The effect of a Formula One driver's age on their performance

A Bayesian analysis

Miro Kaarela, Roope Kausiala, Tatu Timonen

January 25, 2022

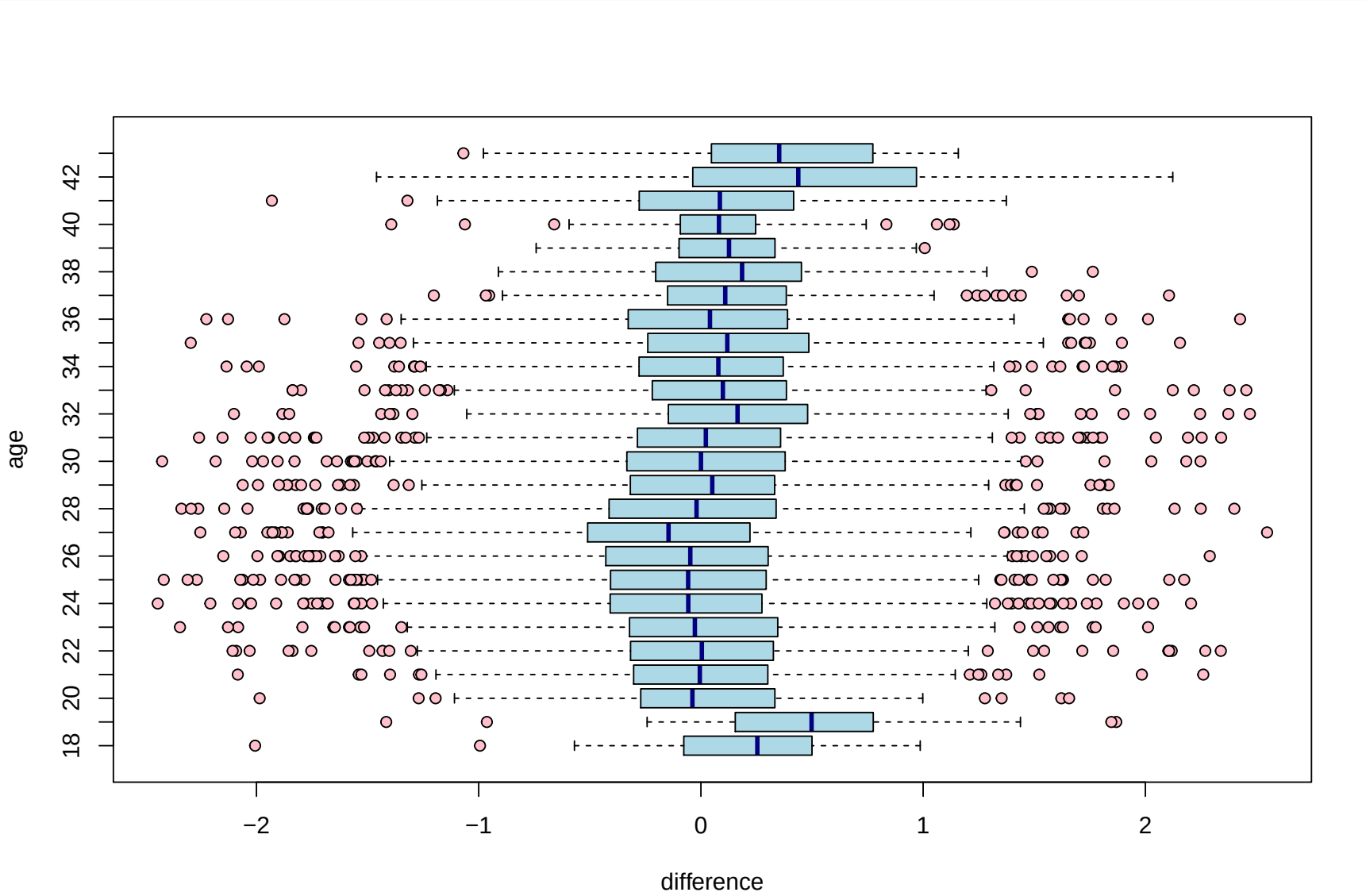
INTRODUCTION

- Formula One (F1) is the most prestigious and popular auto racing league in the world
- It is a well-known fact that motor skills and reflexes of adults deteriorate with age
- In our analysis we explore the effect of F1 drivers' age on their performance

