Hierarchical Bayesian Overdispersion Models for Non-Gaussian Repeated Measurement Data

Aregay Mehreteab

I-BioStat, KULeuven, Belgium

Bayes2013

May 23, 2013

Outline

- introduction
- modeling issues
- application to data
- simulation study
- concluding remarks and further research

A Clinical Trial of Epileptic Seizures

- a double-blind, parallel group multi-center study
- 59 patients were randomized to either antiepileptic drug progabide or to placebo
- follow-up over four successive two week periods
- the number of seizures experienced during the last week
- **Objective**: Reduction in the number of seizures by the treatment

A Case Study in Onychomycosis

- treatment of toenail dermatophyte onychomycosis (TDO) over 12 weeks
- a randomized, double-blind, parallel group, multi-center study
- two oral treatments (in what follows represented as A and B) were compared
- outcomes were recorded from baseline onwards up to 48 weeks
- sample to 146 and 148 subjects for groups A and B, respectively
- severity of infection
- percentage of severe infection decreases

HIV Study

- concerned with diagnostic tests
- information about the prevalence of HIV infection in injecting drug users (IDUs)
- study took place in the 20 Italian regions, in the time frame 1998–2006
- reported by the European Monitoring Center for Drugs and Drug Addiction
- for an elaborate discussion, we refer to Del Fava *et al.* (2011)

Recurrent Asthma Attacks in Children

- a new application anti-allergic drug was given to children who are at a higher risk to develop asthma
- the children were randomly assigned to either drug or placebo
- time between the end of the previous event (asthma attack) and the start of the next event
- the different events are clustered within a subject and ordered over time
- for detail see Duchateau and Janssen (2007) and Molenberghs *et al.* (2010)

Kidney Data Set

- the data were studied in McGilchrist and Aisbett (1991)
- response: time to first and second recurrence of infection, at the point of insertion of catheters
- observation is censored when catheters are removed, other than for reasons of infection
- 38 kidney patients in the study and each subject contributes two observations

Objectives

- to generalize the additive model to the exponential family
- compare the additive to the multiplicative combined model
- impact of misspecification of the GLM and GLMM for hierarchical and overdispersed data

Poisson Multiplicative Model for the Epilepsy Data Set

- accommodates both overdispersion and clustering simultaneously
- Y_{ij}: number of epileptic seizures experienced for patient i during week j
- **Likelihood:**
- $Y_{ij}|b_i, \theta_{ij} \sim \text{Poisson}(\theta_{ij}\kappa_{ij})$,
  $\log(\kappa_{ij}) =$
  $\beta_0 + \beta_{\text{Base}} \cdot \text{lbase}_i + \beta_{\text{Age}} \cdot \text{lage}_i + \beta_{\text{Trt}} \cdot T_i + \beta_{V_4} \cdot V_{4j} + \beta_{BT} \cdot T_i \cdot \text{lbase}_i + b_i$

Multiplicative Model: Bayesian Formulation

- **Prior and hyper-priors:**
  - an independent diffuse normal priors $\beta \sim \mathrm{N}(0; 100000)$
  - $\theta_{ij} \sim \mathrm{Gamma}(\alpha, \beta)$
  - $\beta{=}\alpha$
  - a uniform prior distribution $U(0, 100)$ was considered for α
  - $b_i \sim N(0, \sigma_b^2)$; $\sigma_b^{-2} \sim G(0.01, 0.01)$
- to improve convergence, all of the covariates, were centered about their mean (Breslow and Clayton 1993 and Thall and Vail 1990)

Bernoulli Multiplicative Model for the Onychomycosis Study

- Y_{ij} be the jth binary response for subject i coded as 1 for severe infection and 0 otherwise
- **Likelihood:**
- $Y_{ij}|b_i, \theta_{ij} \sim \text{Bernoulli}(\pi_{ij} = \theta_{ij}\kappa_{ij})$,
  $\text{logit}(\kappa_{ij}) = \beta_1 T_i + \beta_2(1 - T_i) + \beta_3 T_i t_{ij} + \beta_4(1 - T_i)t_{ij} + b_i$,
- $\theta_{ij} \sim \text{Beta}(\alpha, \beta)$, $b_i \sim N(0, \sigma_b^2)$
- $\alpha \sim U(0, 100)$ and $\beta \sim \text{U}(0, 100)$

