Testing the order of finite mixture models by EM-test

Jiahua Chen
Canada Research Chair, Tier I

Department of Statistics
University of British Columbia

Joint work with Pengfei Li

Mixture estimation and applications
Edinburgh, UK

Outline

1. Review
   - Existing results and their limitations
2. EM-test
   - Test statistic
3. Some theory
   - Limiting distribution
   - Hints for the proof
4. Simulation
   - simulation

Finite mixture models: definition

- Let $\{f(x;\theta) : \theta \in \Theta\}$ be a parametric distribution family.
  - Θ: parameter space for θ.

- Then, $f(x;\Psi) = \sum_{h=1}^{m} \alpha_h f(x;\theta_h)$: density function of a finite mixture model.
  - m: order of the mixture model.
  - θ_h: the parameter of the hth sub-population.
  - α_h: the proportion of the hth sub-population in the mixture.

Finite mixture models: definition

- One may put all parameters into a mixing distribution:
  - $\Psi(\theta) = \sum_{h=1}^{m} \alpha_h I(\theta_h \leq \theta)$.
  - $\Psi(\theta)$ is a distribution on Θ with m support points.
  - Each observation X is a random outcome from a subpopulation $f(x; \theta_h)$ so that θ_h itself is a random outcome from $\Psi(\theta)$.
  - Since h is unknown, X is also an incomplete observation.

Testing the order

- Consider the hypothesis test problem on the order m.
- The order m has many possible practical implications:
  - Cluster analysis: it is the number of clusters contained in the data.
  - Linkage analysis: $m > 1$ is an indication of the presence of a subpopulation with linkage.
  - Statistical finance: m represents the number of regimes needed for proper modeling.
  - Reoder and Lindsay (1997): $m = 2, 3$ gives information on whether the mode of inheritance is recessive or dominant.

The simplest case: $\alpha_1 f(x;\theta_1) + \alpha_2 f(x;\theta_2)$

- Suppose we have an iid sample $X_1, \ldots, X_n$.
- The log-likelihood function of the mixing distribution is given by

$$l_n(\alpha_1, \alpha_2, \theta_1, \theta_2) = \sum_i \log\{\alpha_1 f(x_i;\theta_1) + \alpha_2 f(x_i;\theta_2)\}.$$

- Testing $H_0 : m = 1$ against $H_a : m = 2$ can be done through the likelihood ratio statistic:

$$R_n = 2\{\sup l_n(\alpha_1, \alpha_2, \theta_1, \theta_2) - \sup l_n(\alpha_1, \alpha_2, \theta, \theta)\}.$$

  Reject H_0 if R_n is larger than some threshold value.
- The challenge is, however, to find the proper threshold value.

The simplest case: $\alpha_1 f(x;\theta_1) + \alpha_2 f(x;\theta_2)$

- Suppose we have an iid sample $X_1, \ldots, X_n$.
- The log-likelihood function of the mixing distribution is given by

$$l_n(\alpha_1, \alpha_2, \theta_1, \theta_2) = \sum_i \log\{\alpha_1 f(x_i;\theta_1) + \alpha_2 f(x_i;\theta_2)\}.$$

- Testing $H_0 : m = 1$ against $H_a : m = 2$ can be done through the likelihood ratio statistic:

$$R_n = 2\{\sup l_n(\alpha_1, \alpha_2, \theta_1, \theta_2) - \sup l_n(\alpha_1, \alpha_2, \theta, \theta)\}.$$

Reject H_0 if R_n is larger than some threshold value.

- The challenge is, however, to find the proper threshold value.

Likelihood ratio test for regular models

- For regular models, R_n has an asymptotic chisquare distribution under the null model.
- Hence, the threshold value can be taken as some quantiles of the chisquare distribution to ensure the test has at least an asymptotically correct size.