DATA AND ON QUANTIFYING PERFORMANCE

- We utilize the Ergast F1 dataset [1]
- In F1, there is a qualifying event before every race
  - We consider the time difference between a driver and his teammate on their respective best qualifying rounds as a performance measure
- We further pool the obtained time difference data according to drivers' age

TIME DIFFERENCE TO TEAMMATE IN QUALIFYING



SEPARATE

- The distributions for the parameters of the model are different for each age:

$$
\begin{aligned}
t_i &\sim \mathrm{N}(\mu_{\mathrm{age}(i)}, \sigma_{\mathrm{age}(i)}), \\
\mu_{\mathrm{age}(i)} &\sim \mathrm{N}(0, 1), \\
\sigma_{\mathrm{age}(i)} &\sim \mathrm{N}(0, 1).
\end{aligned}
$$

HIERARCHICAL

- As with separate, but the parameters of the distributions yielding the model parameters come from hyperpriors:

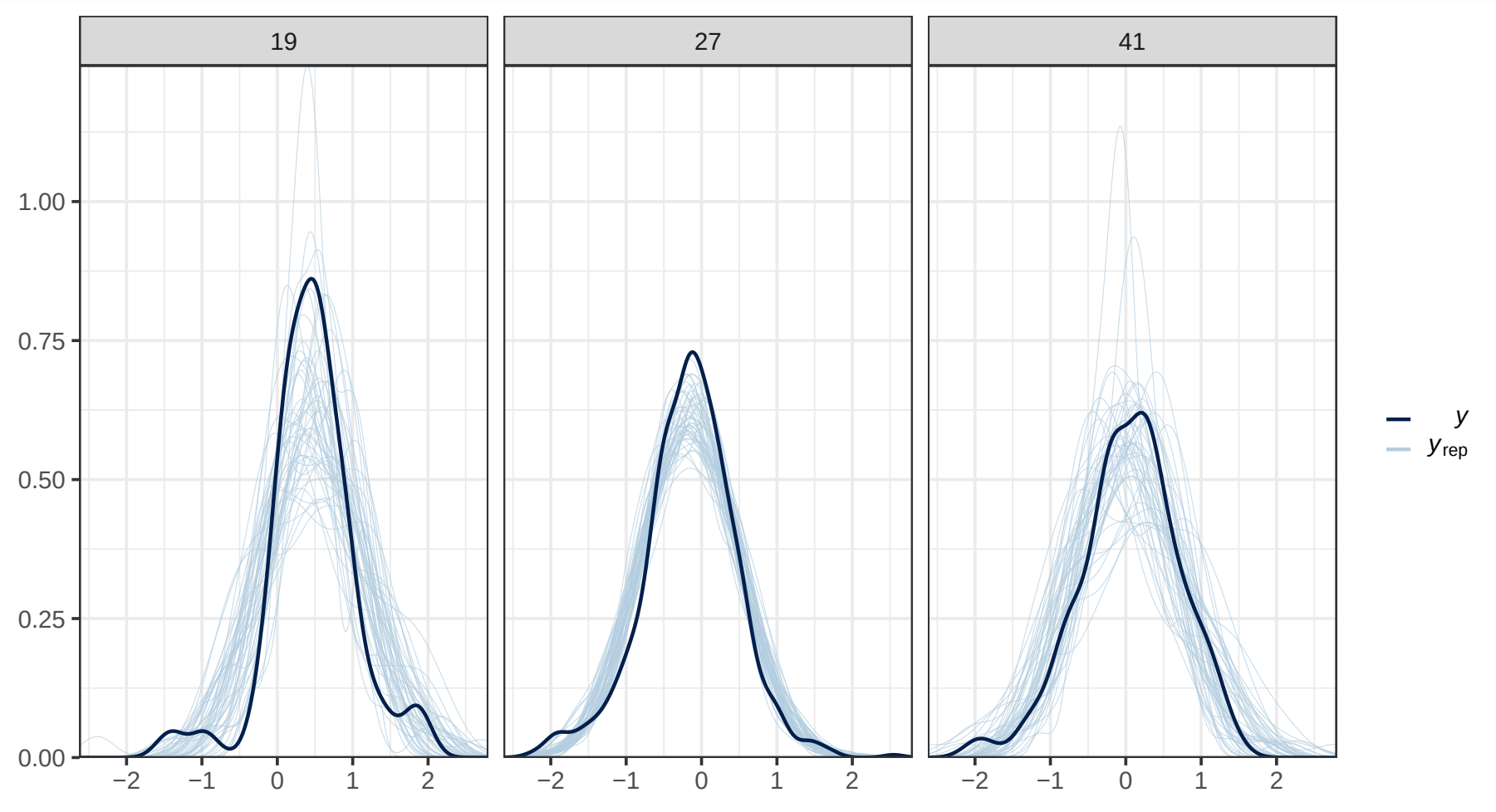
$$
\begin{aligned}
t_i &\sim \mathrm{N}(\mu_{\mathrm{age}(i)}, \sigma_{\mathrm{age}(i)}), \\
\mu_{\mathrm{age}(i)} &\sim \mathrm{N}(\mu_{\mathrm{unknown}_\mu}, \tau_\mu), \\
\mu_{\mathrm{unknown}_\mu} &\sim \mathrm{N}(0, 1), \\
\tau_\mu &\sim \mathrm{N}(0, 1), \\
\sigma_{\mathrm{age}(i)} &\sim \mathrm{N}(\mu_{\mathrm{unknown}_\sigma}, \tau_\sigma), \\
\mu_{\mathrm{unknown}_\sigma} &\sim \mathrm{N}(0, 1), \\
\tau_\sigma &\sim \mathrm{N}(0, 1).
\end{aligned}
$$