Binomial Multiplicative Model for the HIV Study

- **Likelihood:**
- $Y_{ij}|b_i, \theta_{ij} \sim \text{Binomial}(\pi_{ij} = \theta_{ij}\kappa_{ij}, m_{ij})$, $\text{logit}(\kappa_{ij}) = \beta_j + b_i$
- Y_{ij} is the event for subject i at time j,
- π_{ij} is the prevalence and m_{ij} is the number of trials
- $\theta_{ij} \sim \text{Beta}(\alpha, \beta)$, $b_i \sim N(\beta_0, \sigma_b^2)$
- $\alpha \sim U(1, 100)$ and $\beta = \alpha$

Weibull Multiplicative Model for the Asthma and Kidney Data

- Y_{ij} is the time at risk for a particular asthma attack
- **Likelihood:**
- $Y_{ij}|b_i, \theta_{ij} \sim \text{Weibull}(r, \theta_{ij}\kappa_{ij})$,
  $\log(\kappa_{ij}) = \beta_0 + \beta_1 T_i + b_i$
- **Kidney data set:**
  - Y_{ij} is the time to first and second recurrence of infection in kidney patients on dialysis
  - $Y_{ij}|b_i, \theta_{ij} \sim \text{Weibull}(r, \theta_{ij}\kappa_{ij})$,
    $\log(\kappa_{ij}) = \beta_0 + \beta_1 \cdot \text{age}_{ij} + \beta_2 \cdot \text{sex}_i + \beta_3 \cdot D_{i1} + \beta_4 \cdot D_{i2} + \beta_5 \cdot D_{i3} + b_i$
  - we used a truncated Weibull for censored observations and $r = 1$

Additive Model

- **Why:**
  - failure to converge and computationally expensive for multiplicative model
  - to expand the modeler's toolkit, and for quality of fit
- the general family is the same as in the multiplicative, except that the mean now is:
- $\eta_{ij} = h(\mu_{ij}^a) = h[E(Y_{ij}|\boldsymbol{b_i}, \beta)] = \boldsymbol{x_{ij}}'\beta + \boldsymbol{z_{ij}}'\boldsymbol{b_i} + \theta_{ij}$
- the difference is on the specification of the overdispersion random effect θ_{ij}
- $\theta_{ij} \sim N(0, \sigma_\theta^2)$ and $\sigma_\theta^{-2} \sim G(0.01, 0.01)$
- more generally in terms of assuming a normal distribution for θ_{ij} throughout the exponential family

Multiplicative Vs Additive Models

- both additive and multiplicative models allow two separate random effects
- the first one captures subject heterogeneity and a certain amount of overdispersion
- the second one is for the remaining extra-model-variability
- **Binary and Binomial Data:**
  - the multiplicative effect cannot be absorbed into the linear predictor
  - because the logit and probit links do not allow for this

Multiplicative Vs Additive Models

- **For time-to-event and count data:**
  - the link function is logarithmic
  - the multiplicative effect could also be absorbed into the linear predictor
  - affects the intercept but not the other parameters
  - the transformed gamma effect is reasonably symmetric
  - the difference between the multiplicative and additive models may be relatively small

Model fitting

- Bayesian approach using MCMC through R2WinBUGS package
- three chains of 100,000 iterations, a 10,000 burn-in period, and 100 thinning
- convergence was checked using trace plots and estimated potential scale reduction factors, $\widehat{R}$
- **Model selection**: Deviance Information Criteria (DIC)

Epilepsy Data: Posterior Summary Statistics

	GLM		Multiplicative w/o b_i		Additive w/o b_i		GLMM		Multiplicative with b_i		Additive with b_i	
Par.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.
β_0	-2.73	(-3.52,-1.91)	-1.50	(-3.10,0.11)	-1.78	(-3.37,-0.18)	-1.31	(-3.73,1.17)	-1.42	(-3.84,0.99)	-1.28	(-3.73,1.22)
β_{Base}	0.95	(0.87,1.03)	0.90	(0.74,1.08)	0.91	(0.74,1.00)	0.88	(0.59,1.15)	0.88	(0.60,1.17)	0.88	(0.62,1.16)
$\beta_{\mathbf{Age}}$	0.89	(0.66,1.11)	0.55	(0.07,1.04)	0.58	(0.12,1.05)	0.48	(-0.25,1.19)	0.49	(-0.22,1.19)	0.47	(-0.26,1.18)
β_{Trt}	-1.34	(-1.64,-1.04)	-0.91	(-1.47,-0.38)	-0.97	(-1.52,-0.41)	-0.95	(-1.79,-0.17)	-0.94	(-1.77,-0.10)	-0.93	(-1.80,-0.09)
β_{V_4}	-0.16	(-0.27,-0.05)	-0.14	(-0.36,0.08)	-0.09	(-0.32,0.14)	-0.16	(-0.27,0.05)	-0.10	(-0.28,0.07)	-0.12	(-0.28,0.05)
$\beta_{\mathbf{BT}}$	0.56	(0.44,0.69)	0.35	(0.09,0.62)	0.37	(0.10,0.65)	0.35	(-0.06,0.79)	0.34	(-0.10,0.77)	0.34	(-0.09,0.77)
σ_b							0.54	(0.43,0.68)	0.50	(0.37,0.65)	0.51	(0.38,0.65)
σ_θ					0.60	(0.51,0.69)					0.36	(0.29,0.45)
α			2.75	(2.04,3.63)					8.10	(4.95,13.37)		
DIC	1646.98		1168.11		1181.17		1271.62		1152.91		1157.29	

