Lotka-Volterra predator-prey model: SBC

Hudson's Bay Company lynx-hare data

Teemu Säilynoja

2024-07-03

In this notebook, we collect the results of the four other SBC notebooks handling the Lotka-Volterra model.
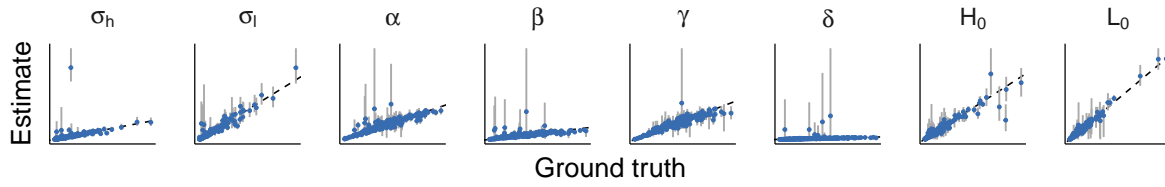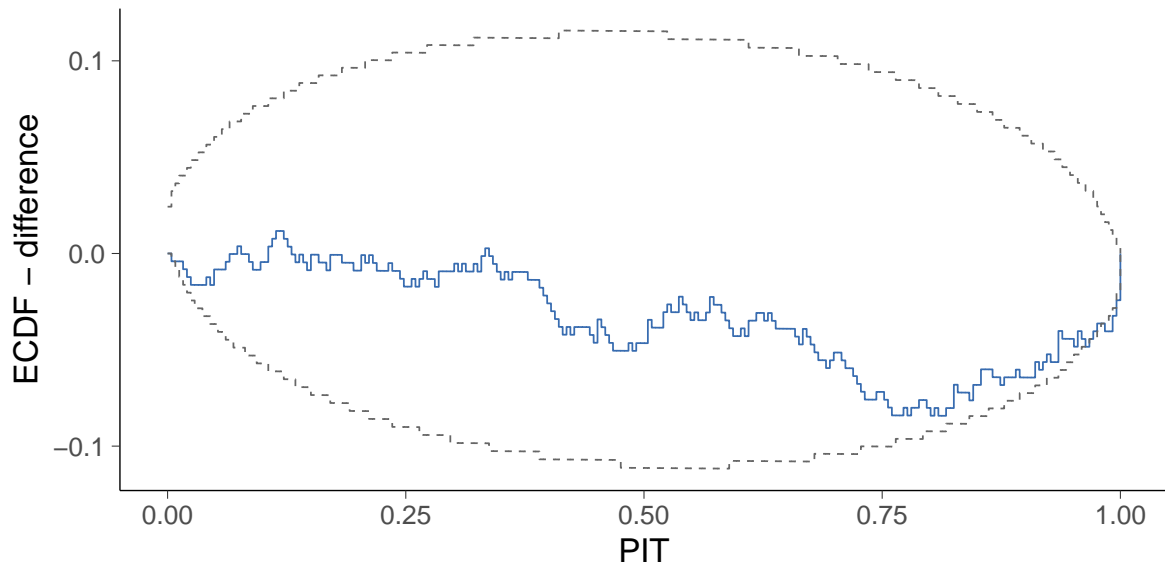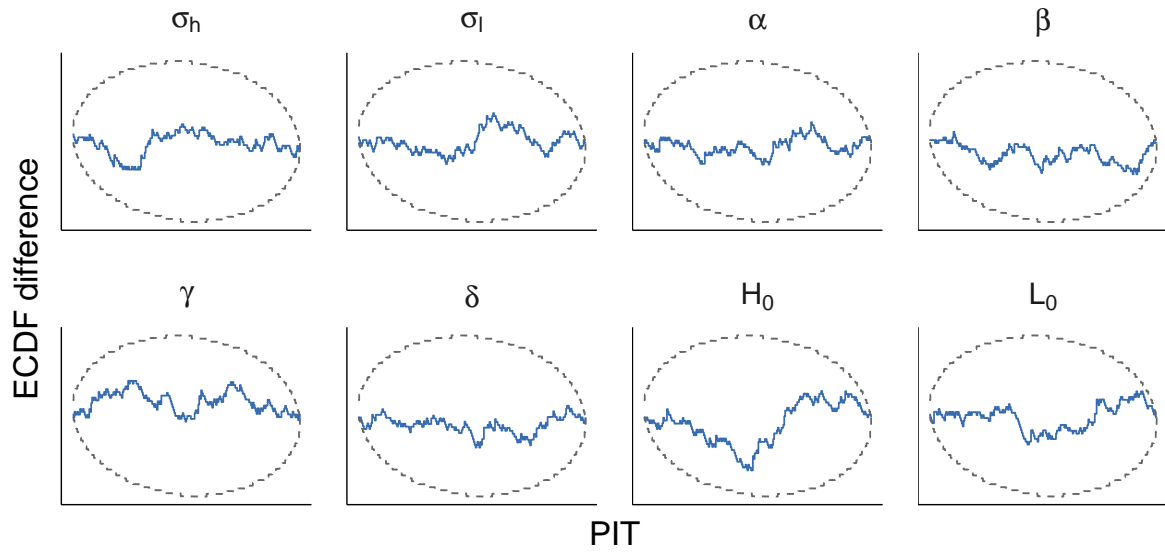
1 Prior SBC

