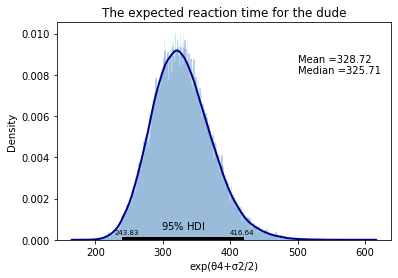
Assignment 5

Import the data and run the Pystan model.

Task A

1. What is the expected reaction time for the individual called “the dude” in [our data set](https://docs.google.com/spreadsheets/d/1y8Hvj8AeIt1Nl7b6yNQ6-z9_XlUAE7ftHM2sSmA7gsM/edit?usp=sharing) (corresponding to ind = 3 in python and ind=4 in STAN/matlab/julia)?
   1. Answer this by providing a histogram and appropriate summaries (e.g. mean, mode, median, 95% credible interval of the expected reaction time).



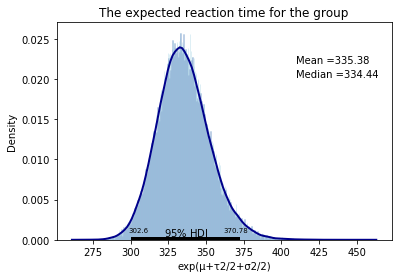
* 1. With only one measurement, how are we able to give a 95% credible interval for “the dude”? This is not possible using frequentist statistics.

Ans - Even with one measurement, the hierarchical model takes the trend for the entire group and uses that to predict the reaction time for one individual, even though he has only one reading.

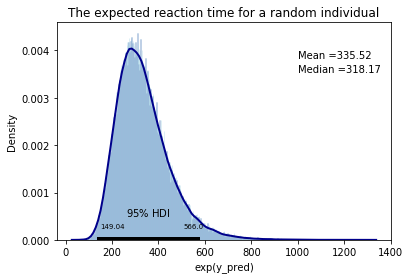
1. What is the group’s reaction time?

Given a random new individual from our group, e.g. someone that simply “forgot” to do the test (you all know who you are ;)), what is the:

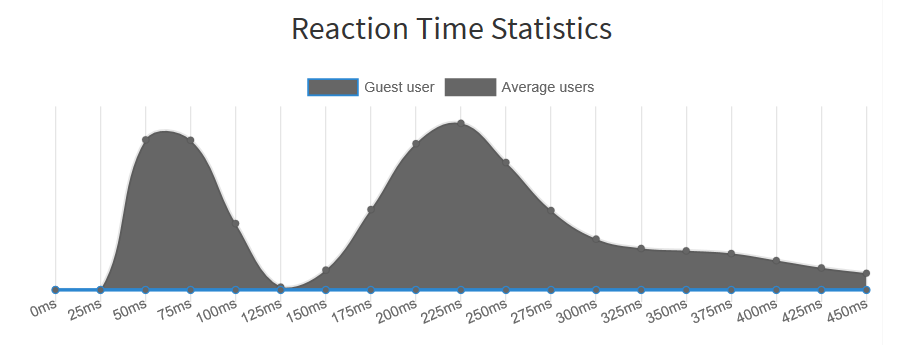
* + 1. Expected reaction time for that random individual?



* + 1. predicted reaction time for a single measurement for that individual? (provide the posterior predictive distribution for the reaction time) In STAN you can do this in the “generated quantities { }” section in your STAN code. In python/matlab/julia: 1) pick a posterior sample from mu, tau and sigma given your data. 2) simulate a new theta given these samples, i.e. theta~N(mu,tau). 3) simulate a reaction time measurements given this theta and sigma (from step 1 above), i.e. logy~N(theta,sigma). 4) Calculate y from logy or zlogy. Repeat 1-4) and you will have posterior predictive samples from a new individual.



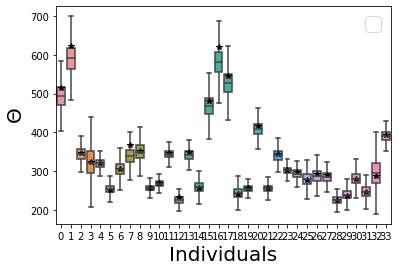
* 1. Compare the posterior predictive histogram with the statistics you can see at the [website](https://www.humanbenchmark.com/tests/reactiontime/statistics).



The curve from 25ms to 125ms looks unexplainable. It could be an automatic code clicking after a programed time, thus achieving very small reaction times in a specific range.

The second curve gives a realistic picture, but the mean reaction time that is 225ms is lower different from our dataset. It can reflect trained responses. The right trail is quite spread out and not ver narrow. This could be explained by users being children or using another device like mobile phone instead of a computer, which tends to record the clicks slower.

1. Provide a figure that compare thetas obtained with our hierarchical Bayesian model above with thetas obtained by treating the participants individually and using the sample means (i.e. theta[j] is in the latter case the sample mean of the j:th participants logarithmic reaction times).



* 1. Can you explain the differences?

From the box plot, it is clear that individuals with more number of observations, have compact ranges which is very near to the mathematical mean, but for those with very few observations, the mathematical mean does not give a picture of the possible reaction times.