Survival of the Rickest

Data Analysis of Survival Rates in Rick and Morty

anonymous

1 General information

For this assignment, no AI tools have been used. Please note that for digit/precision control, the number of digits from the used functions were used as digit control was already performed there. No custom digit control was applied as we have no custom implementations that require this.

2 Introduction

In recent years, "Rick and Morty" has transcended its status as a mere animated television show to become a cultural zeitgeist. The series is renowned for its over-the-top narrative, populated by a cast of zany characters. One distinguishing feature of the show is its tendency for pushing the boundaries, often involving characters in outrageous scenarios and plenty of death. In this project, our objective is to delve into the intriguing world of "Rick and Morty" and unravel the patterns surrounding character survival, from those meeting a swift demise to those enduring until the bitter end.

The sheer diversity in character survival rates within the show necessitates a comprehensive exploration of the myriad variables that influence these outcomes. From the absurd to the profound, the survival trajectories of characters in "Rick and Morty" are anything but predictable. In this project, our models are tasked with navigating the complex landscape of factors that contribute to the fate of characters. Two distinct modeling approaches — normal and hierarchical — have been employed to forecast the survival durations of various species. In addition, we compare the performance of these two temporal models with one that is based on non-temporal predictions.

In this report, we will first visualise our data, and how we derived and implemented the priors and models from this information. The next parts are then all focused on the assessment and comparison of the model performance. This includes, among others, the investigation of convergence diagnostics, posterior predictive checks, a sensitivity analysis to different priors and a model comparison, including the predictive performance assessment. The report will then be concluded with a discussion on the results, what was learned from the data analysis and a self-reflection on what we learned from the project.

3 Data and Analysis Problem

The data was acquired through the Rick and Morty API (https://rickandmortyapi.com/), a comprehensive resource encompassing character information from the first 3 seasons. Our focus centers on key attributes such as their inaugural and final episodes, along with their prevailing survival status. Despite the abundance of online resources, it is noteworthy that our project presents a unique analysis, as there is an absence of prior studies delving into the nuanced aspects of character survival using the available dataset.

Since the data is only available through an API, we had to do some extensive pre-processing to get the desired dataframe. However, since this would go beyond the scope of this report, we have not included the code here.

The data contains information for each episode about the characters who appeared there, their species and gender, and how many episodes they survived up to that episode. This information is also shown in the plot below, with the plot showing the development of the survived episode count per character. Here we can see that some characters have a much higher number of survived episodes, which presumably indicates that they are main characters, and thus should survive all throughout.
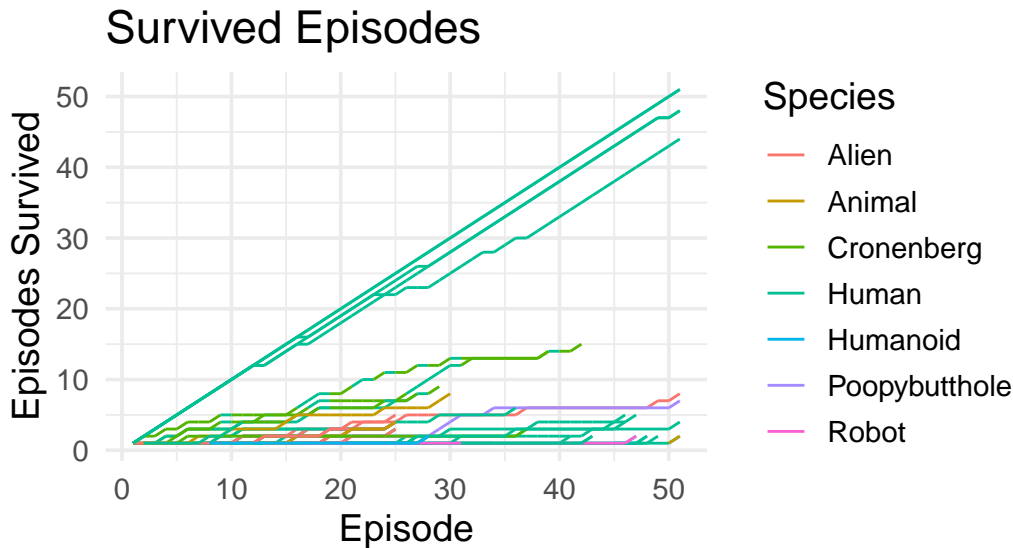
As already motivated in the introduction, our analysis will focus on predicting the survived episodes for different characters of a given species over the episodes covered in this dataset (50 in total). This will be done using two models, one hierarchical and one non-hierarchical. The overall objective of this data analysis is to compare and test the suitability of these two models and to compare their performance. In addition to predicting the survival rate per episode and per character, we also want to compare the performance of the two models mentioned with a model that predicts the total number of survived episodes in a non-temporal fashion per character. This also allows us to verify and compare our modeling approach to find the best possible analysis setup for this problem. The exact models used to accomplish this task will be described in the next section.

4 Model Descriptions

For our project and analysis, we are using three models: all use the lognormal observation model, which allows only positive values which is suited for our problem as the number of survived episodes can only be positive and larger than 0. This is also why we used an intercept of one in the formula of all models. The formula is also what differentiates the models. In our first, non-hierarchical model, we used only the episode number and species as predictors while in the hierarchical model with time series, we added a group effect from the different characters ("names") towards the episode number. This allows us to account for individual changes in the number of survived episodes of certain characters inside the species, for instance due to their appearance in only certain episodes, if they are a friend to the main characters or if they are a main character themselves. Along our two main models, we added a third, hierarchical model without a time series element, because as discussed below we saw that there were potential problems with the episodic nature of our data. In this model, we kept species and gender as our predictors. The choice to add gender as a predictor when it wasn't one for the previous models, comes from the fact that we wanted to have more group separation because there was no time series element. For the group effect, we specified a random intercept (1 being the intercept term) for the two remaining terms, that is species and gender. That means that we want to investigate whether there is some group level patterns that have an effect on the number of episodes survived, in other words, whether a certain species or gender has a specific common survival pattern. The three formulas of the models can now be seen below:

$$episodes\_survived \sim 1 + episode\_number + species$$
$$episodes\_survived \sim 1 + episode\_number + species + (episode\_number|name)$$
$$episodes\_survived \sim 1 + species + gender + (1|species) + (1|gender)$$

How the models were implemented in R will be shown in a later section.