1.1 First attempt

```
SBC diagnostic messages


SBC_results with 250 total fits.
 - 3 (1%) fits resulted in an error.
 - No fits gave warnings.
 - 23 (9%) fits had at least one Rhat > 1.01. Largest Rhat was NA.
 - 18 (7%) fits had tail ESS undefined or less than half of the maximum
 rank, potentially skewing
the rank statistics. The lowest tail ESS was NA.
 If the fits look good otherwise, increasing `thin_ranks` (via
 recompute_SBC_statistics)
or number of posterior draws (by refitting) might help.
 - The lowest bulk ESS was NA
 - 9 (4%) fits had some failed chains.
 - 29 (12%) fits had divergent transitions. Maximum number of divergences
 was 3805.
 - 20 (8%) fits had iterations that saturated max treedepth. Maximum
 number of max treedepth was 3960.
 - 247 (100%) fits had some steps rejected. Maximum number of rejections
 was 7693.
 - Maximum time per chain was NA sec.
```

According to the SBC diagnostics, roughly one in ten of the prior SBC iterations had some divergent transitions, or iterations that saturated the maximum treedepth. Additionally, we received warnings of high $\hat{R}$ values. All of the iterations had some steps rejected. Some rejections are common at the early stages of warm-up stage, but shouldn't continue to the sampling stage after warm-up. Three fits failed completely.

Running 250 SBC iterations took 7h 55m 33.2s.

From the PIT-ECDF plots, we see that there seems to be some tendency to underestimate the joint log-likelihood. We see that some iterations had especially bad parameter recovery.

