2017 BDA II UHasselt

Bayesian Frailty and Non-Linear Model

Group members: Name(Email)
Ezgi Tanriver Ayder (ezgi.tanriver@student.uhasselt.be
Oana Petrof (oana.petrof@student.uhasselt.be)
Olusoji Oluwafemi Daniel (oluwafemi.olusoji@student.uhasselt.be)
Owokotomo Olajumoke Evangelina (olajumoke.owokotomo@student.uhasselt.be)
Julie Schenk (julie.schenk@student.uhasselt.be)

Submission Date: 17/04/2017

Lecturer:
Prof. dr. Emmanuel Lesaffre

Analysis 1: A Bayesian frailty model

The data used in this section involves an experiment where 100 cows are followed up for mastitis infection and their time of infection was measured. Each cow has four udder quarters (the location of the udder. 1=left-front, 2=left-rear, 3=right-front, and 4=right-rear) and as such, there is possible correlation between the infection times of the udder quarters. Another covariate of interest was parity which was dichotomised into primiparous cows (heifer=1) and multiparous cows (heifer=0). Censoring status was also recorded to indicate experiencing infection (1) or not(0). Based on the nature of the data and response, a frailty model which allows for a shared parameter (random effect) was used. Assuming that time to infection follows a log-normal distribution, we have;

$$log(t_{ij}) = \beta^T x_{ij} + b_i + \epsilon_{ij} \qquad i = 1, ..., 100 \quad j = 1, ..., 4 \quad \epsilon_{ij} \sim N(0, \sigma^2) \quad b_i \sim N(0, \sigma_b) \quad (1)$$

The following proof was useful in writing the codes: If $\epsilon_{ij} \sim N(0, \sigma^2)$ then $(\epsilon_{ij}/\sigma) \sim N(0, 1)$

$$\begin{aligned} log(t_{ij}) &= \beta_0 + \beta_1 * heifer_{ij} + \beta_2 * Q1_{ij} + \beta_3 * Q2_{ij} + \beta_4 * Q3_{ij} + b_i + \epsilon_{ij} \\ \lambda_{ij} &= \beta_0 + \beta_1 * heifer_{ij} + \beta_2 * Q1_{ij} + \beta_3 * Q2_{ij} + \beta_4 * Q3_{ij} + b_i \\ \epsilon_{ij} &= log(t_{ij}) - \lambda_{ij} \quad \text{and} \quad \epsilon'_{ij} = (log(t_{ij}) - \lambda_{ij})/\sigma \end{aligned} \quad (2)$$

$f(.)$ and $f_0(.)$ denote the density function respectively and let $S(.)$ and $S_0(.)$ denote the survival function of t and ϵ respectively, δij for censoring indicator and ll for loglikelihood

$$\begin{aligned} f_0(.) &= \frac{1}{\sqrt{2\pi}} e^{\frac{\epsilon'^2_{ij}}{2}} \quad \text{and} \quad S_0(.) = 1 - CDF(\epsilon'_{ij}) \\ ll &= \prod_{i=1}^{100} \prod_{j=1}^{4} \delta ij * \frac{log(f_0(.)}{tij * \sigma} + (1 - \delta ij) * log(S_0(.)) \end{aligned} \quad (3)$$

R2jags package and CODA in R software was used to run the models which enables jags to be called within R, allowing for better pre and post-processing in R.

1. Take three chains and assess the convergence of the chains. Does hierarchical centering improve the rate of convergence? Assess this with appropriate diagnostics and the effective sample size

A uniform prior distribution, U(0,100), was used for the within(σ_ϵ) and between(σ_b) cow variability. A vague normal distribution, $\beta_i \sim N(0, 1.0E-3)$ was used as prior for the fixed effects (which is random in Bayesian framework). Three chains were initiated with 50,000 iterations and burn-in of 25000. The use of multiple chains can reveal simple problems faster, and as well prevent the chain from getting stuck for a long time in an area around a local mode. The initial 25000 iterations were used as burn-in to discard the influence of the initial values on the sampling algorithm. To have an idea on how fast the chain converges and how well they explore the posterior, graphical procedures such as trace plots and autocorrelation function plot were used. Brooks-Gelman-Rubin (BGR) was used as a formal diagnostic tool,

since graphical approaches are not sufficient to prove the posterior is sampled appropriately, Lesaffre and Lawson (2013)

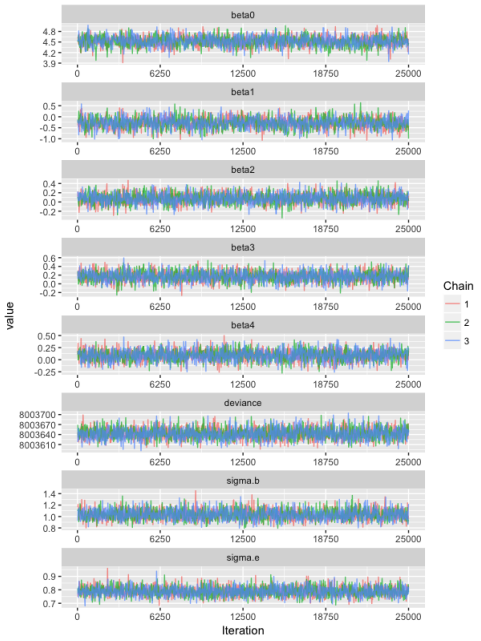

Figure 1: Without Hierarchical Centering

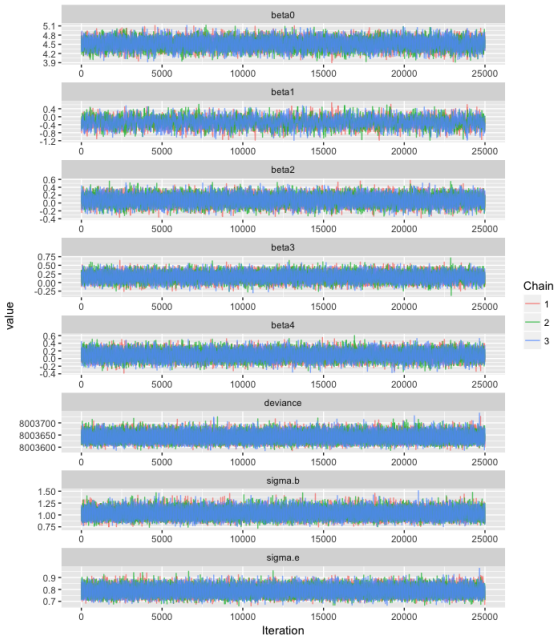

Figure 2: Hierarchical Centering

The trace plots presented above gives an idea of the mixing rate of the three chains in the model with and without hierarchical centering.

