Ein Vergleich von MCMC und numerischer Integration mittels Simulationen bei der Parameterschätzung im Gamma-Frailty-Modell

Diana Pietzner, Oliver Kuß, Andreas Wienke

imebi

Institut für Medizinische Epidemiologie, Biometrie und Informatik
Medizinische Fakultät der Universität Halle-Wittenberg

Contents

- **Model**:
  Cox model,
  shared frailty model
- **Data**:
  HALLUCA study
- **Method:**
  Simulation,
  PROCs NLMIXED, MCMC
- **Results**

Cox model

Proportional Hazard Model (Cox 1972)

$$\mu(t \mid \mathbf{X}) = \mu_0(t)e^{\boldsymbol{\beta}^T \mathbf{X}}$$

- Influence of covariates X on time to event (regression model)
- Parametric or semiparametric model

Cox model

Proportional Hazard Model (Cox 1972)

$$\mu(t \mid \mathbf{X}) = \mu_0(t)e^{\boldsymbol{\beta}^T\mathbf{X}}$$

- Influence of covariates X on time to event (regression model)
- Parametric or semiparametric model
- How to model unobserved heterogeneity?

Frailty models

Proportional Hazard Model (Cox 1972)

$$\mu(t|\mathbf{X}) = \mu_0(t)e^{\boldsymbol{\beta}^T\mathbf{X}}$$

$$\mu(t|\mathbf{X}, Z) = Z\mu_0(t)e^{\boldsymbol{\beta}^T\mathbf{X}}$$

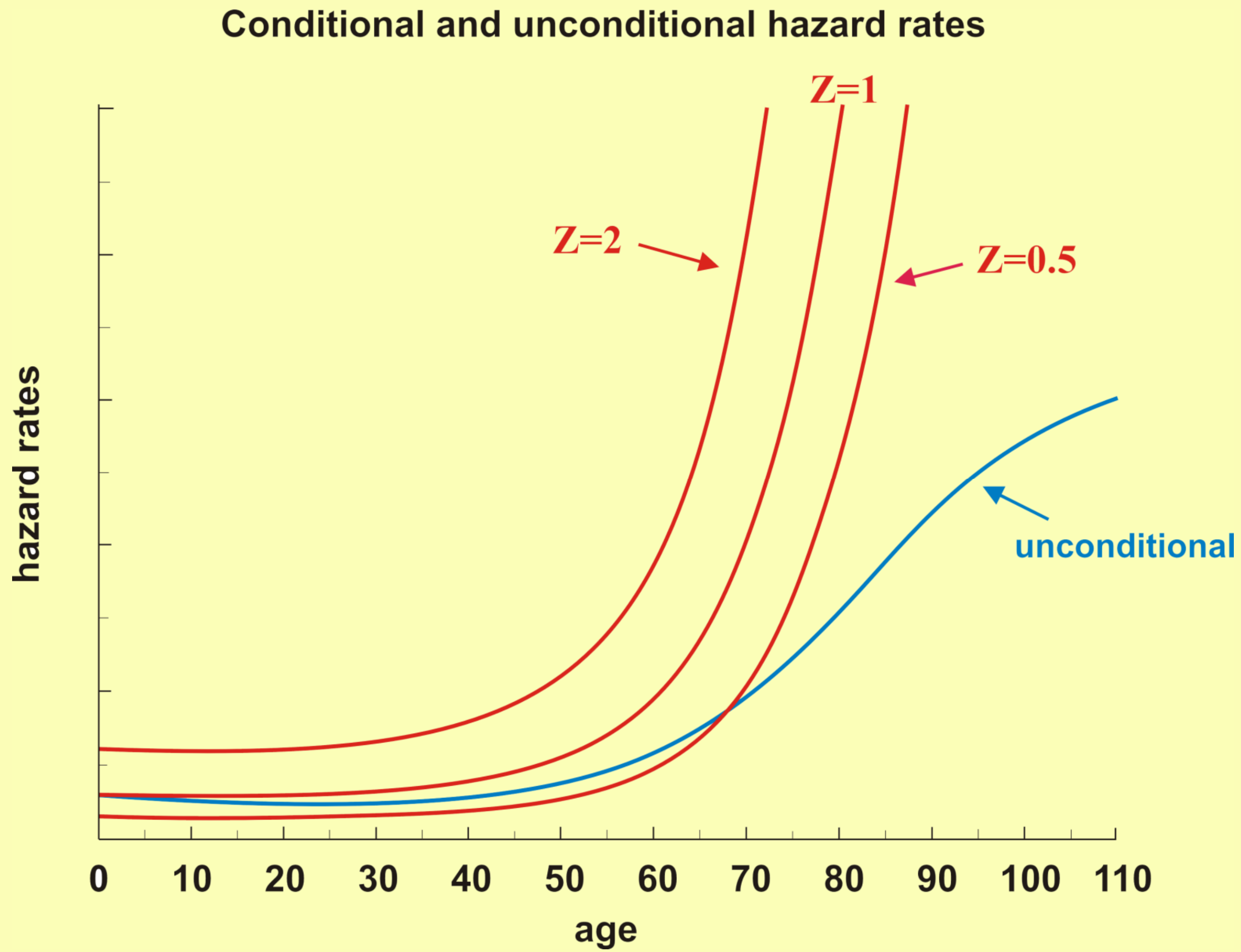
- Z : Frailty
- Vaupel et al. (1979), Lancaster (1979), Beard (1959)

