Empirical Likelihood Based Deviance Information Criterion

Yin Teng

Smart and Safe City Center of Excellence
NCS Pte Ltd

June 22, 2016

Outline

Outline

Empirical Likelihood (EL)-Definition

- $X \in \mathbb{R}$ is a random variable following F_θ^0
- $F_\theta^0 \in \mathcal{F}_\theta$, $\theta = (\theta_1, \ldots, \theta_d) \in \Theta \subseteq \mathbb{R}^d$
- $h(X, \theta) = (h_1(X, \theta), \ldots, h_q(X, \theta))^T$ are known to satisfy

$$E_{F_\theta^0}[h(X, \theta)] = 0. \tag{1.1}$$

- $x = (x_1, \ldots, x_n)$ are n observations of X
- $F(x_i) = P(X \leq x)$ and $F(x_i -) = P(X < x)$
- A non-parametric likelihood of F can be defined as

$$\mathcal{L}(F) = \prod_{i=1}^{n} \{F(x_i) - F(x_i -)\} \tag{1.2}$$

Empirical Likelihood-Definition (Cont'd)

- EL estimates F^0_θ by maximising $\mathcal{L}(F)$ over $\mathcal{F}_\theta$ under constraints depending on $h(x, \theta)$
- Define $\omega_i = F(x_i) - F(x_i-)$, for a given $\theta \in \Theta$. EL computes

$$\hat{\omega}(\theta) = \underset{\omega \in \mathcal{W}_\theta}{\operatorname{argmax}} \sum_{i=1}^{n} \log \omega_i(\theta) \tag{1.3}$$

where

$$\mathcal{W}_\theta = \left\{ \omega : \sum_{i=1}^{n} \omega_i(\theta) h(x_i, \theta) = 0 \right\} \cap \Delta_{n-1}.$$

Here Δ_{n-1} is the $n-1$ dimensional simplex

Empirical Likelihood - Definition (Cont'd)

- F_θ^0 can be estimated by

$$\hat{F}_\theta^0(x) = \sum_{i=1}^{n} \hat{\omega}_i(\theta) \mathbb{1}_{\{x_i \leq x\}}.$$

- The empirical likelihood corresponding to $\hat{F}_\theta^0$ is then given by

$$L(\theta) = \prod_{i=1}^{n} \hat{\omega}_i(\theta). \tag{1.4}$$

Bayesian Empirical Likelihood

- We have prior $\pi(\theta)$
- We can define a posterior as

$$\Pi(\theta|x) = \frac{L(\theta)\pi(\theta)}{\int L(\theta)\pi(\theta)d\theta}. \tag{1.5}$$

1. If $L(\theta) = n^{-n}$, then $\Pi(\theta|x) = \pi(\theta)$
2. Absence of analytical form

Bayesian Empirical Likelihood - Problems

Computational issues

- No analytical form of posterior - Gibbs sampling fails
- Non-convex problem- random walk MH not easy
- Mixed effect models-Parallel tempering mixed slowly

Bayesian model selection

- Bayes factor -Not easy to calculate
- Posterior predictive distribution -The integral is not easily obtained

Outline

Deviance information criterion(DIC)

- $f(y|\theta)$ depends on a parameter vector $\theta \in \Theta \subseteq \mathbb{R}^p$
- $y_1, \ldots, y_n$ are n i.i.d random variables
- A posterior of θ is then given by

$$\Pi(\theta|y) = \frac{p(y|\theta)\pi(\theta)}{\int_{\theta\in\Theta} p(y|\theta)\pi(\theta)d\theta}.$$

- a Bayesian deviance is defined as

$$D(\theta) = -2\log p(y|\theta) - 2\log p(y). \tag{2.1}$$

where $p(y)$ is some fully specified standardizing term which is a function of the data alone (Spiegelhalter u. a., 2002)

Deviance information criterion (DIC)