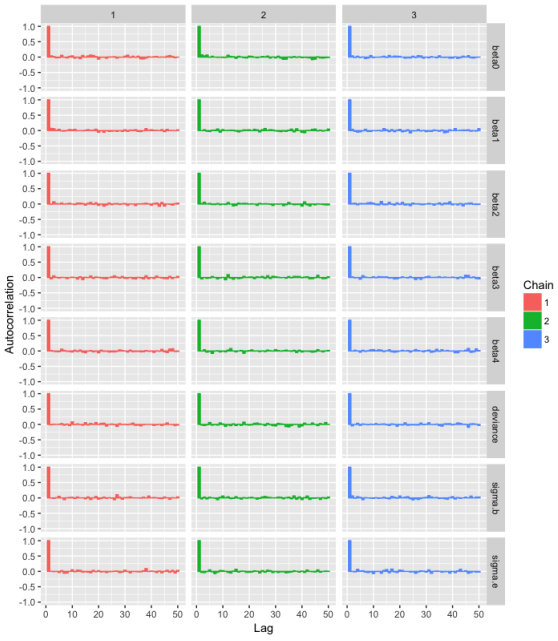

Figure 3: Without Hierarchical Centering

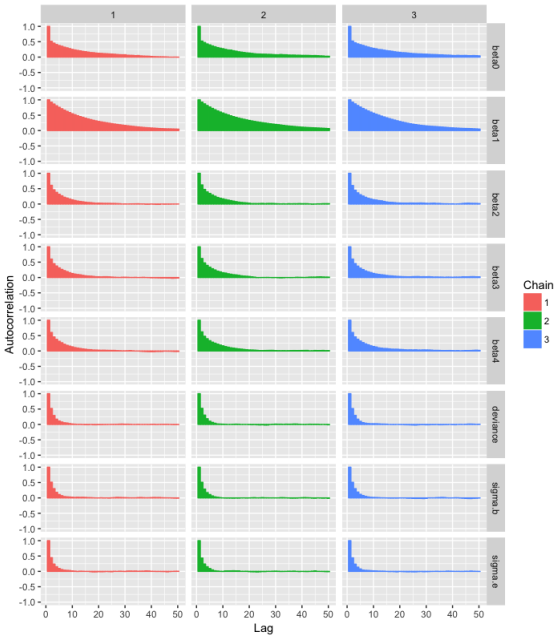

Figure 4: Hierarchical Centering

While convergence is doubtful in the model without hierarchical centering due to poor mixing rate of the chains, the trace plots for the model with hierarchical centering has better mixing rate. Moreover, stationarity can be claimed for the model with hierarchical centering based on the thick pen test. On the other hand, the autocorrelation plot revealed low mixing rate in the model with hierarchal centering for two of the parameters, β_0 and β_1. While, this is not a convergence diagnostic, it simply revealed that the starting point was not easily forgotten and the posterior was explored slowly for these parameters compared to the other parameters.

The Brooks-Gelman-Rubin diagnostic plot below presents a formal test of convergence for both models.

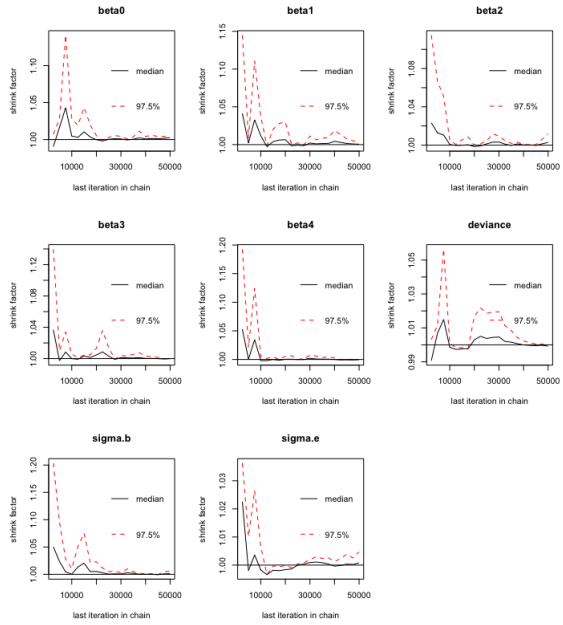

Figure 5: Without Hierarchical Centering

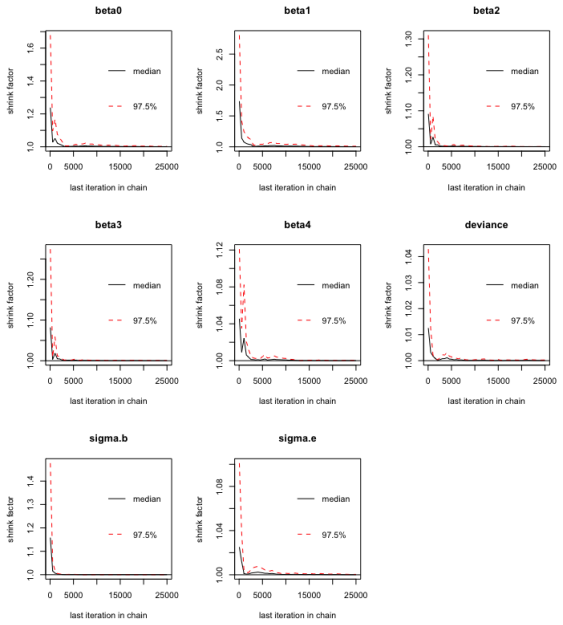

Figure 6: Hierarchical Centering

The graphs in Figure 6, with hierarchical centering, showed that most of the parameter estimates have already stabilized after 5000 iterations and the 97.5% upper bound is about 1, indicating good mixing rate and achieved convergence as compared to Figure 5 with no hierarchical centering. Also, a look at the estimated potential scale reduction factor ($\hat{R}$, see Lesaffre and Lawson, 2013), showed that convergence has been achieved since all $\hat{R}$ values were estimated to be 1.

	beta0	beta1	beta2	beta3	beta4	deviance	sigma.b	sigma.e
No HC	1665	1943	3000	3000	2966	3284	3000	3106
HC	5334	2361	9600	9551	9558	22601	21396	25898

Table 1: Effective Sample size, HC= Hierarchical Centering

The effective sample size was also considered, table 1 revealed that the precision of MCMC estimate of the posterior mean of all the parameters based on 50000 MCMC samples, is as good as taking lower independent samples for the model without hierarchical centering compared that with hierarchical centering. That is, lesser independent samples are required for the model without hierarchical centering compared to that with hierarchical centering. This is due to the high autocorrelation observed in the model with hierarchical centering (see Figure 4). Either ways, for both models, there is an enormous loss of information due to dependence sampling. Subsequently, the model with hierarchical centering will be used for further analysis since convergence was achieved in this model.