Shared frailty models

- Clayton 1978
- Frailty varies between clusters, identical within cluster
- Frailty is fixed over time
- Frailty is unobservable
- Distribution of Z is specified
- Z is non-negative
- Frailty Z is independent of the observed covariates X
- Different notation: $\mu(t|X,Z) = \mu_0(t)e^{\beta^T X + W}$

$$Z = e^W$$

Shared frailty models



Shared Gamma frailty model

Gamma distributed frailty

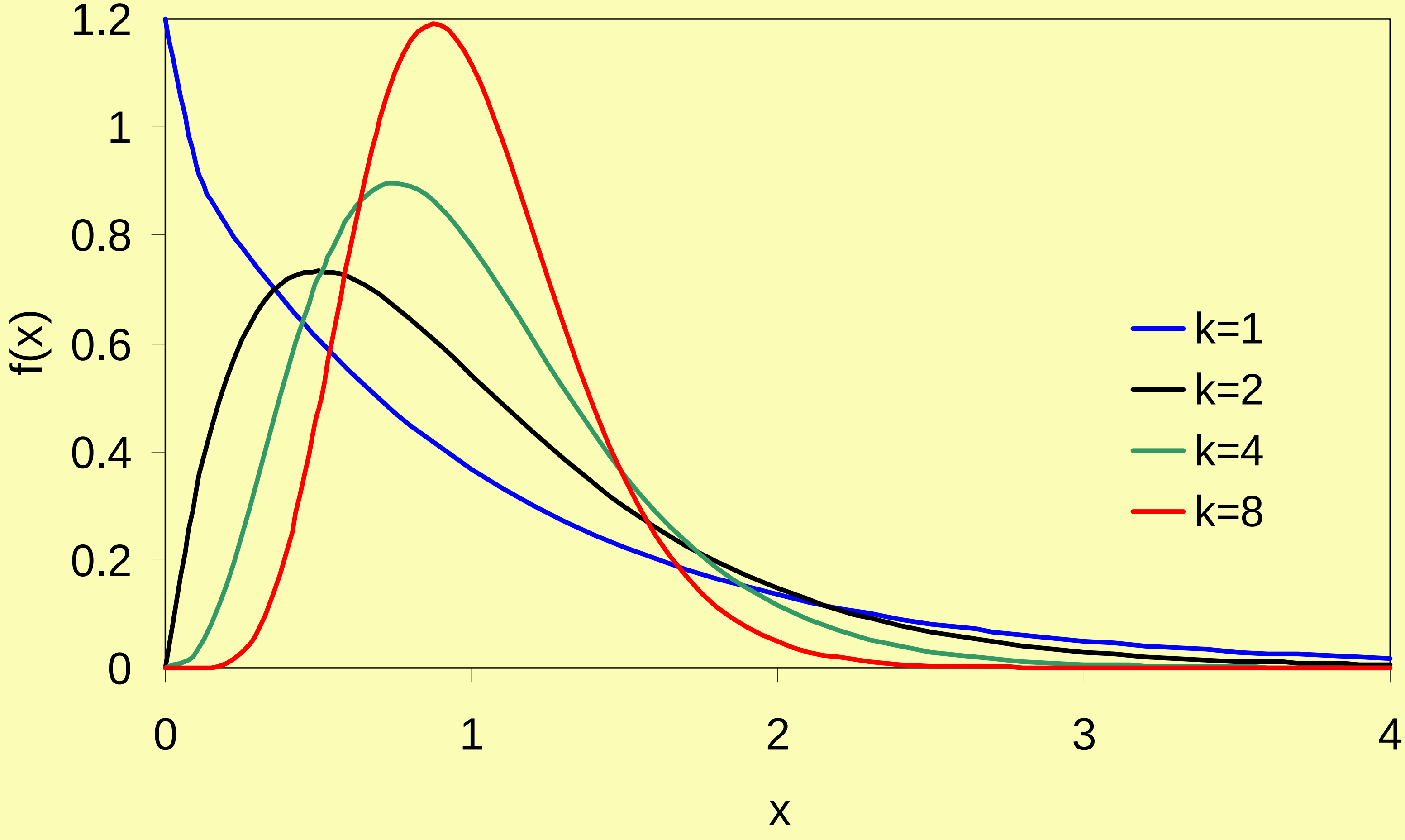
Assumption: $Z \sim \Gamma(k, \lambda)$

$$EZ = \frac{k}{\lambda} \quad V(Z) = \frac{k}{\lambda^2}$$

$$EZ = 1 \quad V(Z) = \frac{1}{\lambda} = \sigma^2$$

$$\Rightarrow Z \sim \Gamma(1/\sigma^2, 1/\sigma^2)$$

Shared Gamma frailty model



Shared frailty models

Shared Gamma frailty model (Clayton 1978)

(without observed covariate X)

For i-th cluster:

$$S(t_1, \dots, t_{n_i} | Z_i) = S_0(t_1)^{Z_i} \dots S_0(t_{n_i})^{Z_i}$$

$$S(t_1, \dots, t_{n_i}) = ES(t_1, \dots, t_{n_i} | Z_i)$$

$$= Ee^{-Z_i (M_0(t_1) + \dots + M_0(t_{n_i}))}$$

$$= \dots$$

$$= (S(t_1)^{-\sigma^2} + \dots + S(t_{n_i})^{-\sigma^2} - (n_i - 1))^{-1/\sigma^2}$$

HALLUCA study

- **Halle Lung Carcinoma**
- Population-based study in Halle and surroundings
- Provision of medical care to lung cancer patients
- April 1996 to September 1999

- 1696 lung cancer patients were identified, in cooperation with the regional cancer registries
- 1349 deaths, medial survival 9.3 months

- Covariates: sex, age, histological type, performance state (ECOG), tumor state, …
- Data from several diagnostic units (=clusters, see Kuss 2008)

Simulation

- Simulation of survival times of 1500 patients
- 25 different clusters
- One continuous covariate X with parameter ß=0.8
