**Chapter 2 Simulations**

# 2.1 Introduction

In this chapter, we carried-out simulations with synthetic patient arrival datasets, and compares the posterior distribution results and prediction rate between Independent Bayes Model (IBM) and Hierarchical Bayes Model (HBM) in scenarios of with different anomalies and different hierarchy structures.

Sections 2.1 details functions used to synthesis simulation datasets. Sections

2.2 evaluates the use of diff t priors for Bayesian models to be used during the simulations. Section 2.3 explores the impact of size of anomaly towards anomaly detection at diff t levels of hierarchy. Section 2.4 and section 2.5 considers how complexity of the hierarchical structure affects anomaly detection at diff t levels of hierarchy.

# Simulation Methods

Synthetic datasets were used for the simulation studies, synthetic datasets refer to data information that’s artificially manufactured rather than generated from real- world events. Synthetic data was preferred over the alternative option, perturbed data, dataset that is generated by adding alteration and noise to real-word data. (Drechsler et al., 2008; Drechsler, 2011) Advantages of synthetic data includes: (1) Disclosure protection, as medical dataset almost always contained sensitive information about individuals.(2) Reduce the effects from possible background noises presented in real- word data. (3) Allows control over the hierarchical structure of datasets, note that real-world hierarchical data are often extremely complex and can have lots of levels, and complexity that comes with it, so we would prefer use of simpler structures for simulation studies.

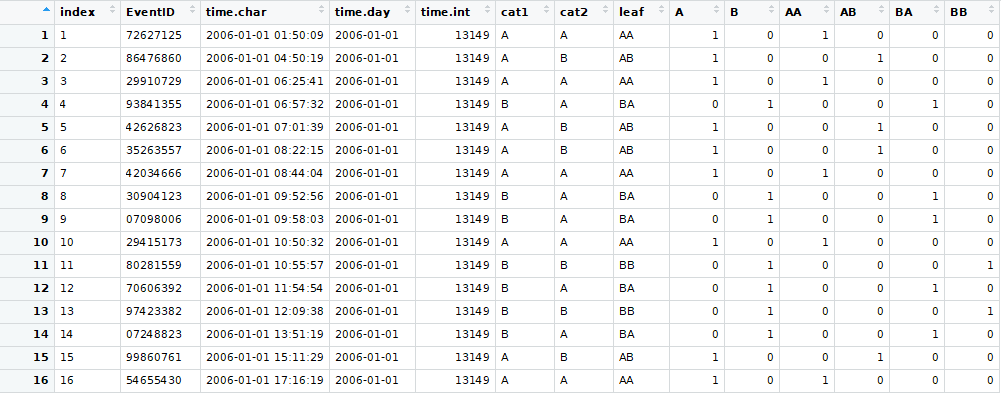


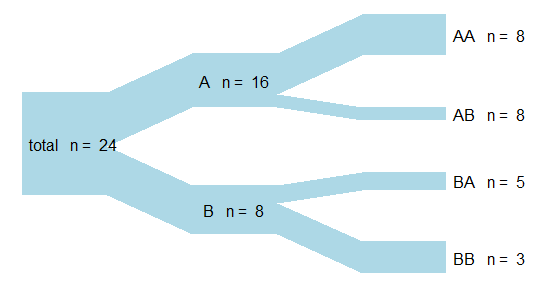
Figure 2.1: Entries of the raw dataset

Several custom R-functions were created to synthesis datasets used in this study, Descriptions of the functions are provided in Appendix A.1.There are several reasons for creating custom simulation synthesis functions. Firstly, a large number of simulated datasets with variable settings were expected to be required to perform our analyses, this required a large number of custom function arguments that existing packages may not contain. Secondly, synthetic datasets was simulated to mimic a health dataset in a pure theoretical setting. Thirdly, an emphasis was placed on the hierarchical nature of certain variable of dataset (for example, ICD codes and geographic location), brunch and leaf variable was designed to be generated with a custom structures during the simulation process. Lastly, as this is an exploration research it was unclear what functionalities were needed, existing statistical data simulation packages may become insuffi t at later stage.

simdata and adddaily.anomaly are the two major function used in the data synthesis process. simdata function were used to synthesise a raw dataset (figure [2.1)](#_bookmark0) that simulate a hospital record that you would typically found in hospital digital archives, where each row correspond to a single entry of hospital event with a unique hospital event identifier. Time, dates and various information about the hierarchical structure are automatically generated by the function. The default number of simulation is 1,000,000, and default period is a 12 year long period between 1 January 2006 and 31 December 2018, and a matrix that contains information of the hierarchical structure are manipulated and used to produce hierarchical structure with diff t characteristics.

The particular matrix used in the simdata function contains the theoretical value of each leaf of a two level hierarchical structure in a proportion out of 1000. As shown with the example in fi [2.2,](#_bookmark1) if the theoretical value at of the leafs (at level 2) are

8 5



*⇒* 8 3

333 208

*⇒* 333 125

Figure 2.2: Conversion of a simple hierarchical structure to a matrix

8, 8, 5, and 3, you would represent this information with a matrix with numbers 333,333,208 and 125. each column correspond to a single group, therefore column sums of the matrix correspond to the count of a level 1 categories, and the sum of all numbers in the matrix equal to the total count. The parent levels are simply the sum of children levels, the complete hierarchical structure can be generated automatically just from numerical information of leafs. Note that the number of column of the matrix correspond to the maximum number of leafs within a level 1 category, for level 1 categories with less than maximum amount of leafs, 0 is used to represent no leaf. The default value of the matrix is 250, 250, 250 and 250, which represent a two level hierarchical of equal proportions, with 2 level 1 categories, each with 2 level 1 categories of equal proportions. This way of storing hierarchical information is simple and efficient but can not be used for hierarchical models that contain 3 or more levels, and may be a little confusing for other people at fi future improvements of the function could be allowing usage of a user-generated hierarchical time series (hts) object from the hts package instead of the matrix.

adddaily.anomaly were used to addanomaly function to add anomaly to daily count data, a hierarchical time-series dataset generated by tabulating the raw dataset using tabulatedata function. The default setting is a point anomaly of 1 day, on 2015-11-15. The amount of anomaly added in terms of percentage of total count, and amount of anomaly assigned to each leaf is varied.

For point anomalies The amount of anomaly assigned to each leaf is transcribed with a binding of 3 vector, a vector of percentage of anomaly to be added to total, a percentage vector of anomaly to be added to level 1 categories and a vector of

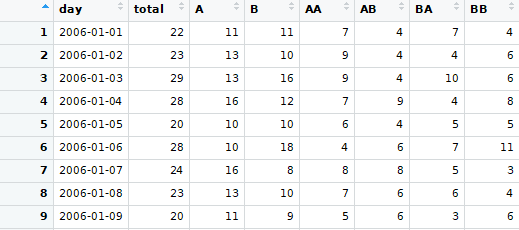


Figure 2.3: Entries of the daily count dataset

percentage of anomaly to be added to level 1 categories. For example, for a hierarchical time series with the structure as the example given in fi [2.2,](#_bookmark1) if we want to add all anomaly to AA, the percentage vector of anomaly for total is *{*1*}* and this will be the case for any hierarchical structure, the percentage vector of anomaly for level 1 categories is *{*1*,* 0*}* and the percentage vector for level 2 categories is *{*1*,* 0*,* 0*,* 0*}*, so

we just need to add a percentage vector *{*1*,* 1*,* 0*,* 1*,* 0*,* 0*,* 0*}*.

0

|  |  |  |  |
| --- | --- | --- | --- |
| *Ptotal* | *Plv*1 | *Plv*2 *P* |  |
| 1 + | 1 0 + 1 | 0 0 0 *⇒* 1 1 0 1 0 | 0 |

Figure 2.4: percentage vector

Details about setups used for the data synthesis process can be found in the start of each section, for each of our simulation tests, and the R codes used can be found in 91-createdata-test.R, 92-createdata-anoamly.R, 93-createdata-proportion.R and 94-createdata-count.R found in GitHub Repository [https://github.com/jungxu](https://github.com/jungxue/research-masters-Jung)e/ [research-masters-Jung](https://github.com/jungxue/research-masters-Jung).

# Simulation 1: Prior selection

## Priors

For Bayesian inference to make sense a prior is essentially required, however there are diff t ways to generate the prior. For simulations, no external informations are available, therefore the informative prior approach is out of the question. However it is possible for us to assign relatively simple weak-informative and non-informative priors by generating probability distributions with vaguely defi parameters. And there- fore, we proposed 6 diff t priors that express diff t hypothesised information about the prior.

A simple Poisson process can be used to model for a random, and mutually independent arrivals rate, with the likelihood model:

*y,t ∼ Poisson*(*µ,t*) *log*(*µ,t*) = *log*(*ρ,t*)

(2.1)

For the Poisson model, our observations *yi,t* (*y* of *i*th category given time *t*) is hypothesised to follow the distribution of a Poisson process, with parameter *µi,t* as the expected number of arrivals per time-interval. For our simulations this time interval is setup to be a 12 year long period between 1 January 2006 and 31 December 2018. Priors can then assigned to *µi, t*, this provides information to the parameter and will have significant impact in our calculations. Samples of the JAG codes used to build our models are presented in appendix A.1.8, and complete JAGS codes and R codes used to obtain our graphical results are available on GitHub repository: [https://github.com/jungxue/research-masters-Jung.](https://github.com/jungxue/research-masters-Jung)

#### Model 1 Non-informative Prior

Model 1 is a Independent Bayesian Model with no priors (IBM) that models each categories independently on all hierarchical levels. This model provides no information for the estimation of distribution parameter *µi,t*, and therefore the prior can be defi

as a non-informative prior, here in this thesis we call it the null prior. The likelihood

model is:

*yi,t ∼ Poisson*(*µi,t*) *log*(*µi,t*) = *log*(*ρi,t*)

(2.2)

#### Model 2 Normal(1,0.3) Prior

Model 2 is a Independent Bayesian Model with *ρ ∼ Normal*(1*,* 0*.*3) priors that models each categories independently on all hierarchical levels. This model does contain information for the estimation of distribution parameter *µ, t*, and therefore the prior can be defi as a weak-informative prior. *T*[*λ>*0] indicates a truncation of distribution at 0 in which *λ* values less than 0 is truncated, therefore *λ* should not contain any negative values. The likelihood model is:

*yi,t ∼ Poisson*(*µi,t*) *log*(*µi,t*) = *log*(*ρi,tλi,t*)

(2.3)

Priors for model parameters:

*λi,t ∼ Normal*(*µi,t, σi,t*)*T*[*λ>*0] (2.4) Hyper-priors for model parameters:

*µi,t ∼ Normal*(1*,* 0*.*1)

*i,t ∼ Normal*(0*.*3*,* 0*.*1)

*σ*2

(2.5)

#### Model 3 Normal(1,0.1) Prior

Model 3 is a Independent Bayesian Model with *ρ ∼ Normal*(1*,* 0*.*1) priors that models each categories independently on all hierarchical levels. This model does contain information for the estimation of distribution parameter *µ, t*, and therefore the prior can be defi as a weak-informative prior. *T*[*λ>*0] indicates a truncation of distribution at 0 in which *λ* values less than 0 is truncated, therefore *λ* should not contain any negative values. The likelihood model is:

*yi,t ∼ Poisson*(*µi,t*) *log*(*µi,t*) = *log*(*ρi,tλi,t*)

(2.6)

Priors for model parameters:

*λi,t ∼ Normal*(*µi,t, σi,t*)*T*[*λ>*0] (2.7) Hyper-priors for model parameters:

*µi,t ∼ Normal*(1*,* 0*.*1)

*i,t ∼ Normal*(0*.*1*,* 0*.*1)

*σ*2

(2.8)

#### Model 4 Gamma(4,3) Prior

Model 4 is a Independent Bayesian Model with *ρ ∼ Gamma*(4*,* 3) priors that models each categories independently on all hierarchical levels. This model does contain information for the estimation of distribution parameter *µ, t*, and therefore the prior can be defi as a weak-informative prior. *T*[*λ>*0] indicates a truncation of distribution at 0 in which *λ* values less than 0 is truncated, therefore *λ* should not contain any negative values. The likelihood model is:

*yi,t ∼ Poisson*(*µi,t*) *log*(*µi,t*) = *log*(*ρi,tλi,t*)

(2.9)

Priors for model parameters:

*λi,t ∼ Gamma*(*αi,t, βi,t*)*T*[*λ>*0] (2.10) Hyper-priors for model parameters:

*αi,t ∼ Normal*(4*,* 0*.*1)

*βi,t ∼ Normal*(3*,* 0*.*1)

(2.11)

#### Model 5 Laplace(1,1) Prior

Model 5 is a Independent Bayesian Model with *ρ ∼ Laplace*(1*,* 1) priors that models each categories independently on all hierarchical levels. This model does contain information for the estimation of distribution parameter *µ, t*, and therefore the prior can be defi as a weak-informative prior. *T*[*λ>*0] indicates a truncation of distribution at 0 in which *λ* values less than 0 is truncated, therefore *λ* should not contain any negative values. The likelihood model is:

*yi,t ∼ Poisson*(*µi,t*) *log*(*µi,t*) = *log*(*ρi,tλi,t*)

(2.12)

Priors for model parameters:

*λi,t ∼ Laplace*(*αi,t, βi,t*)*T*[*λ>*0] (2.13) Hyper-priors for model parameters:

*αi,t ∼ Normal*(4*,* 0*.*1)

*βi,t ∼ Normal*(3*,* 0*.*1)

(2.14)

#### Model 6 Mixture Prior

Lastly, Model 6 is a Independent Bayesian Model with a mixture priors that models each categories independently on all hierarchical levels. It is a mixture of no informa- tion and weak information, in other words, we are proposing that for most of the time there are no variation, and for some of the time there are informations. The idea comes from the prior Berry and Berry (2004) where they proposed a scenario where majority of diff is zero and there are just some variations. However note that their prior is one sided and center at 0, our prior is two sided and center at 1. This model does contain information for the estimation of distribution parameter *µ, t*, and therefore the prior can be defi as a weak-informative prior.*T*[*λ>*0] indicates a truncation of distribution at 0 in which *λ* values less than 0 is truncated, therefore *λ* should not contain any negative values. The likelihood model is:

*yi,t ∼ Poisson*(*µi,t*) *log*(*µi,t*) = *log*(*ρi,tλi,t*)

Priors for model parameters:

(2.15)

*λi,t ∼ Normal*(*spikei,t* + (1 *− spikei,t*)*slabi,t,* 0*.*1)*T*[*λ>*0] (2.16) Hyper-priors for model parameters:

*spikei,t ∼ Binomial*(0*.*9*,* 1)

*slabi,t ∼ Normal*(*µi,t, σi,t*) Hyper-hyper-priors for model parameters:

*µi,t ∼ Normal*(1*,* 0*.*1)