- Bayesian model fit is measured by posterior expectation of the deviance
$$\overline{D(\theta)} = E_\Pi \left[D(\theta) \right] \tag{2.2}$$
- Bayesian model complexity is measured by the effective number of parameters in the model
$$p_D = \overline{D(\theta)} - D(\hat{\theta}_\Pi). \tag{2.3}$$
where $\hat{\theta}_\Pi$ is some posterior estimate of parameter
- DIC is defined as
$$DIC = \overline{D(\theta)} + p_D = 2\overline{D(\theta)} - D(\hat{\theta}_\Pi) = D(\hat{\theta}_\Pi) + 2p_D. \tag{2.4}$$

Empirical likelihood based DIC - Definition

- The empirical likelihood deviance is given by

$$D_{EL}(\theta) = -2\sum_{i=1}^{n} \log \hat{\omega}_i(\theta) - 2n\log n. \tag{2.5}$$

- BayesEL model fit can be defined as BayesEL posterior mean of (2.5)

$$\overline{D_{EL}(\theta)} = E_{\Pi_{EL}}\left[D_{EL}(\theta)\right]. \tag{2.6}$$

- BayesEL model complexity is defined as the effective number of parameters

$$p_D^{EL} = \overline{D_{EL}(\theta)} - D_{EL}(\bar{\theta}_{EL}). \tag{2.7}$$

- $\bar{\theta}_{EL}$ is the posterior mean
- The empirical likelihood based DIC (ELDIC) is defined as

$$ELDIC = \overline{D_{EL}(\theta)} + p_D^{EL}. \tag{2.8}$$

Empirical likelihood based DIC - Definition (Cont'd)

DIC

$$D(\theta) = -2\log p(\theta|y) - 2\log p(y)$$

$$DIC = \overline{D(\theta)} + p_D$$

$$\overline{D(\theta)} = E_{\Pi}(D(\theta))$$

$$p_D = \overline{D(\theta)} - D(\bar{\theta}_{\Pi})$$

EL based DIC

$$D_{EL}(\theta) = -2\log L(\theta) - 2n\log n$$

$$ELDIC = \overline{D_{EL}(\theta)} + p_D^{EL}$$

$$\overline{D_{EL}(\theta)} = E_{\Pi_{EL}}\left[D_{EL}(\theta)\right]$$

$$p_D^{EL} = \overline{D_{EL}(\theta)} - D_{EL}(\bar{\theta}_{EL})$$

Empirical likelihood based DIC - Definition (Cont'd)

Properties of DIC

- Approximate normality of $\Pi(\theta|y)$ is important!
  - positivity of p_D
  - decision theoretical justification
  - Bayesian version of AIC (Akaike, 1974)
- p_D is not invariant to reparameterization!

Properties of ELDIC

- Approximate normality of $\Pi_{EL}(\theta|y)$ is important!
  - positivity of p_D^{EL}
  - consistency of $\bar{\theta}_{EL}$
  - decision theoretical justification
  - Bayesian version of ELAIC (Variyath u. a., 2010)
- p_D^{EL} is not invariant to reparameterization!

Approximate normality

Theorem 2.1

Let

$$J(\hat{\theta}_n) = -\frac{1}{n}\frac{\partial^2}{\partial\theta\partial\theta^T}\log L(\theta)\Big|_{\theta=\hat{\theta}_n}. \tag{2.9}$$

Assume that $J(\hat{\theta}_n)$ and J_0 are invertible. Under some regularity condition, for $\{\theta :\| \theta - \theta_0 \|= O(n^{-1/2})\}$, the posterior distribution of θ has density

$$\Pi_{EL}(\theta|y) \propto \exp\Big\{ -\frac{1}{2}(\theta - \bar{\theta}_{EL})^T J_n(\theta - \bar{\theta}_{EL}) + o_p(1)\Big\} \tag{2.10}$$

where

$$J_n = J_0 + nJ(\hat{\theta}_n), \tag{2.11}$$

$$\bar{\theta}_{EL} = J_n^{-1}\{J_0 m_0 + nJ(\hat{\theta}_n)\hat{\theta}_n\}. \tag{2.12}$$

Furthermore, if $J(\hat{\theta}_n)$ and J_0 are positive definite, $J_n^{1/2}(\theta - \bar{\theta}_{EL}) \xrightarrow{D} N(0, I)$.