- in all models, the treatment is found to be significant

Onychomycosis Data: Posterior Summary Statistics

	GLM		Multiplicative w/o b_i		Additive w/o b_i		GLMM		Multiplicative with b_i		Additive with b_i	
Par.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.
β_1	-0.53	(-0.75,-0.31)	-0.42	(-0.64,-0.19)	-0.60	(-0.94,-0.34)	-1.80	(-2.74,-0.93)	-1.80	(-2.92,-0.83)	-1.83	(-2.85,-0.94)
β_2	-0.56	(-0.77,-0.34)	-0.44	(-0.67,-0.21)	-0.62	(-0.96,-0.36)	-1.66	(-2.58,-0.83)	-1.64	(-2.77,-0.59)	-1.71	(-2.70,-0.85)
β_3	-0.26	(-0.32,-0.20)	-0.26	(-0.33,-0.20)	-0.27	(-0.36,-0.20)	-0.57	(-0.70,-0.46)	-0.74	(-1.05,-0.51)	-0.58	(-0.71,-0.47)
β_4	-0.18	(-0.23,-0.13)	-0.18	(-0.23,-0.13)	-0.19	(-0.26,-0.14)	-0.41	(-0.51,-0.32)	-0.45	(-0.57,-0.35)	-0.42	(-0.52,-0.33)
σ_b							4.14	(3.41,5.00)	4.93	(3.80,6.40)	4.21	(3.49,5.06)
σ_θ					0.56	(0.08,1.80)					0.26	(0.07,0.63)
α/β			13.55	(9.81,19.27)					17.53	(12.27,23.85)		
β	0.56	(0.44,0.69)	0.35	(0.09,0.62)	0.37	(0.10,0.65)	0.35	(-0.06,0.79)	0.34	(-0.09,0.77)	0.34	(-0.10,0.77)
DIC	1819.69		1819.89		1831.79		955.524		947.57		953.60	

- the DIC values for the GLMM, multiplicative and additive models with clustering random effects models are similar
- in all models, the evolution of the treatment and placebo group over time was significant

HIV Data: Posterior Summary Statistics

	GLM		Multiplicative w/o b_i		Additive w/o b_i		GLMM		Multiplicative with b_i		Additive with b_i	
Par.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.
β_0	-1.83	(-1.85,-1.81)	0.30	(0.09,0.66)	-1.98	(-2.37,-1.58)	-2.13	(-2.44,-1.82)	-1.09	(-1.59,-0.60)	-2.03	(-2.47,-1.62)
β_1	0.17	(0.14,0.19)	-1.16	(-1.55,-0.85)	-0.15	(-0.74,0.41)	0.02	(-0.01,0.05)	-0.06	(-0.26,0.13)	-0.10	(-0.26,0.06)
β_2	0.11	(0.84,0.14)	-1.21	(-1.59,-0.90)	-0.18	(-0.06,0.01)	-0.03	(-0.06,0.01)	-0.09	(-0.29,0.10)	-0.15	(-0.31,0.02)
β_3	0.15	(0.12,0.18)	-0.95	(-1.35,-0.64)	-0.11	(-0.68,0.42)	0.043	(0.01,0.07)	-0.11	(-0.31,0.83)	-0.09	(-0.26,0.08)
β_4	0.08	(0.05,0.11)	-0.89	(-1.29,-0.57)	-0.10	(-0.67,0.45)	-0.01	(-0.04,0.03)	-0.11	(-0.31,0.08)	-0.10	(-0.26,0.06)
β_5	0.072	(0.04,0.10)	-0.96	(-1.36,-0.65)	-0.17	(-0.71,0.35)	-0.01	(-0.04,0.02)	-0.21	(-0.41,-0.03)	-0.15	(-0.32,0.01)
β_6	0.03	(-0.01,0.06)	-0.97	(-1.37,-0.65)	-0.19	(-0.76,0.37)	-0.04	(-0.07,-0.01)	-0.23	(-0.44,-0.04)	-0.17	(-0.33,-0.01)
β_7	-0.01	(-0.03,0.03)	-0.88	(-1.29,-0.55)	-0.18	(-0.73,0.34)	-0.22	(-0.42,-0.03)	-0.29	(-0.53,-0.08)	-0.16	(-0.33,0.01)
β_8	-0.01	(-0.03,0.03)	-0.68	(-1.10,-0.32)	-0.23	(-0.76,0.34)	-0.08	(-0.11,-0.04)	-0.27	(-0.46,-0.08)	-0.19	(-0.37,-0.03)
σ_b							0.87	(0.64,1.22)	1.08	(0.78,1.52)	0.88	(0.64,1.23)
σ_θ					0.87	(0.78,0.97)					0.25	(0.22,0.28)
α			1.14	(1.01,1.34)					13.19	(9.99,17.05)		
DIC	45576.50		1612.09		1614.61		3816.21		1595.95		1597.27	