5 Description of Priors

To get a better grasp of what our priors should be we watched the first three seasons of Rick and Morty. For the weakly informative priors, we set priors for both the episode number and the different species in order to account for the various predictors in the model formulas. All of the priors are normal priors, also because our observation model is lognormal, with zero mean and a logarithmic standard deviation that depends on the specific predictor. For the episode number, through our watching and research we looked at how much the survival count of a character can rise overall during an episode and decided that this value is probably pretty low as most characters do not appear in a specific episode, meaning that their survival count also does not increase. Therefore, we set this prior to

$$episode\_number \sim normal(0, log(2))$$

For the priors of the different species, it was clear from our research that the survived episodes might deviate between the species. In general, robotic characters, i.e. robots and humanoids, have usually a very low survival rate of only one or two episodes. For this reason, we will take these two as a reference for the priors of the other species. In contradiction to the robotic characters, humans usually have a much higher survival rate as they appear in many more episodes and also include the main characters, giving them a higher standard deviation. Also, animals tend to have a higher survival rate, even though they appear in less episodes. The same applies to the "Cronenberg" and "Poopybutthole" species, which appear in even less episodes. All of this gives us the following priors for the species:

$$human \sim normal(0, log(8))$$
$$animal \sim normal(0, log(4))$$
$$cronenberg \sim normal(0, log(3))$$
$$poopybutthole \sim normal(0, log(3))$$

As for the final model, we added priors for the genders. In the Rick and Morty universe, there are three genders (male, female, genderless) as well as an 'unknown' category. Although our prior information on the mortality of each gender is not very informed, we can state that after removing the episode number feature the significance of gender in the distribution heightens. The reasoning for that is that certain episodes include more gender biases depending on the location of the episode (e.g. an all-female planet), but now it is not grouped per episode anymore. From watching the episodes we could see that a lot of the military characters that are killed off in mass are male, so we can infer that male characters have a higher mortality rate, but for the other genders we will simply assume a slightly lower mortality. As we always need one of the categories to be the baseline for the priors, we have chosen female as our baseline. The three additional priors in the formula look like the following:

$$male \sim normal(0, log(2))$$
$$genderless \sim normal(0, log(4))$$
$$unknown \sim normal(0, log(4))$$

The concrete R implementations will be shown in the next section.

6 R-Implementation of the Priors and Models

In the following, the concrete implementation of the models and respective prior is shown as introduced and explained in the previous two sections. We chose brms as our model fit interface. The prior is shown first, which is then followed by the code for the models. Please refer to the cleaned BRMS code for further details on the concrete implementation. Please note that we use BRMS instead of Stan.

```r
# Logarithmic prior used in the models for the different predictors
priors <- c(
  prior(normal(0, log(2)), coef = "episode_number"),
  prior(normal(0, log(8)), coef = "speciesHuman"),
  prior(normal(0, log(4)), coef = "speciesAnimal"),
  prior(normal(0, log(3)), coef = "speciesCronenberg"),
  prior(normal(0, log(3)), coef = "speciesPoopybutthole")
)
```

```r
# Non-Hierarchical lognormal model
f1 <- brms::brm(
  episodes_survived ~ 1 + episode_number + species,
  data = rickmorty,
  family = "lognormal",
  prior = priors,
  file = "model/f1"
)
```

```r
# Hierarchical model with the inclusion of the group effect
f2 <- brms::brm(
  episodes_survived ~ 1 + episode_number + species + (episode_number|name),
  data = rickmorty,
  family = "lognormal",
  prior = priors,
  file = "model/f2"
)
```

```r
# Non-time series hierarchical model with creation of the categorical data
# and respecrive prior
rickmorty_cat <- rickmorty[c('name', 'species', 'gender', 'last_episode')]
rickmorty_cat <- distinct(rickmorty_cat)

priors_cat <- c(
  prior(normal(0, log(8)), coef = "speciesHuman"),
  prior(normal(0, log(4)), coef = "speciesAnimal"),
  prior(normal(0, log(3)), coef = "speciesCronenberg"),
  prior(normal(0, log(3)), coef = "speciesPoopybutthole"),
  prior(normal(0, log(2)), coef = "genderMale"),
  prior(normal(0, log(4)), coef = "genderGenderless"),
  prior(normal(0, log(4)), coef = "genderunknown")
)
```

```r
f3 <- brms::brm(
  last_episode ~ 1 + species + gender + (1|species) + (1|gender),
```

```
    data = rickmorty_cat,
    family = "lognormal",
    prior = priors_cat,
    file = "model/f3"
  )
```

7 Investigation of Convergence Diagnostics

In this section, the divergence diagnostics of the two models, i.e. the R-hat and Efficient Sample Size (ESS) values and the number of divergences will be investigated. Below are the values for the non-hierarchical model:

```
 Family: lognormal
  Links: mu = identity; sigma = identity
Formula: episodes_survived ~ 1 + episode_number + species
   Data: rickmorty (Number of observations: 1606)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000
```

Population-Level Effects:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS
Intercept	-0.06	0.08	-0.20	0.10	1.00	3255
episode_number	0.03	0.00	0.02	0.03	1.00	5427
speciesAnimal	-0.01	0.13	-0.26	0.24	1.00	3988
speciesCronenberg	0.91	0.12	0.68	1.13	1.00	3634
speciesHuman	0.48	0.07	0.34	0.62	1.00	2861
speciesHumanoid	-0.51	0.14	-0.78	-0.23	1.00	3841
speciesPoopybutthole	0.22	0.16	-0.08	0.52	1.00	4462
speciesRobot	-0.79	0.18	-1.14	-0.42	1.00	4392

	Tail_ESS
Intercept	3204
episode_number	3171
speciesAnimal	3477
speciesCronenberg	2979
speciesHuman	2960
speciesHumanoid	2994
speciesPoopybutthole	2773
speciesRobot	3366

Family Specific Parameters:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	1.05	0.02	1.02	1.09	1.00	5753	3144

```
Draws were sampled using sample(hmc). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

Divergent Transitions: 0

The R-hat value all have a perfect value of 1 and the ESS values also look very good for the given data. No divergent transitions can be seen. As a result, we conclude that the first model has converged very well (Rhat values < 1.05). In addition, as the ESS are very high, we can see that the model uses a very big portion of the

dataset for sampling and that the model effectively samples the posterior, as no divergent transitions can be observed.

Below are the values for the second, hierarchical time series model:

```
 Family: lognormal
  Links: mu = identity; sigma = identity
Formula: episodes_survived ~ 1 + episode_number + species + (episode_number | name)
   Data: rickmorty (Number of observations: 1606)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Group-Level Effects:
~name (Number of levels: 377)
                              Estimate Est.Error l-95% CI u-95% CI Rhat
sd(Intercept)                     0.40      0.03     0.35     0.46 1.01
sd(episode_number)                0.02      0.00     0.02     0.02 1.01
cor(Intercept,episode_number)    -0.28      0.10    -0.46    -0.06 1.02
                              Bulk_ESS Tail_ESS
sd(Intercept)                      508     1102
sd(episode_number)                 266      488
cor(Intercept,episode_number)      216      479

Population-Level Effects:
                      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
Intercept                -0.20      0.07    -0.34    -0.07 1.00     1035
episode_number            0.01      0.00     0.01     0.02 1.01      736
speciesAnimal            -0.09      0.12    -0.33     0.15 1.00     1651
speciesCronenberg         0.00      0.08    -0.17     0.17 1.00     2071
speciesHuman              0.10      0.07    -0.03     0.23 1.00     1725
speciesHumanoid          -0.06      0.11    -0.27     0.16 1.00     2324
speciesPoopybutthole     -0.13      0.20    -0.51     0.27 1.00     1880
speciesRobot             -0.12      0.25    -0.61     0.37 1.00     2288
                      Tail_ESS
Intercept                 1864
episode_number            1845
speciesAnimal             2193
speciesCronenberg         2746
speciesHuman              2090
speciesHumanoid           2562
speciesPoopybutthole      2782
speciesRobot              2748

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     0.25      0.01     0.24     0.26 1.00     3802     3155

Draws were sampled using sample(hmc). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).