Consistency of the posterior mean

Theorem 2.2

Assume $m_0 = O_p(1)$ *and* $J_0 = O_p(1)$*, under some regularity assumption,* $\bar{\theta}_{EL} \xrightarrow{p} \theta_0$*.*

Decision theoretic justification

- $y_{rep} = (y_{rep,1}, \ldots, y_{rep,n})$ is a replicate of $y = (y_1, \ldots, y_n)$ from the same data generating process.
- y_{rep} follows $F^0_{rep,\theta}$, which is actually same as F^0_θ.
- Let

$$\mathcal{W}_{y_{rep},\theta} = \{\omega : \sum_{i=1}^{n} \omega_i h(y_{rep,i}, \theta) = 0\} \cap \Delta_{n-1}.$$

- Given θ and y_{rep}, an empirical likelihood of y_{rep} is

$$L_{rep}(\theta) = \prod_{i=1}^{n} \hat{\omega}_{rep,i}$$

  where $\hat{\omega}_{rep} = (\hat{\omega}_{rep,1}, \ldots, \hat{\omega}_{rep,n})$ and

$$\hat{\omega}_{rep} = \underset{\omega \in \mathcal{W}_{y_{rep},\theta}}{\operatorname{argmax}} \sum_{i=1}^{n} \log \omega_i.$$

Decision theoretic justification (Cont'd)

- The loss function of using the observed data y to predict y_{rep}

$$\mathcal{L}(y_{rep}, y) = -2 \log L_{rep}(\tilde{\theta}(y)) - 2n \log n = -2 \sum_{i=1}^{n} \log n\hat{\omega}_{rep,i} \tag{2.13}$$

- $\tilde{\theta}(y)$ is a summary of θ based on the observed data y.
- The BayesEL posterior predictive distribution function

$$F(y_{rep}|y) = \int \Pi_{EL}(\theta|y) F^0_{rep,\theta} d\theta. \tag{2.14}$$

- The risk of predicting y_{rep} by using y

$$\mathcal{R}(y) = E_{y_{rep}|y}(\mathcal{L}(y_{rep}, y)) = \int \mathcal{L}(y_{rep}, y) dF(y_{rep}|y). \tag{2.15}$$

Decision theoretic justification (Cont'd)

Theorem 2.3

If $\tilde{\theta}(y) = \bar{\theta}_{EL}$, assuming that y_{rep} and y are generated from the same mechanism, we get

$$E_y(\mathcal{R}(y)) = E_y E_{y_{rep}|y}\left(\mathcal{L}(y_{rep}, y)\right) = E_y(ELDIC) + o(1). \qquad (2.16)$$

Under a diffused prior (i.e. $J_0 = o_p(1)$). By Theorem 2.1, we have $\bar{\theta}_{EL} \approx \hat{\theta}_n$ and $p_D^{EL} \approx p$. Thus

$$ELDIC = D_{EL}(\bar{\theta}_{EL}) + 2p_D^{EL} \approx -2 \log L(\hat{\theta}_n) + 2p.$$

ELDIC reduces to empirical likelihood based AIC(Variyath u. a., 2010).

Prior varying with the size of samples

Consider the following three classes of priors.