- Gompertz baseline hazard:

$$\mu_0(t) = \lambda e^{\varphi t}, \lambda = 0.002, \varphi = 0.002$$

- Gamma frailty with variance parameter $\sigma^2 = 0.5$
- Simulation of survival times by reversing the survival function

$$T_i = M_0^{-1}\left(-(\ln U_i)\frac{1}{Z_i}e^{-\beta X}\right), \quad U_i \sim U[0,1]$$

(see Bender et al. 2005)

NLMIXED: Liu's method

- NLMIXED calculates ML-Estimates for data with normal random effect using Gaussian quadrature
- Liu, Yu 2008:
- Transformation of log-likelihood to

$$l(t_i) = \int_{-\infty}^{\infty} l_i^A + l_i^B + l_i^C \phi(e_i) de_i,$$

$\phi(e_i)$ density of standard normal distribution

- Starting estimators:

$$\beta_0 = 0, \lambda_0 = 0.05, \varphi_0 = 0.05, \sigma_0^2 = 1$$

PROC NLMIXED

```
proc nlmixed data=frailty11 qpoints=50;
    ods output ParameterEstimates=MLEst55;
    parms lambda phi=0.002 to 0.05 by .002 sigmasq=1 beta=0;
     bounds lambda sigmasq >=0;
     basehaz=lambda*exp(phi*time);
     cumbasehaz=lambda/phi*(exp(phi*time)-1);
     expa=exp(a);
     *log Gamma density;
     loggammaden=(1/sigmasq-1)*a - 1/sigmasq * expa - 1/sigmasq* log(sigmasq) –lgamma
    (1/sigmasq);
     lognormalden=-a*a/2; /* log standard normal density */
     mu= a+ beta*x;
     loglik2=-exp(mu) * cumbasehaz;
     if delta=1 then loglik=log(basehaz) + mu + loglik2; /*log likelihood for failure */
     if delta=0 then loglik=loglik2; /*log likelihood for censoring */
     if lastgroup=1 then loglik=loglik + loggammaden + a - lognormalden;
     model time ~ general(loglik);
     random a ~ normal(0,1) subject=group;
     title"Gamma frailty-Modell mit Gompertz baseline (Liu-Methode)";
       by i;
 run;
```