Divergent Transitions: 0
```

The values for the hierarchical model are again very good with no divergent transitions, almost perfect R-hat and very good ESS values. The only major difference to the non-hierarchical model can be seen with the group level effects, where slightly worse ESS and R-hat values are given. The same applies to the bulk ESS values for the intercept and episode number in the population level effects, but they are still very good. Therefore, we can say that also the hierarchical model has converged very well with R-hat values smaller than 1.05, ESS values that indicate that the model uses a bulk of the dataset for sampling and no divergent transitions so that the model effectively samples the posterior.

```
Warning: There were 147 divergent transitions after warmup. Increasing
adapt_delta above 0.8 may help. See
http://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup


 Family: lognormal
  Links: mu = identity; sigma = identity
Formula: last_episode ~ 1 + species + gender + (1 | species) + (1 | gender)
   Data: rickmorty_cat (Number of observations: 382)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000


Group-Level Effects:
~gender (Number of levels: 4)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.53      0.48     0.02     1.87 1.00     1187     1763


~species (Number of levels: 7)
              Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.53      0.45     0.02     1.72 1.01      924     1105


Population-Level Effects:
                    Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
Intercept               3.16      0.58     1.98     4.38 1.00     1516
speciesAnimal           0.18      0.59    -1.15     1.41 1.01     1427
speciesCronenberg       0.06      0.63    -1.26     1.40 1.00     1697
speciesHuman           -0.05      0.64    -1.48     1.43 1.00     1474
speciesHumanoid        -0.05      0.82    -1.88     1.61 1.00     1040
speciesPoopybutthole    0.19      0.59    -1.11     1.38 1.00     2137
speciesRobot            0.49      0.83    -1.30     2.35 1.00     1681
genderGenderless       -0.18      0.68    -1.58     1.35 1.00     1440
genderMale             -0.02      0.44    -0.92     0.93 1.00     1653
genderunknown          -0.27      0.62    -1.52     1.24 1.00     1700
                    Tail_ESS
Intercept               1357
speciesAnimal           1065
speciesCronenberg       1310
speciesHuman             882
speciesHumanoid          912
speciesPoopybutthole    1990
speciesRobot            1115
genderGenderless        1263
genderMale              1324
genderunknown           1304


Family Specific Parameters:
       Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
```