After convergence has been ascertained, an idea of the posterior density was sought for. The density curve of the posterior mean of each parameter is presented in Figure 7 and 8. These plots revealed that the density plots for the model with hierarchical centering was quite the same and more smoothed as compared to that without hierarchical centering, for the 3 chains. Furthermore, in both plots, the betas tend to be more symmetric compared to the deviance and the sigmas which were right skewed.

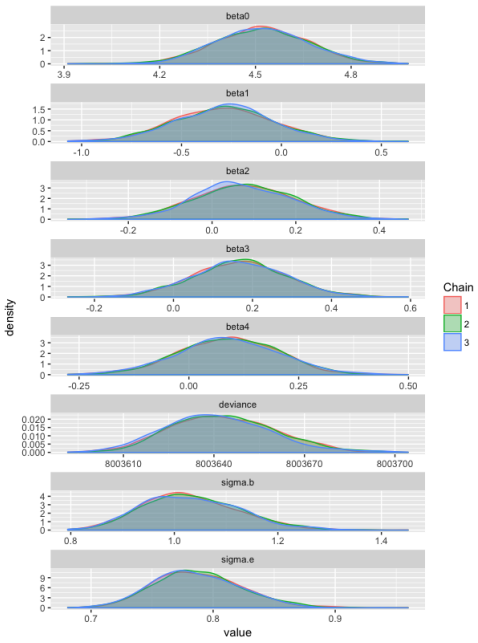

Figure 7: Density plot Without Hierarchical Centering

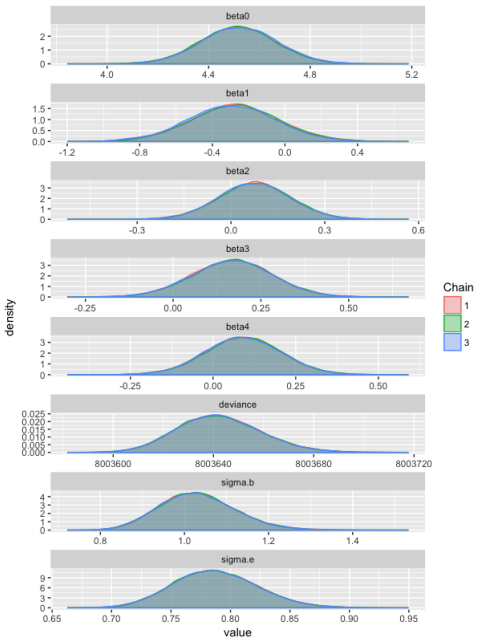

Figure 8: Density plot with Hierarchical Centering

2. Check graphically the assumed normality of the random effects

The posterior mean of the random intercept is shown in Figure 9. The shape of the histogram for the random intercept case doubt on the normality assumption of the random intercept, these deviation might be due to reason like omission of an important covariate, or due to the shrunk nature of random effect. The Normal Q-Q plot revealed that the sample quantile of the random intercept deviate at the left tail from the theoretical quantiles of the normal distribution and as such it can be said that the true random intercept distribution is probably

not normal.

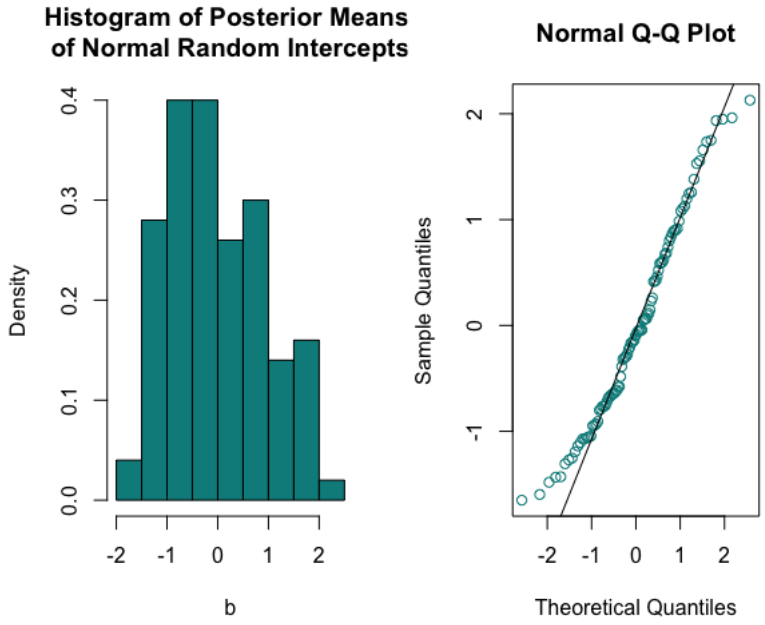

Figure 9: Distribution of Posterior Means for Random intercept

3. Evaluate the appropriateness of the model by making use of posterior predictive checks (PPC). In case the PPCs do not indicate a good fit, check if a t-distribution rather than a normal distribution for the random effects would provide a better fit.

A way to access the goodness of fit of a model is via PPC. It is a method where one sample from the assumed model and compare the extremeness of the sampled values to the observed via a statistics, under the assumed model. In other words, one compares values of a statistic evaluated at both the observed and sampled values from the posterior predictive density, a form of p-value is computed which compares how extreme is the statistic computed using sampled values from the posterior predictive density to the observed values. Here, six different statistic are considered; *minimum, maximum, skewness, kurtosis, Kolmogorov-Smirnov statistic and the Sinharay and Stern statistic* (see Lesaffre and Lawson, 2013 for more details).

Statistics	Posterior mean(p_D)
Minimun	0.1504
Maximum	0.4334
Skewness	0.1142
Kurtosis	0.6023
Kolmogorov-Smirnov	0.1495
Sinharay and Stern	0.1543

Table 2: Posterior predictive p-value for random intercept for different goodness of fit tests.

The results of the posterior predictive checks are presented in Table 2 above. Since none of the PPC estimate is low (*i.e.* < 0.05), it is concluded that the normal random effect fits the data well, hence there might be no need for use of t-distributed random effects.

4. Vary the prior specifications of the model parameters to assess the sensitivity of the results.

The choice of the prior distribution is crucial and important in Bayesian framework, since the posterior distribution might be driven by the choice of prior. Hence, the choice of the prior should be considered with care, especially when the likelihood does not dominate the prior. When the likelihood does dominate the prior, then the posterior will generally be invariant over a wide range of prior. Since results can be affected by different specifications of priors, a sensitivity analysis was carried out to assess this, Junaidi *et al.* (2011).