MCMC: Bayesian Estimation

- Let $\theta = (\beta, \lambda, \varphi, \sigma^2)$ vector of unknown parameters
- Prior $p(\theta)$ may contain prior information about the unknown parameters

- **Posterior mean**

$$\hat{\theta} = E(\theta | t) = \int \theta \cdot p(\theta | t) d\theta$$

with $p(\theta | t) = p(\theta) \cdot L(t)$ posterior density and $p(\theta)$ prior, $L(t)$ likelihood

- **Posterior median**

$$\hat{\theta} : P(\hat{\theta}(t) | t) = 0.5$$

MCMC: Bayesian Estimation

- Choice of prior distribution is subjective:
  choose priors with variance $\mathrm{var}(\theta) = 20$ and expectations

$$E_{p(\beta)}(\beta) = 0,\ E_{p(\lambda)}(\lambda) = 0.05,\ E_{p(\varphi)}(\varphi) = 0.05,\ E_{p(\sigma^2)}(\sigma^2) = 1$$

- Prior distribution for β, λ, φ

$$\beta \sim N(0,20)$$

$$\lambda \sim \Gamma((0.05)^2/20, 0.05/20)$$

$$\varphi \sim \Gamma((0.05)^2/20, 0.05/20)$$

- Frailty Z as hyperprior

$$Z \sim \Gamma(1/\sigma^2, 1/\sigma^2)$$

$$\sigma^2 \sim \Gamma(1/20, 1/20)$$

MCMC: Metropolis algorithm

- **Markov chain**: sequence of samples from a target distribution

  $$\theta_1 \leftarrow \theta_2 \leftarrow \theta_3 \leftarrow \ldots \leftarrow \theta_{k-1} \leftarrow \theta_k$$

  each sample depends on the previous one only (memoryless)

- **Monte Carlo** integration:
  approximation of an expectation

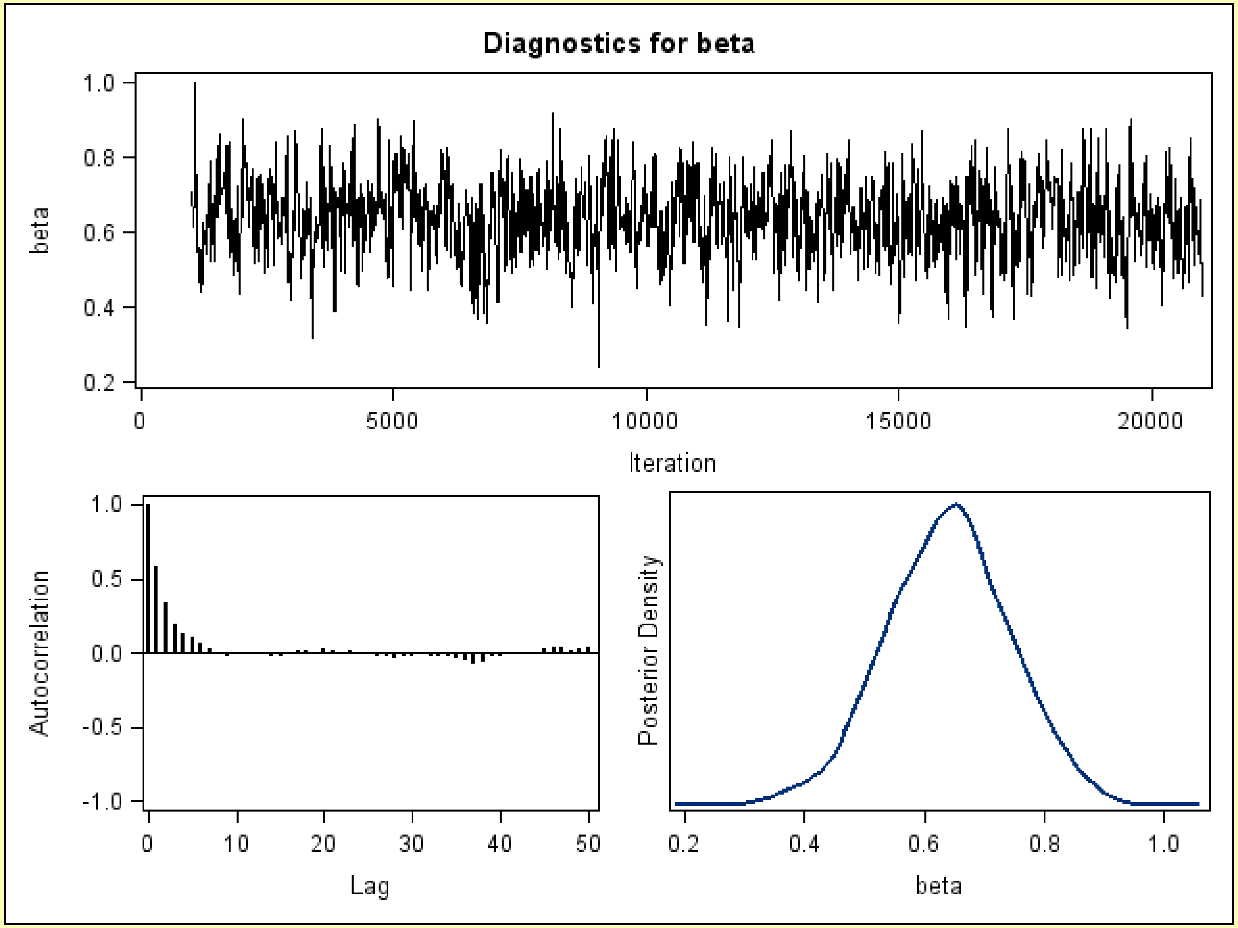
  $$\int \theta \cdot p(\theta | t)\, d\theta = \frac{1}{k} \sum_{i=1}^{k} \theta_i$$