*i,t ∼ Normal*(0*.*1*,* 0*.*1)

*σ*2

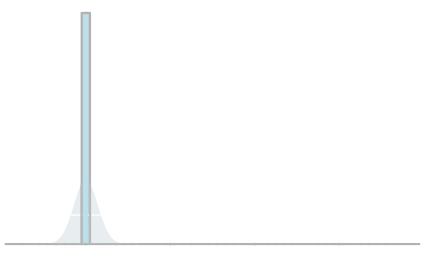
(2.17)

(2.18)

10.0

PRGHO 1: 1XOO SULRU

0.8



PRGHO 2: 1RUPDO(1,0.3) SULRU

7.5

0.6

5.0

GHQVLW\

0.4

2.5

GHQVLW\

0.2

0.0

1.25

1.00

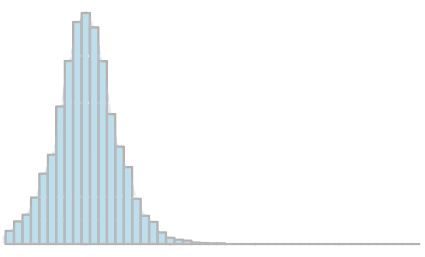
0 1 2 3 4 5

\1

PRGHO 3: 1RUPDO(1,0.1) SULRU

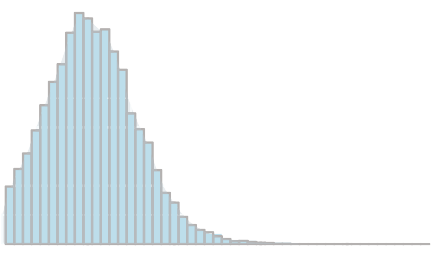
0.0

0.6



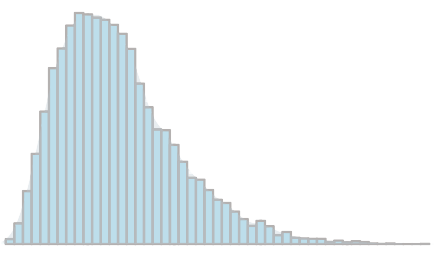
GHQVLW\

0 1 2 3 4 5



\2

PRGHO 4: \*DPPD(4,3) SULRU



0.75

GHQVLW\

0.4

0.50

0.25

0.2

0.00

0 1 2 3 4 5

\3

0.0

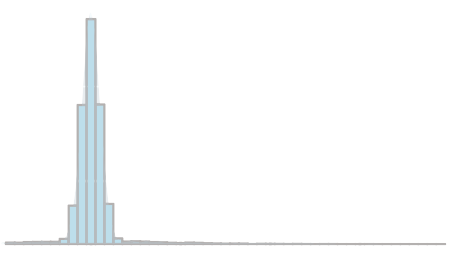
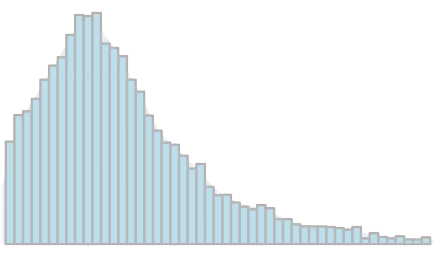
0 1 2 3 4 5

\4

PRGHO 5: /DSODFH(1,1) SULRU

PRGHO 6: 0.5 0L[WXUH SULRU

3



0.4

GHQVLW\

GHQVLW\

2

0.2

1

0.0

0 1 2 3 4 5

\5

0

0 1 2 3 4 5

\6

Figure 2.5: Histo-density plot of the prior probability distribution of proposed prior density models. The red dashed line indicates the center point of these distributions, which is 1.

Figure [2.5](#_bookmark2) show the distribution of *λ*, because we assign prior distribution to our *μ* by directly multiplying by *λ*, the distribution of *λ* are the equivalent to the distribution of our prior. As it is presented, our priors have diﬀerent distributions that suggest a diﬀerent believe. Model 1 is a non-informative prior, that contain no information, and hence no distribution is observed. Model 2 is a *Normal*(1*,* 0*.*3) prior that seem reasonably, however becaus eit’s left tail was truncated at 0, the symmetrical property has been lost. Model 3 is a *Normal*(1*,* 0*.*1) that has a smaller variance, hence the distribution is a little bit taller, indicating a slight more observations tend to be close to the mean. Model 4 is a *Gamma*(4*,* 3) prior, It skew towars the right hand

side, this trans lates into we are observing more positive anomalies. Modle 5 is a Laplace(4,3)prior, it also skew to the right but with a heavier tail and a much higher distribution at xtreme positive values, This translates into information at we are consistently obsering a large abomaly. Lastly we have the mixture model, Information from

## Simulation setups

In section 2.3.2 we presents the results of the posterior distribution calculations for the synthesised data, with anomalies added on category AA on level 2 (leaf) of the hierarchy. Default setting for the data synthesis process is used and a 25% anomaly is added. Note the 25% anomaly meant 25% of the count of total category on top of the hierarchy, not an addition of 25% on Category AA.

## Results

In section 2.3.2 we presents the results of the posterior distribution calculations for the synthesised data, with anomalies added on category AA on level 2 (leaf) of the hierarchy. Default setting for the data synthesis process is used and a 255 anomaly is added. Note the 25% anomaly menat 25% of the count of total category on top of the ierarchy, not an addition of 25% on Catergory AA.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| M1:M2 | -2.126 | -1.574 | -0.599 | -0.868 | -0.158 | 0.113 |
| M1:M3 | -4.456 | -3.295 | -2.069 | -2.032 | -0.650 | 0.189 |
| M1:M4 | -2.992 | -1.729 | -1.174 | -1.303 | -0.800 | 0.106 |
| M1:M5 | -2.645 | -1.745 | -0.764 | -1.172 | -0.590 | -0.124 |
| M1:M6 | -3.383 | -2.200 | -1.645 | -1.683 | -1.047 | -0.258 |
| M2:M3 | -4.555 | -2.288 | -1.884 | -1.164 | 0.713 | 1.440 |
| M2:M4 | -2.393 | -1.009 | -0.552 | -0.434 | -0.155 | 2.232 |
| M2:M5 | -2.382 | -0.494 | -0.071 | -0.304 | 0.267 | 0.782 |
| M2:M6 | -3.331 | -1.571 | -0.626 | -0.815 | -0.122 | 1.639 |
| M3:M4 | -2.450 | -0.285 | 0.895 | 0.730 | 1.787 | 3.658 |
| M3:M5 | -1.020 | -0.578 | -0.168 | 0.860 | 2.345 | 3.677 |
| M3:M6 | -1.995 | -1.322 | -0.412 | 0.349 | 1.649 | 4.198 |
| M4:M5 | -1.884 | -0.905 | 0.501 | 0.131 | 0.913 | 2.282 |
| M4:M6 | -2.858 | -1.537 | -0.829 | -0.380 | 1.029 | 2.039 |
| M5:M6 | -3.259 | -0.929 | -0.672 | -0.511 | 0.378 | 1.454 |

Table 2.1: DIC (with 1,000,000 iteration) summary distribution cross comparison results for model 1 to model 6

DIC is used to comapre between models

Table [2.1](#_bookmark3) gives the comparison of of DIC distributions. A important note is that initial dic results with 1000 iterations are very unstable, showiing signs of slow convergence , this meant that the dic values for each model diff by a large margin each time dic is ran for diff t jags modles, therrefor it it is very hard to decide which model is the best.

so the number o iteration was increased and increasing iteration to 1,000,000 does stioll yields unstable DIC. The problem has been raised

and other research students from UoA stats lab also encoiunter

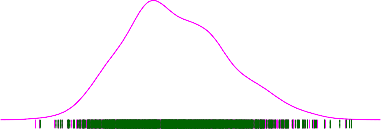
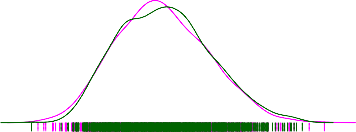
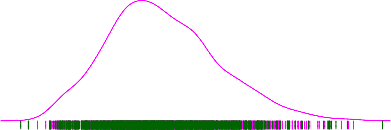
Quantifying uncertainty in marginal likelihood estimation via the stepping-stone sampling algorithm John Siryj (Renate Meyer, Patricio Maturana Russel)

However, what has been observed is that at least for the numerous trials, model 1 always yields inferior DIC comapre to all other models

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| model1 | 21.0615 | 0.0000 | 21.0615 | 21.0615 | 21.0615 |  | 0.0000 |
| model2 | 24.4976 | 4.7805 | 15.7987 | 24.2398 | 34.1927 | 1.0000 | 3119.0000 |
| model3 | 24.3373 | 4.9322 | 15.7804 | 23.9781 | 34.8648 | 1.0000 | 3009.0000 |
| model4 | 24.1116 | 4.8248 | 15.5574 | 23.7915 | 34.4035 | 1.0000 | 3587.0000 |
| model5 | 24.4041 | 4.9549 | 15.9001 | 24.1139 | 34.7204 | 1.0000 | 3152.0000 |
| model6 | 24.8840 | 4.9712 | 16.4075 | 24.5082 | 35.4618 | 1.0000 | 2895.0000 |

Table 2.2: Posterior distributions of diff t models for Total, with added anomalies, and calculated with independent Bayes model

10 20 30 40 50



Density

0.00 0.02 0.04 0.06 0.08

0.00 0.02 0.04 0.06 0.08

model3

model5

0.00 0.02 0.04 0.06 0.08

model4

model6

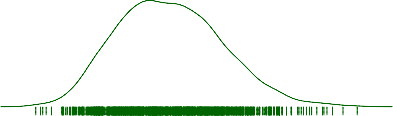
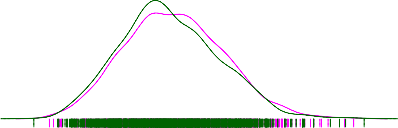
10 20 30 40 50

10 20 30 40

0.00 0.02 0.04 0.06 0.08

model1

10 20 30 40 50



0.00 0.02 0.04 0.06 0.08

0.00 0.02 0.04 0.06 0.08

model2

10 15 20 25 30 35

10 20 30 40 50

Figure 2.6

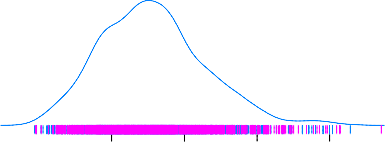
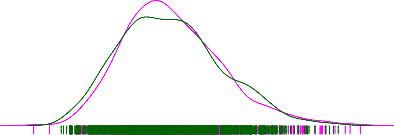
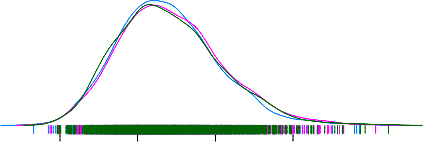
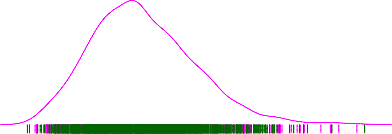
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| model1 | 10.5463 | 0.0000 | 10.5463 | 10.5463 | 10.5463 |  | 0.0000 |
| model2 | 12.3481 | 3.5241 | 6.4608 | 12.0826 | 19.7636 | 1.0000 | 2833.0000 |
| model3 | 12.2433 | 3.5058 | 6.3331 | 11.9460 | 20.1100 | 1.0000 | 2955.0000 |
| model4 | 12.0331 | 3.2668 | 6.4515 | 11.7501 | 19.1090 | 1.0000 | 3322.0000 |
| model5 | 12.3475 | 3.5362 | 6.4799 | 12.0049 | 19.9832 | 1.0000 | 2503.0000 |
| model6 | 12.9588 | 3.6536 | 6.8353 | 12.6056 | 21.2272 | 1.0000 | 3000.0000 |

Table 2.3: Posterior distributions of diff t models for A , with added anomalies, and calculated with independent Bayes model

0.12

0.12

0 5 10 15 20 25 30



Density

0.04 0.08

0.12

0.00

0.04 0.08

model3

model5

0.05 0.10

model4

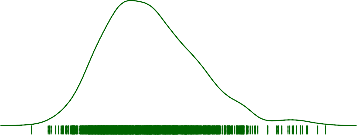
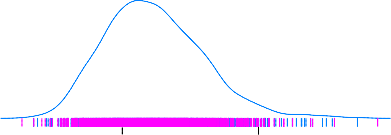
0.00

0.04 0.08

model6

5 10 15 20 25 30

5 10 15 20 25 30



0.00 0.05 0.10 0.15

0.00

model1

0.12

0.00

0 5 10 15 20 25

5 10 15

0.00 0.04 0.08

model2

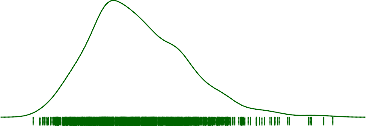
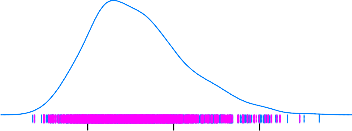
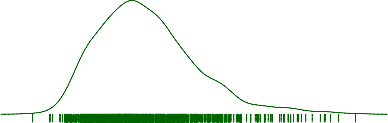
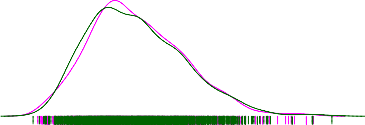
0 10 20 30

Figure 2.7

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| model1 | 5.2477 | 0.0000 | 5.2477 | 5.2477 | 5.2477 |  | 0.0000 |
| model2 | 7.4081 | 2.7444 | 2.8634 | 7.1136 | 13.3165 | 1.0000 | 2273.0000 |
| model3 | 7.3234 | 2.7423 | 2.8910 | 7.0234 | 13.3408 | 1.0000 | 2372.0000 |
| model4 | 6.9192 | 2.4111 | 3.0404 | 6.6744 | 12.3585 | 1.0000 | 3000.0000 |
| model5 | 7.6640 | 2.8689 | 3.0550 | 7.2466 | 14.0954 | 1.0000 | 1886.0000 |
| model6 | 7.8524 | 2.7397 | 3.4434 | 7.5279 | 14.0425 | 1.0000 | 2856.0000 |

Table 2.4: Posterior distributions of diff t models for AA, with added anomalies, and calculated with independent Bayes model

0 5 10 15 20 25



Density

0.00 0.05 0.10 0.15

0.00 0.05 0.10 0.15

model3

model5

0.00 0.05 0.10 0.15

model4

model6

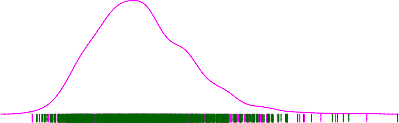
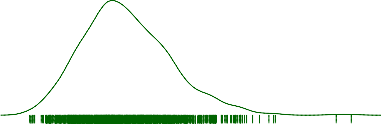
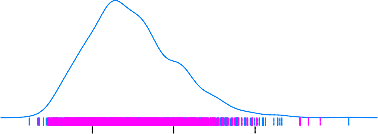
0 5 10 15 20 25

0 5 10 15 20 25

0.0 0.1 0.2 0.3

model1

0 5 10 15 20



0.00 0.05 0.10 0.15

0.00 0.05 0.10 0.15

model2

2 4 6 8

0 5 10 15 20 25

Figure 2.8

10.2 10.6 11.05 10 20

SIMULATIONS

2000

Trace plot for category total of model 1 to model 6

2000

Figure 2.10: Trace plot for category A of model 1 to model 6

5 10 20

5 10 20

5 10 20

5 10 20

20.6 21.0 21.410 20 30 40 10 20 30 40 10 20 30 40 10 20 30 40 5100 20 30 40

CHAPTER 2.