Parameter	N(0,10E-3)	N(0,10E-6)	N(0,10E-9)
β_0	4.52(0.150)	4.525(0.1497)	4.520(0.151)
β_1	-0.292(0.246)	-0.294(0.247)	-0.284(0.249)
β_2	0.075(0.115)	0.073(0.115)	0.075(0.116)
β_3	0.167(0.116)	0.166(0.114)	0.168(0.117)
β_4	0.09(0.116)	0.089(0.114)	0.091(0.117)

Table 3: Varying Prior for Fixed Effect, Posterior mean(posterior standard deviation)

Sensitivity of the fixed effect parameter to prior assumptions was checked by using different prior distributions. This was done by changing the prior variance from 1.0E-3 to 1.0E-6 and finally to 1.0E-9. Results of this exercise are reported in Table 3. It can be observed that varying the prior variance had little impact on all the fixed parameter estimates and their standard errors. This might be to the fact that the likelihood dominated the prior .

Parameter	U(0,100)	U(0,1.0E6)	U(0,1.0E9)	IG(1.0E-2,1.0E-2)	IG(1.0E-4,1.0E-4)
σ_ϵ	0.788(0.035)	0.788(0.036)	-	0.786(0.035)	0.786(0.035)
σ_b	1.032(0.091)	1.033(0.092)	1.033(0.092)	-	-

Table 4: Varying Prior for Variances, Posterior Mean (Posterior standard deviation)

Due to the known impact of the inverse gamma distribution on the level 2 variance of hierarchical models *i.e.* variance of the random effects (resulting posterior distribution is highly dependent on chosen prior parameter values), in this case only the uniform distribution is considered as a prior distribution for the level 2 variance, while both a uniform and inverse gamma distribution are considered as prior for the level 1 variance. Results obtained for both variance parameters remained consistent under the range of priors considered (Table 4), hence the choice of prior had little effect on the variance estimates.

5. Remove the unimportant covariates and check with DIC that no important covariates have been deleted. Is the model with covariates any better than the model without covariates?

In Bayesian context, the posterior mean of the deviance ($\overline{D(\theta)}$) presents a natural way to compare models but it lacks a way of accounting or penalizing for model complexity. To correct for this, the Deviance Information Criterion (DIC) was proposed. DIC is defined as $(\overline{D(\theta)} + p_D)$, where p_D is the effective number of parameters which is a Bayesian measure of model complexity, Lesaffre and Lawson (2013).

The DIC presented in Table 5 is corrected for the fact that a non-conventional likelihood was used. In other words, the DICs are corrected for the use of Poisson zeros trick. The DIC reported is computed using $DIC = DIC(0) - 2nC$, where DIC(0) is the DIC obtained from the model with Poisson zeros trick, n is the sample size and C(10000) is the constant added.

Model	Covariate	$\overline{D(\theta)}$	P_D	DIC
1	No Heifer	3641.911	142.4	3784.332
2	No Quarter	3640.419	133.3	3773.728
3	No Covariate	3640.807	135.8	3776.641
4	Full Model	3641.837	139.8	3781.7

Table 5: DIC obtained after each covariate has been removed

Comparing the DICs of the full model with that of the model without covariates, revealed that the model without covariate is better since it has a lower DIC value. However, this is not a clear win, since the difference between the DIC of both models is about 5 (using the rule of thumb of difference in DIC lesser than 5, indicates no clear winner).

To further investigate the relevance of each covariate in the model and to ascertain that no important covariate was removed from the model, a step-wise removal of each variable from the full model was carried out, although the credible interval for each parameter already suggested that none of the covariates are significant. The results of the exercise showed that removal of quarter from the full model reduced the DIC by 8, while the removal of heifer increases the full model DIC by 3. However, both reductions are lower than 10, which implies that, none of the models can be entirely ruled out when compared against one another.

6. Check with DIC and pseudo-Bayes factor the assumption that the time to

infection of four quarters of any cow are independent.

The model which assumes dependent and independent time of infection for the four quarter was fitted, after convergence has been ascertained, the DIC was computed. The DIC for the frailty model is much lower than that of the independent (a difference of 243 >> 10 between both models), indicating that the frailty model is a clear winner here. Hence it can be concluded that the time to infection of the four udder quarters are not independent.

Another approach to evaluating the evidence for or against a model based on the data is the Pseudo Bayes Factor (PSBF), which is essentially a variant of the Bayes factor (BF), but PSBF does not suffer from the setback of BF (sensitivity to the prior distribution). PSBF is defined as the ratio of the products of Conditional Predictive Ordinate (CPO) for both models, Lesaffre and Lawson (2013). For this report, the CPO was computed using the approach presented in Ntzoufras (2009) and Codgon (2014). This involves computing the Posterior Predictive Ordinate, PPO (evaluating the posterior predictive density at every observation), taking its inverse, and then doing some post processing to get the Logarithm of the Pseudo Marginal Likelihood (LPML). The log(PSBF) is then the difference between the LPML for both models. The result of this exercise was a negative log(PSBF), in favor of the frailty model. This conclusion is consistent with the DIC.

Model	DIC	$\sum log(CPO)$	$log(PSBF_{12})$
Independent(M_1)	4023	-2010.98	
Frailty($M2$)	3780	-1872.281	-138.699

Table 6: Model selection statistics

7. Compute summary statistics for the median survival time for each sub-group and the parameters of the final model.

Based on the the full model with heifer and quarters, one can create 8 subgroups for each combination of heifer and quarters. Since the time to event is lognormal, the median of log-normal distribution can be written as e^{μ}, for each subgroup we have the following expression for median using equation (2);

1. heifer = 0 and quater =1 $\Rightarrow e^{\beta_0+\beta_2}$
2. heifer = 0 and quater =2 $\Rightarrow e^{\beta_0+\beta_3}$
3. heifer = 0 and quater =3 $\Rightarrow e^{\beta_0+\beta_4}$
4. heifer = 0 and quater =4 $\Rightarrow e^{\beta_0}$
5. heifer = 1 and quater =1 $\Rightarrow e^{\beta_0+\beta_1+\beta_2}$
6. heifer = 1 and quater =2 $\Rightarrow e^{\beta_0+\beta_1+\beta_3}$
7. heifer = 1 and quater =3 $\Rightarrow e^{\beta_0+\beta_1+\beta_4}$

8. heifer = 1 and quater =1 $\Rightarrow e^{\beta_0+\beta_4}$

Modelling this median for each subgroup on the same number of iteration , the following summary statistics were obtained for the converged chain Table 8. In general, the multiparous cows have higher median survival time compared to the primiparous cows.

subgroup	mean	sd	2.5%	97.5%	MCse
Q1,Multiparous	100.371	15.207	73.964	133.514	0.1635
Q2,Multiparous	110.231	16.863	80.760	146.767	0.1813
Q3,Multiparous	101.976	15.518	74.927	135.877	0.1651
Q4,Multiparous	93.216	14.142	68.401	123.941	0.2002
Q1,Primiparous	75.908	16.992	48.042	114.533	0.2853
Q2,Primiparous	83.362	18.757	52.882	125.986	0.3157
Q3,Primiparous	77.136	17.387	48.932	116.729	0.2926
Q4,Primiparous	70.504	15.874	44.862	106.765	0.2516

Table 7: Summary statistics of median of each subgroup. QI= I^{th}Udder Quarter, I=1,2,3,4. Monte-Carlo Standard Error(MCse) obtained via Time Series approach.