MCMC: assessing convergence

- Trace plots to determine whether Markov chain has reached stationarity
- plots the trace of samples versus the simulation index
- caution: trace plot is necessary but nor sufficient condition for stationarity
- alternative: statistical diagnostic tests

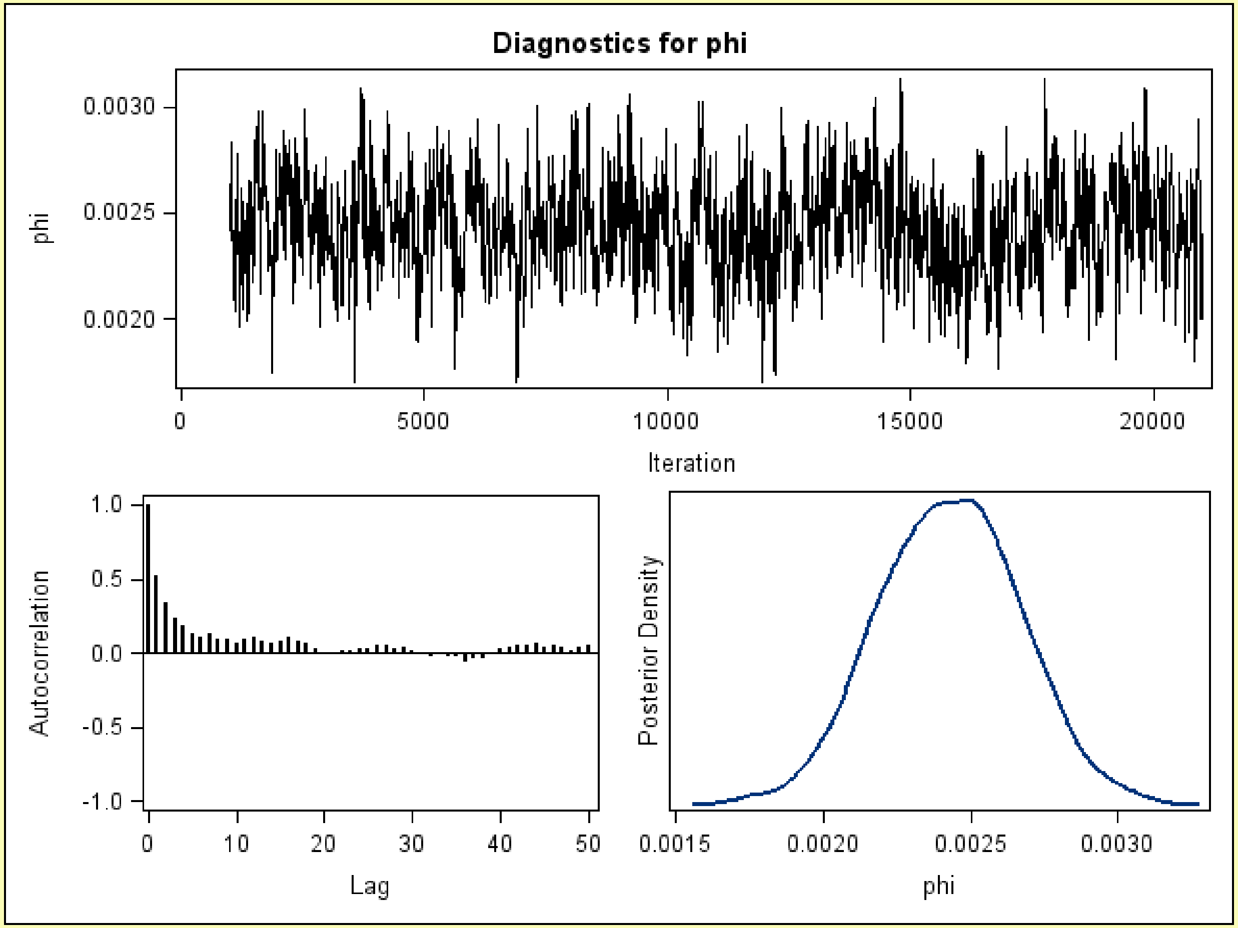
Example: Trace Plot

for parameter β



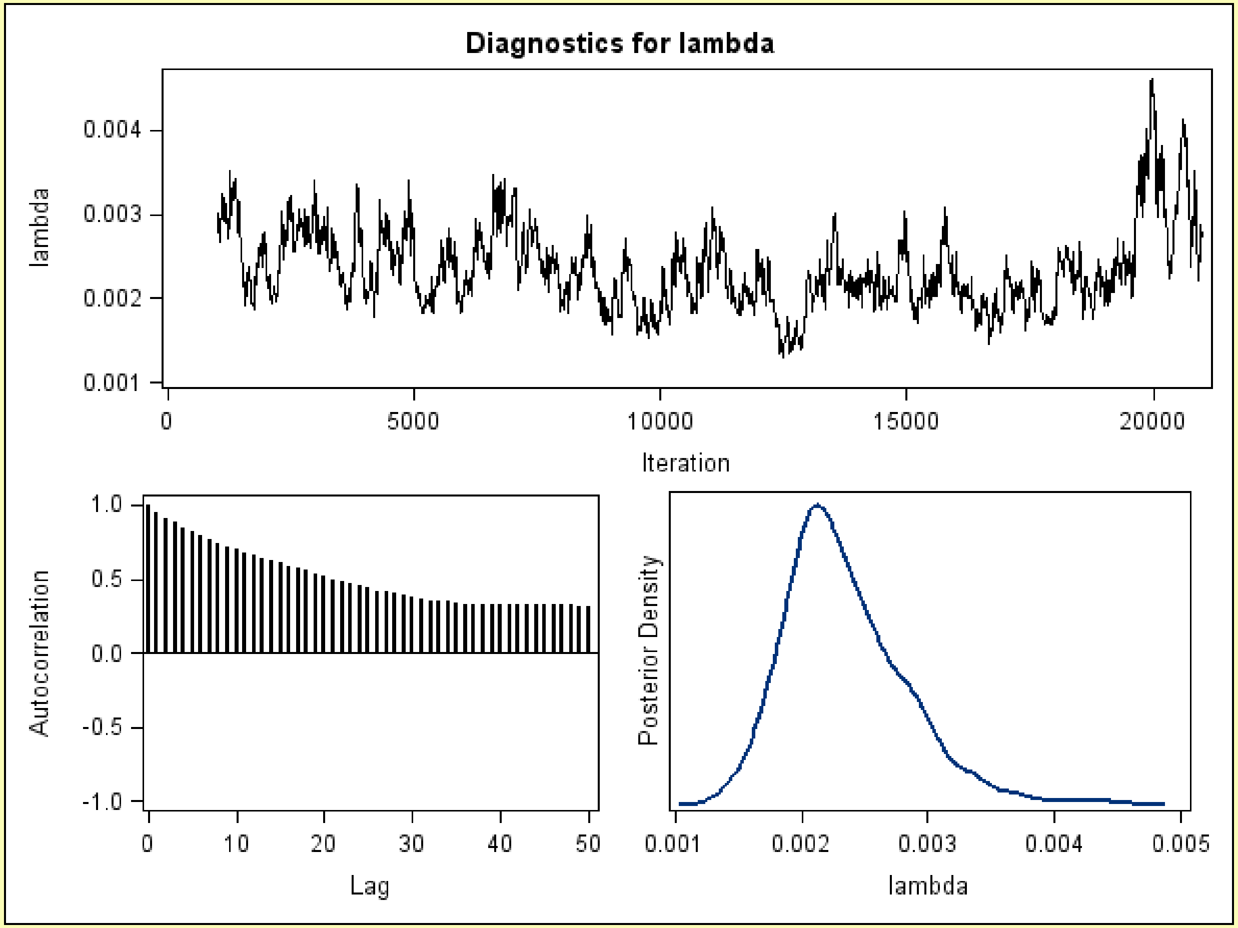
Example: Trace Plot

for parameter φ



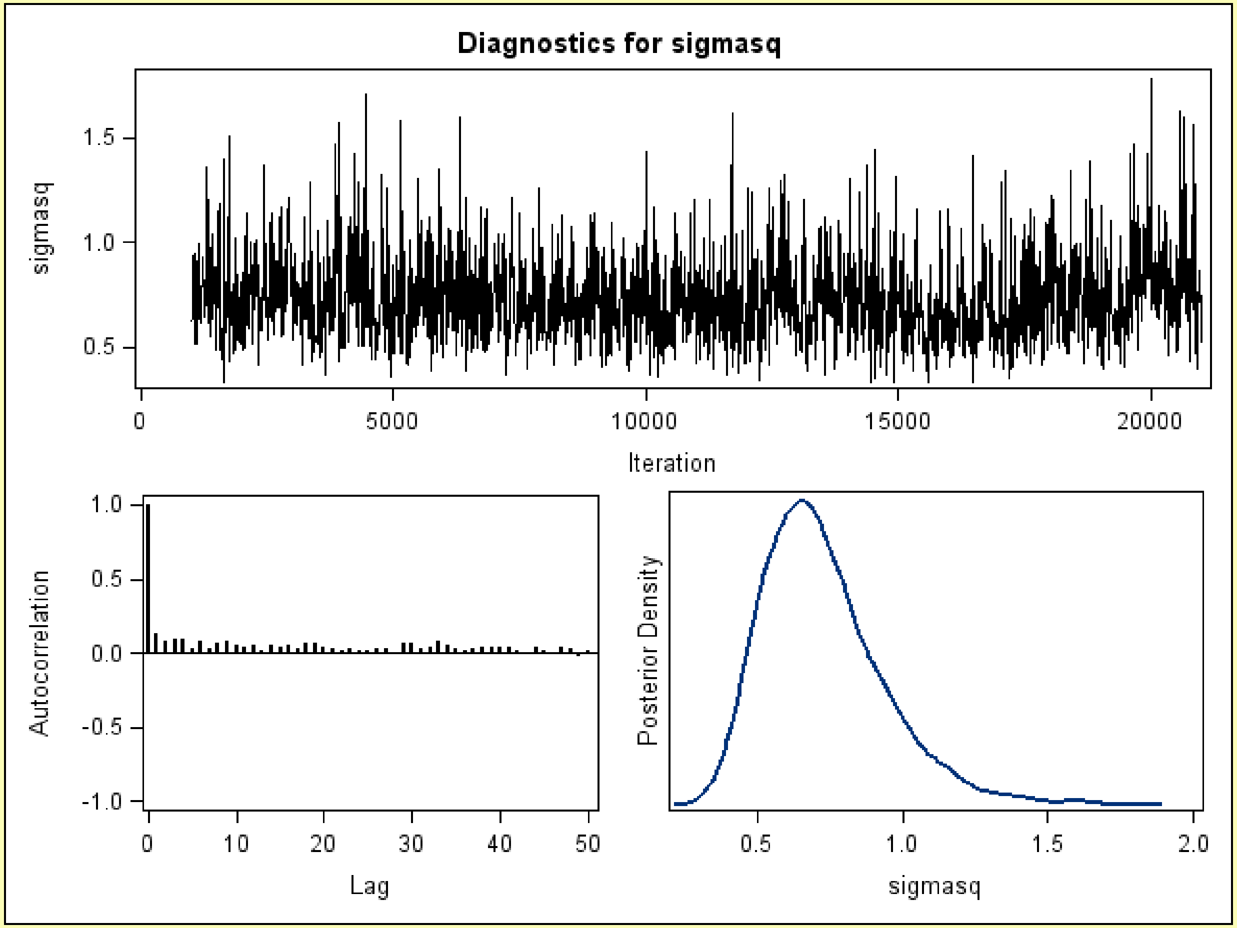
Example: Trace Plot

for parameter λ



Example: Trace Plot

for parameter σ^2



PROC MCMC

```
proc mcmc data=frailty11 seed=332786 nmc=20000 thin=10 nbi=1000 ntu=1000
    monitor=(beta lambda phi sigmasq);
    ods select all;
    ods output PostSummaries=BayesMeanEst11;
    array z[25];
    parms phi .05;
    parms lambda .05;
    parms beta 0;
    parms z: 0;
    parms sigmasq 1;
    prior beta ~ normal(0, var=20);
    prior phi ~ gamma(.05*.05/20,iscale=.05/20);