Finite mixture model is not regular

- Even the simplest finite mixture model, $\alpha_1 f(x;\theta_1) + \alpha_2 f(x;\theta_2)$, is not regular:
  - When $\alpha_1 = 0$, any θ_1 value parameterizes the same distribution. This is the loss of identifiability (type I).
  - When $\theta_1 = \theta_2$, any (α_1, α_2) parameterize the same distribution. This is the loss of identifiability (type II).
  - The null model is not an interior point in the set of alternative models.

Some existing results

- For $(1-\alpha)N(0,1)+\alpha N(\theta,1)$, Hartigan (1985) found that $R_n \to \infty$ as $n \to \infty$.
- Bickel and Chernoff (1993) and Liu and Shao (2004) showed that this R_n goes to infinite at rate $\log\log n$.
- Ghosh and Sen (1985) considered the model

$$\alpha_1 f(x;\theta_1)+\alpha_2 f(x;\theta_2).$$

Under restriction that Θ is compact, and $|\theta_1-\theta_2|>\epsilon>0$. R_n has a limit in the form of Gaussian process.

Remedy through a restriction on mixing proportion

- Chen and Cheng (1995) and Lemdani and Pons(1999) considered the same model

$$\alpha_1 f(x;\theta_1) + \alpha_2 f(x;\theta_2).$$

- They placed a condition $\min(\alpha_1, \alpha_2) > \epsilon > 0$, in addition to Θ is compact.
- Under these restrictions, R_n has $0.5\chi_0^2 + 0.5\chi_1^2$ limiting distribution.

Breakthroughs for a few simple cases

- Straightforward result is also possible under two-component model $\alpha_1 f(x;\theta_1) + \alpha_2 f(x;\theta_2)$:
  - Chernoff and Lander (1995) obtained limiting distributions for R_n when $f(x;\theta)$ is binomial, mostly chisquare mixtures.
  - Chen and Chen (2001) and many others under restriction that Θ is compact, worked out the limiting distribution of R_n, Gaussian process based.
  - Chen and Chen (2005) found the limiting distribution of R_n when $f(x;\theta) = N(\theta, \sigma^2)$ with σ^2 being a structure parameter, Gaussian process with a spike.
- Progress for more general cases will be presented later.

Existing results on non-likelihood-based methods

- One may avoid much of the trouble of the likelihood ratio test through other approaches.
- Neyman and Scott (1966) designed a $C(\alpha)$ test which is a generalized score test.
- $C(\alpha)$ test can be regarded as a test for over-dispersion which is also studied in biostatistics (e.g. Dean and Lawless, 1989).
- Lindsay (1989) discussed moment based approaches. The result is applicable to general order, when f belongs to an exponential family.

The research in this area is abundant

- There are many books published on finite mixture models:
  - Titterington, Makov and Teicher (1995),
  - McLachlan and Peel (2000),
  - Lindsay (1995) and so on.
- Semi or Non-parametric mixture models have drawn a lot attention recently. For instance,
  - Cruz-Medina and Hettmansperger (2004, JRSSC),
  - Hunter, Wang and Hettmansperger (2007, AOS),
  - Hall and Zhou (2003, AOS).

Testing the order of the mixture model in general

- Dacunha-Castelle and Gassiat (1999) worked out the limiting distribution of the likelihood ratio test under general model set-ups;
  Yet
  - the limiting distribution must be characterized by supermum of the squared and truncated Gaussian process;
  - the result is obtained under the restriction that Θ is compact;
  - the result excludes some important distribution families such as exponential distribution.

Existing result related to order of the mixture model

- Reoder (1994) and Lindsay and Reoder (1997) developed a graphic method.
- They noticed that the density ratio changes signs m times.
- The graph of the density ratio obtained by some nonparametric estimator reveals some order information.
- Clearly, the method does not quantify the evidence of the order.