1.2 Using Pathfinder for initialization

SBC diagnostic messages

```
SBC_results with 250 total fits.
 - No fits had errors.
 - No fits gave warnings.
 - 13 (5%) fits had at least one Rhat > 1.01. Largest Rhat was NA.
 - 8 (3%) fits had tail ESS undefined or less than half of the maximum
 rank, potentially skewing
the rank statistics. The lowest tail ESS was NA.
 If the fits look good otherwise, increasing `thin_ranks` (via
 recompute_SBC_statistics)
or number of posterior draws (by refitting) might help.
 - The lowest bulk ESS was NA
 - No fits had failed chains.
 - 28 (11%) fits had divergent transitions. Maximum number of divergences
 was 3998.
 - 11 (4%) fits had iterations that saturated max treedepth. Maximum
 number of max treedepth was 3978.
 - 250 (100%) fits had some steps rejected. Maximum number of rejections
 was 7445.
 - Maximum time per chain was 942.225 sec.
```
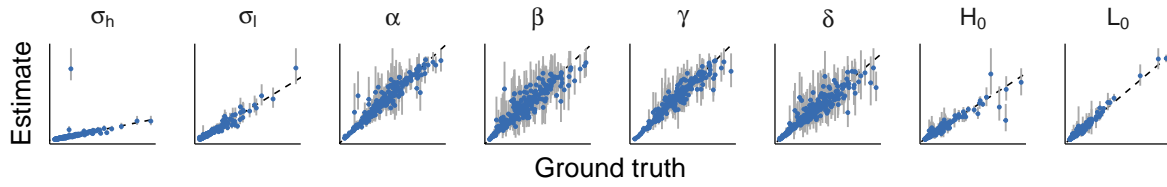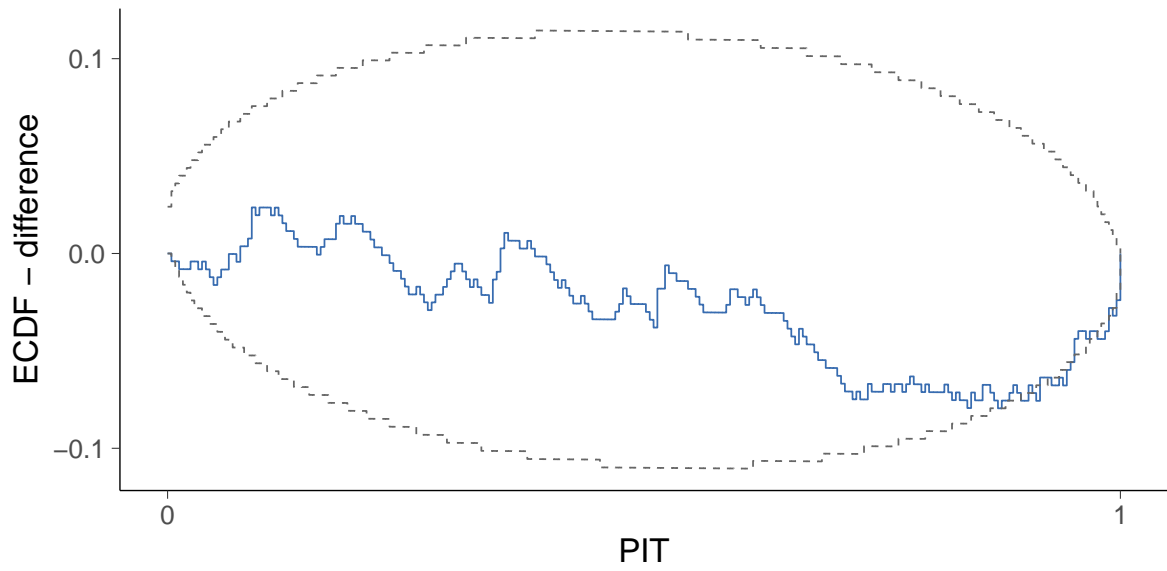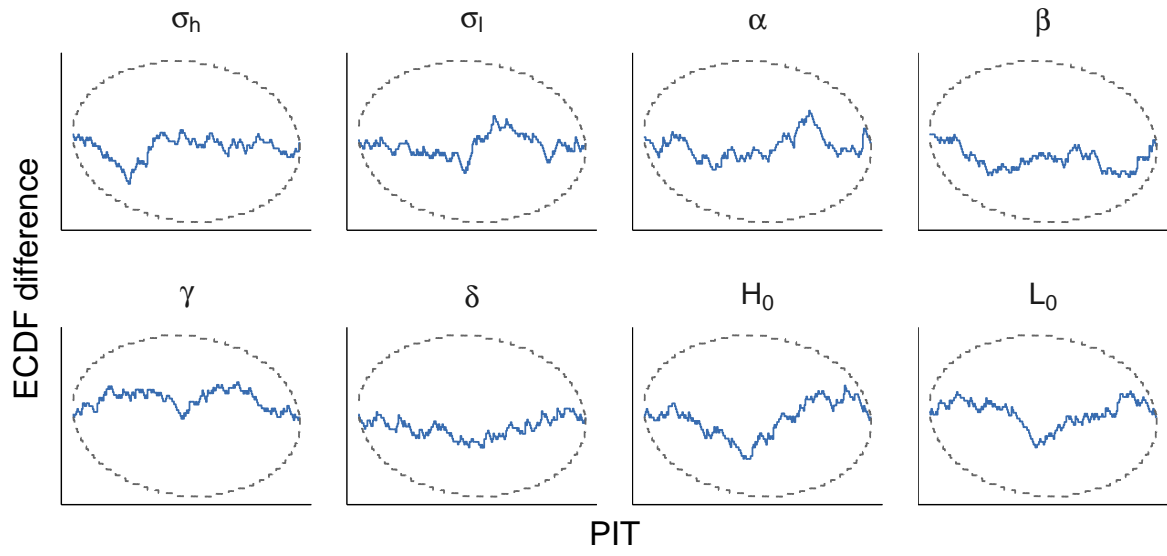
We attempt at improving the inference by initializing the Markov chains with Pathfinder. We don't anymore have fits that would fail completely, and have in general improved the convergence diagnostics. Still the inference is not without problems. This time, running 250 SBC iterations took 6h 31m 52.5s.