    prior lambda ~ gamma(.05*.05/20,iscale=.05/20);
    prior z: ~ egamma(1/sigmasq,iscale=1/sigmasq);
    hyperprior sigmasq ~ gamma(1/20,iscale=1/20); /*var=20*/
    t=time;
    bx=beta*x+z[group];
    llike = delta * (log(lambda) + phi *t + bx) - lambda/phi * (exp(phi*t)-1) * exp(bx);
    model t ~ general(llike);
    by i;
run;
```

Mean of Estimators (NLMIXED and MCMC)

Evaluation of 1000 samples of size 1500

comparison of ML-estimates and Bayesian posterior mean estimates

		NLMIXED[1]		MCMC	
	True value	**Mean of Estimator**	**Empirical Standard Deviation**	**Mean of Estimator**	**Empirical Standard Deviation**
β	0.8	0.7967	0.1098	0.7962	0.1075
λ	$0.2 * 10^{-2}$	$0.2017 * 10^{-2}$	$0.03 * 10^{-2}$	$0.2574 * 10^{-2}$	$0.89 * 10^{-2}$
φ	$0.2 * 10^{-2}$	$0.2016 * 10^{-2}$	$0.03 * 10^{-2}$	$0.2002 * 10^{-2}$	$0.02 * 10^{-2}$
σ^2	0.5	0.4870	0.1431	0.5591	0.2964

[1]For 158 samples there was no convergence (NLMIXED)