Existing result related to order of the mixture model

- Reoder (1994) and Lindsay and Reoder (1997) developed a graphic method.
- They noticed that the density ratio changes signs m times.
- The graph of the density ratio obtained by some nonparametric estimator reveals some order information.
- Clearly, the method does not quantify the evidence of the order.

Existing result related to order of the mixture model

- Chen, Chen and Kalbfleisch(2001, 2004, JRSSB) developed a modified likelihood ratio test.
- More discussions on this approach will be given later.
- The method does not work for $H_0 : m = m_0$ against $m > m_0$ when $m_0 > 2$.

Review summary

- In spite of abundant research, testing $H_0 : m = m_0$ for general m_0 remains a challenging problem.
  - The result of Dacunha-Castelle and Gassiat (1999) is hard to implement;
  - Graphic methods do not provide quantitative assessments;
  - MLRT does not work for $m_0 > 2$.
- There are rooms to new methods for testing $H_0 : m = m_0$.

EM-test for homogeneity $(m_0 = 1)$

- Li, Chen and Marriot (2009) and Chen and Li (2009) introduced EM-test for $m = m_0 = 1$ against $m > m_0$.
- The method has some distinct advantages compared to most existing methods:
  - The test statistic has simple limiting distributions under one-parameter mixture models, and under normal mixture models.
  - The result holds without requiring (a) Θ is compact; nor (b) $E_0[f(x;\theta)/f(x;\theta_0)]^2 < \infty$.
- In the latest development, we applied EM-test to test $m = m_0$ for one-dim θ.

EM-TEST FOR HOMOGENEITY $(m_0 = 1)$

- Li, Chen and Marriot (2009) and Chen and Li (2009) introduced EM-test for $m = m_0 = 1$ against $m > m_0$.
- The method has some distinct advantages compared to most existing methods:
  - The test statistic has simple limiting distributions under one-parameter mixture models, and under normal mixture models.
  - The result holds without requiring (a) Θ is compact; nor (b) $E_0[f(x;\theta)/f(x;\theta_0)]^2 < \infty$.
- In the latest development, we applied EM-test to test $m = m_0$ for one-dim θ.

EM-test for general order $m = m_0$

- We now present an EM-test for

$$H_0 : m = m_0 \text{ versus } H_A : m > m_0$$

where m_0 is an arbitrarily given positive integer.

  - Caution: although the result is neat, the preparation is long and tedious.

EM-test: hypotheses

Testing problem

Given a random sample $X_1, \ldots, X_n$ from $f(x; \Psi)$,

$$\text{we wish to test } H_0 : m = m_0 \text{ versus } H_A : m > m_0$$

where m_0 is any given positive integer.

EM-test: initialization

Compute the MLE of Ψ, $\hat{\Psi}_0$ under H_0:

$$\hat{\Psi}_0 = \sum_{h=1}^{m_0} \hat{\alpha}_{0h} I(\hat{\theta}_{0h} \leq \theta)$$

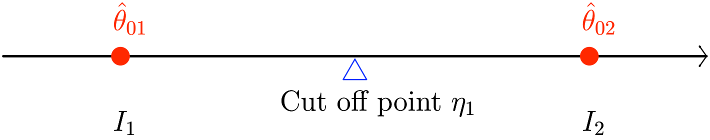
with $\hat{\theta}_{01} \leq \hat{\theta}_{02} \leq \cdots \leq \hat{\theta}_{0m_0}$.

EM-test: motivation

- $\hat{\Psi}_0$ is the best representation of the population by m_0 subpopulations.
- If the order of the population is $m > m_0$, replacing each subpopulation by two should improve the fit.
- The degree of improvement provides evidence for $> m_0$ sub-populations.

