CI lengths simulation

Rike

2023-11-21

Simulation results

For each run, we simulated n observations from a standard normal distribution and subsequently calculated the following two interval lengths:

- $l_1 = Q\left((X_i)_{i=1}^n, 0.9\right) - Q\left((X_i)_{i=1}^n, 0.1\right)$ "directional method"
- $l_2 = 2 \cdot Q\left((|X_i|)_{i=1}^n, 0.8\right)$ "absolute method"

where $(X_i)_{i=1}^n$ is the simulated data and $Q(\cdot, \tau)$ is the quantile function at level τ. In a finite sample setting, there are different ways of calculating Q: we use methods 1, 7, 8 from the `quantile()` function (`type` argument). See short note at the bottom of this page.

In the previous IMF application, we have observed that l_2 is usually larger, using the default quantile calculation method `type=7`.

To investigate whether this is a more universal result, we simulate data with $n \in [7, 50000]$ (evenly spaced on a log scale) and 10000 iterations each. For each n, we record the proportion of instances where $l_2 > l_1$, dependent on the quantile calculation method. The simulation results are shown in the plots on the following page, separately for the two confidence levels.

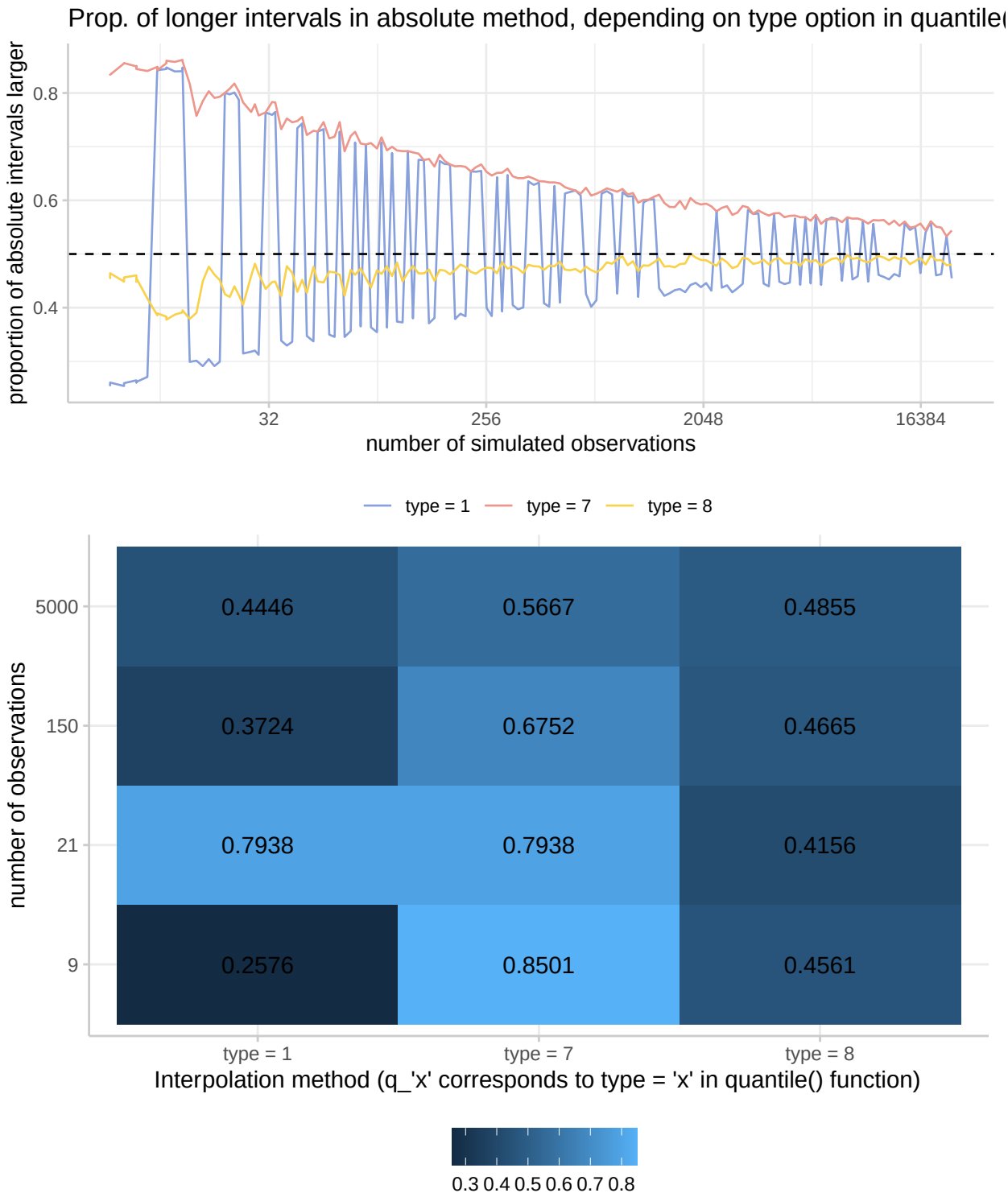
Short note on the different types

The following is taken from (Hyndman & Fan 1996: *Sample Quantile in Statistical Packages*). Fore more details, we refer to that text.

- `type = 1` does not interpolate between observations. As the inverse of the empirical distribution function, it always "jumps to the upper value". Thus, severe discontinuties can arise (see following plots)
- `type = 7` and `type = 8` linearly interpolate between observed data points, but do so in different ways, i.e. they choose different midpoints between different values for the cumulative probabilities (see Hyndman)

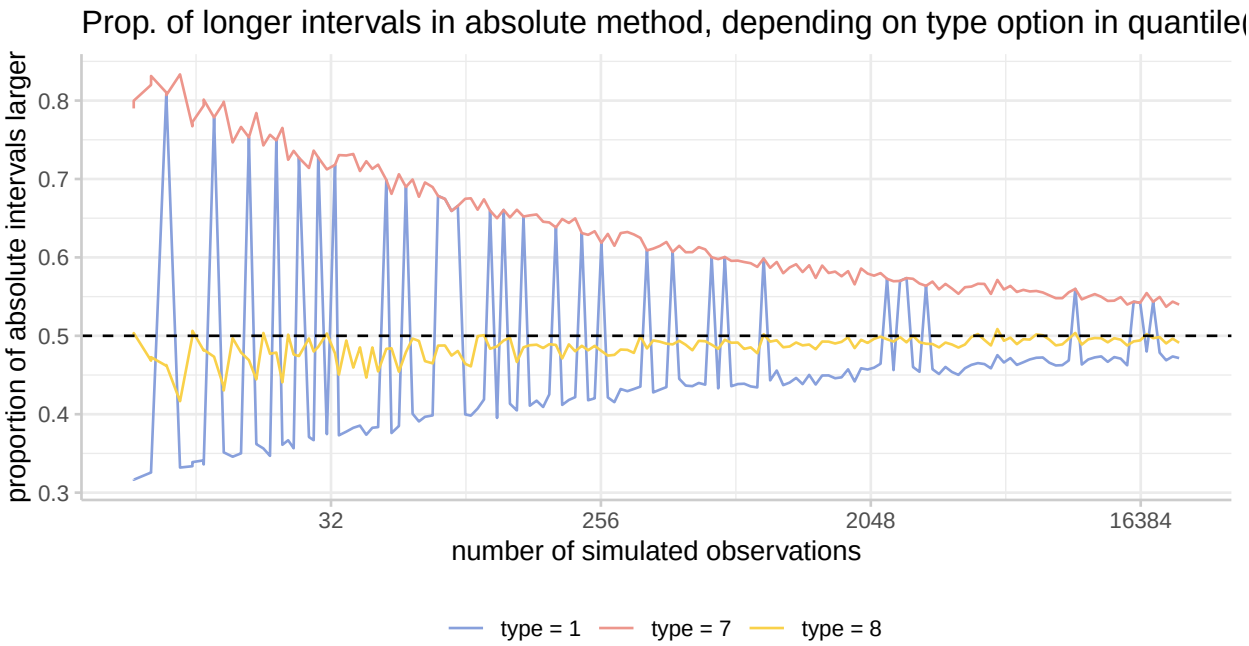
For confidence level 80%

```
simdat <- data.table::fread(here("simulations", "results", "simdat_props80.csv"))
```