```
sigma      0.69       0.03       0.64      0.74 1.00      4297       2176
```

Draws were sampled using sample(hmc). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).


Divergent Transitions: 147

For the third, non-time series model, the Rhat values for the species and gender predictions are still very good
(< 1.05). The ESS values are still very high for the new categorical dataset, indicating that the model uses a
bulk of the dataset for sampling. For the third model like the previous two, we can conclude that convergence
has been reached. Unfortunately there are some divergent transitions which would indicate that the model does
struggle with effectively sampling the posterior. As the third model was more of an experiment than our main
analysis, we did not take additional steps to improve the number of divergent transitions.

This was our first time running the models for all three, therefore we did not have to conduct any iterative
improvements to achieve this level of convergence.


8 Posterior Predictive Checks

In this section, posterior predictive checks of the models will be performed in order to check their specification
and configuration. For this, a normal posterior predictive check plot and one with grouping to get more insights
are created for all models.
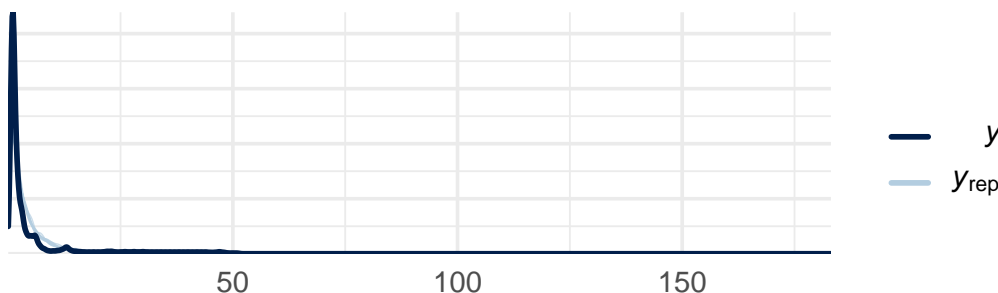
In the PPC for the non-hierarchical model, it can be seen that posterior is overall very well fitted to the original
values. Therefore, we can say that the main features of the original data are replicated. The only differences
can be seen at spikes around the bend at the beginning of the data, which are not modeled. This may, however,
also be a good behavior as the model is not so overfitted by not adapting all the small features of the original
data.

Grouped PPC for non–hierarchical model
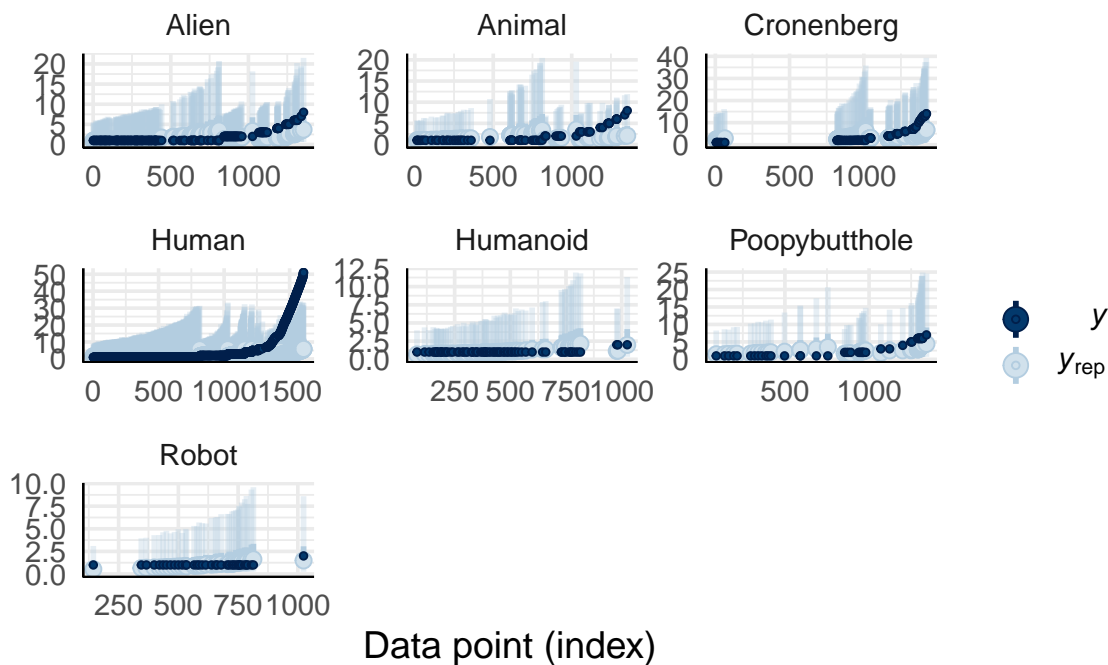
When looking at the checks for the different species, we can see that the first model does not replicate the values very well. Not only do the posterior draws not fit the original data well and most of the draws are outside of the original values, the standard deviation is also very high for all species investigated. Therefore, we can assume that the first model in its configuration is not well suited for the given data and we needed to change some configurations of the model. This is why the hierarchical model was also investigated, the configuration of which was explained before. The results of its PPC can be seen below:

PPC for hierarchical model
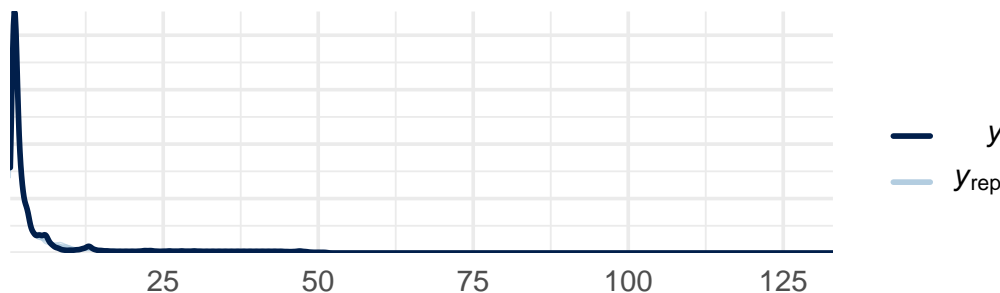