1. For $\pi_1(\theta)$, $m_{0,n} = O_p(1)$, $J_{0,n} = o_p(n)$
2. For $\pi_2(\theta)$, $m_{0,n} = \theta_0 + o_p(1)$, $J_{0,n} = O_p(n)$
3. For $\pi_3(\theta)$, $m_{0,n} = \theta_0 + o_p(1)$, $J_{0,n} = O_p(n^{1+\alpha})$, where $\alpha > 0$

Note that

- $\pi_1(\theta)$, $J_{0,n}$ increases in a slower rate than n.
- $\pi_2(\theta)$ is shrinking to θ_0 at the same rate of n
- $\pi_3(\theta)$ is shrinking to θ_0 at a faster rate than n.

Prior varying with the size of samples

Theorem 2.4

Assume that $\lambda_{(1)}J(\hat{\theta}_n))$ *and* $\lambda_{(1)}(J(\hat{\theta}_n)^{-1})$ *are bounded. Under the same assumptions in Theorem 2.1, under the priors* $\pi_1(\theta)$, $\pi_2(\theta)$ *and* $\pi_3(\theta)$, $\bar{\theta}_{EL} \xrightarrow{p} \theta_0$.

Theorem 2.5

Under the same assumption of Theorem 2.4, let p *be the dimension of the parameter, the following statements hold.*

1. *If the prior is* $\pi_1(\theta)$, $p_D^{EL} \xrightarrow{p} p$, *as* $n \to \infty$
2. *If the prior is* $\pi_2(\theta)$, *then* $p_D^{EL} < p$, *as* $n \to \infty$.
3. *If the prior is* $\pi_3(\theta)$, *then* $p_D^{EL} \xrightarrow{p} 0$, *as* $n \to \infty$

An alternative definition of p_D^{EL}

- Recall that p_D is not invariant to reparameterization!
- Gelman u. a. (2003) defined an alternative measure

$$p_V = \frac{V_\Pi[D(\theta)]}{2}.$$

- p_D^{EL} is also not invariant to reparameterization!
- We define

$$p_V^{EL} = \frac{V_{\Pi_{EL}}(D_{EL}(\theta))}{2}$$

An alternative definition of p_D^{EL}(Cont'd)

Theorem 2.6

Assume $J_0 = o_p(1)$, under the same assumptions in Theorem 2.4, we have

$$p_V^{EL} = p_D^{EL} + o_p(1) \tag{2.17}$$

Corollary 2.7

Define $ELDIC^{(2)} = D_{EL}(\bar{\theta}_{EL}) + 2p_V^{EL}$, then we have

$$E_y\left(\mathcal{R}(y)\right) = E_y(ELDIC^{(2)}) + o(1) \tag{2.18}$$

Priors and P_D^{EL}

- The predictor vector $x_i = (x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5})^T$ was generated independently from $\mathcal{N}(0, I)$, where I is the identity matrix.
- The coefficient vector $\beta = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)$ was generated independently from $\mathcal{N}(0, 0.25I)$.
- 100 independent observations were generated from a linear regression model given by

$$y_i = x_i^T \beta + \epsilon_i \quad , i = 1, 2, \ldots, 100,$$

  where ϵ_i follows $N(0, 2.5^2)$.
- Three priors of β are considered. They are $N(0, 100)$, $N(\beta_0, 0.01)$ and $N(\beta_0, 0.001)$ respectively. In $\pi_2(\beta)$ and $\pi_3(\beta)$, β_0 is the true value of β.

Priors and P_D^{EL}

Suppose $\omega = (\omega_1, \ldots, \omega_n)$ is the vector of jumps of the estimated joint distribution of y and x at the i^{th} observation. We define the set of constraints as

$$\mathcal{W}_\beta = \left\{ \omega : \sum_{i=1}^{n} \omega_i (y_i - x_i^T \beta) = 0, \sum_{i=1}^{n} \omega_i x_{ij} (y_i - x_i^T \beta) = 0, j = 1, \ldots, 5 \right\} \cap \Delta_{99}.$$