- as expected, the 95% credible interval obtained from the GLM are narrower than those obtained from the other models

Asthma attack study: Posterior Summary Statistics

	GLM		Multiplicative w/o b_i		Additive w/o b_i		GLMM		Multiplicative with b_i		Additive with b_i	
Par.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.
β_0	-4.26	(-4.32,-4.19)	-3.94	(-4.03,-3.83)	-4.06	(-4.15,-3.96)	-4.36	(-4.48,-4.25)	-4.22	(-4.37,-4.07)	-4.26	(-4.39,-4.13)
β_1	-0.10	(-0.18,-0.01)	-0.08	(-0.20,0.04)	-0.08	(-0.20,0.05)	-0.10	(-0.26,0.07)	-0.09	(-0.26,0.08)	-0.09	(-0.27,0.08)
σ_b							0.50	(0.43,0.58)	0.48	(0.40,0.56)	0.47	(0.39,0.56)
σ_θ					0.68	(0.59,0.76)					0.44	(0.31,0.56)
α			3.42	(2.71,4.32)					9.15	(4.87,20.82)		
DIC	18679		18638		18551		18556		18519		18490	

- using a GLM model for these data will lead to a significant effect of the treatment while the other models prove insignificant for treatment effect

Kidney Study: Posterior Summary Statistics

	GLM		Multiplicative w/o b_i		Additive w/o b_i		GLMM		Multiplicative with b_i		Additive with b_i	
Par.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.	Mean	Cred. I.
β_0	-3.79	(-4.82,-2.85)	-3.77	(-4.81,-2.78)	-3.77	(-4.83,-2.78)	-3.76	(-4.92,-2.61)	-3.73	(-4.92,-2.70)	-3.76	(-4.92,-2.65)
β_1	0.00	(-0.02,0.03)	0.00	(-0.02,0.03)	0.00	(-0.02 ,0.03)	0.00	(-0.02,0.03)	0.00	(-0.02,0.03)	0.00	(-0.02,0.03)
β_2	0.04	(-0.75,0.82)	0.06	(-0.75,0.85)	0.12	(-0.78,1.02)	0.11	(-0.84,1.11)	0.12	(-0.83,1.05)	0.16	(-0.86,1.17)
β_3	0.52	(-0.26,1.31)	0.50	(-0.30,1.27)	0.50	(-0.39,1.35)	0.52	(-0.41,1.45)	0.53	(-0.45,1.51)	0.51	(-0.49,1.47)
β_4	-1.37	(-2.55,-0.26)	-1.31	(-2.56,-0.16)	-1.2	(-2.52,0.10)	-1.06	(-2.48,0.40)	-1.03	(-2.47,0.45)	-1.02	(-2.47,0.45)
β_5	-1.59	(-2.24,-0.89)	-1.60	(-2.25,-0.92)	-1.62	(-2.31,-0.89)	-1.63	(-2.41,-0.85)	-1.63	(-2.41,-0.84)	-1.63	(-2.40,-0.82)
σ_b							0.46	(0.03,0.96)	0.44	(0.02,0.94)	0.40	(0.02,0.94)
σ_θ											0.35	(0.01,0.84)
α			48.68	(4.45,98.05)					51.34	(5.60,97.77)		
DIC	672.78		672.21		671.24		671.56		671.56		671.74	