The PIT-ECDF plots look very similar, and we still see some cases of bad parameter recovery.

1.3 Conclusion

Even with the Pathfinder initialization, the inference isn't calibrated and we can't trust the model to in general give us trustworthy inference results.

One could try to improve the priors, or the computational aspects of the model. Perhaps adjusting the tolerance of the ODE solver, or increasing the adapt delta and max treedepth parameters of the NUTS algorithm could improve the results.

As the inference is relatively slow, iterating on the model is quite slow.

2 Posterior SBC

Next, we condition our analysis on the historical Hudson Bay Company data of lynxes and hares.

2.1 First attempt

SBC diagnostic messages

```
SBC_results with 500 total fits.
 - 6 (1%) fits resulted in an error.
 - No fits gave warnings.
 - 154 (31%) fits had at least one Rhat > 1.01. Largest Rhat was NA.
 - 149 (30%) fits had tail ESS undefined or less than half of the maximum
 rank, potentially skewing
the rank statistics. The lowest tail ESS was NA.
 If the fits look good otherwise, increasing `thin_ranks` (via
 recompute_SBC_statistics)
or number of posterior draws (by refitting) might help.
 - The lowest bulk ESS was NA
 - 23 (5%) fits had some failed chains.
 - 79 (16%) fits had divergent transitions. Maximum number of divergences
 was 998.
 - 14 (3%) fits had iterations that saturated max treedepth. Maximum
 number of max treedepth was 1000.
 - 494 (100%) fits had some steps rejected. Maximum number of rejections
 was 2026.
 - Maximum time per chain was NA sec.
```
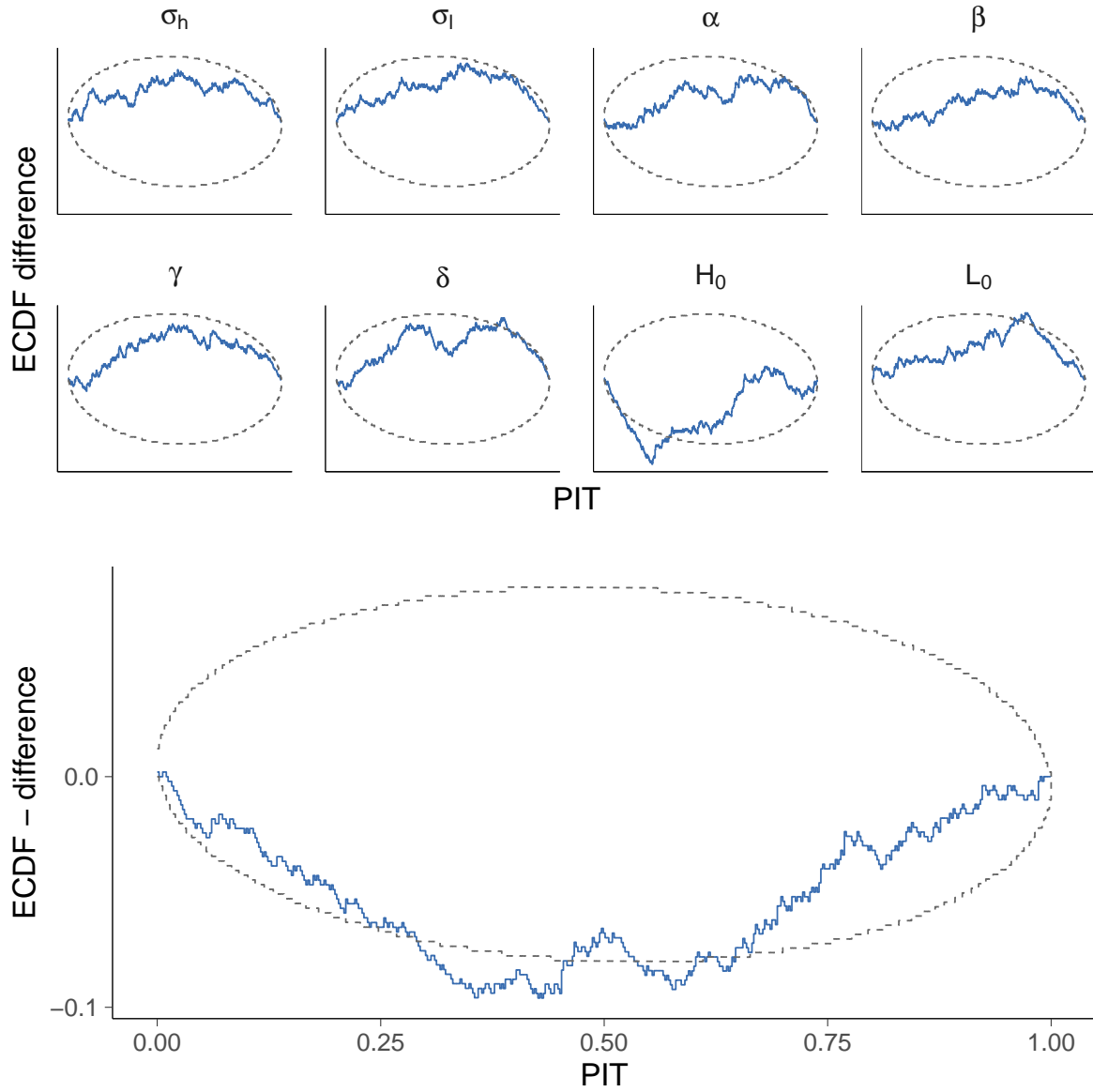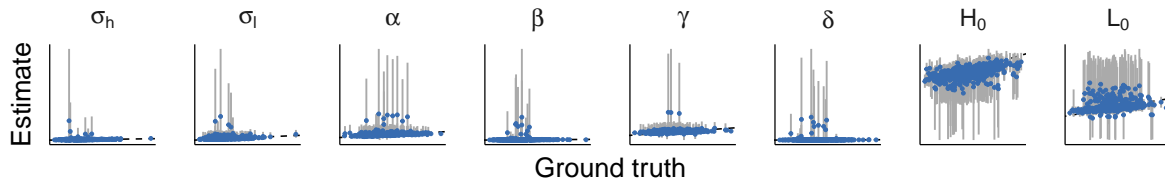
We still observe issues in the inference. Large $\hat{R}$ values are more common than with prior SBC, and we have multiple fits and chains that fail to produce samples.