Given β, the empirical likelihood is given by

$$L(\beta) = \prod_{i=1}^{n} \hat{\omega}_i \tag{2.19}$$

where

$$\hat{\omega} = \arg \max_{\omega \in \mathcal{W}_\beta} \sum_{i=1}^{n} \log \omega_i.$$

Priors and P_D^{EL}

Table: The means and standard deviations (sd) of $\overline{D_{EL}}(\theta)$, $D_{EL}(\bar{\theta}_{EL})$, $p_D^{EL}$and p_V^{EL} for 1000 repetitions under priors $N(0, 100)$, $N(\beta_0, 0.01)$ and $N(\beta_0, 0.001)$ respectively.

Priors	$\overline{D_{EL}(\theta)}$		$D(\bar{\theta}_{EL})$		p_D^{EL}		p_V^{EL}	
	Mean	sd	Mean	sd	Mean	sd	Mean	sd
$\mathcal{N}(0, 100)$	7.65	2.92	2.62	2.92	5.04	0.12	5.14	0.32
$\mathcal{N}(\beta_0, 0.1)$	6.48	2.80	3.29	2.77	3.19	0.17	3.06	0.68
$\mathcal{N}(\beta_0, 0.001)$	8.85	5.77	8.76	5.77	0.09	0.02	0.26	0.23

Variables selection for over-dispersed Poisson regression

- We have marginal distribution of y such that

$$E(y) = \mu \qquad Var(y) = \mu(1 + \mu\lambda)$$

- We consider a four covariates generalized linear model such that

$$\log(\mu) = \beta_0 + x_1\beta_1 + x_2\beta_2 + x_3\beta_3 + x_4\beta_4$$

  with $\beta = c(0.5, 0.5, 0.6, 0, 0)$.

- The covariance structure is $cov(x_i, x_j) = (0.5)^{|i-j|}$ and λ has four levels, $(0, 1/8, 1/6, 1/4)$

Variables selection for over-dispersed Poisson regression

Similar to linear model example, we use $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_4)$ to denote model and $x_{i\boldsymbol{\gamma}}$ to denote the i^{th} observation for covariate vector of model $\boldsymbol{\gamma}$. Let $\beta_{\boldsymbol{\gamma}}$ be the corresponding coefficient vector. Then the set of constraints is defined as

$$\mathcal{W}_{\beta_{\boldsymbol{\gamma}}} = \Big\{\omega : \sum_{i=1}^{n} \omega_i (y_i - \exp\{\beta_0 + x_{i\boldsymbol{\gamma}}\beta_{\boldsymbol{\gamma}}\}) = 0, \sum_{i=1}^{n} \omega_i x_{ij} (y_i - \exp\{\beta_0 + x_{i\boldsymbol{\gamma}}\beta_{\boldsymbol{\gamma}}\}) = 0, \\ j = 1, 2, 3, 4\}\Big\} \cap \Delta_{99}$$

The empirical likelihood under model $\boldsymbol{\gamma}$ is given by

$$L(\beta_{\boldsymbol{\gamma}}) = \prod_{i=1}^{n} \hat{\omega}_i \qquad (2.20)$$

where

$$\hat{\omega} = \arg\max_{\omega \in \mathcal{W}_{\beta_{\boldsymbol{\gamma}}}} \sum_{i=1}^{n} \log \omega_i.$$

Variables selection for over-dispersed Poisson regression

Table: Comparison of $ELDIC^{(1)}$, $ELDIC^{(2)}$, $DIC^{(1)}$ and $DIC^{(2)}$ based on % time the model selected (1) TM; (2) TM+1; (3) TM+2 with different over-dispersed parameter.