2.3. SIMULATION 1:

PRIOR SELECTION

3000

4000

5000

Iteration number

Figure 2.9:

3000

4000

5000

Iteration number

37

model1 model2 model3 model4 model5 model6

model1 model2 model3 model4 model5 model6

2000 3000 4000 5000

4.8 5.2 5.6

5 10 20

5 101520

5 10 15 20 5 101520

PRGHO1

PRGHO2

PRGHO3

PRGHO4

,WHUDWLRQ QXPEHU

Figure 2.11: Trace plot for category AA of model 1 to model 6

**$XWRFRUUHODWLRQ SORW IRU FDWHJRU\ 7RWDO RI DOO PRGHOV**

0 5 10 15 20 25 30

0.10

●

●

● ●

●

●

●

●

●

● ●

● ●

●

●

● ●

●

●

●

● ●

●

● ● ● ●

●

● ●

● ● ● ●

● ● ● ● ●

● ●

● ● ●

●

●

●

● ● ● ●

● ●

● ●

●

● ●

●

● ●

● ●

● ● ●

● ●

●

●

●

● ● ● ● ● ● ● ● ●

● ● ● ● ●● ●

●● ●

● ●

● ● ● ● ● ● ●

●

●

● ●

● ● ●

●

●

●

●

●

● ●

●

● ●

●

●

●

●

●

●

●

●

● ● ● ●

●

●

●

●

●

●

●

●

●

●

●

●

●

● ●● ● ●

●

●

●

●

●

●

●

●

●

●

●

●

●

●

●

●

●

●

●

●

●

●

●

●

●

●●

●

● ● ●

●

● ●

●

●

● ● ●

●

● ●

● ●

●

● ● ● ● ●

●

●

●

●

●

● ●

●

●

● ●

● ●

●

●

● ●● ●

●

● ● ●

●

●

● ●

● ● ● ●

●

● ● ● ●

● ● ● ● ●

●

●

●

● ● ●

● ●

●

●

●

●● ●

●

● ●

●

● ●

● ●

● ● ● ● ●● ● ● ● ●

●

●

●

●

● ● ●

●

●

● ● ● ●

●

●

●

●

● ● ●

●

●

●

● ● ● ●

●

●

● ● ● ● ●

● ● ●

●

●

●

● ●

●

●

● ● ● ● ●

●●

● ● ●

●

●

●

●

●

●

●

●

●

●

●

●

● ●

● ● ●

● ● ●

● ●

●

●

●

● ●●

●

● ●

● ● ● ● ● ● ● ● ● ●

●

●

●●

●

● ●

●

● ● ●

●

● ●

● ●

● ●

● ● ●

● ● ●

● ●

● ●

●

●

● ● ●

● ●

●

●

●

●

● ●

●

●

0.05

PRGHO5

PRGHO6

0.00 í0.05 í0.10

0.10

$XWRFRUUHODWLRQ

0.05

PRGHO3

PRGHO4

0.00 í0.05 í0.10

0.10

0.05

PRGHO1

PRGHO2

0.00 í0.05 í0.10

0 5 10 15 20 25 30

/DJ

5 101520

PRGHO5

PRGHO6

Figure 2.12: Trace plot for category AA of model 1 to model 6

Autocorre lation plot tor category A of all models

10 15 20 25 30

,cg:

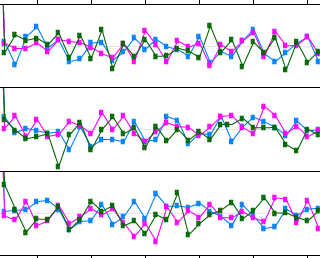
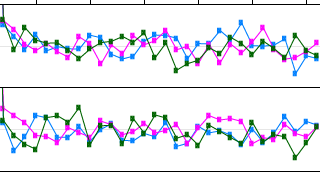
"'

§

s

<(

**0.10**



Ii

..,

w

i

0.05

0.00

-0.05

**-0.10**

**0.10**

0.05

0.00

-0.05

**-0.10**

**0.10**

0.05

0.00

-0.05

**-0.10**

10 15 20 25 30

Lag

### Figure 2.13: Trace plot for category AA of model 1to model 6

Autocorrelation plot for category AA of all models

**10 15 20 25 30**

0.2



w

i

-g

E

"1''

-g

E

..,

1-g'

E

0.0

,cg: 1"''

t

<"( 0.2

-0.2

0.0

-0.2

10 15 20 25 30

Lag

**10 15 20 25 30**

### Figure 2.14: Trace plot for category AA of model 1to model 6

* + 1. **Discussion**

...suggest non-informatiove prior is

And because all of other models tried are not backed by purely theoretical there are no reason to believe one model is better than another.

However I knew that the random number generation used during synthesiing process generatisd normal distribution for all categories. Therefore using a normal prior seem to be the simpliest option. And it is decided that model 2 is our best prior to use for all other simulation testings.

chandola2009anomaly

There are some diff in Anomaly detection techniques, such as supervision and parametrisation.

#### (add these term? important concept but might take too long)

Therefore, the goal of supervised learning is to learn a function that, given a sample of data and desired outputs, best approximates the relationship between input and output observable in the data. Unsupervised learning, on the other hand, does not have labeled outputs, so its goal is to infer the natural structure present within a set of data points.

https://towardsdatascience.com/supervised-vs-unsupervised-learning-14f68e32ea8d <http://home.cc.umanitoba.ca/> gersteia/MBIO7040/06-introStats.html Supervised vs unsupervised

later on we will go into the concept of bayesia, howerver here what we need to know is that bayesian is a form of supervised/unsupervised......

Parametric vs non-parametric

In the literal meaning of the terms, a parametric statistical test is one that makes assumptions about the parameters (defining properties) of the population distribution(s) from which one’s data are drawn, while a non-parametric test is one that makes no such assumptions. In this strict sense, ”non-parametric” is essentially a null category, since virtually all statistical tests assume one thing or another about the properties of the source population(s).

Strength and weakness

no acf, true for simualtion as each day is gererated randomly and assume complete independence

# 2.4 Simulation 2: Size of anomaly

## 2.4.1 Simulation setups

Examining the Counts on the day that we add anomaly using R codes, we can see that anomaly Different anomalies is added to category AA of the synthetic data....etc etc...

> daily 1 . df[ anomalydateloc ,]

day total A B AA AB BA BB

3606 2015 -11 -15 19 7 12 2 5 5 7

> daily 1 . S 10 . df[ anomalydateloc ,] day total A B AA AB BA BB

3606 2015 -11 -15 21 9 12 4 5 5 7

> daily 1 . S 25 . df[ anomalydateloc ,] day total A B AA AB BA BB

3606 2015 -11 -15 24 12 12 7 5 5 7

> daily 1 . S 50 . df[ anomalydateloc ,] day total A B AA AB BA BB

3606 2015 -11 -15 30 18 12 13 5 5 7

> daily 1 . S 100 . df[ anomalydateloc ,] day total A B AA AB BA BB

3606 2015 -11 -15 40 28 12 23 5 5 7

> daily 1 . S 250 . df[ anomalydateloc ,] day total A B AA AB BA BB

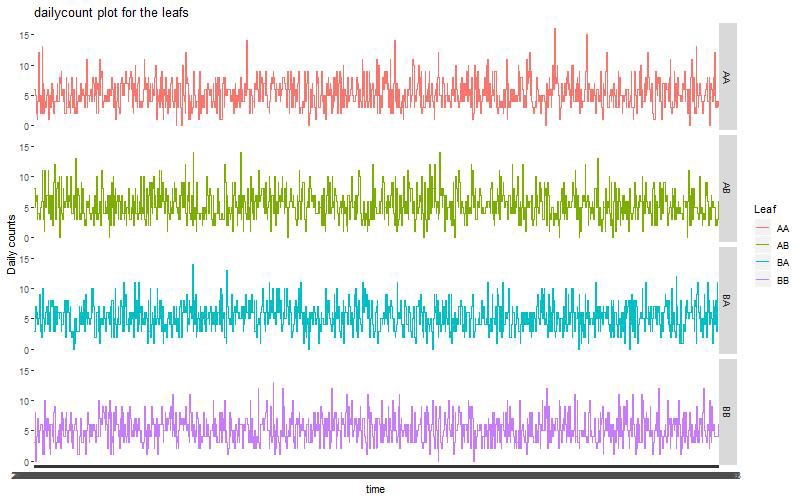
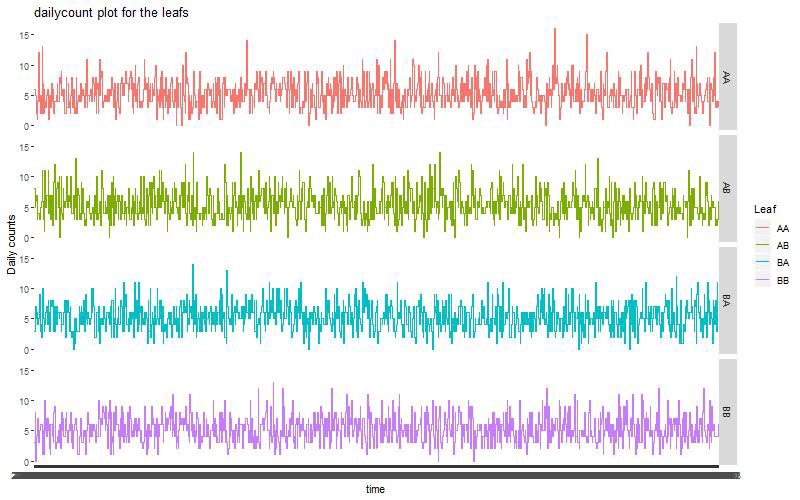
3606 2015 -11 -15 72 60 12 55 5 5 7

> daily 1 . S 500 . df[ anomalydateloc ,] day total A B

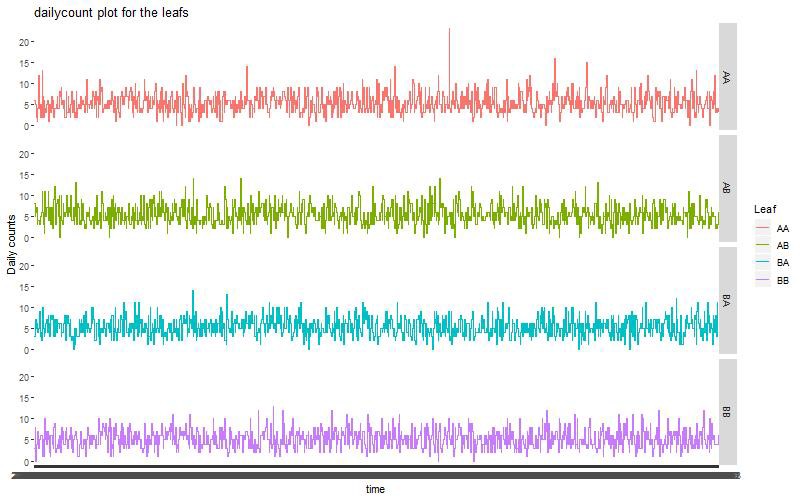
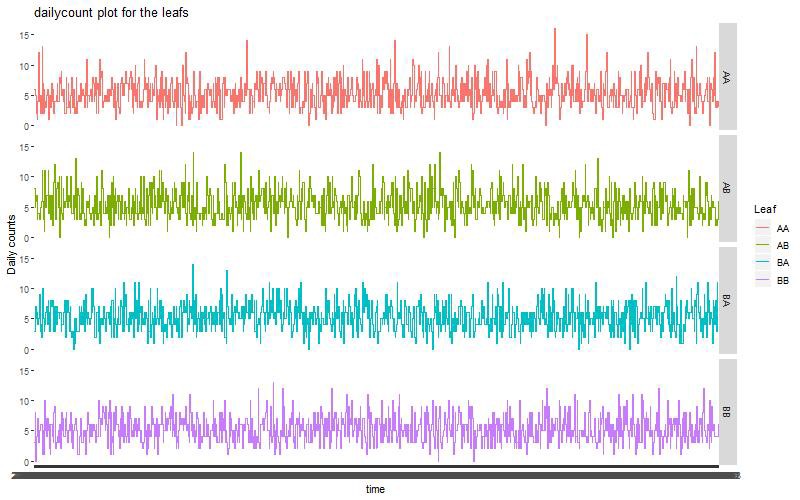
3606 2015 -11 -15

AA AB BA BB

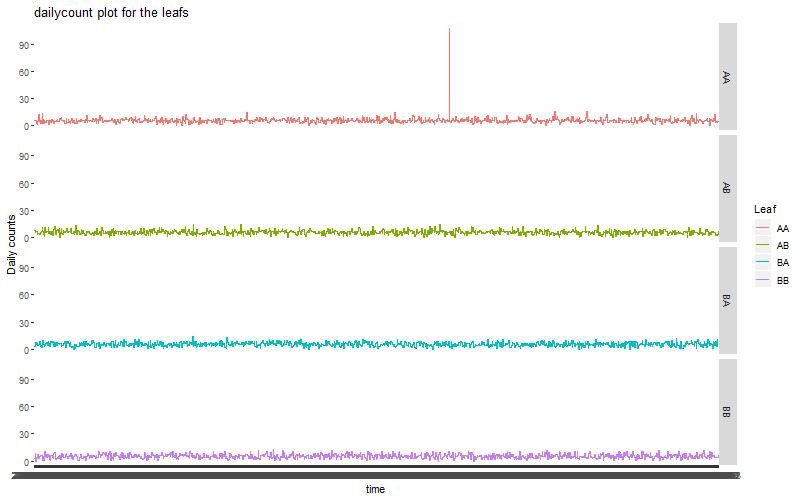
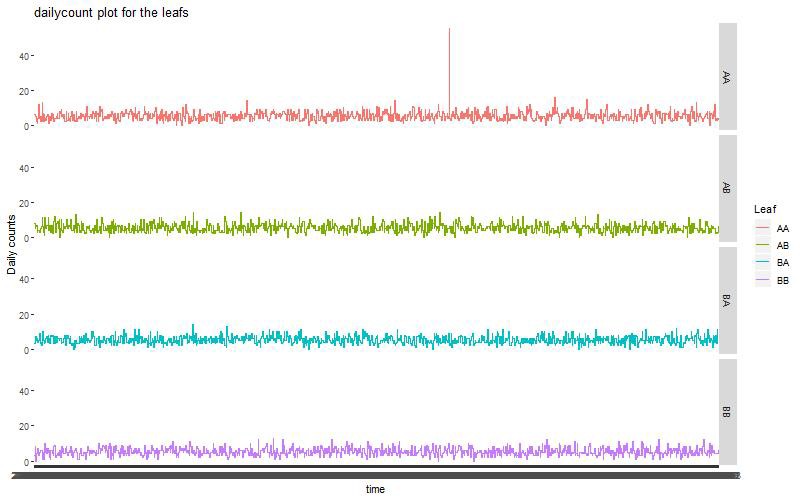
124 112 12 107 5 5 7