EM-test: graphic illustration

- Divide Θ into m_0 intervals ($I_h's$) with the cut-off points being $\eta_h = (\hat{\theta}_{0h} + \hat{\theta}_{0\,h+1})/2, \quad h = 1, \ldots, m_0 - 1.$

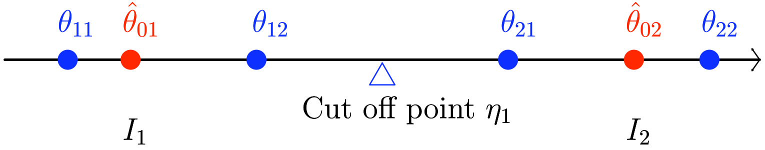


EM-test: initialization

- For some $\boldsymbol{\beta}_0 = (\beta_{01}, \ldots, \beta_{0m_0})^\tau$ with $\beta_{0h} \in (0, 0.5]$, we create a class of mixing distributions of order $2m_0$:

$$\Omega_{2m_0}(\boldsymbol{\beta}_0) = \left\{ \sum_{h=1}^{m_0} \alpha_h \{\beta_{0h} I(\theta_{h1} \le \theta) + (1 - \beta_{0h}) I(\theta_{h2} \le \theta)\} : \theta_{h1}, \theta_{h2} \in I_h \right\}.$$

  - I_h are determined by $\hat{\Psi}_0$, $\boldsymbol{\beta}_0$ is pre-chosen.
  - only $(\alpha_h, \theta_{h1}, \theta_{h2})$ are free parameters.



EM-test: Step 1 (constrained maximization)

- Define a modified log-likelihood function on $\Omega_{2m_0}(\beta_0)$

$$pl_n(\Psi) = l_n(\Psi) + \sum_{h=1}^{m_0} p(\beta_{0h})$$

where

- $l_n(\Psi)$: log-likelihood function.
- $p(\cdot)$ is a penalty function that is maximized at 0.5 with the maximum value 0.

EM-test: Step 1 continued

- For each β_0, we compute

$$\Psi^{(1)}(\beta_0) = \arg\max\{pl_n(\Psi) : \Psi \in \Omega_{2m_0}(\beta_0)\}$$

  where the maximization is with respect to $(\alpha_h, \theta_{h1}, \theta_{h2})$.

- Define the test statistic

$$M_n^{(1)}(\beta_0) = 2\{pl_n(\Psi^{(1)}(\beta_0)) - l_n(\hat{\Psi}_0)\}.$$

  - $M_n^{(1)}(\beta_0)$ reflects how much the likelihood has improved when each subpopulation is split into two.

EM-test: Step 2 (EM-iteration)

- Starting from $\Psi^{(1)}(\boldsymbol{\beta}_0)$, we use EM-iteration to update the value of Ψ for pre-given times.
  - In further iterations, $(\alpha_h, \theta_{h1}, \theta_{h2}, \beta_h)$ are all updated without constraints.
  - That is, $\Psi^{(k)}(\boldsymbol{\beta}_0)$ for $k > 1$ are not confined in $\Omega_{2m_0}(\boldsymbol{\beta}_0)$.
- Denote the updated values by $\Psi^{(K)}(\boldsymbol{\beta}_0)$.
- Define the test statistic

$$M_n^{(K)}(\boldsymbol{\beta}_0) = 2\{pl_n(\Psi^{(K)}(\boldsymbol{\beta}_0)) - l_n(\hat{\Psi}_0)\}.$$

EM-TEST: STEP 3 (MULTIPLE CHOICES OF $\boldsymbol{\beta}_0$)

- Let B be a finite set of numbers from $(0, 0.5]$. For instance, let $\mathrm{B} = \{0.1, 0.3, 0.5\}$.
- Create B^{m_0} by direct product.
- Use each point in B^{m_0} as a choice of $\boldsymbol{\beta}_0 = (\beta_{01}, \ldots, \beta_{0m_0})^\tau$.
- Compute $M_n^{(K)}(\boldsymbol{\beta}_0)$ for each $\boldsymbol{\beta}_0 \in \mathrm{B}^{m_0}$.