In this first plot for the hierarchical model, we can see that the fit looks overall very good and similar to the non-hierarchical model. However, at the bend where the spikes can be observed, a slightly better behavior can be observed, including more of the original data. Overall, this model replicates the original data very well.

Grouped PPC for hierarchical model


Data point (index)
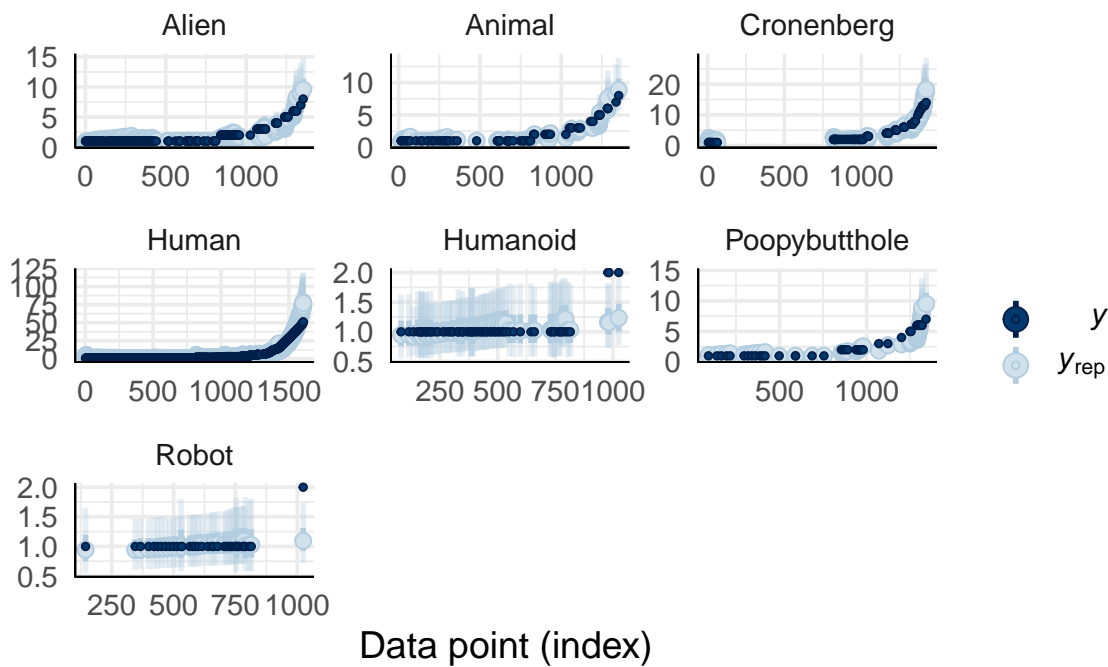
With the species-wise PPC, more differences and details, especially when being compared to the first model, can be observed. Overall, this model fits the original data much better, including most of the data in the posteriors and having a significantly smaller standard deviation. Therefore, we can assess that the inclusion of the group effects with the hierarchical model led to an overall much better performance when being compared to the non-hierarchical model and to an overall better configuration.

PPC for hierarchical non–time series model
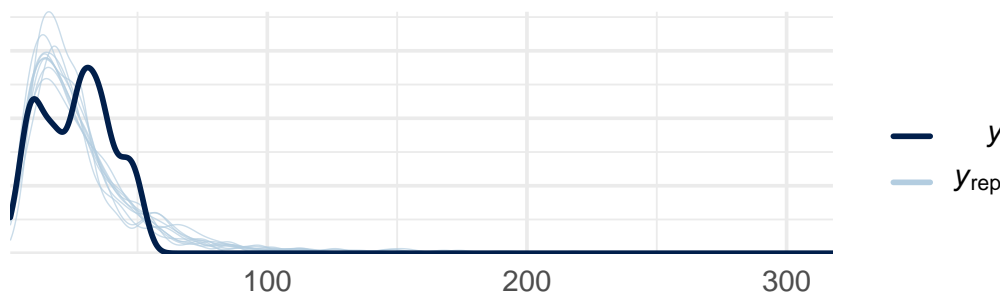
This plot indicates that the third model only does a decent job of predicting the data. The peaks are in the incorrect spot and it doesn't follow the shape of the observed data. This may have something to do with the lack of episode information in the third model.

Grouped PPC for hierarchical non–time series model
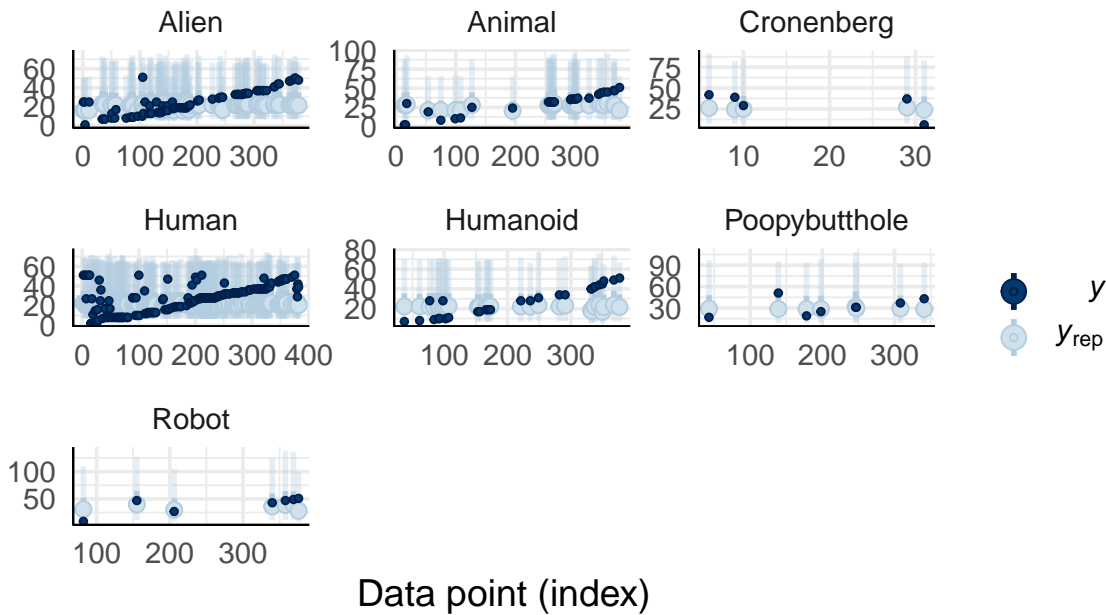
We can clearly see from the grouped intervals plot that the third model struggles to fit the observed data, with most of the observations falling in a different place.

9 Sensitivity Analysis

For the sensitivity analysis, we fitted the three models a second time with very simple and uninformative priors, i.e. where the standard deviation of each predictor is the same and equal to $log(2)$. Therefore the priors look like the following for the temporal models:

$$episode\_number \sim normal(0, log(2))$$
$$human \sim normal(0, log(2))$$
$$animal \sim normal(0, log(2))$$
$$cronenberg \sim normal(0, log(2))$$
$$poopybutthole \sim normal(0, log(2))$$

And likewise for the non-temporal model:

$$human \sim normal(0, log(2))$$
$$animal \sim normal(0, log(2))$$
$$cronenberg \sim normal(0, log(2))$$
$$poopybutthole \sim normal(0, log(2))$$
$$male \sim normal(0, log(2))$$
$$genderless \sim normal(0, log(2))$$
$$unknown \sim normal(0, log(2))$$

After fitting the models, the same diagnostics as above, i.e. the convergence diagnostics and posterior predictive checks were run for the new models:

9.1 Convergence Diagnostics

The convergence diagnostics are first given for the non-hierarchical model, then for the hierarchical temporal model and lastly for the non-temporal hierarchical model.

