Supplementary Information

Hidden Markov Models of Evidence Accumulation in Speeded Decision Tasks

Šimon Kucharský, N.-Han Tran, Karel Veldkamp, Maartje Raijmakers, & Ingmar Visser

July 23, 2021

1 Parameter recovery using maximum a posteriori estimates

The 1,000 data sets generated during the prior predictive simulation were used to fit the model coded in Stan (Carpenter et al., 2017), utilizing the `optimize` function conducting L-BFGS-B optimization routine to find the maximum a posteriori estimates (MAP) of the parameters. Initial values for the parameters were generated by randomly drawing from their prior distribution. Regardless, the log-likelihood frequently underflowed right at the beginning of the routine, got stuck during optimization, or converged at a local maximum. Thus, the fitting routine was repeated for each data set. If the optimization converged to an optimum, we checked whether label switching occurred: We calculated the percentage of trials where the model state classification corresponded to the true state. If the percentage was below 50%, label switching was assumed and the model was refitted (by construction of the priors, label switched optimum is not a global optimum). The model was repeatedly run until the optimization converged and did not label switch, or until the number of attempts to fit the model exceeded 50 attempts. If the latter occurred, the fit was classified as unsuccessful and removed from the results. Out of the total of 1,000 simulations, 986 succeeded. Consequently, 14 data sets were not fitted successfully using MAP estimation.

Figure 1 shows the scatter plot between the true (x-axis) and estimated (y-axis) values for the nine free parameters in the model: the drift for the correct choice under the controlled state ($\nu_1^{(1)}$), the drift for the correct choice under the guessing state ($\nu_1^{(2)}$), the standard deviation of drifts (σ), the decision boundary under the controlled ($\alpha^{(1)}$) and guessing ($\alpha^{(2)}$) state, the non-decision time (τ), the initial probability of the controlled state (π_1), the probability of dwelling in the controlled (ρ_{11}) and the guessing (ρ_{22}) state. The correlations for the LBA parameters range from high ($r = 0.74$ for $\nu_1^{(1)}$) to nearly perfect ($r = 0.98$ for τ) and the points lie close to the identity line, suggesting good recovery of the LBA parameters. An exception is the parameter σ, which shows a pattern of underestimating the true values, if the true value is relatively high.

As for the parameters characterizing the evolution of the latent states, the recovery of the initial state probability is sub optimal ($r = 0.22$). Overall, the parameter recovery results using MAP are very similar to those using posterior expectation.

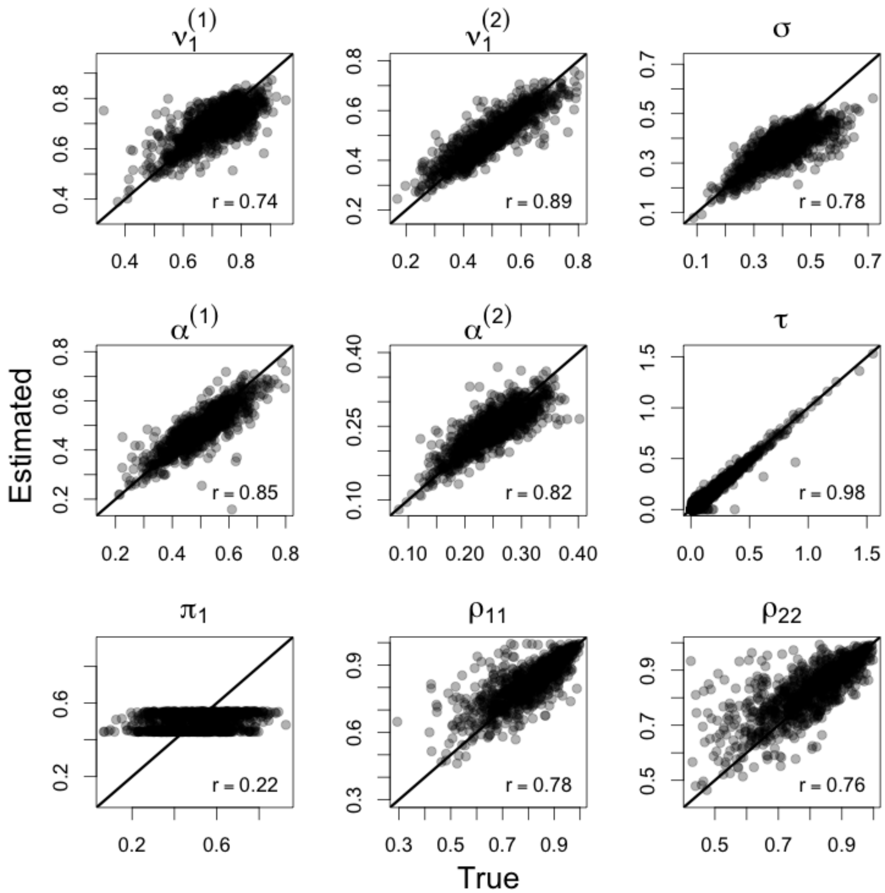

Figure 1. Parameter recovery using maximum a posteriori estimates. Correlation plots between the true values (x-axis) and the estimated values (y-axis). The slope line shows the identity function.

2 Sensitivity analysis

Figure 7 of the manuscript shows the distribution of posterior contraction and posterior z-score for each individual parameter. The following table shows the descriptive statistics of the same.

	Contraction		Z-Score	
parameter	mean	sd	mean	sd
$\nu_1^{(1)}$	0.592	0.228	0.018	0.984
$\nu_1^{(2)}$	0.799	0.135	-0.009	1.009
$\alpha^{(1)}$	0.777	0.175	0.016	1.032
$\alpha^{(2)}$	0.705	0.153	-0.009	0.981
σ	0.609	0.158	-0.013	0.976
τ	0.978	0.013	0.006	0.989
π_1	0.041	0.067	-0.034	0.998
ρ_{11}	0.610	0.377	-0.021	0.933
ρ_{22}	0.611	0.310	0.037	0.985

References

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., . . . Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, *76*(1), 1–32. doi: 10.18637/jss.v076.i01