We do see, that the inference is faster. Running 500 SBC iterations took 5h 46m 4.9s.

The PIT-ECDF plots are telling of great issues with the calibration of the posterior inference. The join log-likelihood is in general over estimated, while especially the estimates of the initial populations seem to have calibration issues.
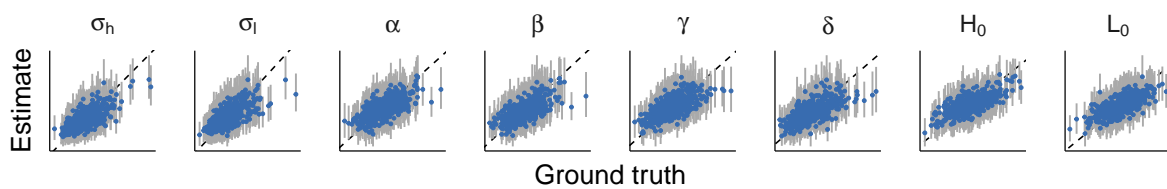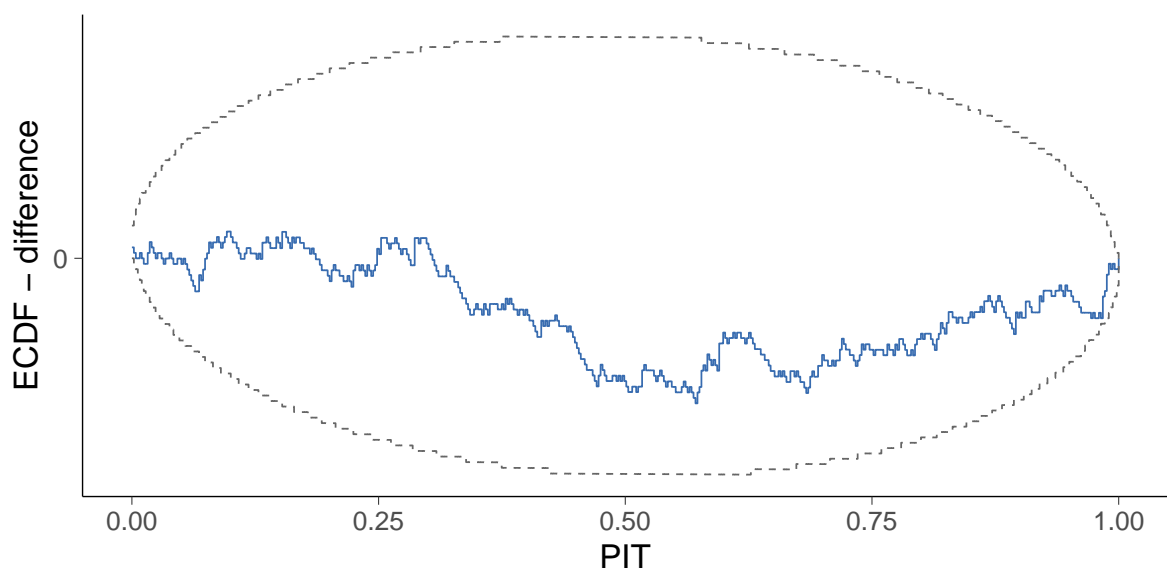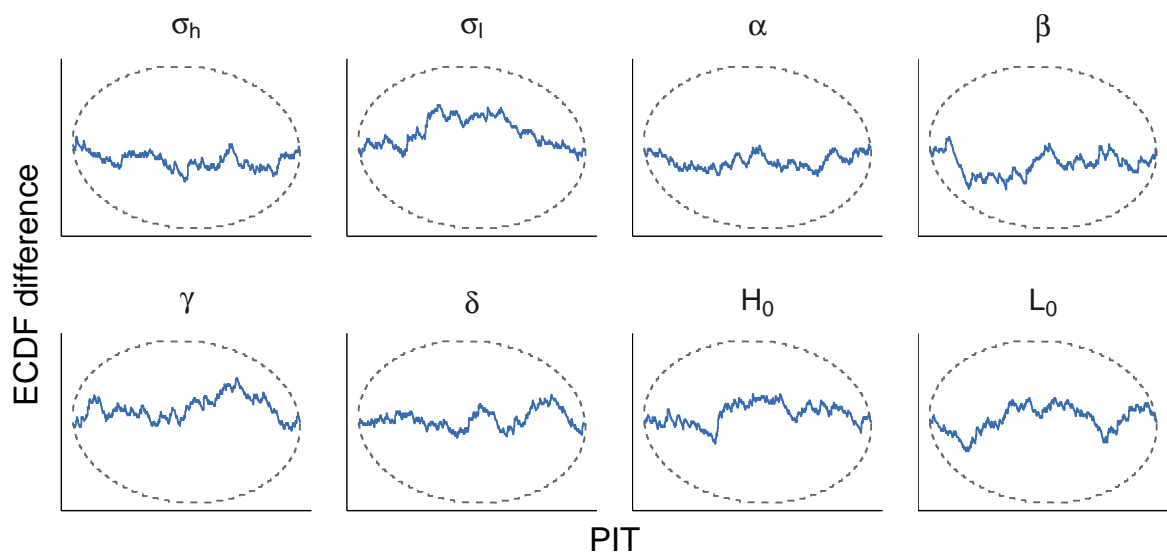
2.2 Using Pathfinder for initialization

```
SBC diagnostic messages


SBC_results with 500 total fits.
 - No fits had errors.
 - No fits gave warnings.
 - 6 (1%) fits had at least one Rhat > 1.01. Largest Rhat was 1.012.
 - All fits had tail ESS > half of the maximum rank.
 - The lowest bulk ESS was 687
 - No fits had failed chains.
 - No fits had divergent transitions.
 - No fits had iterations that saturated max treedepth.
 - 500 (100%) fits had some steps rejected. Maximum number of rejections
 was 25.
 - Maximum time per chain was 32.053 sec.
```

Again, we employ Pathfinder to initialize the Markov chains. The high $\hat{R}$ values as well as some individual posterior plots indicate multi-modality. We still have a warning about large $\hat{R}$ values, but now even the largest $\hat{R}$ is only 1.012. Additionally we have no warnings on divergent transitions or max treedepths. We still have rejections in the sampling, but the maximum number of rejections was 25 and was likely to take place in the early stages of the warm-up.

We also observe a massive boost in sampling efficiency, as running 500 SBC iterations only took 2h 31m 51.4s.

All of the PIT-ECDF plots look good, and the parameter recovery is decent.

2.3 Conclusion

After posterior SBC, the inference with Pathfinder initialization looks good enough for us to use the model.