Anomaly= 10 Anomaly= 2

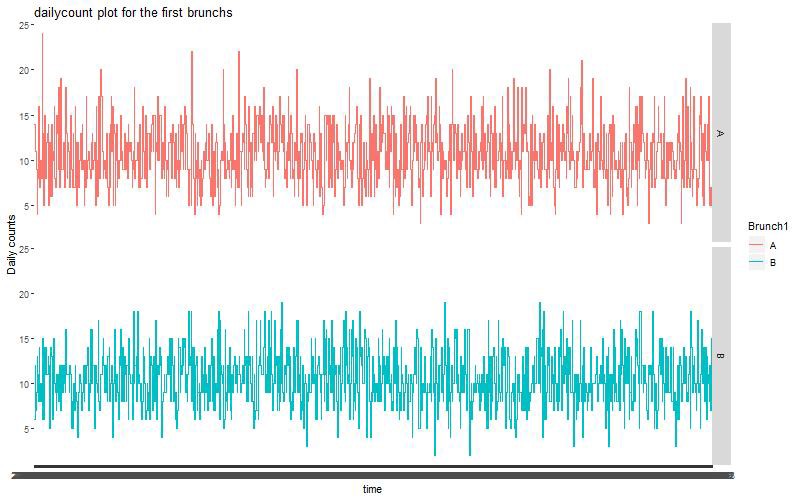
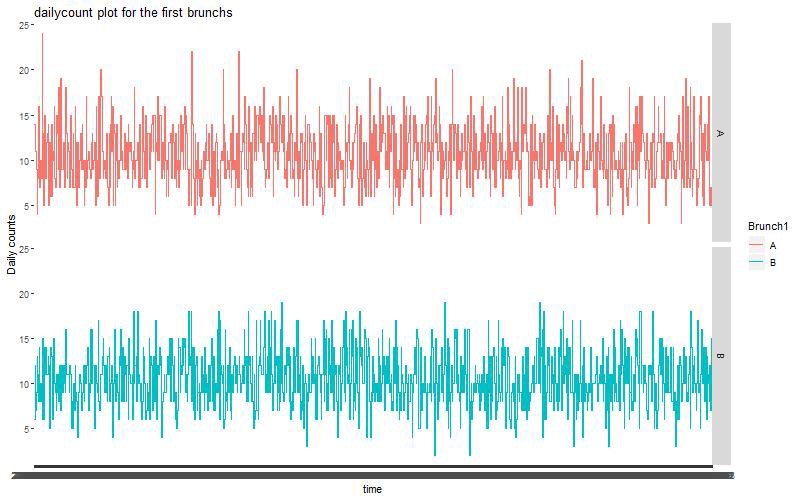


Anomaly= 50 Anomaly= 100

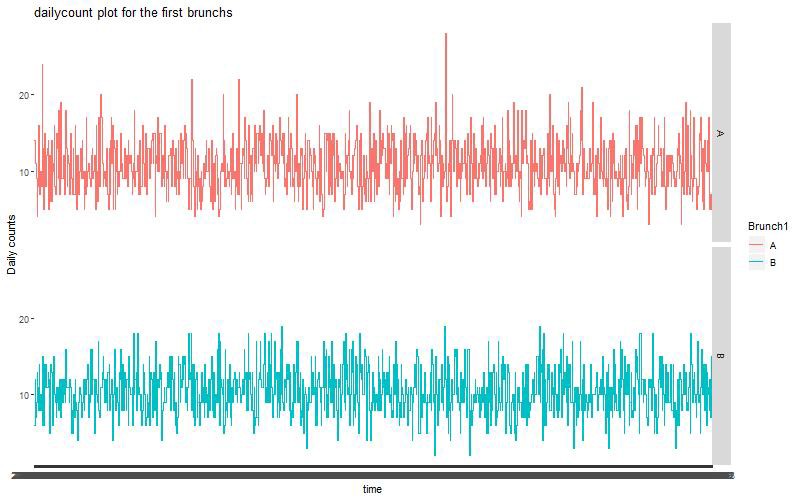
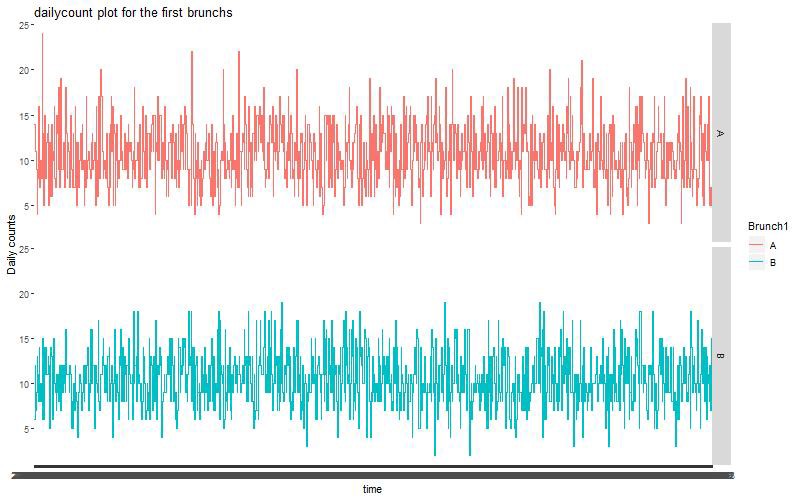


Anomaly’= 250 Anomaly = 500

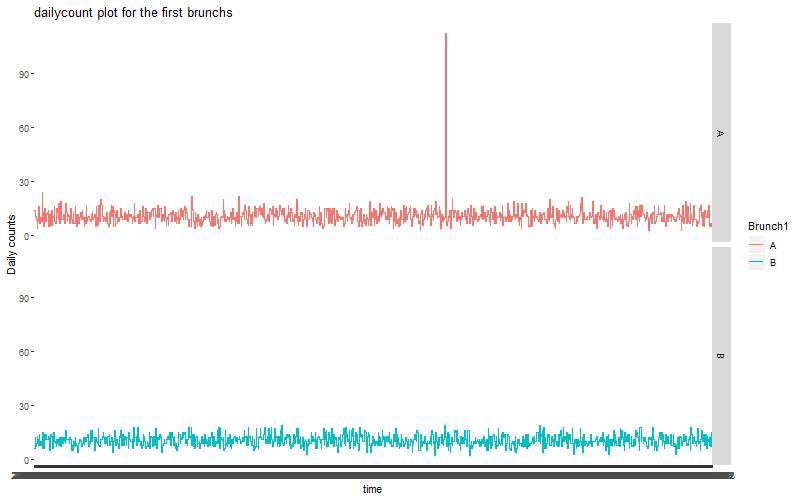
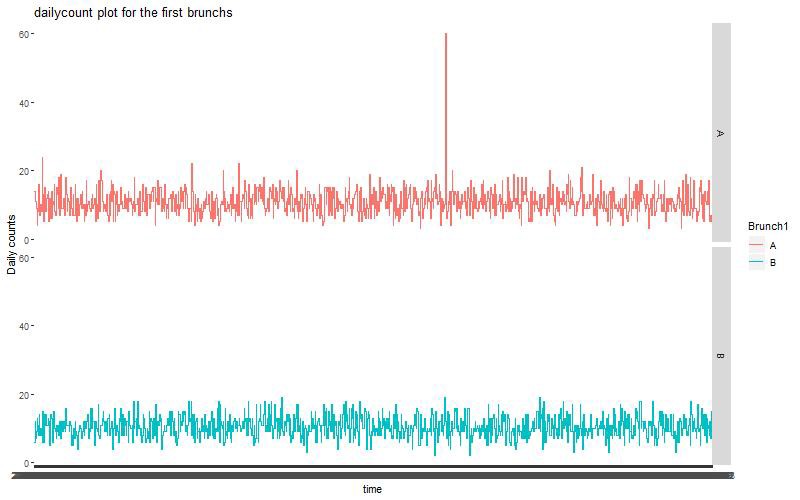
Figure 2.15: Visual inspection of diff t increment of anomalies at level 2 of hierarchy

Anomaly= 10 Anomaly= 2



Anomaly= 50 Anomaly= 100



Anomaly’= 250 Anomaly = 500

Figure 2.16: Visual inspection of diff t increment of anomalies at level 1 of hierarchy

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **2.4.2** | **Results** |  | | | | | |
|  | Anomaly |
|  | (% of total) | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
|  | 0 | -0.466 | 0.188 | 0.437 | 0.545 | 0.955 | 1.556 |
|  | 10 | -1.910 | -0.693 | 0.876 | 1.144 | 1.894 | 6.640 |
|  | 25 | -6.274 | -3.079 | -1.966 | -1.984 | -0.658 | 1.825 |
|  | 50 | -4.352 | -1.185 | -0.038 | -0.385 | 0.823 | 2.420 |
|  | 100 | -5.074 | -1.096 | 0.323 | -0.383 | 1.194 | 1.873 |
|  | 250 | -1.813 | -1.657 | -0.917 | -0.181 | 0.546 | 3.681 |
|  | 500 | -4.416 | -1.678 | -0.613 | -0.875 | 0.430 | 1.401 |

Table 2.5: DIC comparisons between Independent Bayes model and Hierarchical Bayes model (HBM -IBM) for diff t anomalies sizes

see appendix

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| Anomaly0 | 19.1481 | 4.2840 | 11.7152 | 18.8136 | 28.3845 | 1.0000 | 3000.0000 |
| Anomaly10 | 21.4084 | 4.7272 | 13.1528 | 21.0041 | 31.6607 | 1.0000 | 3000.0000 |
| Anomaly25 | 24.4356 | 4.9374 | 16.0687 | 24.0399 | 34.7248 | 1.0000 | 2975.0000 |
| Anomaly50 | 30.2932 | 5.3613 | 20.8844 | 29.8973 | 41.9325 | 1.0000 | 2912.0000 |
| Anomaly100 | 40.4338 | 6.1531 | 29.3330 | 40.0987 | 53.8314 | 1.0000 | 2889.0000 |
| Anomaly250 | 71.9683 | 8.4372 | 56.5531 | 71.6763 | 89.4346 | 1.0000 | 3000.0000 |
| Anomaly500 | 122.4365 | 10.7971 | 101.9952 | 122.2718 | 144.5457 | 1.0000 | 3132.0000 |

Table 2.6: Posterior distributions of diff t models for Total, with diff t incre- ments of added anomalies, and calculated with independent Bayes model

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| Anomaly0 | 19.3114 | 4.4361 | 11.5960 | 18.9555 | 28.9686 | 1.0000 | 2934.0000 |
| Anomaly10 | 21.3883 | 4.5971 | 13.3033 | 21.1152 | 31.2900 | 1.0000 | 2799.0000 |
| Anomaly25 | 24.3800 | 4.9083 | 15.8281 | 24.1057 | 34.9855 | 1.0000 | 3000.0000 |
| Anomaly50 | 30.5664 | 5.4857 | 20.7071 | 30.2800 | 42.3242 | 1.0000 | 2713.0000 |
| Anomaly100 | 40.5794 | 6.2480 | 29.1623 | 40.2459 | 53.0399 | 1.0000 | 2704.0000 |
| Anomaly250 | 72.0177 | 8.2580 | 57.1126 | 71.8129 | 88.7105 | 1.0000 | 2723.0000 |
| Anomaly500 | 121.9641 | 10.3188 | 102.7605 | 121.6646 | 142.4118 | 1.0000 | 2173.0000 |

Table 2.7: Posterior distributions of diff t models for Total, with diff t incre- ments of added anomalies, and calculated with Hierarchical Bayes model

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| Anomaly0 | 7.1758 | 2.7278 | 3.0586 | 6.7391 | 13.6749 | 1.0000 | 2844.0000 |
| Anomaly10 | 9.3215 | 3.1367 | 4.1286 | 8.9552 | 16.3998 | 1.0000 | 2630.0000 |
| Anomaly25 | 12.3410 | 3.5655 | 6.4488 | 11.9857 | 20.3604 | 1.0000 | 3038.0000 |
| Anomaly50 | 18.4421 | 4.3162 | 10.8431 | 18.1170 | 27.7122 | 1.0000 | 2839.0000 |
| Anomaly100 | 28.3032 | 5.3432 | 19.1588 | 27.8652 | 40.0187 | 1.0000 | 3008.0000 |
| Anomaly250 | 58.6906 | 7.3628 | 45.1079 | 58.6335 | 73.7809 | 1.0000 | 2118.0000 |
| Anomaly500 | 104.7585 | 9.5672 | 87.1657 | 104.5580 | 123.4846 | 1.0000 | 2235.0000 |

Table 2.8: Posterior distributions of diff t models for A , with diff t increments of added anomalies, and calculated with Independent Bayes model

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| Anomaly0 | 7.1740 | 2.7263 | 2.8494 | 6.8467 | 13.2835 | 1.0000 | 2559.0000 |
| Anomaly10 | 9.1731 | 3.0292 | 4.1612 | 8.8307 | 15.8837 | 1.0000 | 2961.0000 |
| Anomaly25 | 12.4274 | 3.6049 | 6.3125 | 12.0656 | 20.2620 | 1.0000 | 2703.0000 |
| Anomaly50 | 18.5048 | 4.3162 | 10.9271 | 18.2437 | 28.0090 | 1.0000 | 2580.0000 |
| Anomaly100 | 28.6123 | 5.2341 | 19.2090 | 28.1992 | 39.9182 | 1.0000 | 2370.0000 |
| Anomaly250 | 59.6236 | 7.2375 | 46.4370 | 59.4535 | 74.8256 | 1.0000 | 1940.0000 |
| Anomaly500 | 107.9043 | 9.7689 | 89.7498 | 107.4554 | 129.0404 | 1.0000 | 1338.0000 |

Table 2.9: Posterior distributions of diff t models for A , with diff t increments of added anomalies, and calculated with Hierarchical Bayes model