```
 Family: lognormal
  Links: mu = identity; sigma = identity
Formula: episodes_survived ~ 1 + episode_number + species
   Data: rickmorty (Number of observations: 1606)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000
```

Population-Level Effects:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS
Intercept	-0.04	0.08	-0.19	0.10	1.00	3476
episode_number	0.03	0.00	0.02	0.03	1.00	5837
speciesAnimal	-0.02	0.13	-0.28	0.22	1.00	3720
speciesCronenberg	0.89	0.12	0.66	1.11	1.00	3429
speciesHuman	0.47	0.07	0.33	0.61	1.00	2516
speciesHumanoid	-0.52	0.14	-0.79	-0.24	1.00	4217
speciesPoopybutthole	0.21	0.15	-0.10	0.50	1.00	4346
speciesRobot	-0.80	0.18	-1.15	-0.44	1.00	4385

	Tail_ESS
Intercept	3724
episode_number	3482
speciesAnimal	2619
speciesCronenberg	2957
speciesHuman	2845
speciesHumanoid	3023
speciesPoopybutthole	3083
speciesRobot	2867

Family Specific Parameters:

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat	Bulk_ESS	Tail_ESS
sigma	1.05	0.02	1.02	1.09	1.00	6249	2921

```
Draws were sampled using sample(hmc). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).
```

Divergent Transitions: 0

```
 Family: lognormal
  Links: mu = identity; sigma = identity
Formula: episodes_survived ~ 1 + episode_number + species + (episode_number | name)
   Data: rickmorty (Number of observations: 1606)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000
```

Group-Level Effects:

~name (Number of levels: 377)

	Estimate	Est.Error	l-95% CI	u-95% CI	Rhat
sd(Intercept)	0.40	0.03	0.35	0.46	1.01

```
sd(episode_number)                        0.02        0.00        0.02        0.02 1.03
cor(Intercept,episode_number)            -0.28        0.10       -0.46       -0.09 1.04
                              Bulk_ESS Tail_ESS
sd(Intercept)                      591     1514
sd(episode_number)                 208      525
cor(Intercept,episode_number)      150      484

Population-Level Effects:
                      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
Intercept                -0.20      0.07    -0.33    -0.07 1.00      821
episode_number            0.01      0.00     0.01     0.02 1.01      652
speciesAnimal            -0.08      0.12    -0.31     0.16 1.00     1672
speciesCronenberg         0.01      0.08    -0.15     0.16 1.00     2215
speciesHuman              0.10      0.06    -0.03     0.22 1.00     1835
speciesHumanoid          -0.06      0.11    -0.28     0.16 1.00     1893
speciesPoopybutthole     -0.11      0.20    -0.51     0.26 1.00     2255
speciesRobot             -0.12      0.24    -0.60     0.34 1.00     1994
                      Tail_ESS
Intercept                 1851
episode_number             862
speciesAnimal             2226
speciesCronenberg         2640
speciesHuman              2583
speciesHumanoid           2785
speciesPoopybutthole      2920
speciesRobot              1764

Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     0.25      0.01     0.24     0.26 1.00     3725     3152

Draws were sampled using sample(hmc). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).


Divergent Transitions: 0


Warning: There were 151 divergent transitions after warmup. Increasing
adapt_delta above 0.8 may help. See
http://mc-stan.org/misc/warnings.html#divergent-transitions-after-warmup


 Family: lognormal
  Links: mu = identity; sigma = identity
Formula: last_episode ~ 1 + species + gender + (1 | species) + (1 | gender)
   Data: rickmorty_cat (Number of observations: 382)
  Draws: 4 chains, each with iter = 2000; warmup = 1000; thin = 1;
         total post-warmup draws = 4000

Group-Level Effects:
~gender (Number of levels: 4)
            Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.37      0.33     0.01     1.26 1.01     1095     1880
```

```
~species (Number of levels: 7)
             Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sd(Intercept)     0.34      0.32     0.01     1.19 1.01      511      526


Population-Level Effects:
                      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS
Intercept                 3.14      0.42     2.23     3.97 1.00     1071
speciesAnimal             0.16      0.37    -0.82     0.85 1.01      664
speciesCronenberg         0.06      0.43    -0.82     0.86 1.00     1180
speciesHuman             -0.03      0.36    -0.75     0.90 1.01      675
speciesHumanoid          -0.04      0.55    -1.16     1.20 1.00      736
speciesPoopybutthole      0.19      0.39    -0.66     0.92 1.01      799
speciesRobot              0.48      0.59    -0.67     1.66 1.00      969
genderGenderless         -0.13      0.44    -0.94     0.79 1.01     2323
genderMale               -0.03      0.36    -0.74     0.82 1.00     1767
genderunknown            -0.22      0.40    -0.96     0.73 1.00     1487
                      Tail_ESS
Intercept                 1071
speciesAnimal              195
speciesCronenberg          826
speciesHuman               167
speciesHumanoid            225
speciesPoopybutthole       340
speciesRobot               714
genderGenderless          1951
genderMale                1780
genderunknown             1362


Family Specific Parameters:
      Estimate Est.Error l-95% CI u-95% CI Rhat Bulk_ESS Tail_ESS
sigma     0.69      0.03     0.64     0.74 1.00     1833     1164


Draws were sampled using sample(hmc). For each parameter, Bulk_ESS
and Tail_ESS are effective sample size measures, and Rhat is the potential
scale reduction factor on split chains (at convergence, Rhat = 1).