Moreover, the second udder quarter in both multiparous and primiparous cows have the highest median survival time, while the fourth udder quarter has the lowest survival time in both type of cows. These estimated median survial time are obtained with good accuracy, since all MCse are lesser than 5% of the standard deviation of the posterior mean.

The final chosen model was the one that contains all the variables, since there was no clear winner between the full and model without covariates. The parameter estimates of the final model is displayed in Table 8 with the MCse indicating good accuracy of estimates. It is observed that heifer and udder quarter have no significant influence on the survival probability of the midpoint time of infection.

The exponential of coefficient illustrates the effect of covariate on survival time. For example, for a given quarter, the survival time of primiparous cow is decreased by $e^{-0.280} = 0.756$ as compared to multiparous cows and for a given parity, the survival time of Quarter 1,2 and 3 is increased as compared to quarter 4 since the parameter estimates are positive.

parameter	mean	sd	2.5%	97.5%	MCse
beta0	4.518	0.149	4.227	4.811	0.0020376
beta1	-0.280	0.246	-0.769	0.207	0.0050508
beta2	0.076	0.115	-0.148	0.301	0.0011732
beta3	0.169	0.115	-0.058	0.395	0.0011795
beta4	0.092	0.114	-0.133	0.315	0.0011714

Table 8: Summary statistics of the final model

Analysis 2: A Bayesian growth model

The data used in this section contains the age (in days) and dry weight of eye lens (in milligrams) of 72 rabbit. However, one of the observation contained a negative measurement for lens, this is discarded since it is not biologically plausible, hence an obvious recording error. Thus, for further computation 71 observations were used.

1. Thoroughly explore the dataset and provide descriptive statistics

The summary statistics for the rabbit data are presented in Table 9. The average age of the rabbits in the sample is around 240 days with an average dry weight of eye lens of 143.37 milligrams. The minimal dry weight of the eye lens is 21.66 while the maximal dry weight of eye lens is 246.7.

Statistics	Age	Lens
Min	2.00	21.66
Median	193.5	157.64
Mean	239.97	143.37
Max	860	246.7

Table 9: Summary statistics of the rabbit data

From figures 10 and 11, there is a nonlinear positive relationship between the variables Age and Lens, as well as between Age and log(Lens). Moreover, the slope of this relationship is positive in both case .

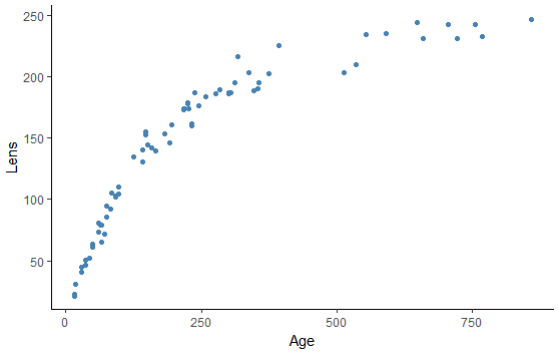

Figure 10: Scatter plot of Age and Lens

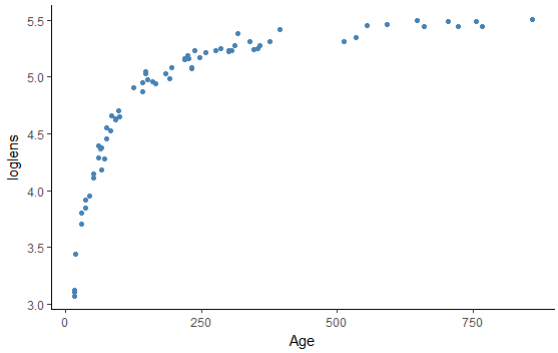

Figure 11: Scatter plot of Age and log(Lens)

2. Dudzinkski and Mykytowycz (1961) has suggested the following nonlinear model:

$$Lens_i = \alpha \; exp\Big(- \frac{\beta}{Age_i + \gamma}\Big) + \epsilon_i \qquad i = 1, ..., 71 \tag{4}$$

With α, β and γ unknown parameters and ϵ a normally distributed error term with constant variance (σ^2).

3. Fit the suggested nonlinear model 4, using vague priors for all models parameters. Take three chains and assess the convergence of the chains. Also report on the efficiency of the MCMC method.

Taking a N(0,1.0E-6) prior for α, β and γ with an IG(0.001,0.001) prior for $sigma$, the model above was fitted using 50,000 iterations with 25,000 burn-ins, without thinning. Convergence of the Markov chains was graphically checked with the trace plots while the dependence in the chain was examined with autocorrelation plot. While convergence was achieved, the autocorrelation was observed to be high, hence a thinning interval of 10 was applied to reduce autocorrelation.

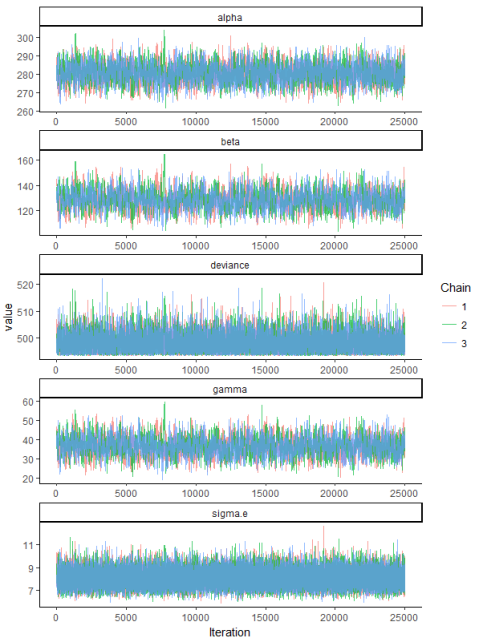

Figure 12: Trace plots for α, β, deviance, γ and σ without thinning

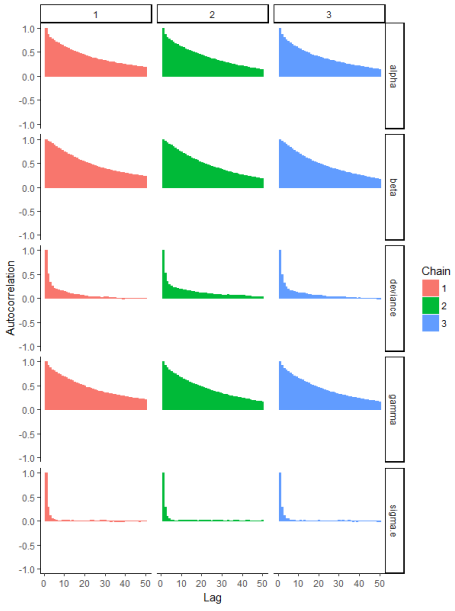

Figure 13: Autocorrelation for α, β, deviance, γ and σ without thinning

From the trace plots in Figure 14, horizontal strip and the individual moves are hardly noticeable, indicating that the chains are mixing well and there is stationarity for all the parameters, despite the application of thinning. Moreover, the Monte Carlo errors for all the parameters were less than 5% of their corresponding posterior standard deviation values. The corresponding autocorrelation for the parameters is seen in Figure 15. Autocorrelation plots is not a direct indication for convergence but it can be a useful tool (Lesaffre and Lawson, 2013). Autocorrelation shows the process of mixing rate which is measured by the autocorrelation at different lag.

