.DE

Macro Alfó
Dipartimento di Statistica, Probabilità e Statistiche Applicate, Università degli Studi,
“La Sapienza” di Roma, Piazzale Aldo Moro, 5I-00185, Rome, Italy
E-mail address: marco.alfo@uniromal.it

Bernard Garel
Groupe de Statistique et Probabilités (GSP), Institut National Polytechnique de Toulouse et
LSP—Institut de Mathématiques de Toulouse, Statistics and Probability,
LEN7, ENSEEIHT, BP 7122, F31071 Toulouse cédex 7, France
E-mail address: garel@enseeiht.fr

Valentin Patilea
Centre for Research in Economics and Statistics, Ecole Nationale de la Statistique et de 1’
Analyse de 1’Information, Campus de Ker-Lann, Rue Blaise Pascal BP 37203,
35172 Bruz cedex, France
*E-mail address:*patilea@ensai.fr

Günther Walther
Department of Statistics, Stanford University, Stanford, CA 94305, USA
*E-mail address:*gwalther@stanford.edu