Divergent Transitions: 151
```

Above the tests with uninformative priors were conducted again for prior sensitivity analysis. We can see that none of the three models are that sensitive to the choice of priors, as the values are pretty much the same. This is a good thing, as we can see that model performance is not easily influenced by the choice of prior.
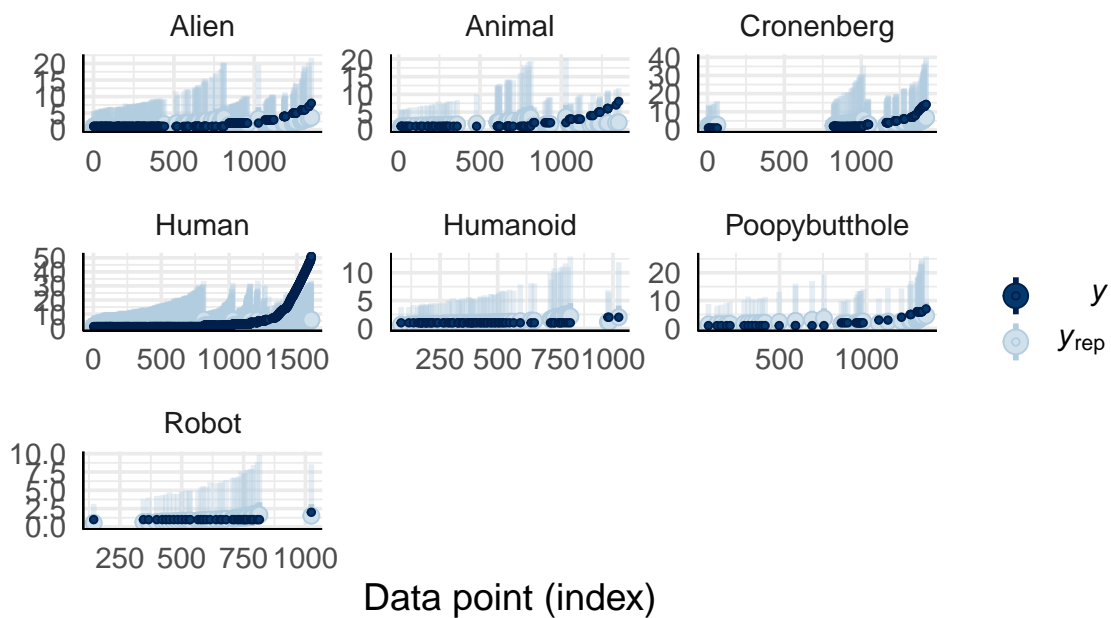

9.2 Posterior Predictive Checks

The posterior predictive checks (pp_check not grouped and grouped each) are first given for the non-hierarchical model, then for the hierarchical temporal model and lastly for the non-temporal hierarchical model.
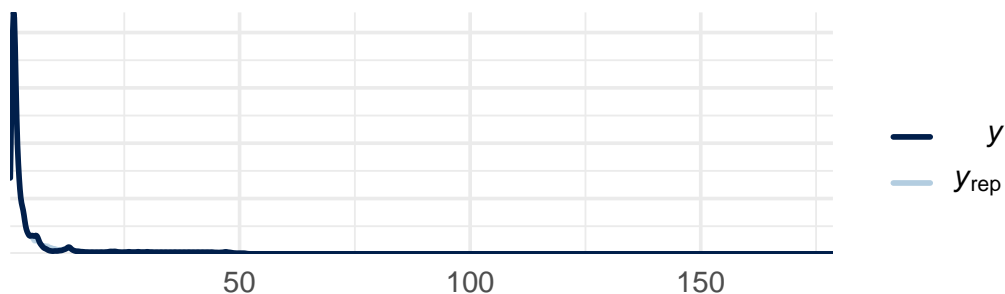
PPC for uninformative non−hierarchical model
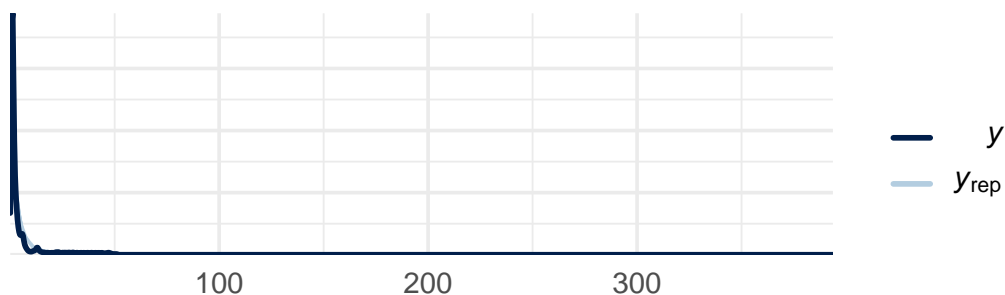
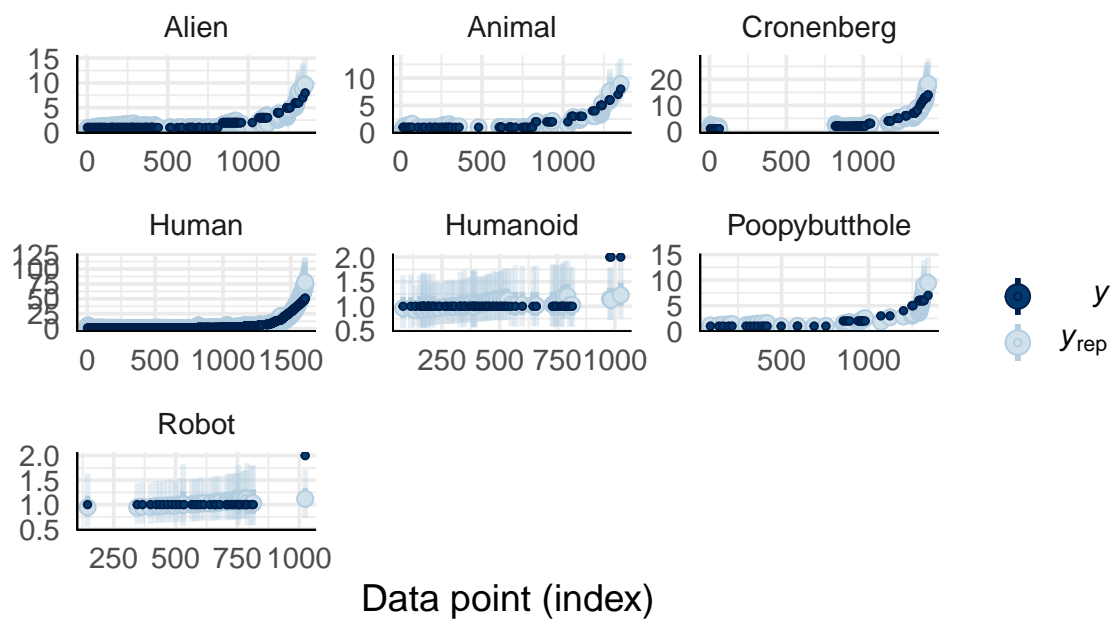Using all posterior draws for ppc type 'intervals_grouped' by default.

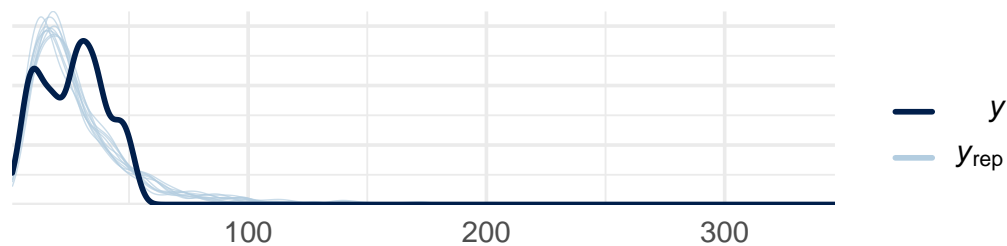# Grouped PPC for uninformative non−hierarchical model

Data point (index)

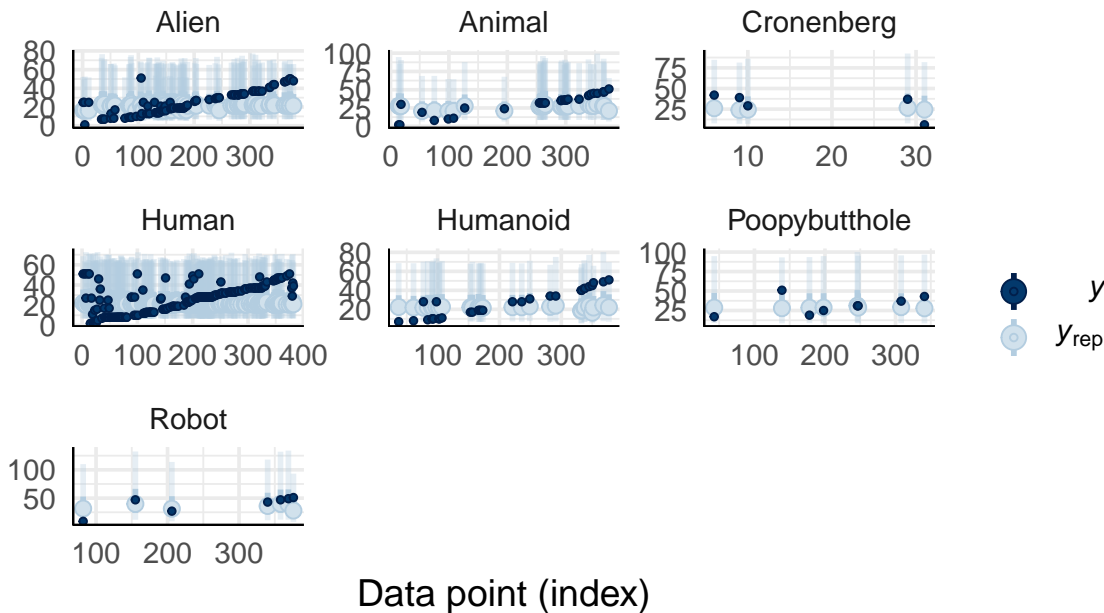# PPC for uninformative hierarchical model

Grouped PPC for uninformative hierarchical model

Data point (index)

PPC for uninformative hierarchical non−time series model

Grouped PPC for uninformative hierarchical non–time series model

Also for the posterior predictive checks, no significant difference between the uninformative and informative priors can be observed. This, again, shows that the models are not very sensitive to the choice of prior.

## 10 Model Comparison and Predictive Performance Assessment

In this part, the three models will be compared using LOO-CV. For this, the predictive performance and then the k-hat values for all models are plotted and compared, after calculating loo. The performance assessment will only be done using the LOO-Compare function as the comparison of accuracy would not be sensible here, as we are comparing regression models and not classification models, so doing this task would not be possible for our problem.