SEPARATE WITH ADDITIONAL TEAMMATE PARAMETERS

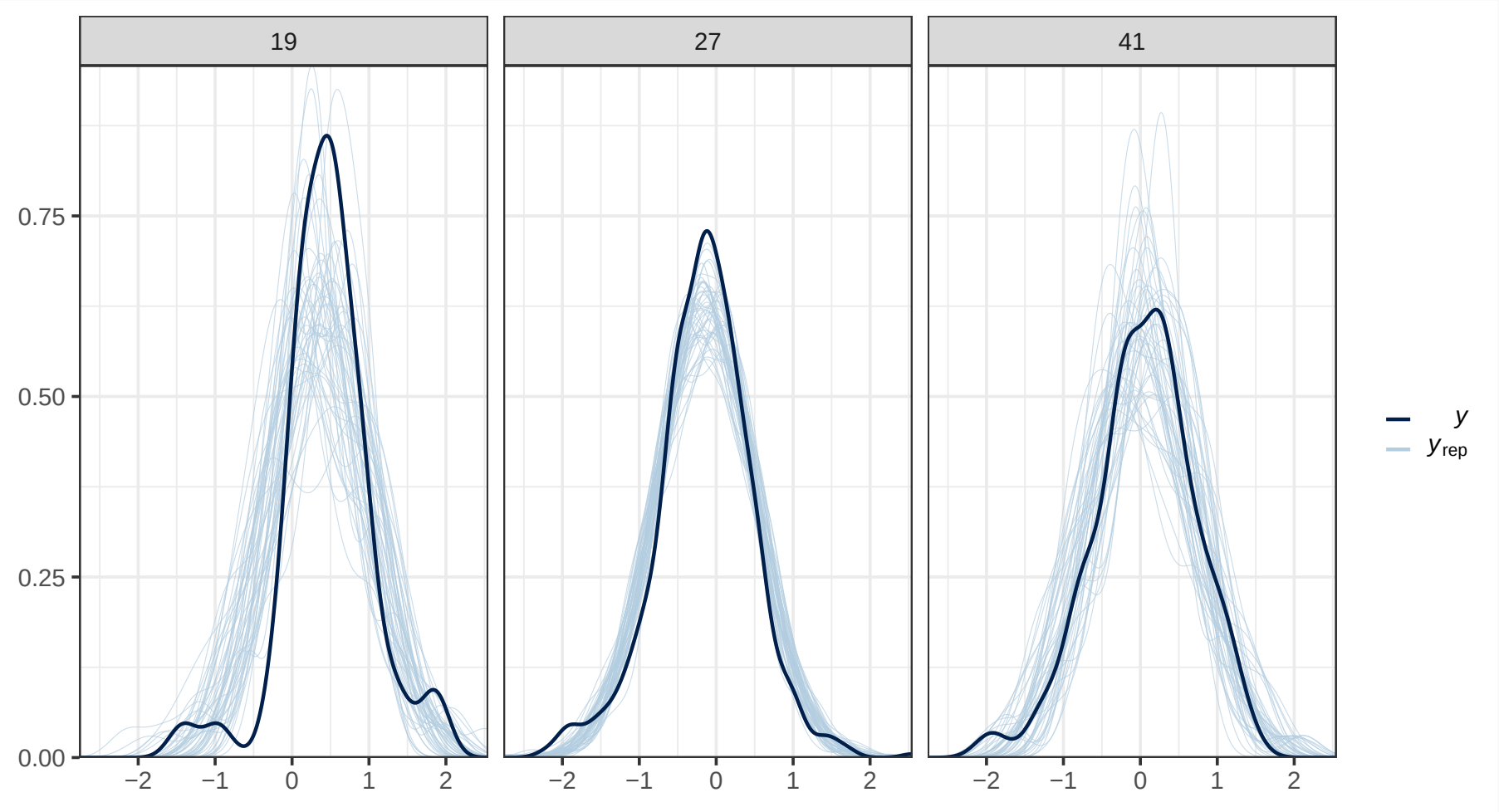
- Consider a parameter shifting the model mean:

$$
\begin{aligned}
t_i &\sim \mathrm{N}(\mu_{\text{age}(i)} + \alpha_{\text{teammate(i)}}, \sigma_{\text{age}(i)}), \\
\mu_{\text{age}(i)} &\sim \mathrm{N}(0, 1), \\
\sigma_{\text{age}(i)} &\sim \mathrm{N}(0, 1), \\
\alpha_{\text{teammate(i)}} &\sim \mathrm{N}(0, 0.5).
\end{aligned}
$$