λ	Model	$ELDIC^{(1)}$	$ELDIC^{(2)}$	$DIC^{(1)}$	$DIC^{(2)}$
0	TM	0.676	0.660	0.712	0.684
	TM+1	0.942	0.940	0.964	0.954
	TM+2	1.000	1.000	1.000	1.000
1/8	TM	0.620	0.604	0.548	0.508
	TM+1	0.944	0.930	0.896	0.888
	TM+2	1.000	1.000	1.000	1.000
1/6	TM	0.624	0.606	0.548	0.518
	TM+1	0.954	0.950	0.902	0.880
	TM+2	1.000	1.000	1.000	1.000
1/4	TM	0.592	0.590	0.440	0.418
	TM+1	0.920	0.928	0.876	0.842
	TM+2	0.994	0.992	0.998	0.998

Analysis of gene expression data

- Data is from $n = 118$ microarray experiments collected and analyzed by (Wille u. a., 2004)
- To reveal the correlation structure of these genes, Wille u. a. (2004) proposed a modified graphical gaussian modeling approach where the dependence between two genes was investigated only conditioning on a third one rather than all the other genes at a time.
- Drton u. Perlman (2007) employed a multiple testing based graphical model selection approach to analyze 13 genes from the MEP pathway in Wille u. a. (2004).
- We apply ELDIC on the same genes as Drton u. Perlman (2007).
- Prior settings: $\beta_\gamma \sim double\ exponential(0, 100)$.

Analysis of gene expression data - Cont'd

- These 13 genes are well ordered and observations for each one are standardized.
- For each gene, or say node, a linear model is applied.
- All ancestors of the given node are considered as potential covariates in the linear model.
- Our goal is to select a directed acyclic graphic (DAG) model, which can best illustrate the correlation structure among these genes.

Analysis of gene expression data - Cont'd

- Suppose $k \in \{4, 5, \ldots, 13\}$ indicate the number of the gene and g_k indicate the k^{th} gene.
- Given the g_k, the model to fit this gene is denoted by $\boldsymbol{\gamma}^k = (\gamma_1, \ldots, \gamma_{k-1})$, where γ_i, is 1 when g_i is in the model and 0 otherwise.
- Let $g_{\boldsymbol{\gamma}^k}$ be the covariates, $\beta_{\boldsymbol{\gamma}^k}$ be the corresponding coefficient vector, and$\epsilon_{\boldsymbol{\gamma}^k}$ be the error which has mean 0 and variance $\sigma^2_{\boldsymbol{\gamma}^k}$.
- The model $\boldsymbol{\gamma}^k$ is then given by

$$g_k = \beta_{\boldsymbol{\gamma}^k} g_{\boldsymbol{\gamma}^k} + \epsilon_{\boldsymbol{\gamma}^k} \tag{2.21}$$

- Prior settings: $\beta_\gamma \sim double\ exponential(0, 100)$

Analysis of gene expression data - Cont'd

Suppose that $(g_k^{(i)}, g_{\boldsymbol{\gamma}^k}^{(i)})$ is the i^{th} observations of $(g_k, g_{\boldsymbol{\gamma}^k})$. Let ω_i is the weight in the empirical likelihood for the data $(g_k, g_{\boldsymbol{\gamma}^k})$. The set of constraints for empirical likelihood given node k is defined as

$$\mathcal{W}_{\beta_{\boldsymbol{\gamma}^k}} = \Big\{\omega : \sum_{i=1}^{n} \omega_i \left(g_k^{(i)} - [g_{\boldsymbol{\gamma}^k}^{(i)}]^T \beta_{\boldsymbol{\gamma}^k}\right) = 0, \sum_{i=1}^{n} \omega_i g_j^{(i)} \left(g_k^{(i)} - [g_{\boldsymbol{\gamma}^k}^{(i)}]^T \beta_{\boldsymbol{\gamma}^k}\right) = 0 \\ j = 4, \ldots, k\Big\} \cap \Delta_{118}. \quad (2.22)$$