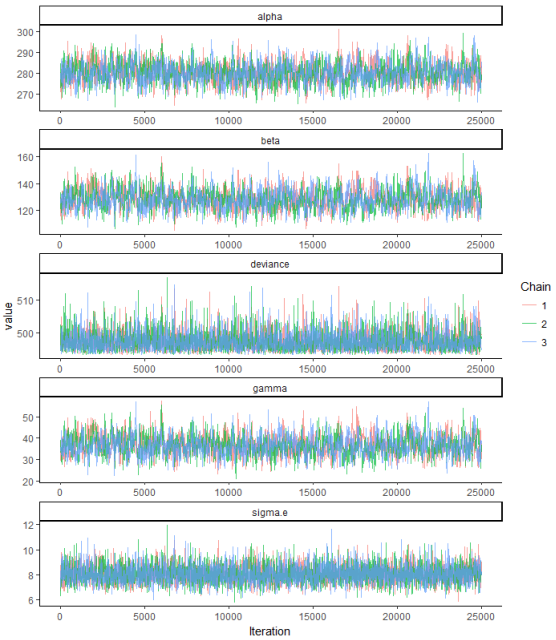

Figure 14: Trace plots for α, β, deviance, γ and σ with thinning

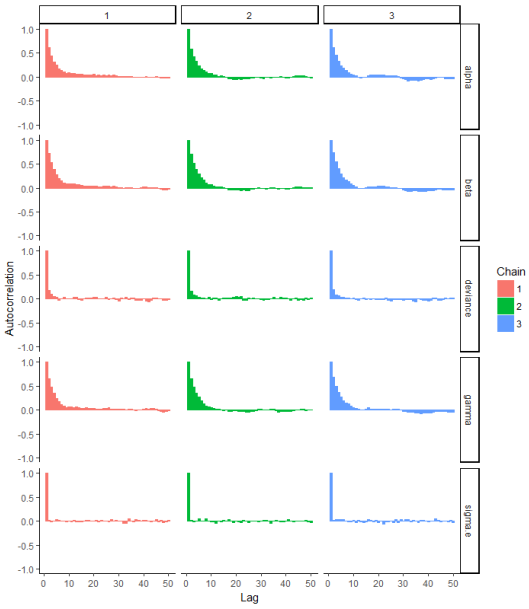

Figure 15: Autocorrelation for α, β, deviance, γ and σ with thinning

In order to check for convergence formally, a formal diagnostic test needs to be performed. Here, BGR diagnostic was used. The estimated potential scale reduction factor ($\hat{R}$) was equal to 1 for all parameters and the corresponding 97.5 % upper confidence interval for α, β and γ was equal to 1.01, additionally for deviance and σ_ϵ, it was found to be equal to 1. Thus, convergence is verified since the upper limit was observed to be close to 1 (Lesaffre and Lawson, 2013).The posterior density showed that α, β, γ are symmetric, while σ is skewed.

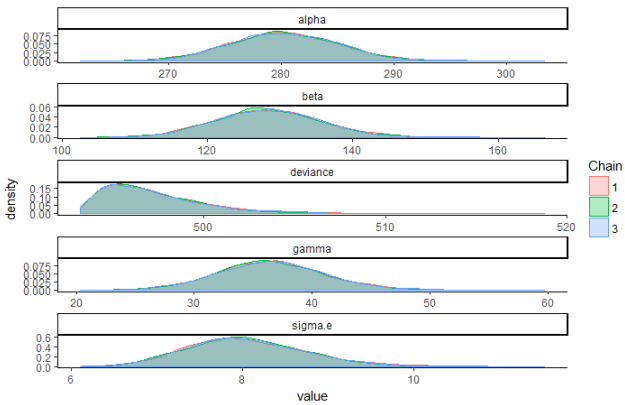

Figure 16: Density plots for deviance, α, β, γ, and σ

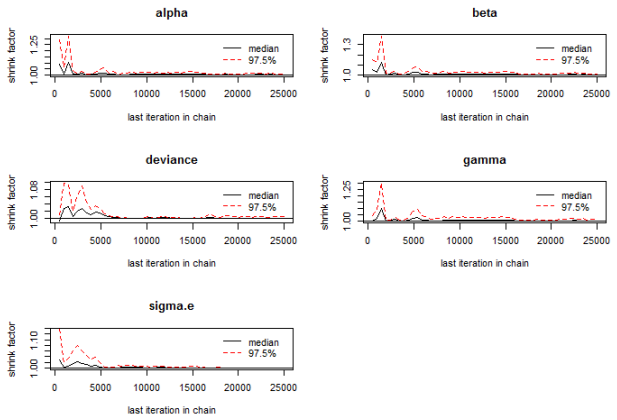

Figure 17: BGR plot of posterior α, β, deviance, γ and σ

The efficiency of the MCMC method can be evaluated by the MC error and the effective

sample size. The MC error should be smaller than 5% of the respective standard deviation of that parameter, indicating good accuracy for the parameter estimates. From Table 10, the MC error is indeed lower than 5% of the standard deviation for all the parameters. This suggests stability of the MCMC chain and accuracy of these posterior parameter estimates. Moreover, in terms of information loss, far lesser number of samples were needed for each parameter if independent sampling was carried out compared to dependence sampling adopted.

Parameter	**SD**	**MC Error**	**5% of SD**	**ESS**
α	4.730	0.0845	0.237	1435
β	7.266	0.142	0.363	1242
Deviance	2.885	0.0309	0.144	4394
γ	4.559	0.0856	0.228	1321
σ	0.701	0.00567	0.035	6963

Table 10: MC error and standard error of posterior parameters. SD=Standard Deviation, ESS=Effective Sample Size

4. Check graphically the assumption(s) made for the error terms

In order to graphically check the assumption made for the error term, *i.e.* error term is normally distributed with mean 0 and variance σ^2, a histogram plot and Q-Q plot of the error terms are constructed (see Figure 18). Based on the histogram, it appears that the assumption of normality made for the error term might be realistic. However, the heavy tails observed in both the Q-Q plot and histogram.