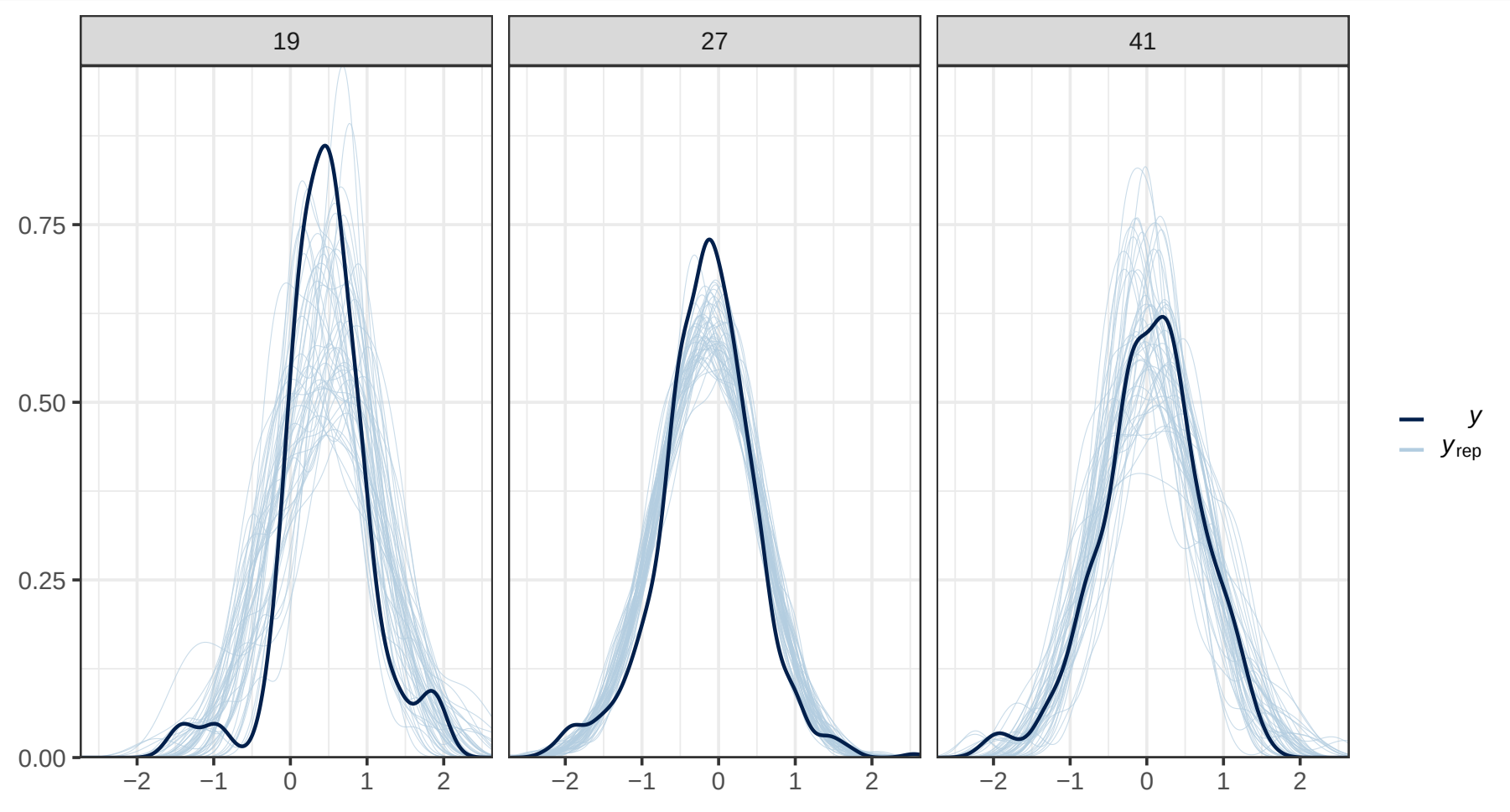
RESULTS WITH THE SEPARATE MODEL



RESULTS WITH THE HIERARCHICAL MODEL



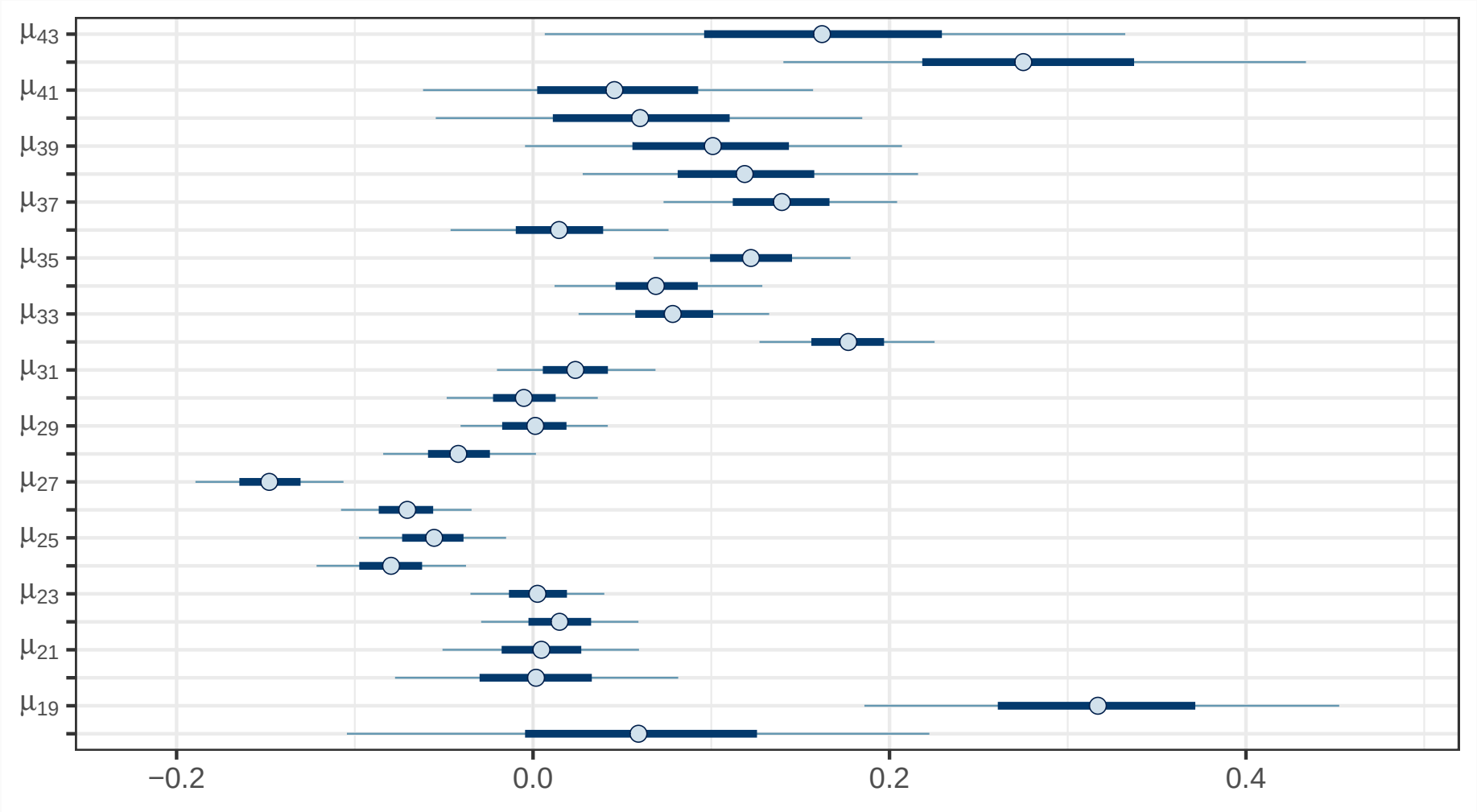