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| Anomaly0 | 2.3062 | 1.4859 | 0.4009 | 2.0034 | 6.0143 | 1.0000 | 2401.0000 |
| Anomaly10 | 4.2976 | 2.1384 | 1.2137 | 3.8992 | 9.5420 | 1.0000 | 2461.0000 |
| Anomaly25 | 7.4243 | 2.8041 | 2.8709 | 7.0744 | 13.8003 | 1.0000 | 2214.0000 |
| Anomaly50 | 13.3455 | 3.5920 | 7.3160 | 13.0175 | 21.0238 | 1.0000 | 2428.0000 |
| Anomaly100 | 22.4374 | 4.5194 | 14.6914 | 22.0898 | 31.9623 | 1.0000 | 2029.0000 |
| Anomaly250 | 48.8502 | 6.1544 | 37.2238 | 48.5964 | 61.2842 | 1.0000 | 1568.0000 |
| Anomaly500 | 100.5322 | 11.0472 | 79.0267 | 100.3259 | 122.1376 | 1.0200 | 1115.0000 |

Table 2.10: Posterior distributions of diff t models for AA, with diff t increments of added anomalies, and calculated with Independent Bayes model

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| Anomaly0 | 2.1538 | 1.4960 | 0.3505 | 1.7838 | 6.1010 | 1.0000 | 2073.0000 |
| Anomaly10 | 4.3301 | 2.1448 | 1.1920 | 3.9651 | 9.5508 | 1.0000 | 2046.0000 |
| Anomaly25 | 7.6847 | 2.8131 | 3.0945 | 7.3379 | 13.8861 | 1.0000 | 2149.0000 |
| Anomaly50 | 13.6061 | 3.6369 | 7.2589 | 13.3337 | 21.4732 | 1.0000 | 2054.0000 |
| Anomaly100 | 23.0253 | 4.5067 | 14.9764 | 22.7288 | 32.6668 | 1.0000 | 1679.0000 |
| Anomaly250 | 51.9695 | 6.3876 | 40.2516 | 51.8076 | 65.5484 | 1.0100 | 951.0000 |
| Anomaly500 | 96.0807 | 9.1704 | 79.5825 | 95.8225 | 115.7111 | 1.0000 | 951.0000 |

Table 2.11: Posterior distributions of diff t models for AA, with diff t increments of added anomalies, and calculated with Hierarchical Bayes model