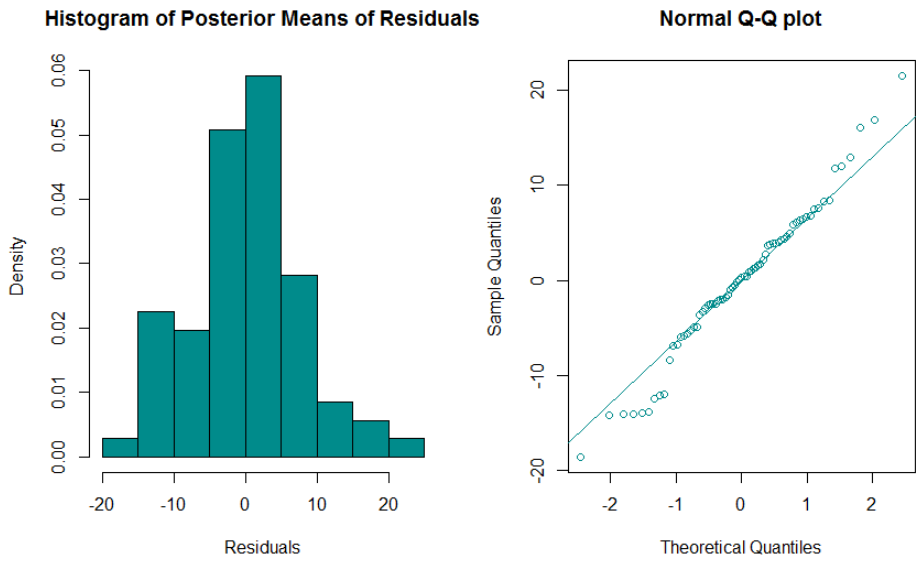

Figure 18: Histogram and Q-Q plot for the residuals for the nonlinear model 4

5. Evaluate the appropriateness of the model by making use of posterior predictive checks PPC

In order to evaluate the appropriateness of the model, one can sample from the assumed model and contrast them with the observed summary measures via the use of discrepancy measures. To test that the residuals are normally distributed, discrepancy measures listed in Table 11 were computed alongside and their posterior mean of their p-values. Every of these were higher than 0.05, indicating that the assumption of normality on the residual is not a bad idea and thus an appropriateness of the model.

Statistics	Posterior mean(p_D)
Minimun	0.4788
Maximum	0.2659
Skewness	0.4811
Kurtosis	0.4225
Kolmogorov-Smirnov	0.5909
Sinharay and Stern	0.3421

Table 11: Posterior predictive p-value for different goodness of fit tests.

6. Use the conditional predictive ordinates to check for outlying observations

Conditional predictive ordinates (CPO) is a cross-validatory measure of extremeness of y_i as well as to highlight extreme observations. Figure 19 represent the inverse-CPO which can be used to spot outlying observations. According to Ntzoufras (2009), inverse-CPO's larger than 40 can be considered as possible outliers, and higher than 70 as extreme values. Based on this, the index plot of inverse-CPO clearly points out one extreme value which is rabbit 54. Moreover, there are two other rabbits (60 and 61) that are potential outliers as well. The plot of the relative difference further confirmed that rabbit 54 and 61 are outliers, since their PPOs are relatively small and clustered together.

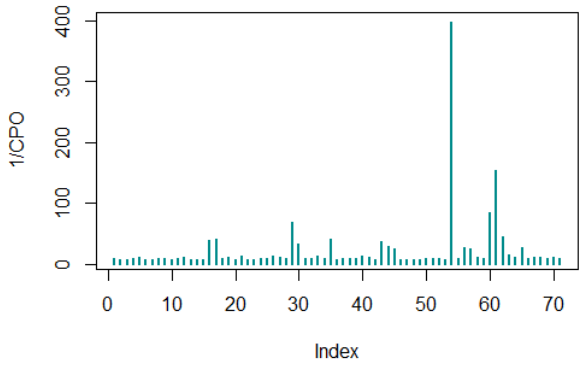

Figure 19: Index plot of 1/ $\hat{CPO}_i$

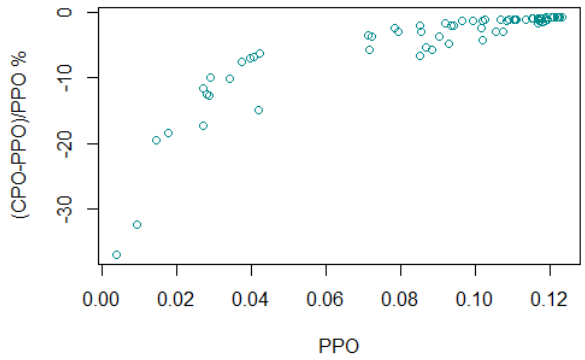

Figure 20: Relative difference of CPO vs PPO

7. Another model was assumed:

$$Lens_i = exp\left(\alpha + exp\left(-\frac{\beta}{Age_i + \gamma}\right) + \epsilon_i\right) \quad i = 1, ..., 71 \quad (5)$$

which can be reformulated as;

$$log(Lens_i) = \alpha + exp\left(-\frac{\beta}{Age_i + \gamma}\right) + \epsilon_i \quad (6)$$

The second model was fitted, using same vague priors above for all model parameters, with three chains. For each chain 50,000 iterations, a burn in of 25,000 were considered and there was no need for thinning. Convergence was graphically checked by using trace plots as presented in Figure 22, where a horizontal thick line indicated that the chains are mixing well. For all the parameters, a good mixing rate was observed from the autocorrelation plot (Figure 21) and there was no severe autocorrelation in any of the chains for all the parameters. Convergence was also formally tested with the BGR diagnostics and convergence was confirmed, since the upper limit was observed to be close to 1 (Lesaffre and Lawson, 2013).