RESULTS WITH THE SEPARATE MODEL WITH α



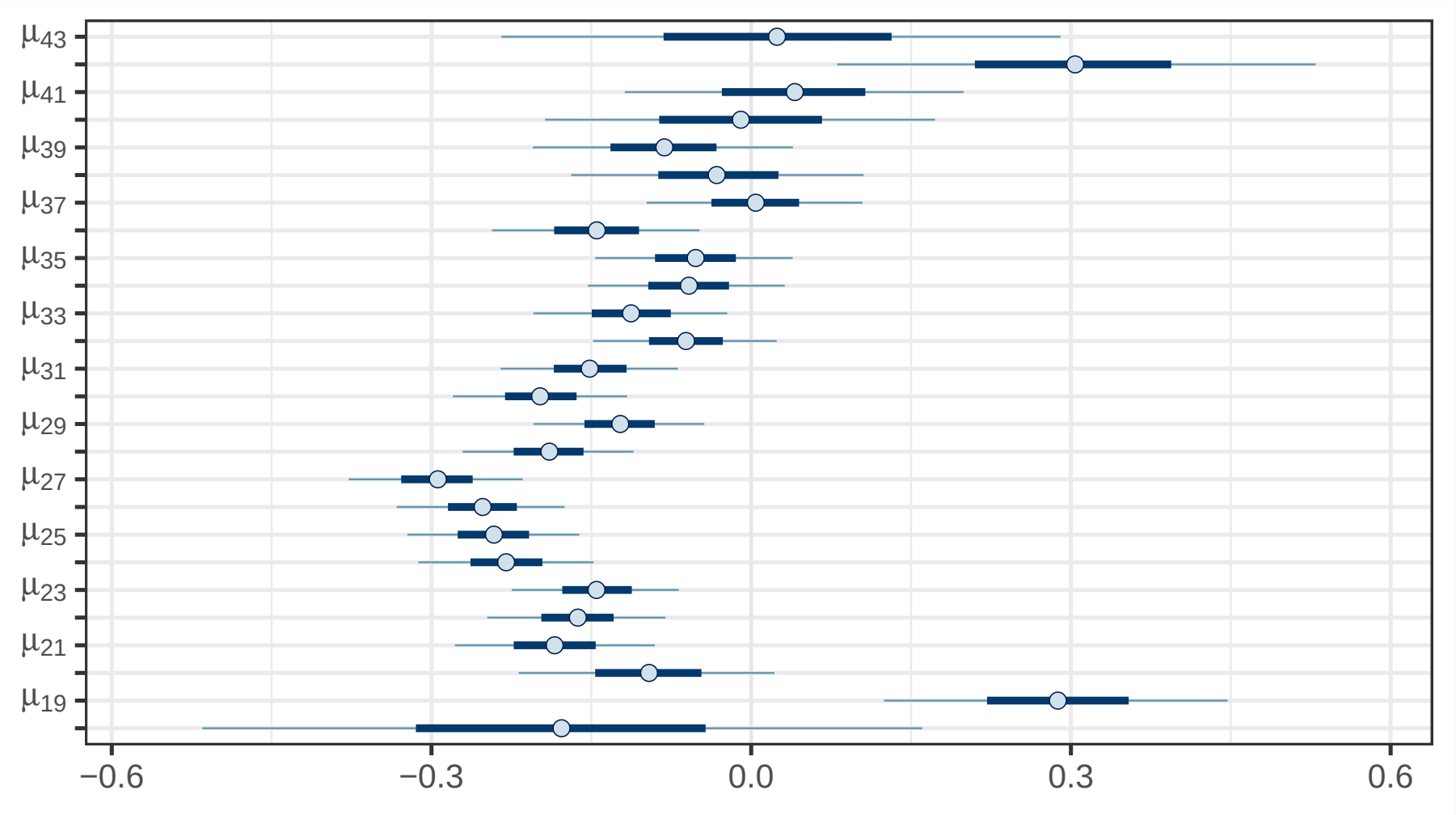
SUMMARY ON THE MODEL DIAGNOSTICS

- The convergence diagnostics were good for all models
- The separate model with α had the best cross-validation score (performed with *loo*)
    - This is expected since it has additional information