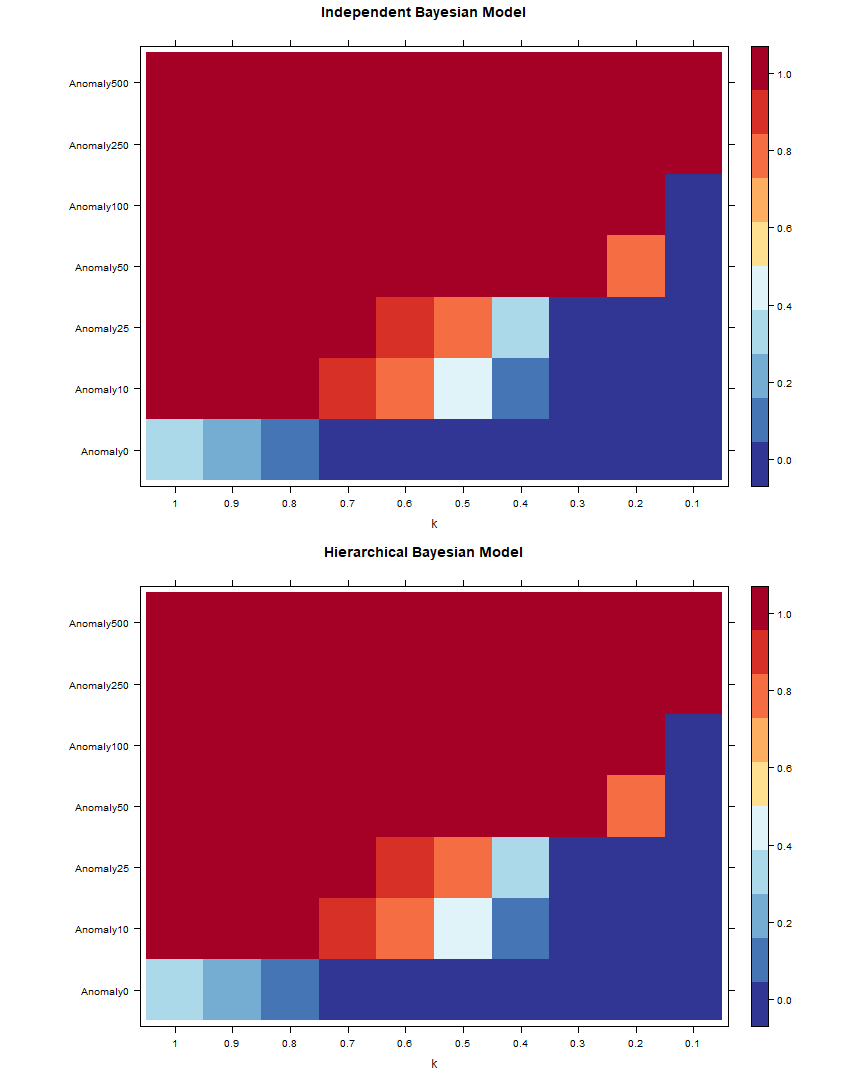


Figure 2.17: total

## 2.4.3 Discussion

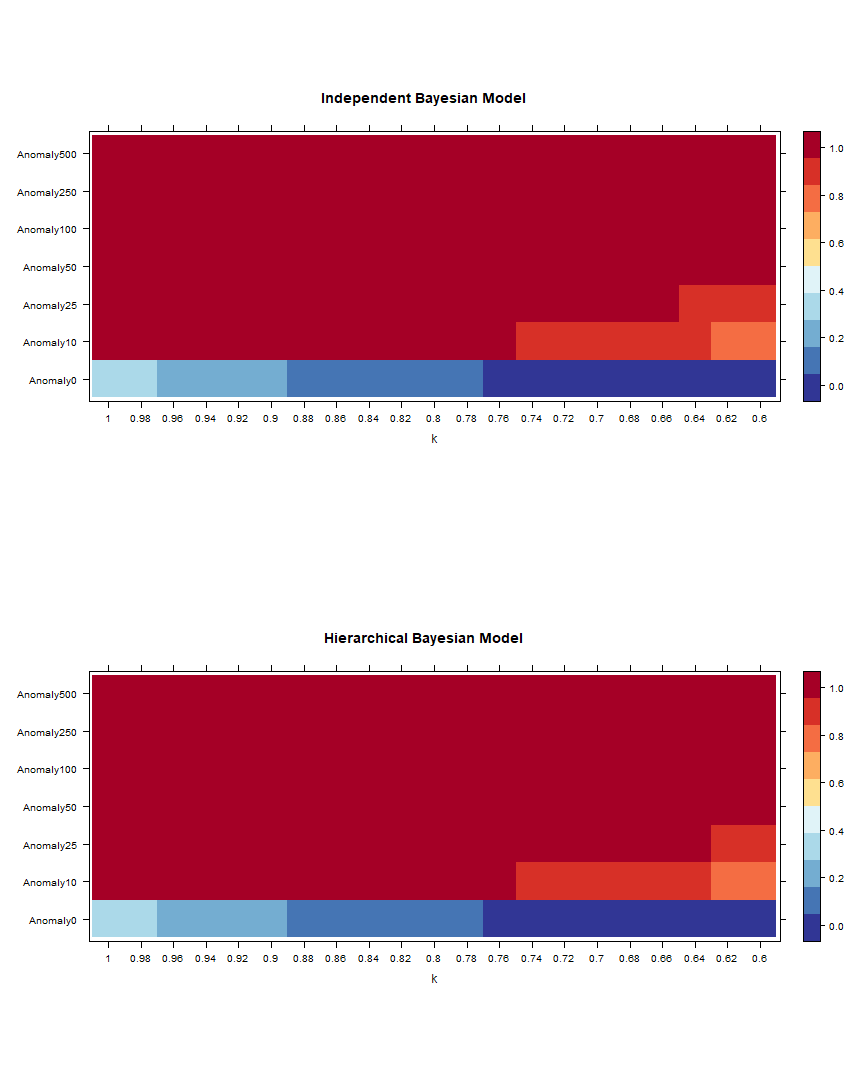


Figure 2.18: total

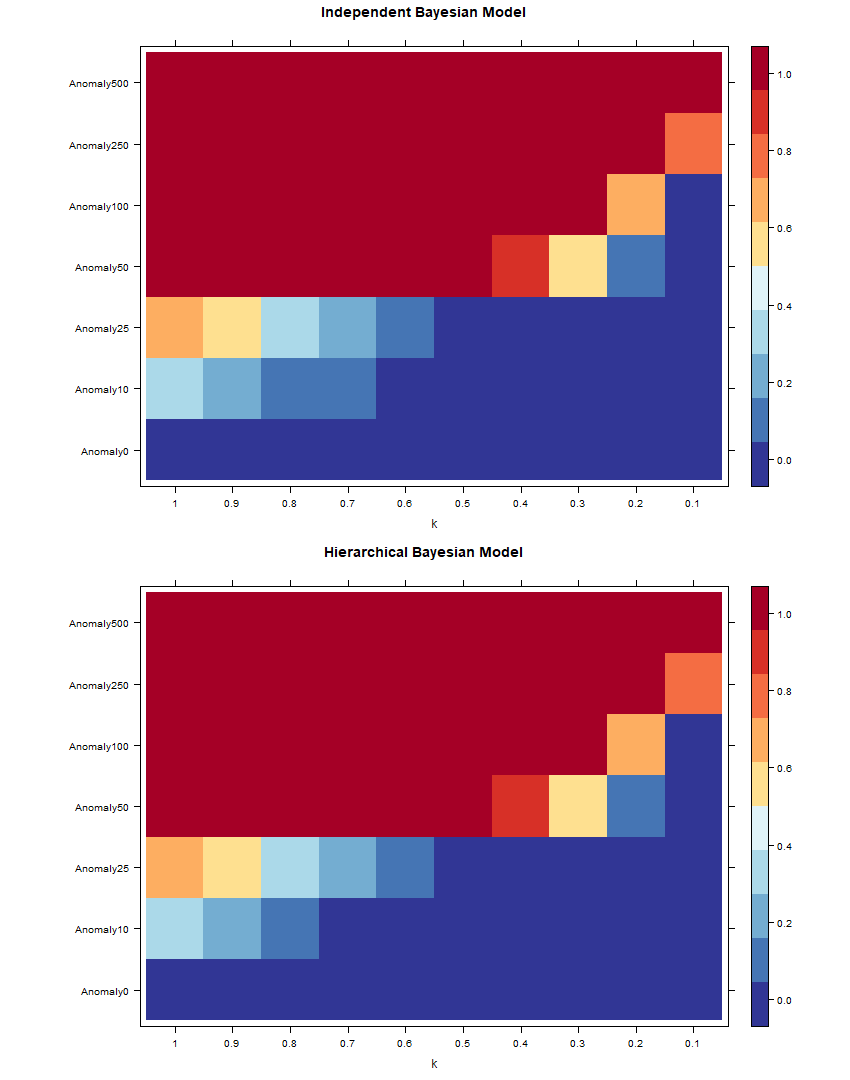


Figure 2.19: A

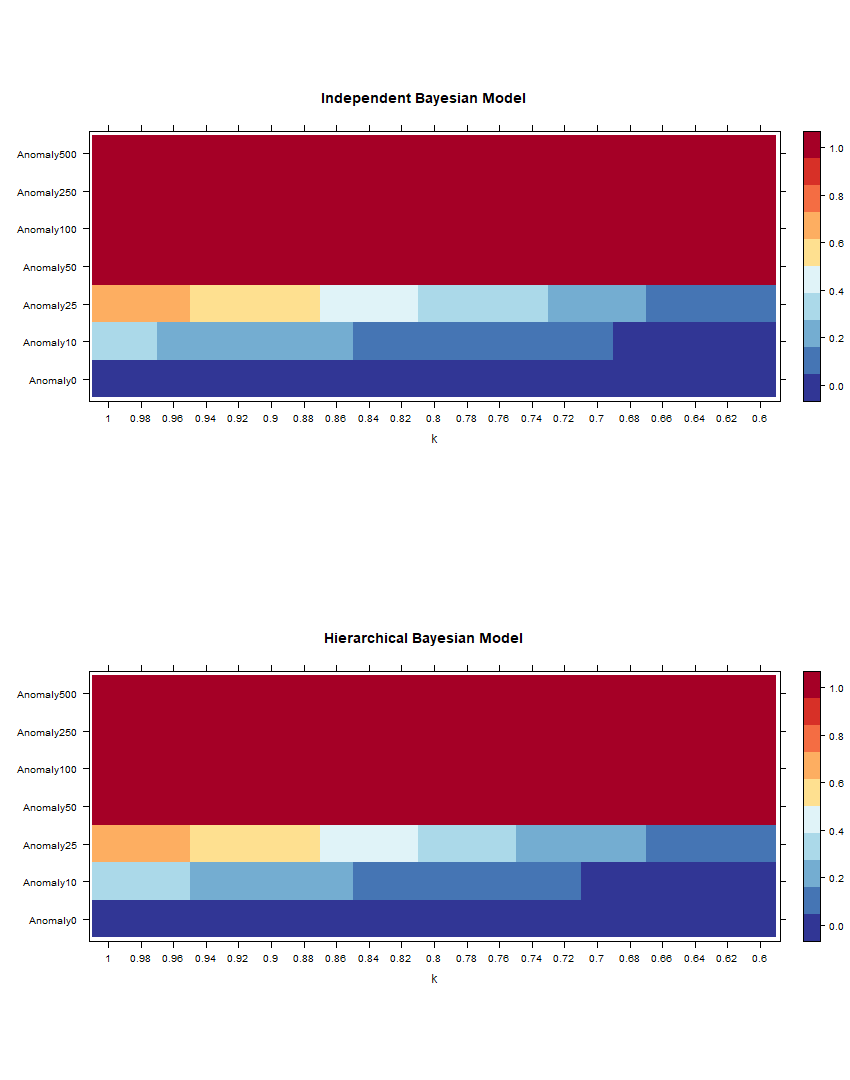


Figure 2.20: A

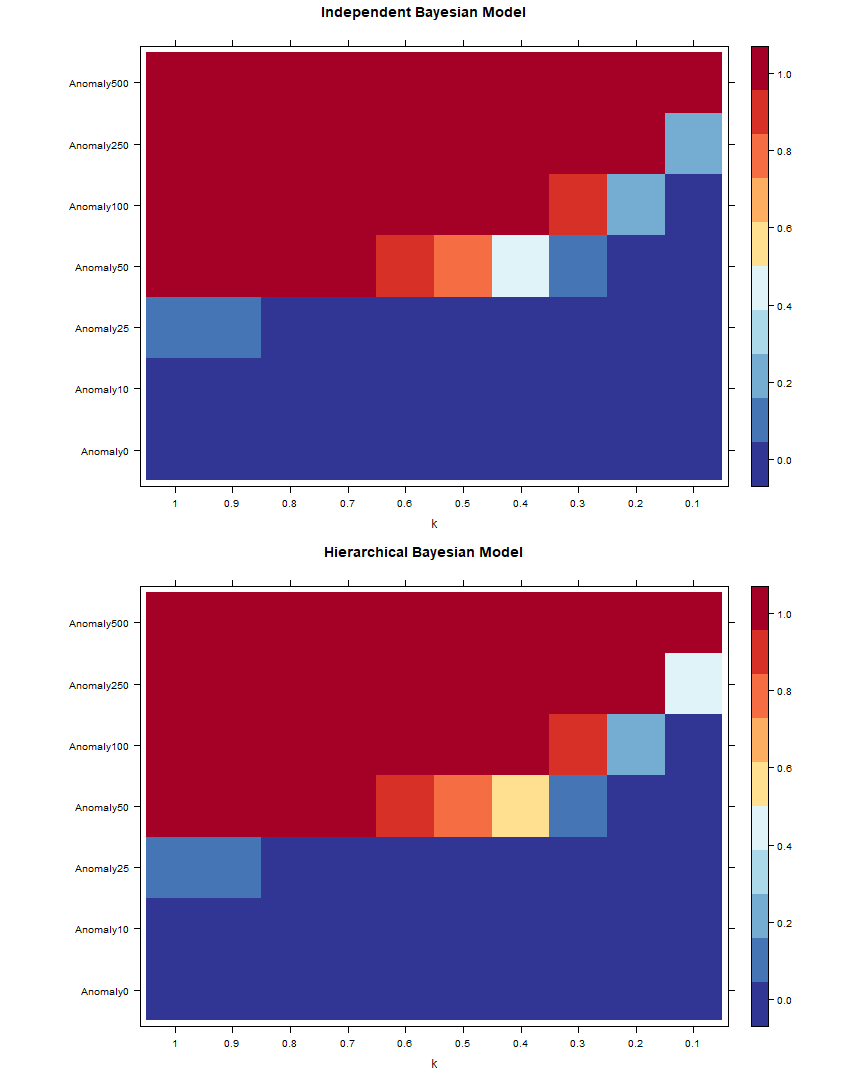


Figure 2.21: AA

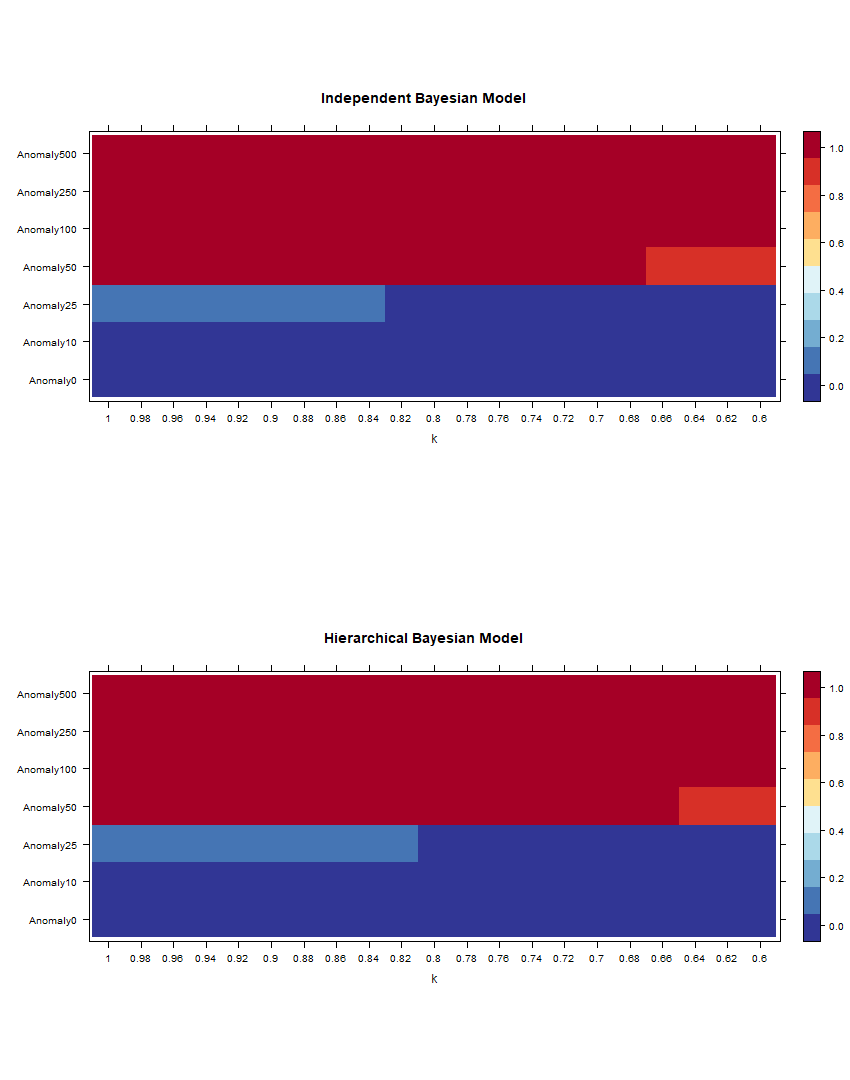
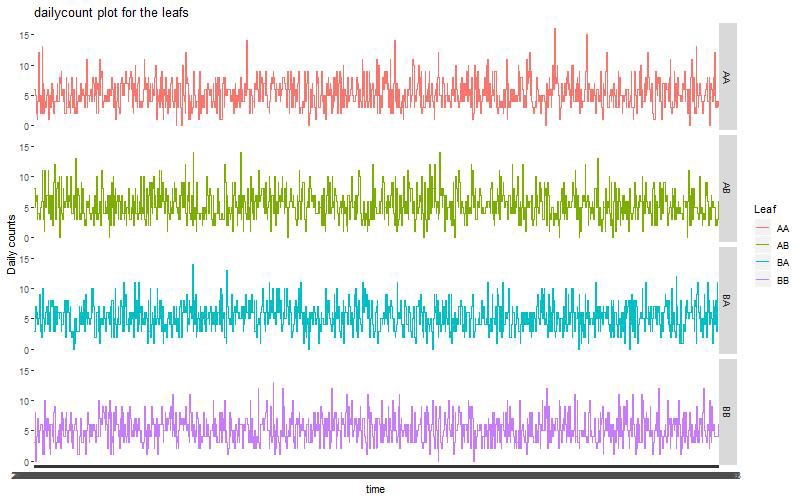
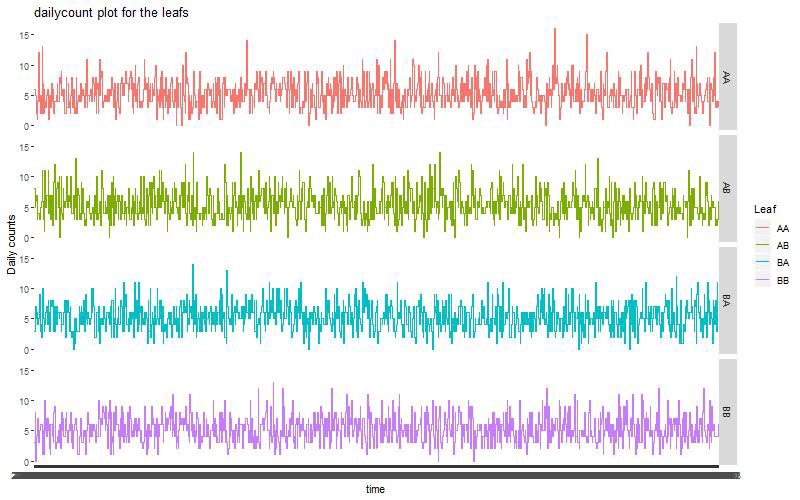


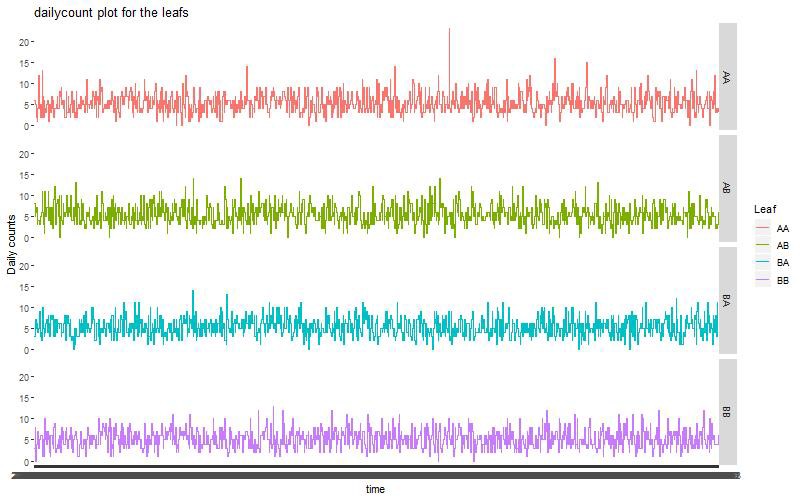
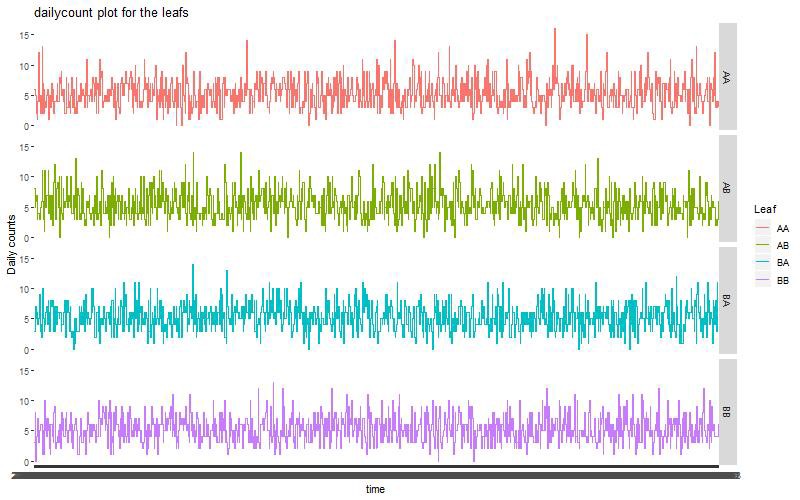
Figure 2.22: AA

# Simulation 3:

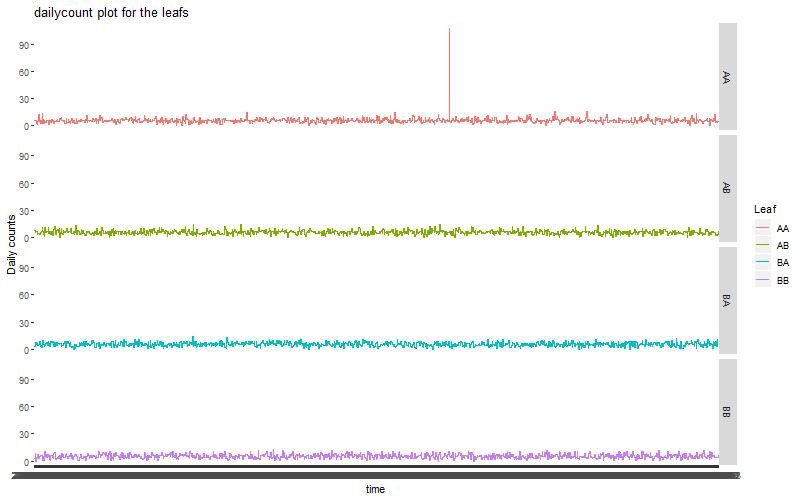
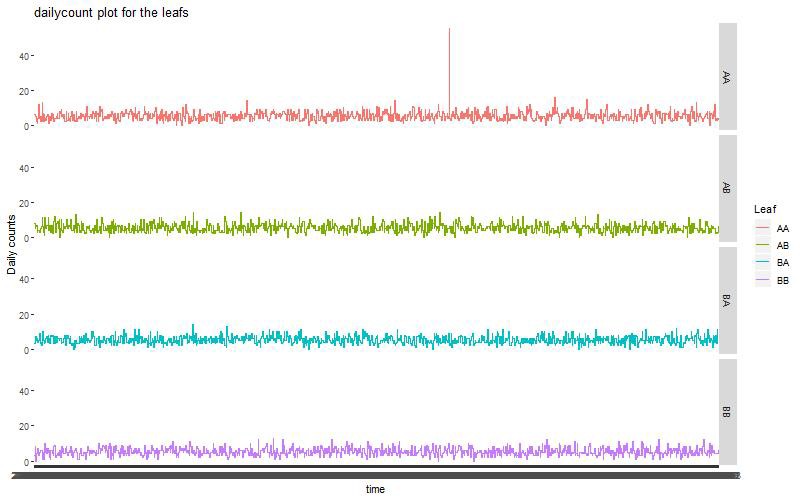
## Simulation setups

Anomaly= 10 Anomaly= 2



Anomaly= 50 Anomaly= 100



Anomaly’= 250 Anomaly = 500

Figure 2.23: Visual inspection of diff t increment of anomalies at level 2 of hierarchy

## Results

still unstable, but multiple testing shows that mean and median mstly negative numbers, which meant hierarchical is better

¿ xtable(postsumtotalh,digits = 4,type = ”latex”, fi = ”plots/sim2/MCMCsumatotalh.tex”,caption

= cap4,label = ”tab:pstprototal”) ¿ xtable(postsumA ,digits = 4,type = ”latex”,

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. |
| 1 | -7.033 | -1.960 | -0.809 | -1.527 | 0.294 | 1.172 |
| 2 | -4.435 | -2.722 | -0.994 | -0.591 | 0.214 | 6.308 |
| 3 | -2.710 | -1.013 | 0.554 | 0.521 | 1.440 | 4.575 |
| 4 | -3.418 | -0.451 | 0.370 | 0.305 | 1.422 | 3.356 |
| 5 | -2.715 | -1.447 | -1.065 | -0.318 | 1.022 | 2.791 |
| 6 | -3.082 | -0.733 | -0.129 | 0.121 | 1.042 | 4.216 |
| 7 | -4.280 | -2.843 | -1.286 | -1.581 | -0.213 | 0.790 |
| 8 | -4.349 | -2.215 | -0.967 | -0.422 | 1.065 | 7.971 |
| 9 | -4.931 | -1.649 | -0.335 | -0.303 | 1.014 | 4.038 |
| 10 | -5.247 | -1.578 | -0.234 | -0.668 | 0.618 | 3.247 |
| 11 | -2.597 | -1.126 | 1.002 | 0.466 | 1.655 | 3.751 |
| 12 | -3.340 | -1.786 | -0.102 | 0.223 | 1.838 | 4.623 |