Final step

- Define the EM-test statistic, for a pre-specified K, as

$$EM_n^{(K)} = \max\{M_n^{(K)}(\beta_0) : \beta_0 \in B^{m_0}\}.$$

- Reject the null hypothesis when $EM_n^{(K)}$ exceeds some critical value.
- The key issue is: the limiting distribution of this statistic.

Final step

- Define the EM-test statistic, for a pre-specified K, as

$$EM_n^{(K)} = \max\{M_n^{(K)}(\beta_0) : \beta_0 \in B^{m_0}\}.$$

- Reject the null hypothesis when $EM_n^{(K)}$ exceeds some critical value.
- The key issue is: the limiting distribution of this statistic.

EM-test: limiting distribution (1)

Theorem 1

Under some regularity conditions on $f(x;\theta)$ and $p(\beta)$, and assume $0.5 \in \mathrm{B}$,

$$EM_n^{(K)} \quad \rightarrow \quad \sup_{\mathbf{v}\geq 0}(2\mathbf{v}^\tau \mathbf{w} - \mathbf{v}^\tau \Omega \mathbf{v}) = \sum_{h=0}^{m_0} a_h \chi_h^2$$

for some $a_h \geq 0$ and $\sum_{h=0}^{m_0} a_h = 1$, under Ψ_0 and fixed K.

- $\mathbf{w} = (w_1, \ldots, w_{m_0})^\tau$: a 0-mean multivariate normal random vector with correlation matrix $\Omega = (\omega_{ij})$.
- $\mathbf{v} = (v_1, \ldots, v_{m_0})^\tau$ and $\{\mathbf{v} \geq 0\} = \{v_1 \geq 0, \ldots, v_{m_0} \geq 0\}$.
- The weights $(a_0, \ldots, a_{m_0})$ depend on Ω.
- Ω can be calculated based on Ψ_0 or $\hat{\Psi}_0$.

EM-test: limiting distribution (2)

Theorem 1 (continued)

In particular,

1. when $m_0 = 1$, $a_0 = a_1 = 0.5$;
2. when $m_0 = 2$, $a_0 = (\pi - \arccos \omega_{12})/(2\pi)$, $a_1 = 0.5$, and $a_0 + a_2 = 0.5$;
3. when $m_0 = 3$, $a_0 + a_2 = a_1 + a_3 = 0.5$ and

$$\begin{aligned} a_0 &= (2\pi - \arccos \omega_{12} - \arccos \omega_{13} - \arccos \omega_{23})/(4\pi), \\ a_1 &= (3\pi - \arccos \omega_{12:3} - \arccos \omega_{13:2} - \arccos \omega_{23:1})/(4\pi), \end{aligned}$$

where

$$\omega_{ij:k} = \frac{(\omega_{ij} - \omega_{ik}\omega_{jk})}{\sqrt{(1 - \omega_{ik}^2)(1 - \omega_{jk}^2)}}.$$

EM-test: limiting distribution (2)

Theorem 1 (continued)

The regularity conditions DO NOT include

- Θ is compact;
- $E_0[f(X;\Psi)/f(X;\Psi_0)]^2 < \infty$.

This is in sharp comparison over most existing methods.

EM-test: tuning parameters

Tuning parameter values

- Initial-value set for β_{0h}: $\mathrm{B} = \{0.1, 0.3, 0.5\}$;
- Iteration number K: 2 or 3;
- Penalty function on β: $p(\beta) = C \log(1 - |1 - 2\beta|)$.

EM-test: tuning parameters

Choice of C

- Any $C > 0$ suffices for asymptotic consideration, its value has some influence on the precision of the test.
- Based on carefully designed computer experiments, we developed some simple data dependent empirical formulas for C.