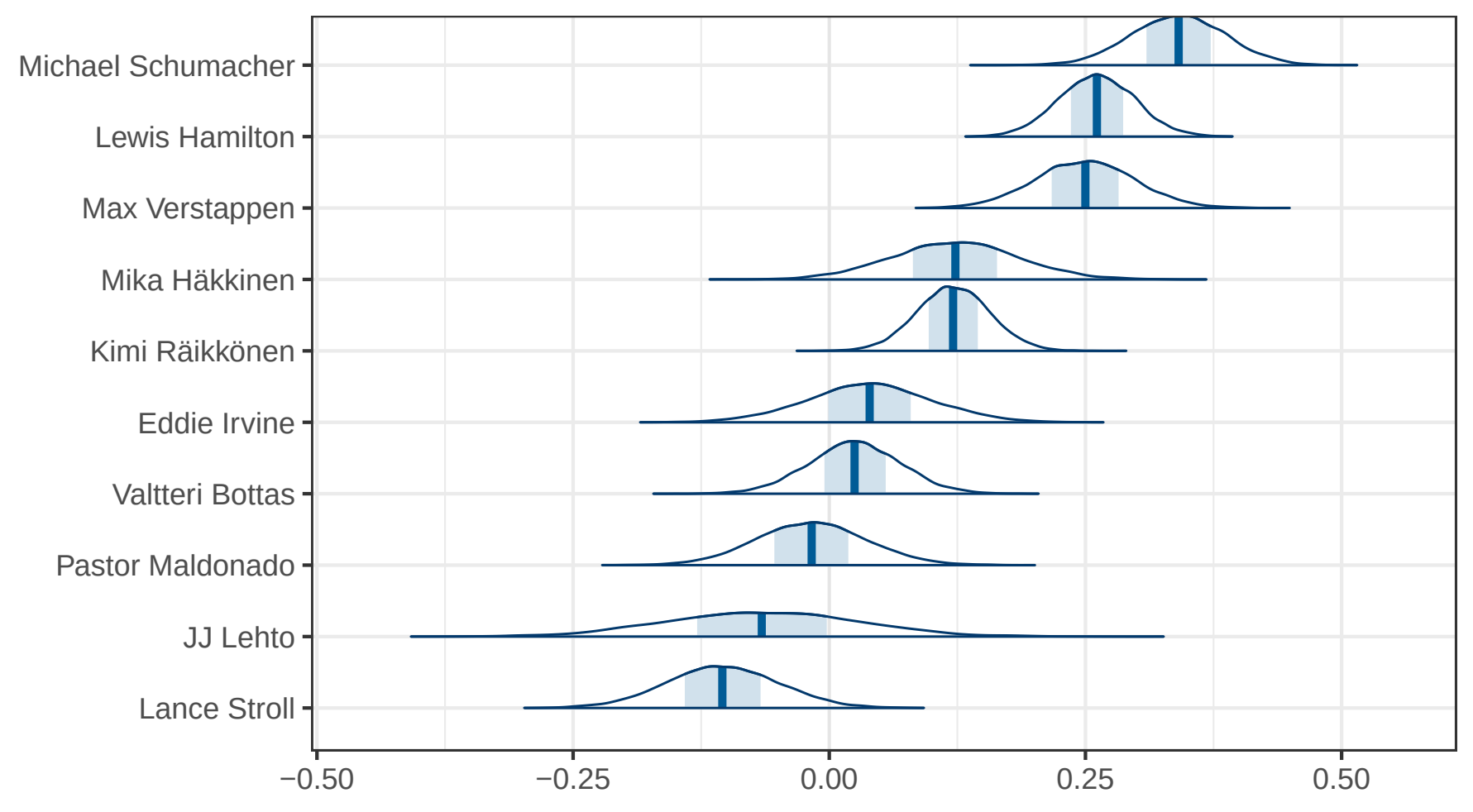
$\mu_{\text{AGE}(i)}$ DISTRIBUTIONS FOR THE HIERARCHICAL MODEL



$\mu_{\text{AGE}(i)}$ DISTRIBUTIONS FOR THE SEPARATE MODEL WITH α



α PARAMETERS FOR AN ARBITRARY SET OF DRIVERS



CONCLUSION

- It appears there is a degree of dependence between a driver's age and their performance
- The separate model with α, and the hierarchical model provide the best fit
- The models indicate that the age of 27 is where a driver reaches their zenith
  - This agrees to a decent extent with previous studies (e.g., [2]) on the effect of age to performance in other competitive regimes
  - $> 90\%$ probability that μ of age 27 is the smallest (best)

REFERENCES I

T. Brieuc.
Ergast developer api.
http://ergast.com/mrd/.

J. J. Thompson, M. R. Blair, and A. J. Henrey.
Over the hill at 24: persistent age-related cognitive-motor decline in reaction times in an ecologically valid video game task begins in early adulthood.
PloS one, 9(4):e94215, 2014.