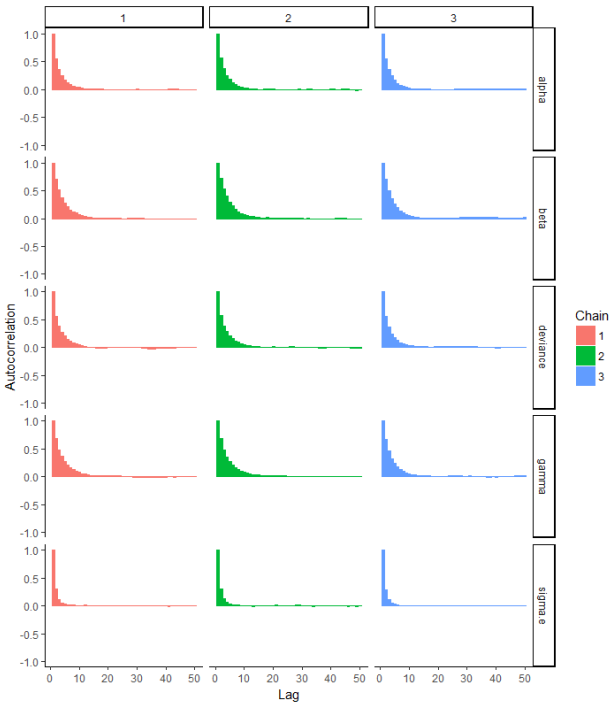

Figure 21: Autocorrelation for second growth model of the Markov chains of α, β, deviance, γ and σ

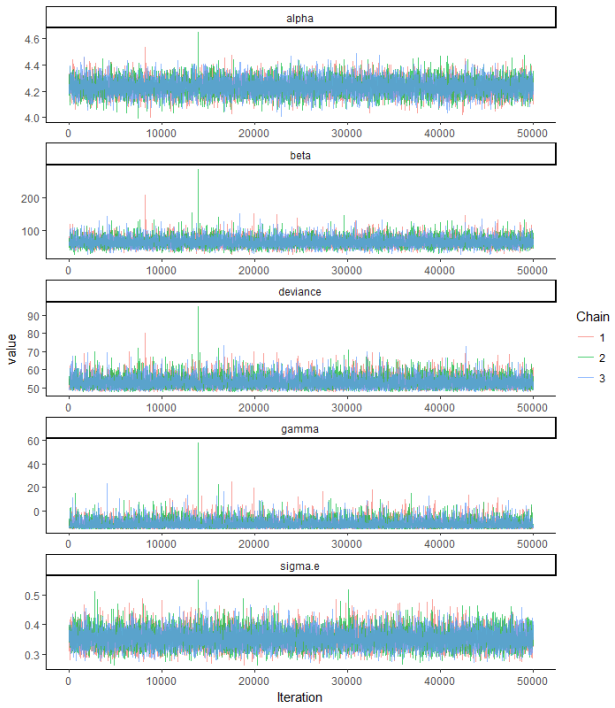

Figure 22: Trace plot for second growth model of the Markov chains of α, β, deviance, γ and σ

8. Determine which of the two models is the best

In order to determine which growth model is best for the data, the DIC was applied. The DIC could not be employed directly here because the response used in the two models are

different, however the DIC can be corrected since the DIC for the first model is based on the likelihood of Lens while in the second model DIC is based on the likelihood of log(Lens). Equation 7 will be employed for this transformation to the original scale for comparability Lesaffre and Lawson (2013).

$$DIC(\lambda) = DIC_\lambda - 2(\lambda - 1)\sum log(y_i) \tag{7}$$

where, DIC_λ is the DIC computed based on the transformed response, lambda is 0 in this case, since the transformation employed is logarithm and $\sum log(y_i) = \sum log(Lens_i)$.

$$\begin{aligned} DIC(\lambda) &= 58.30181 - 2(0 - 1)\sum log(lens) \\ DIC(\lambda) &= 58.30181 + 2(342.9358) = 744.1735 \end{aligned} \tag{8}$$

The DIC for Model 1 is 501.4706, while that of Model 2 after correction for the transformation of the response is 744.1735, which indicates that Model 1 is preferred over the Model 2, since it has a lower DIC value compared to that of Model 2.

9. Give all necessary posterior summary measures for the parameters of both models

Posterior summary measures were obtained after convergence was confirmed for all parameters in both models. As it can be seen in Table 13, the posterior means for all parameters were estimated accurately since for all obtained standard deviations are lesser than 5% of their MC errors. Moreover, the parameter estimates in Model 2 were lower in magnitude as compared to Model 1. Based on the equal tail credible intervals, it is observed that all parameters in both models were statistically significant, since all these intervals are not containing 0.

Model 1					
Parameter	**Mean**	**SD**	**MC error**	**2,50%**	**97,50%**
α	280.475	4.730	0.0845	271.618	290.114
β	128.763	7.266	0.142	115.297	143.677
γ	36.768	4.5587	0.0857	36.625	46.262
σ	8.056	0.7013	0.00567	6.826	9.579
Deviance	497.310	2.885	0.0309	493.719	504.479
Model 2					
α	4.237	0.0628	0.000343	4.115	4.362
β	63.962	15.110	0.0981	39.333	98.291
γ	-10.967	4.015	0.0261	-14.899	-0.408
σ	0.353	0.0307	0.000114	0.298	0.419
Deviance	52.976	3.319	0.0192	48.619	61.233

Table 12: Posterior summary statistics for growth model 1 and 2

10. Compare graphically the estimated average curves of the two models

In order to obtain the estimated average curves of Model 1 and Model 2, the average curve was estimated in the model statement for estimating the parameters of both models. As it can be observed in Figure 23, Model 1 seems to fit the data better, compared to Model 2 presented in Figure 24, which was also supported by the computed DIC .

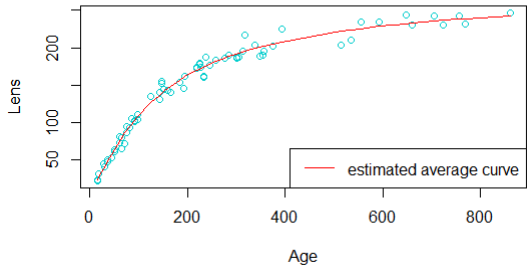

Figure 23: Estimated average curved for growth model 1

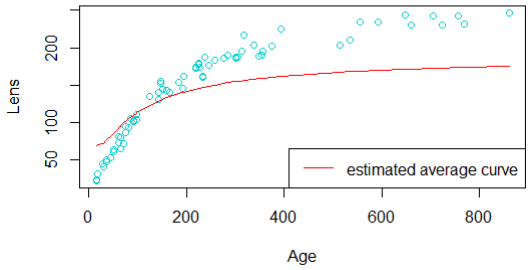

Figure 24: Estimated average curved for growth model2

References

Junaidi, Stojanovski, E. and Nur, D., (2011). *Prior Sensitivity Analysis for a Hierarchical Model*. Proceedings of the Fourth Annual ASEARC Conference, University of Western Sydney,Paramatta, Australia.

Lesaffre, E. and Lawson, A. B. (2013). *Bayesian Biostatistics(First Edition)*. John Wiley & Sons, Hoboken, New Jersey.

Ntzoufras, I. (2009). *Bayesian Modeling Using WinBUGS*. John Wiley & Sons, Hoboken, New Jersey.

Codgon, P. (2014). *Applied Bayesian Modelling(Second Edition)*. John Wiley & Sons, West Sussex.

Ibrahim, J. G., Chen, M-H., Sinha, D. (2001). *Bayesian Survival Analysis*. Springer, New York.