Summary

MCMC	NLMIXED
•slow •problems in estimating λ •convergence for all 1000 simulations	•faster •better estimators •in some cases no convergence
•how to choose priors? •which prior information is available?	•how to find starting estimators?
•problems when there are more/very little clusters (uncatched exception)	

Conclusion

- 2 powerful methods to handle parametric frailty models in SAS
- problems concerning convergence, speed and provision of prior information

Literature

Beard, R.E. (1959) Note on some mathematical mortality models. In: G.E.W.Wolstenholme and M.O'Connor (eds.).The Lifespan of Animals. Ciba Foundation Colloquium on Ageing. Little, Brown, Boston, 302-311

Bender R, Augustin T, Blettner M. (2005) Generating survival times to simulate Cox proportional hazards models. *Stat. Med.*; 24:1713–1723.

Bollmann A, Blankenburg T, Haerting J, Kuss O, Schutte W, Dunst J, Neef H. (2004) Survival of patients in clinical stages I-IIIb of non-small-cell lung cancer treated with radiation therapy alone. Results of a population-based study in Southern Saxony-Anhalt. *Strahlenther. Onkol.*; 180:488–496.

Clayton, D.G. (1978) A model for association in bivariate life tables and its application in epidemiologicalstudies of familial tendency in chronic disease incidence. Biometrika 65, 141 - 151

Cox, D.R. (1972) Regression models and life tables. JRSS (B) 34, 187 – 220

Kuss O, Blankenburg T, Haerting J (2008) A Relative Survival Model for Clustered Responses. Biom J 50; 408 – 418

Lancaster, T. (1979) Econometric methods for the duration of unemployment. Econometrica 47,939 – 956

Liu, L.,Yu,Z. (2008) A likelihood reformulation method in non-normal random effects models. *Statist. Med.* 2008; **27**:3105–3124

Vaupel, J.W., Manton, K.G., Stallard, E. (1979) The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography 16, 439 – 454