EM-test: tuning parameters

- For Poisson, Binomial and Normal Mixture models, we recommend

$$C = \begin{cases} \dfrac{0.5\exp(5 - 10.6\omega_{12} - 123/n)}{1 + \exp(5 - 10.6\omega_{12} - 123/n)} & m_0 = 2; \\ \dfrac{0.5\exp(3.3 - 5.5\omega_{12} - 5.5\omega_{23} - 165/n)}{1 + \exp(3.3 - 5.5\omega_{12} - 5.5\omega_{23} - 165/n)} & m_0 = 3. \end{cases} \quad (1)$$

- For exponential mixture models, we recommend

$$C = \begin{cases} \dfrac{\exp(2.3 - 8.5\omega_{12})}{1 + \exp(2.3 - 8.5\omega_{12})} & m_0 = 2; \\ \dfrac{\exp(2.2 - 5.9\omega_{12} - 5.9\omega_{23})}{1 + \exp(2.2 - 5.9\omega_{12} - 5.9\omega_{23})} & m_0 = 3. \end{cases} \quad (2)$$

The EM-idea is developed from existing methods

- If we severely restrict the alternative model

$$H_a : \alpha_1 f(x; \theta_1) + \alpha_2 f(x; \theta_2),$$

  to those with $\alpha_1 = 0.3$, then R_n has $0.5\chi^2_1 + 0.5\chi^2_0$ limiting distribution.

- This result does not require compact Θ, nor $E_0\{f(X; \Psi)/f(X; \Psi_0)\}^2 < \infty$.
- All aforementioned challenges disappear, though we may not like the restriction.

The EM-idea is developed from existing methods

- If we severely restrict the alternative model

$$H_a : \alpha_1 f(x;\theta_1) + \alpha_2 f(x;\theta_2),$$

  to those with $\alpha_1 = 0.3$, then R_n has $0.5\chi_1^2 + 0.5\chi_0^2$ limiting distribution.

- This result does not require compact Θ, nor $E_0\{f(X;\Psi)/f(X;\Psi_0)\}^2 < \infty$.
- All aforementioned challenges disappear, though we may not like the restriction.

The EM-idea is developed from existing methods

- If we slightly relax the previous restriction on

$$H_a : \alpha_1 f(x;\theta_1) + \alpha_2 f(x;\theta_2),$$

  to those with $0.3 < \alpha_1 < 0.7$, then R_n still has $0.5\chi_1^2 + 0.5\chi_0^2$ limiting distribution.
- We may still dislike the restriction, yet this is practically what Chen and Cheng (1995) and Lemdami and Pons(1999) did.

The EM-idea is developed from existing methods

- We may replace the previous restriction by adding a penalty term $C\log\{4\alpha(1-\alpha)\}$ to the log-likelihood.
- The modified likelihood ratio statistic still has $0.5\chi_1^2 + 0.5\chi_0^2$ limiting distribution, though some additional conditions on $f(x;\theta)$ are needed.
- The common truth is: put a restriction on α explicitly or implicitly.

The EM-idea is developed from existing methods

- We may replace the previous restriction by adding a penalty term $C \log\{4\alpha(1-\alpha)\}$ to the log-likelihood.
- The modified likelihood ratio statistic still has $0.5\chi_1^2 + 0.5\chi_0^2$ limiting distribution, though some additional conditions on $f(x;\theta)$ are needed.
- The common truth is: put a restriction on α explicitly or implicitly.

The truth behind the EM-idea

- Start with pre-chosen α-values (called $\boldsymbol{\beta}_0$ now);
- Overcome the shortcoming caused by fixed α via EM-iteration;
- Limit the number of EM-iterations to retain simple limiting behavior;
- Enhance the performance with multiple initial $\boldsymbol{\beta}_0$-values.

Advantages of the EM-test

- Simple limiting distributions for general $H_0 : m = m_0$;
- Least restrictive on Θ and $f(x; \theta)$;
- Better match between the limiting distribution and the finite sample distribution.
- High efficiency (supported by simulation results).