Table 2.12: DIC comapreisons for brunching number at A

fi = ”plots/sim2/MCMCsumbA.tex”,caption = cap2,label = ”tab:pstprototal”) ¿ xtable(postsumAh ,digits = 4,type = ”latex”, file = ”plots/sim2/MCMCsumbAh.tex”,caption

= cap5,label = ”tab:pstprototal”) ¿ xtable(postsumAA,digits = 4,type = ”latex”, fi = ”plots/sim2/MCMCsumcAA.tex”,caption = cap3,label = ”tab:pstprototal”) ¿

xtable(postsumAAh,digits = 4,type = ”latex”, fi = ”plots/sim2/MCMCsumcAAh.tex”,caption

= cap6,label = ”tab:pstprototal”)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| brunch1 | 21.2367 | 4.7008 | 12.9239 | 20.9013 | 31.1209 | 1.0000 | 2787.0000 |
| brunch2 | 21.4198 | 4.6439 | 13.5516 | 20.9515 | 31.6895 | 1.0000 | 2904.0000 |
| brunch3 | 30.4154 | 5.6673 | 20.1950 | 30.2022 | 42.1275 | 1.0000 | 2877.0000 |
| brunch4 | 13.1540 | 3.5731 | 6.9948 | 12.7660 | 21.1359 | 1.0000 | 3000.0000 |
| brunch5 | 26.4296 | 5.0584 | 17.2806 | 26.1034 | 37.3137 | 1.0000 | 2815.0000 |
| brunch6 | 22.3157 | 4.7457 | 14.0816 | 22.0346 | 32.1219 | 1.0000 | 2881.0000 |
| brunch7 | 17.2269 | 4.1264 | 10.2189 | 16.7961 | 26.4545 | 1.0000 | 2587.0000 |
| brunch8 | 19.1497 | 4.3697 | 11.4783 | 18.8391 | 28.4958 | 1.0000 | 3000.0000 |
| brunch9 | 27.4028 | 5.2818 | 18.0539 | 26.9926 | 38.6413 | 1.0000 | 2854.0000 |
| brunch10 | 18.2805 | 4.2982 | 10.6334 | 18.0157 | 27.3239 | 1.0000 | 3028.0000 |
| brunch11 | 22.2082 | 4.7100 | 13.9406 | 21.8632 | 32.7639 | 1.0000 | 3000.0000 |
| brunch12 | 23.2141 | 4.8762 | 14.6652 | 22.8895 | 33.4077 | 1.0000 | 3090.0000 |

Table 2.13: Posterior distributions of diff t models for Total, with diff t brunch number at A, and calculated with independent Bayes model

## Discussion

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| mu[1] | 21.1342 | 4.6356 | 13.0372 | 20.8723 | 31.0156 | 1.0000 | 2899.0000 |
| mu[2] | 21.4399 | 4.6658 | 13.3765 | 21.0743 | 31.3564 | 1.0000 | 2888.0000 |
| mu[3] | 30.4796 | 5.5918 | 20.4537 | 30.1087 | 42.6573 | 1.0000 | 3069.0000 |
| mu[4] | 13.0428 | 3.6218 | 6.9370 | 12.7050 | 21.1728 | 1.0000 | 2798.0000 |
| mu[5] | 26.4792 | 5.0783 | 17.5450 | 26.1769 | 37.5123 | 1.0000 | 2903.0000 |
| mu[6] | 22.5513 | 4.7242 | 14.3816 | 22.2968 | 32.5958 | 1.0000 | 3000.0000 |
| mu[7] | 17.2175 | 4.1172 | 9.9473 | 16.9176 | 25.9870 | 1.0000 | 3000.0000 |
| mu[8] | 19.2860 | 4.3666 | 11.7869 | 19.0001 | 28.9699 | 1.0000 | 2712.0000 |
| mu[9] | 27.4363 | 5.2930 | 18.1801 | 27.1123 | 38.8600 | 1.0000 | 2719.0000 |
| mu[10] | 18.3153 | 4.2436 | 10.9404 | 18.0915 | 27.1782 | 1.0000 | 2961.0000 |
| mu[11] | 22.2699 | 4.7638 | 13.7242 | 21.8257 | 32.6838 | 1.0000 | 3000.0000 |
| mu[12] | 23.3112 | 4.9101 | 14.5832 | 22.9914 | 33.6660 | 1.0000 | 3006.0000 |

Table 2.14: Posterior distributions of diff t modelsfor Total, with diff t brunch number at A, and calculated with Hierarchiacl Bayes model

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| brunch1 | 10.2222 | 3.2257 | 4.9491 | 9.9313 | 17.3630 | 1.0000 | 3000.0000 |
| brunch2 | 7.2101 | 2.5820 | 3.0653 | 6.9089 | 13.0457 | 1.0000 | 3135.0000 |
| brunch3 | 13.3965 | 3.5903 | 7.1940 | 13.0422 | 21.4547 | 1.0000 | 2843.0000 |
| brunch4 | 5.2344 | 2.2966 | 1.7544 | 4.9212 | 10.8095 | 1.0100 | 2659.0000 |
| brunch5 | 12.2653 | 3.5442 | 6.3801 | 11.9468 | 20.0809 | 1.0000 | 2831.0000 |
| brunch6 | 12.4534 | 3.5617 | 6.4566 | 12.0511 | 20.3078 | 1.0000 | 2826.0000 |
| brunch7 | 11.3345 | 3.3140 | 5.6904 | 11.0780 | 18.6164 | 1.0000 | 2763.0000 |
| brunch8 | 11.3301 | 3.4455 | 5.6405 | 10.9504 | 18.8542 | 1.0000 | 3000.0000 |
| brunch9 | 15.2998 | 3.9278 | 8.5925 | 15.0144 | 24.1089 | 1.0000 | 3000.0000 |
| brunch10 | 10.2282 | 3.2404 | 4.9712 | 9.8590 | 17.4283 | 1.0000 | 2719.0000 |
| brunch11 | 12.2384 | 3.5230 | 6.2765 | 11.9505 | 19.9642 | 1.0000 | 3000.0000 |
| brunch12 | 16.2605 | 4.1472 | 9.1019 | 15.9211 | 25.5411 | 1.0000 | 2614.0000 |

Table 2.15: Posterior distributions of diff t models for A , with diff t brunch number at A, and calculated with independent Bayes model

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| mu[1] | 10.3658 | 3.2745 | 4.9563 | 9.9946 | 17.5591 | 1.0000 | 2480.0000 |
| mu[2] | 7.1699 | 2.7092 | 2.8318 | 6.7969 | 13.3853 | 1.0000 | 2763.0000 |
| mu[3] | 13.5106 | 3.6263 | 7.2349 | 13.2756 | 21.5433 | 1.0000 | 2721.0000 |
| mu[4] | 4.9821 | 2.1601 | 1.7617 | 4.6571 | 9.9815 | 1.0000 | 2848.0000 |
| mu[5] | 12.2966 | 3.5260 | 6.2470 | 11.9827 | 20.1420 | 1.0000 | 2691.0000 |
| mu[6] | 11.9993 | 3.3669 | 6.1917 | 11.7362 | 19.3448 | 1.0000 | 2515.0000 |
| mu[7] | 11.1468 | 3.3194 | 5.5362 | 10.8140 | 18.2800 | 1.0000 | 2791.0000 |
| mu[8] | 10.9205 | 3.2059 | 5.5478 | 10.6286 | 17.9246 | 1.0000 | 2596.0000 |
| mu[9] | 15.1743 | 3.7615 | 8.6417 | 14.8581 | 23.3316 | 1.0000 | 2371.0000 |
| mu[10] | 9.4961 | 2.9241 | 4.6206 | 9.1667 | 16.0461 | 1.0000 | 2730.0000 |
| mu[11] | 11.7279 | 3.3009 | 5.9251 | 11.4201 | 18.8472 | 1.0000 | 2389.0000 |
| mu[12] | 15.4642 | 3.7186 | 9.0081 | 15.1212 | 23.5432 | 1.0000 | 2361.0000 |

Table 2.16: Posterior distributions of diff t modelsfor A , with diff t brunch number at A, and calculated with Hierarchiacl Bayes model

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| brunch1 | 10.3390 | 3.3313 | 4.9841 | 9.9711 | 18.0423 | 1.0000 | 3015.0000 |
| brunch2 | 2.2864 | 1.4796 | 0.4436 | 1.9694 | 5.9125 | 1.0000 | 2383.0000 |
| brunch3 | 4.3819 | 2.1286 | 1.2499 | 4.0705 | 9.5828 | 1.0000 | 2021.0000 |
| brunch4 | 2.2971 | 1.5500 | 0.3511 | 1.9390 | 6.0469 | 1.0000 | 1776.0000 |
| brunch5 | 4.3729 | 2.0611 | 1.2799 | 4.0873 | 9.4255 | 1.0000 | 1568.0000 |
| brunch6 | 5.1595 | 2.0347 | 1.8054 | 4.9425 | 9.5605 | 1.0000 | 1448.0000 |
| brunch7 | 3.2759 | 1.6852 | 0.7437 | 3.0322 | 7.3844 | 1.0000 | 1259.0000 |
| brunch8 | 3.2901 | 1.6666 | 0.7674 | 3.0726 | 7.0195 | 1.0000 | 1255.0000 |
| brunch9 | 3.1983 | 1.5386 | 0.8286 | 3.0263 | 6.7680 | 1.0000 | 1424.0000 |
| brunch10 | 4.3662 | 1.5571 | 1.6687 | 4.1853 | 7.6161 | 1.0100 | 1110.0000 |
| brunch11 | 2.1852 | 1.3714 | 0.2855 | 1.9361 | 5.4025 | 1.0000 | 1077.0000 |
| brunch12 | 2.9119 | 1.4034 | 0.7174 | 2.6996 | 6.1571 | 1.0000 | 947.0000 |

Table 2.17: Posterior distributions of diff ent models for AA, with diff t brunch number at A, and calculated with independent Bayes model

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | mean | sd | 2.5% | 50% | 97.5% | Rhat | n.eff |
| mu[1] | 10.3563 | 3.2788 | 4.9509 | 10.0156 | 17.7970 | 1.0000 | 2540.0000 |
| mu[2] | 2.1951 | 1.5255 | 0.3203 | 1.8362 | 5.9132 | 1.0000 | 1978.0000 |
| mu[3] | 4.5592 | 2.2061 | 1.2605 | 4.2379 | 9.5741 | 1.0000 | 1910.0000 |
| mu[4] | 2.2663 | 1.5396 | 0.3203 | 1.9350 | 5.9029 | 1.0000 | 1540.0000 |
| mu[5] | 4.4613 | 1.9943 | 1.3097 | 4.2350 | 8.8343 | 1.0000 | 1354.0000 |
| mu[6] | 5.2129 | 2.0225 | 1.8571 | 5.0210 | 9.6588 | 1.0000 | 1023.0000 |
| mu[7] | 3.3193 | 1.6606 | 0.7695 | 3.0899 | 7.1315 | 1.0000 | 1053.0000 |
| mu[8] | 3.3147 | 1.6493 | 0.7960 | 3.0817 | 7.1798 | 1.0200 | 1023.0000 |
| mu[9] | 3.2724 | 1.6032 | 0.7633 | 3.0402 | 6.8838 | 1.0000 | 711.0000 |
| mu[10] | 4.4763 | 1.6416 | 1.7625 | 4.3287 | 8.0546 | 1.0100 | 737.0000 |
| mu[11] | 2.2327 | 1.2427 | 0.3545 | 2.0587 | 4.9807 | 1.0000 | 952.0000 |
| mu[12] | 3.0880 | 1.4240 | 0.8039 | 2.9188 | 6.2747 | 1.0000 | 816.0000 |

Table 2.18: Posterior distributions of diff t modelsfor AA, with diff t brunch number at A, and calculated with Hierarchiacl Bayes model