Given $\beta_{\boldsymbol{\gamma}^k}$, the empirical likelihood is given by

$$L(\beta_{\boldsymbol{\gamma}^k}) = \prod_{i=1}^{n} \hat{\omega}_i$$

where

$$\hat{\omega} = \underset{\omega \in \mathcal{W}_{\beta_{\boldsymbol{\gamma}^k}}}{\operatorname{argmax}} \sum_{i=1}^{n} \log \omega_i$$

Analysis of gene expression data

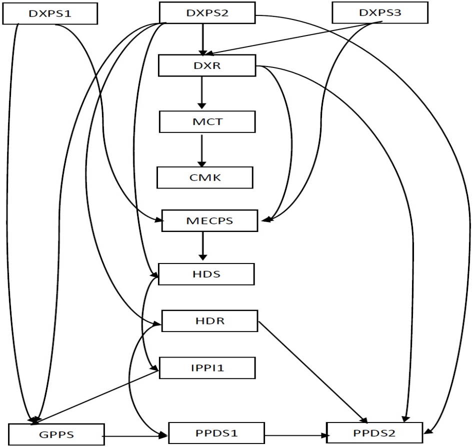

Figure: Graphic Model for gene data

Application to gene expression data (RJMCMC)

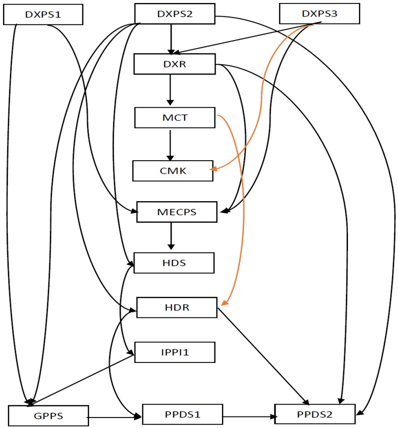

Figure: Graphic Model for gene data

Conclusion and discussion

- The proposed ELDIC has similar form to the classical DIC
- The model with minimum value of ELDIC is selected
- Heuristically, the model with the minimum ELDIC has the smallest posterior predictive risk
- The proposed BayEL based estimate for the effective number of parameters in the model is valid
- Application of ELDIC on the other models remains an open problem
- The magnitude of significant difference between values of ELDICs is still not clear

Thank you !

Bibliography I

[Akaike 1974] AKAIKE, Hirotugu: A new look at the statistical model identification. In: *Automatic Control, IEEE Transactions on* 19 (1974), Nr. 6, S. 716–723

[Drton u. Perlman 2007] DRTON, Mathias ; PERLMAN, Michael D.: Multiple testing and error control in Gaussian graphical model selection. In: *Statistical Science* (2007), S. 430–449

[Gelman u. a. 2003] GELMAN, Andrew ; CARLIN, John B. ; STERN, Hal S. ; RUBIN, Donald B.: *Bayesian data analysis*. Chapman & Hall/CRC, 2003

[Spiegelhalter u. a. 2002] SPIEGELHALTER, David J. ; BEST, Nicola G. ; CARLIN, Bradley P. ; VAN DER LINDE, Angelika: Bayesian measures of model complexity and fit. In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64 (2002), Nr. 4, S. 583–639

[Variyath u. a. 2010] VARIYATH, Asokan M. ; CHEN, Jiahua ; ABRAHAM, Bovas: Empirical likelihood based variable selection. In: *Journal of Statistical Planning and Inference* 140 (2010), Nr. 4, S. 971–981

Bibliography II

[Wille u. a. 2004] WILLE, Anja ; ZIMMERMANN, Philip ; VRANOVÁ, Eva ; FÜRHOLZ, Andreas ; LAULE, Oliver ; BLEULER, Stefan ; HENNIG, Lars ; PRELIC, Amela ; ROHR, Peter von ; THIELE, Lothar u. a.: Sparse graphical Gaussian modeling of the isoprenoid gene network in Arabidopsis thaliana. In: *Genome Biol* 5 (2004), Nr. 11, S. R92