Simulation study: Goal

We wish to demonstrate that

- the limiting distribution approximates the sample distribution of the EM-test well.
- the EM-test is efficient at detecting alternative models.

POISSON MIXTURES AGAINST $H_0 : m = 2$

TABLE: 1. Parameters in Poisson mixture models.

Model	Probability			Support		
N1	0.5	0.5		3	12	
N2	0.75	0.25		3	12	
N3	0.5	0.5		3	15	
N4	0.75	0.25		3	15	
N5	0.5	0.5		3	18	
N6	0.75	0.25		3	18	
A1	0.2	0.4	0.4	1	3	9
A2	0.4	0.2	0.4	1	5	9
A3	0.2	0.4	0.4	1	3	12
A4	0.4	0.2	0.4	3	7	12

Poisson mixtures against $m = 2$

Table: 2. Type I errors (%) of the EM-test and the MLRT under Poisson mixture models against $H_0 : m = 2$ ($n = 200$)

Level	Model	$EM_n^{(2)}$	$EM_n^{(3)}$	MLRT	Model	$EM_n^{(2)}$	$EM_n^{(3)}$	MLRT
10%	N1	10.3	10.4	11.4	N2	9.8	9.9	11.0
5%	N1	5.1	5.1	6.2	N2	5.0	5.1	6.4
1%	N1	1.1	1.1	1.4	N2	1.5	1.5	1.3
10%	N3	9.5	9.6	12.8	N4	9.8	9.9	12.7
5%	N3	5.0	5.0	6.9	N4	5.1	5.1	6.8
1%	N3	1.0	1.0	1.5	N4	1.1	1.1	1.6
10%	N5	10.1	10.2	14.3	N6	9.9	10.0	13.3
5%	N5	5.3	5.3	8.1	N6	5.1	5.1	7.8
1%	N5	1.2	1.2	2.0	N6	1.4	1.4	1.8

Poisson mixtures against $m_0 = 2$

Table: 3. Powers (%) of the EM-test and the MLRT under Poisson mixture models against $H_0 : m = 2$ at the 5% level

Model	$EM_n^{(2)}$	$EM_n^{(3)}$	MLRT	Model	$EM_n^{(2)}$	$EM_n^{(3)}$	MLRT
			$n = 100$				
A1	31.8	31.9	33.4	A2	29.9	30.1	30.8
A3	41.6	41.7	46.0	A4	21.3	21.4	20.9
			$n = 200$				
A1	55.0	55.0	54.5	A2	50.2	50.3	47.8
A3	69.6	69.6	71.3	A4	35.0	35.1	34.8

Simulation conclusions

- The type I errors of the EM-test are close to the nominal levels.
- The asymptotic distribution is a good approximation to the sample distribution of the EM-test.
- The type I errors of the MLRT are somehow larger than the nominal levels, especially when two support points are far away from each other.
- Under almost all alternative models, the power of the MLRT is slightly higher than the EM-test.
- This is probably because the Type I errors of the MLRT are also slightly higher than those of the EM-test.

Summary

- Designed an easy-to-implement EM-test for testing the order of a finite mixture.
  - Easy to calculate the statistics.
  - Easy to calculate the *p*-values for the statistics based on the limiting distribution.
  - Widely applicable (without the two restrictive conditions for the MLRT).

- R package for testing $m = 1, 2, 3$ under Binomial, Normal, Poisson and Exponential mixture models, is available.

Key reference

- Chen, H., Chen, J., and Kalbfleisch, J. D. (2004) Testing for a finite mixture model with two components. *Journal of the Royal Statistical Society, Series B*, **66**, 95-115.
- Chen, J. and Li, P. (2009) Hypothesis test for normal mixture models: The EM approach. *The Annals of Statistics*. **37**, 2523-2542.
- Li, P., Chen, J., and Marriott, P. (2009) Non-finite Fisher information and homogeneity: The EM approach. *Biometrika*, **96**, 411-426.

Thank you

Questions are welcome