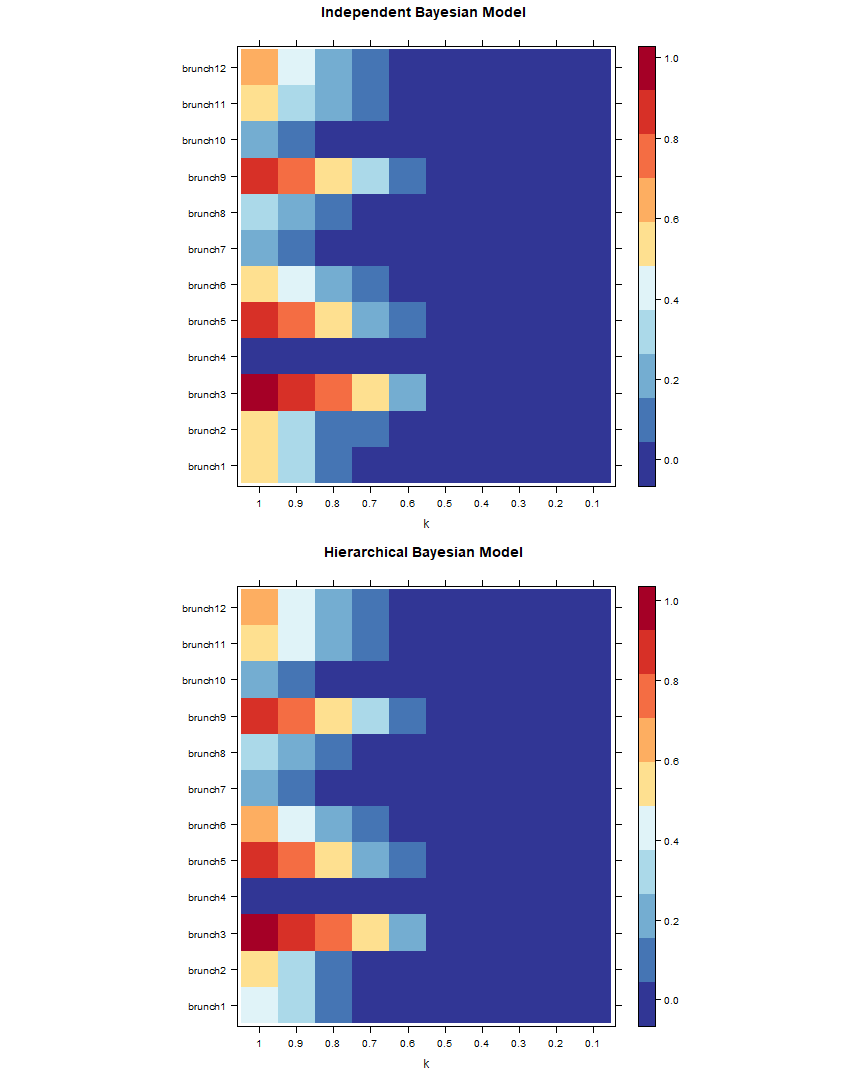


Figure 2.24

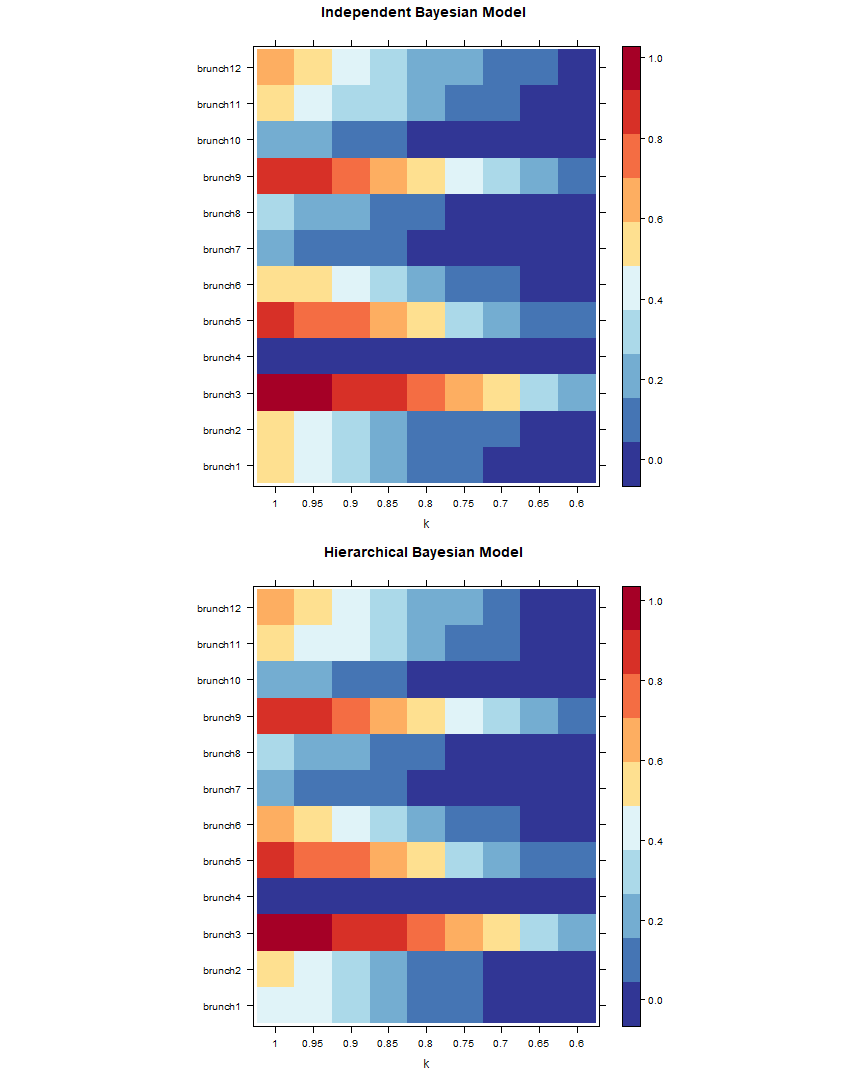


Figure 2.25

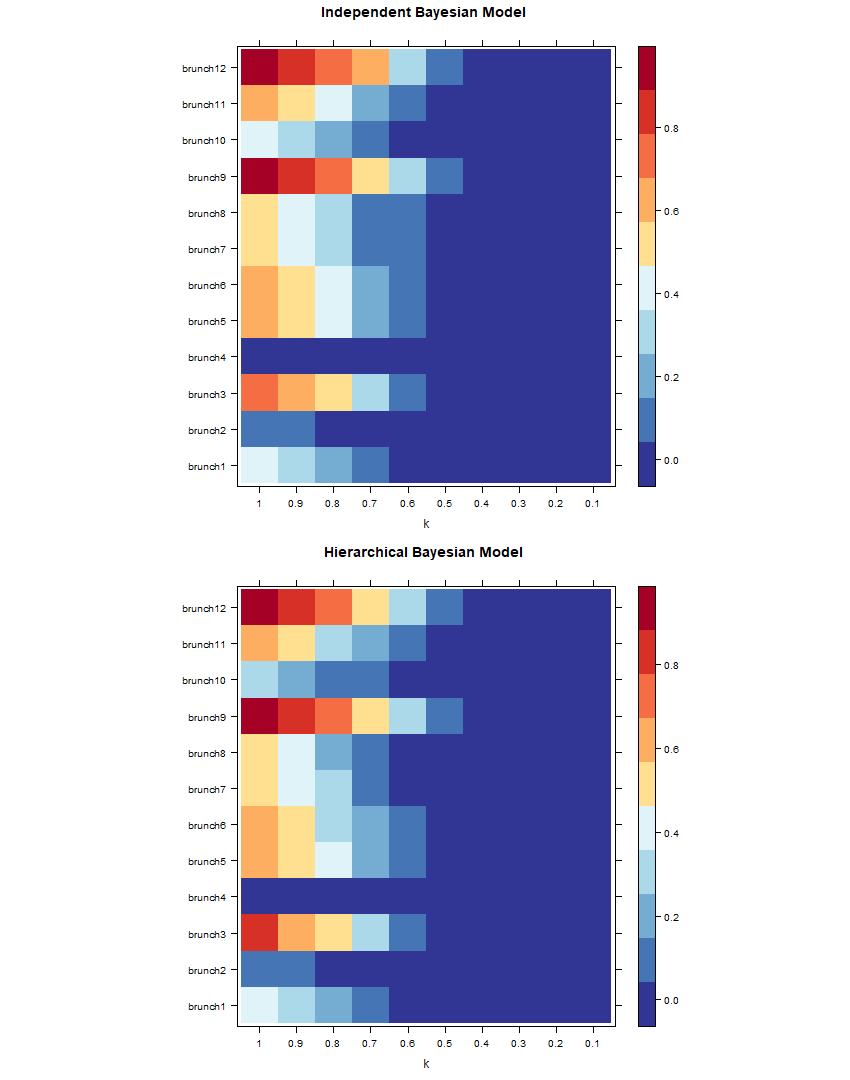


Figure 2.26

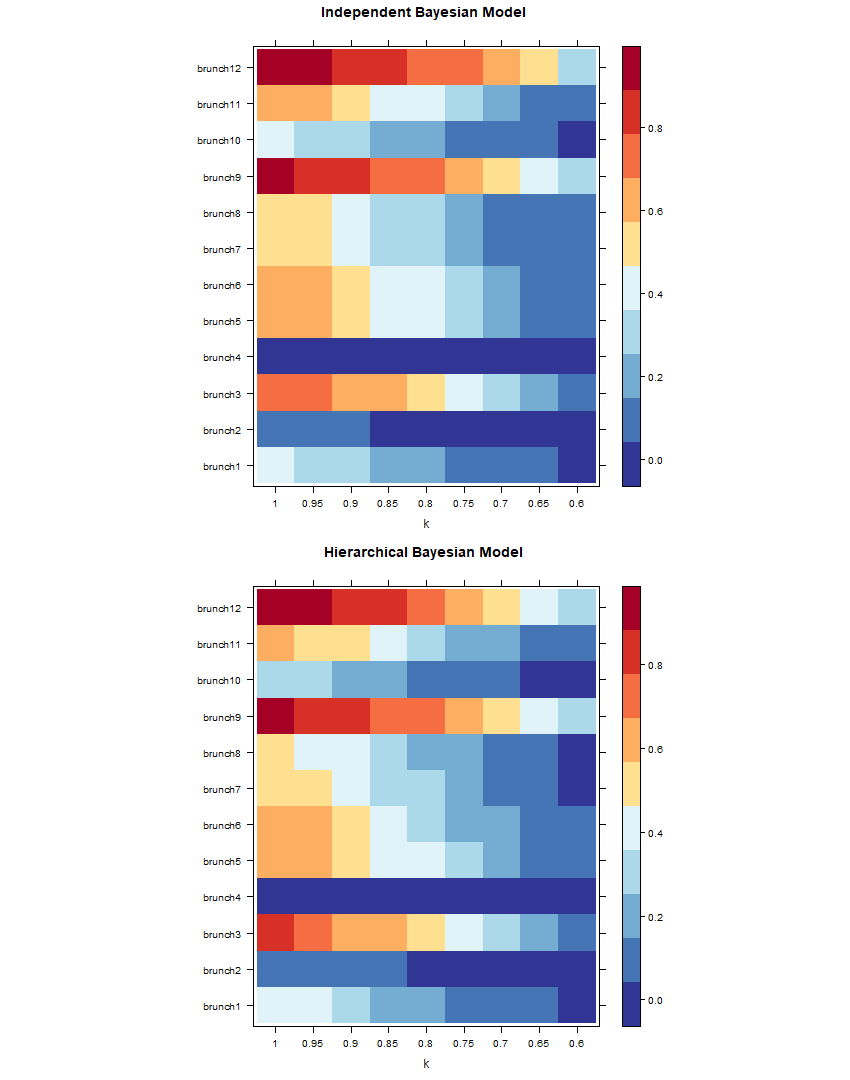


Figure 2.27

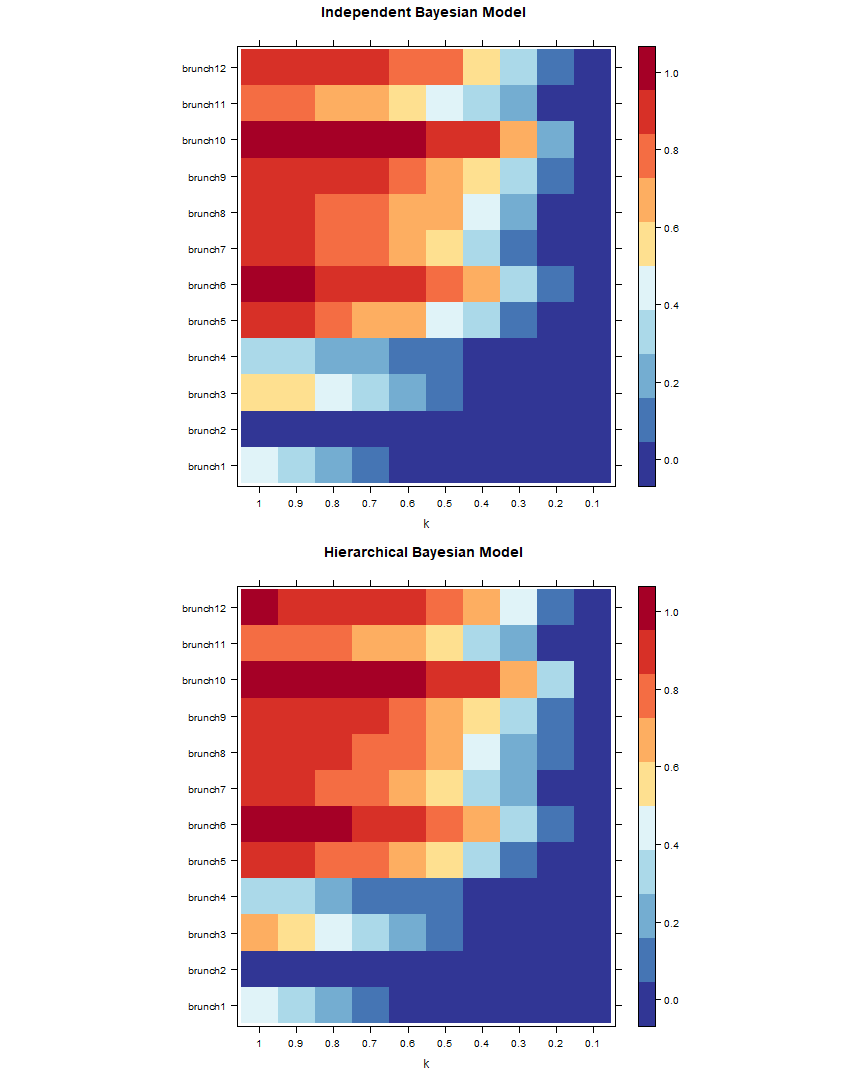


Figure 2.28

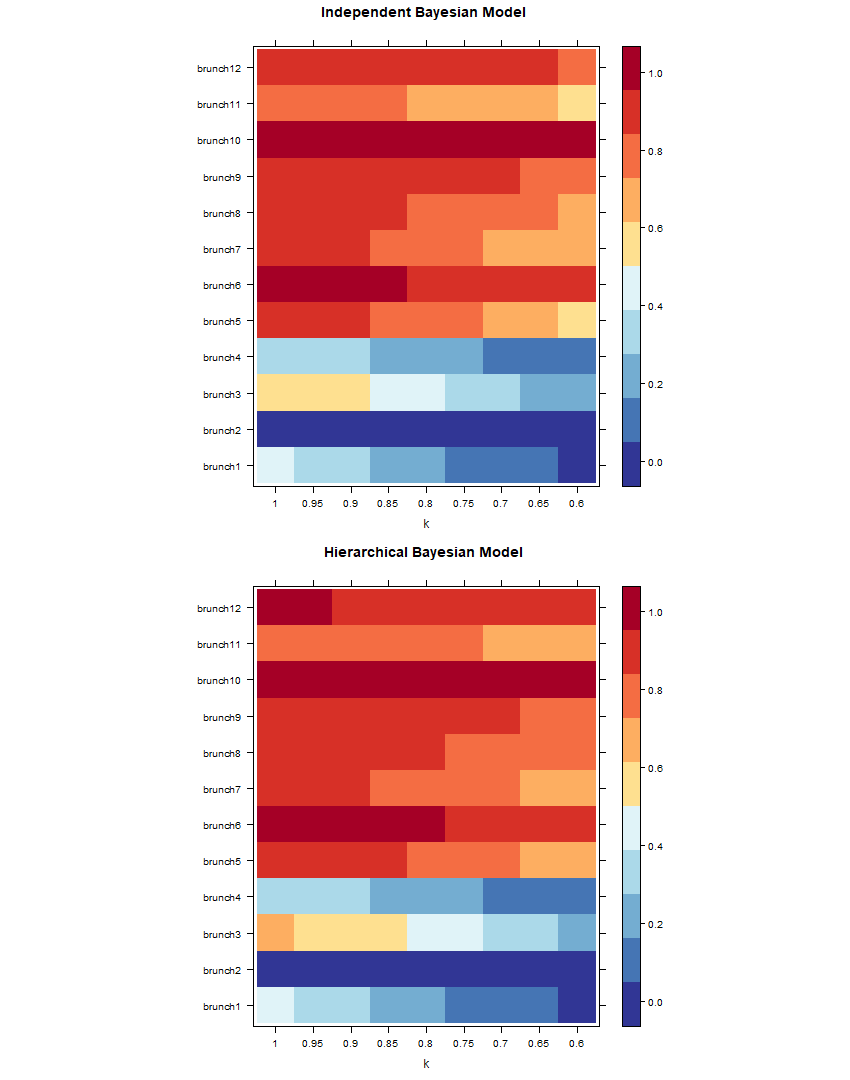


Figure 2.29