```
Warning: Found 195 observations with a pareto_k > 0.7 in model 'f2'. It is
recommended to set 'moment_match = TRUE' in order to perform moment matching
for problematic observations.
```

```
   elpd_diff se_diff
f2     0.0      0.0
f1 -2112.8     62.9
```

It should be noted that f3 cannot be compared to the other two models as we removed the episode information and they thus have a mismatched number of datapoints. However, from comparing the predictive performance of the two first models it can be seen that the hierarchical model gets a much better, i.e. perfect score of 0.0 both for the elpd difference as well as for the standard error difference. For the non-hierarchical model on the other hand, the score is much lower, achieving only a score of -2130 for the elpd difference and a score of 63.7 for the standard error difference. As the standard error difference is also very small for both models and the values are very large (the values are larger than 2 standard errors), the results which model is the better one are also very reliable. Therefore, overall, we can conclude that according to the LOO-compare function, the hierarchical model has a much better predictive performance. However, significantly different values are achieved when looking at the k-hat values:

PSIS diagnostic plot
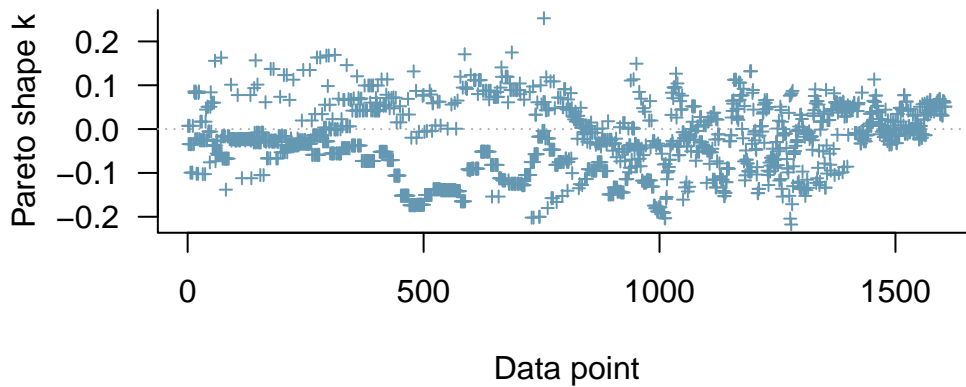
For the first model, the k-hat values are all very good, being all smaller than 0.5. This indicates that the estimation of the corresponding component of the LOO-CV distribution with the PSIS computation is done with a high accuracy and that it is therefore reliable. However, very different results can be observed for the second, hierarchical model:

PSIS diagnostic plot
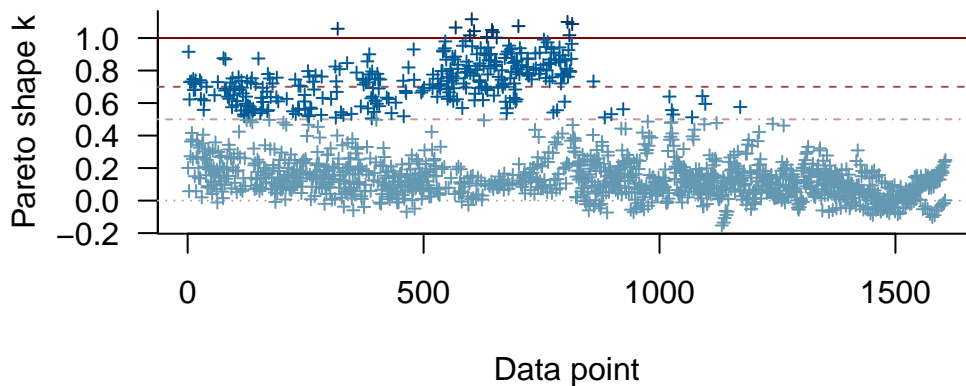
Here, many more observations have a k-hat value of over 0.5 or even over 0.7 and 1.0. To be precise, there are 202 observations in total with k-hat values over 0.7, while still, most of the values are under 0.5, as 202 only represents about 10 percent of the total observations. For all the higher values, it can be said that we get a lower accuracy or that it is not even possible to give a good estimate for the component if it is over 0.7 or even 1.0. These high values mostly occur when the models are misspecified, we have mistakes in data preprocessing or we have outliers. However, as hierarchical models are so complex and have more parameters to give a higher flexibility, these values often occur as PSIS-LOO starts to fail. In addition, these observations with higher k-hat values can be in groups where only few observations exist, which all leads to the conclusion that these values are probably expected with this model and the given data, especially because we also have many groups with only a couple of observations. Examples for this are, for instance, higher values of survival rates for humans, where there are mostly only a couple of values available, for instance for Rick and Morty, as they appear in (mostly) every episode and have therefore a much higher survival rate than everybody else by a big margin. We can also infer that the many high k-hat values are probably due to issues with the data due to the fact that the activation of the "moment_match" option in the second model does not make any difference on the final k-hat values.

Another explanation for this is that since the data is split in episodes, the characters that only appear in a few episodes would end up having the higher k hat values. Therefore we wanted to inspect the impact of removing the time series aspect (i.e. the episode numbers) entirely.

PSIS diagnostic plot
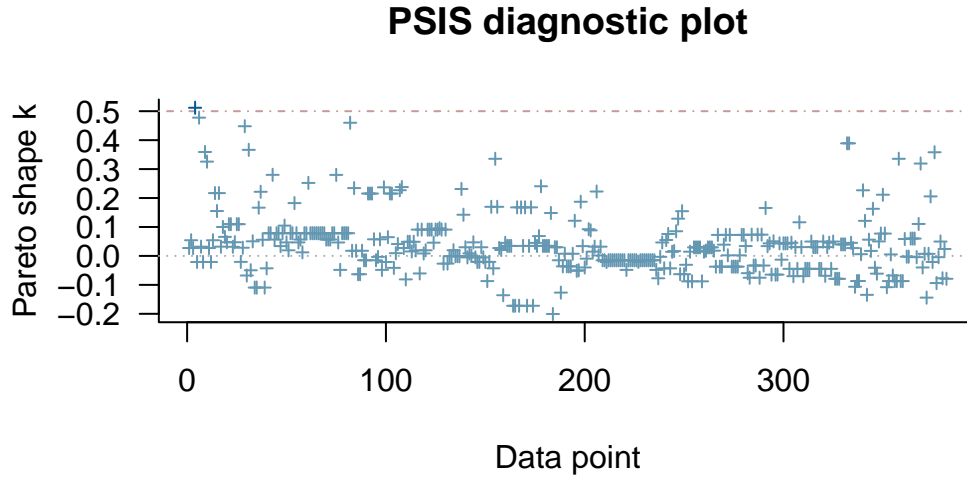
This demonstrates that the k hat values are indeed all lower for the third model, supporting the hypothesis that the episode information impacts the model performance negatively.

11 Discussion

Overall, the second model exhibits commendable predictive capabilities; however, there is room for enhancement, particularly in discerning between main characters and side characters. This distinction is crucial, as main characters, surviving across numerous episodes, are somewhat outliers compared to the more abundant side characters. A refinement in this aspect could significantly bolster the model's accuracy, preventing it from being unduly influenced by the sheer number of side characters and could potentially improve the bad k-hat values, all of which could be seen in the previous analysis.

Furthermore, the analysis reveals a conspicuous limitation in the first model, notably its inadequacy in predicting character survival ratings. This deficiency is attributed to the model's inherent simplicity, rendering it incapable of accommodating the intricate variability associated with character survivability. As a result, substantial improvement may be unattainable, given the model's inherent limitations.

In addition, our data analysis showed that using a non-temporal model does not necessarily improve the predictive capabilities of the model. This may also be due to required improvements in the model itself. However, overall, we assess that the approach with using temporal data for predicting the survival rates is the best approach for our data with the second model giving us overall the best predictive performance.

We also recognise that compared to the true complexity of Rick and Morty as a show, the dataset only having species, gender, character names and episode numbers as features is a fairly simplistic approach. For future analysis, we, or other accomplished researchers, could delve deeper into the Rick and Morty universe to collect additional data on patterns in episodes to add to the features of the dataset, in addition to new data points, eventually covering all seasons.

12 Conclusion

Upon comprehensive analysis, a discernible pattern emerges, indicating that characters, when considered in aggregate, tend to exhibit either early demise or sustained survival over a significant number of episodes. This trend is particularly pronounced in main characters, whose extended presence often defines their significance

within the narrative. Overall the second model seems to perform the best of the ones compared in this project and with a few adjustments, it could be even better, especially regarding its k-hat values. The first model, on the other hand, due to its simplistic nature, fails to capture the intricacies of the survivability of the characters. The third additional model performed in a good way, indicating some possible problems in the episodic nature of the original data especially for the k-values, however as it is not time series data it may not be suitably significant for future analysis. Also, this model showed some issues with its predictive performance that need to be addressed in order to be able to use it for future analysis. Overall, our project has shown that it is possible to predict the survival trajectories of characters in Rick and Morty with very good predictive performance using a temporal hierarchical model, bringing more structure to the complex relationships in this particular show.

13 Self Reflection

Thank you for reading our report! We would like to take this time to think back on our hard work and see how far we have come. Struggling through this project together reminded us that we are not alone, not only in this academic journey but our life's journey. Moreover, this project provided us with valuable insights into our strengths and areas for growth. The project pushed us to really use our communication, collaboration, and time management skills. Navigating through complex tasks and tight timelines, we honed our abilities to convey ideas effectively, work seamlessly as a team, and efficiently allocate our time resources. Overall, the project showed us our room for growth and illuminated the significance of continual learning and adaptability in navigating diverse challenges.