- all the models perform similarly

Simulation Study: Motivation

- to investigate the performance of the two models in terms of
  - computation time
  - parameter estimation
  - 95% coverage probability and DIC values
- to study the impact of misspecification of the GLM and GLMM models

Exponential Model for Time-to-event Data

- we simulated data according to both models
- $\beta_0 = -4.36, \beta_1 = -0.098$
- different level of clustering and overdsipersion was considered
- the shape parameter $r = 1$
- sample size and cluster size were equal to 60 and 10
- 100 datasets, from both additive and multiplicative model
- GLM, GLMM, additive and multiplicative models were fitted
- the bias, MSE, 95% coverage probability, DIC values and computation time were calculated

- **For high and moderate overdispersion:**
  - misspecification of the GLM leads to invalid inference of the intercept and the slope
  - misspecification of the GLMM leads to invalid inference of the intercept and σ_b
  - misspecifcation of the GLMM does not cause serious flaws in inference for the slope
- **For low overdispersion:**
  - misspecification of the GLM and GLMM does not affect estimation and inference
  - as σ_b increases, the impact of misspecification of the GLM increases
- there is a difference between the additive and multiplicative models in the estimation and inference of the intercept

Bernoulli Model for Binary Data

- $Y_{ij}|b_i, \theta_{ij} \sim \text{Binomial}(\pi_{ij}, m_{ij} = 1)$.
- $\beta_1 = -1.804$, $\beta_2 = -1.659$,$\beta_3 = -0.574$, and $\beta_4 = -0.411$
- covariates: time and treatment
- we considered a sufficiently large sample size with 300 subjects, each of them measured at 10 time points
- one hundred data sets were generated and the GLM, GLMM, additive, and multiplicative models were fitted
- the bias, MSE, 95% coverage probability, DIC values and computation time were calculated

- **Additive overdispersion:**
  - for high overdispersion, misspecification of the GLM causes serious flaws in inference for all parameters and
  - misspecification of the GLMM produces invalid inferences for all parameters
  - for moderate overdispersion, it only affects the intercept
  - neither the intercept nor the slope for low overdispersion
  - for moderate and low overdispersion, misspecification of the GLMM does not affect the inference of the parameters, except for the between subject variation
  - for high overdispersion, using the additive or multiplicative model affects the inference about all of the parameters

- **Multiplicative overdispersion:**
  - misspecification of the GLM affects only the inference of intercepts but not for the slopes
  - misspecification of the GLMM causes flaws in inference for the intercepts and σ_b
- there is a difference between the additive and multiplicative models in the estimation and inference of the intercept

Binomial Model

- similar to the Bernoulli case except now $Y_{ij}|b_i, \theta_{ij} \sim \text{Binomial}(\pi_{ij}, m_{ij} = 20)$
- for convenience, we assumed the number of trials to be fixed for all observations
- the sample size and cluster size were equal to 60 and 10

- **Additive overdispersion:**
  - for high overdispersion,misspecification of the GLM and GLMM leads to invalid inferences of the parameters and
  - as the overdispersion level decreases, the impact of misspecification of these two models reduces
- **Multiplicative overdispersion:**
  - misspecification of the GLM affects only the inference of intercepts but not for the slopes
  - misspecification of the GLMM causes flaws in inference for the intercepts and σ_b
- the additive and the multiplicative models perform similarly
- except that there are some differences in the estimation and inferences of the intercepts and σ_b

Computation Time

	Fitting model					
	Weibull		Bernoulli		Binomial	
Generating model	Add	Mult	Add	Mult	Add	Mult
Additive	55:30:07	72:20:08	83:50:53	149:18:21	108:18:55	150:33:09
Multiplicative	60:31:15	63:39:01	78:56:16	152:33:01	109:51:13	149:34:18

- in all scenarios, the additive model converges faster than the multiplicative model, especially for binary data

Concluding Remarks

- **misspecification of the GLM**:
  - causes serious flaws in inference
- **misspecification of the GLMM**:
  - does not strongly affect inferences of the slopes in time-to-event outcomes
  - it does so for binary and binomial hierarchical data with high overdispersion

Concluding Remarks

- the Bayesian approach converged well for some data sets, i.e., the HIV and onychomycosis studies
- difficulties were encountered with a likelihood approach implemented in the SAS procedure NLMIXED, for the multiplicative model
- the multiplicative exhibits more convergence issues and, higher computational expense
- both models can be used as useful alternatives

Further Research

- to explore the effect of sample size and cluster size, especially for the binary data.

Thank You!