For confidence level 50%

```
simdat <- data.table::fread(here("simulations", "results", "simdat_props50.csv"))
```



We observe that the relative length of the two interval methods vastly differs between the 3 types. `type = 1` is highly discontinuous and "switches around" substantially with rising n. The two interpolating types `type = 7` and `type = 8` both approach 0.5 with increasing n. Importantly, for `type = 8`, the interval lenghts for the two methods are overall closer to 0.5 and thus more balanced.

What does this mean for the IMF project

So far, we used `type = 7`, the default for the `quantile()` function - here, in $\sim$ 80 percent of cases for both the 50% and 80% interval, the absolute method produced longer intervals. We reran the analysis for `type = 8` - now this is only the case for $\sim$ 60 percent of cases.

Importantly, the coverage improves slightly for the directional method and coverage rates are closer to nominal levels (coverage for the absolute method specifically remain mostly unchanged).

More detailed investigation should perhaps be done, but I think this could be a real argument for using `type = 8` interpolation method and, additionally, perhaps considering changing to directional errors as a default after all.

(Scores overall and in particular ordering between forecast sources also remain mostly the same, so are not included here.)

GDP Growth

Inflation

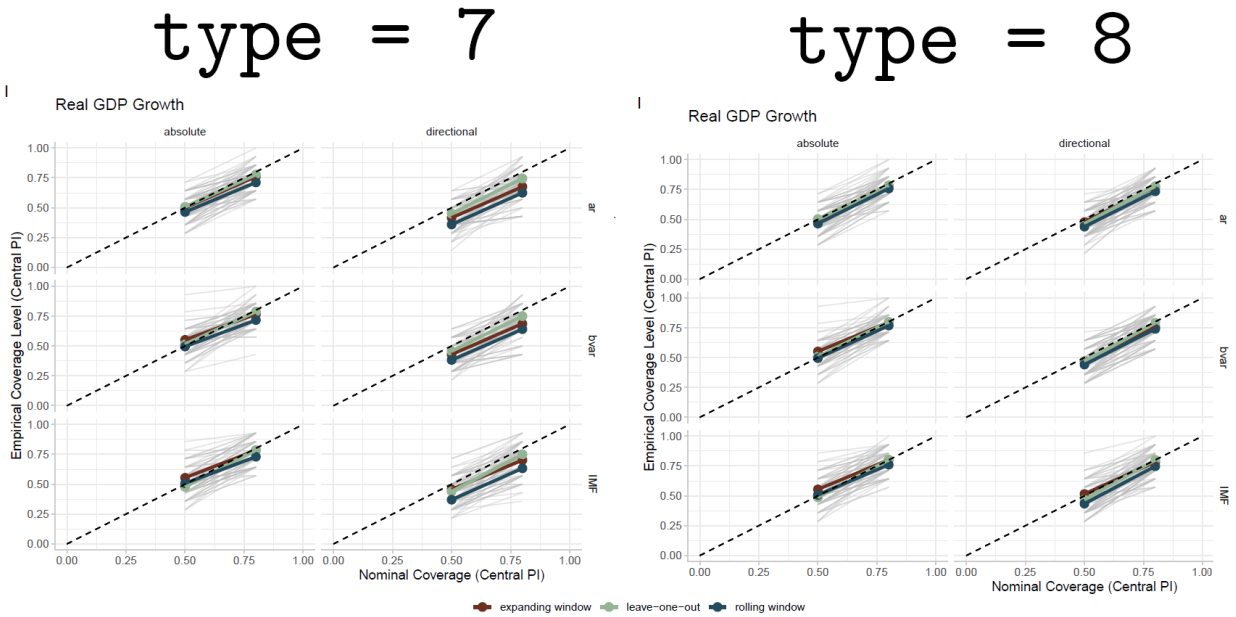


Figure 1: Changes in central interval coverage - GDP Growth.

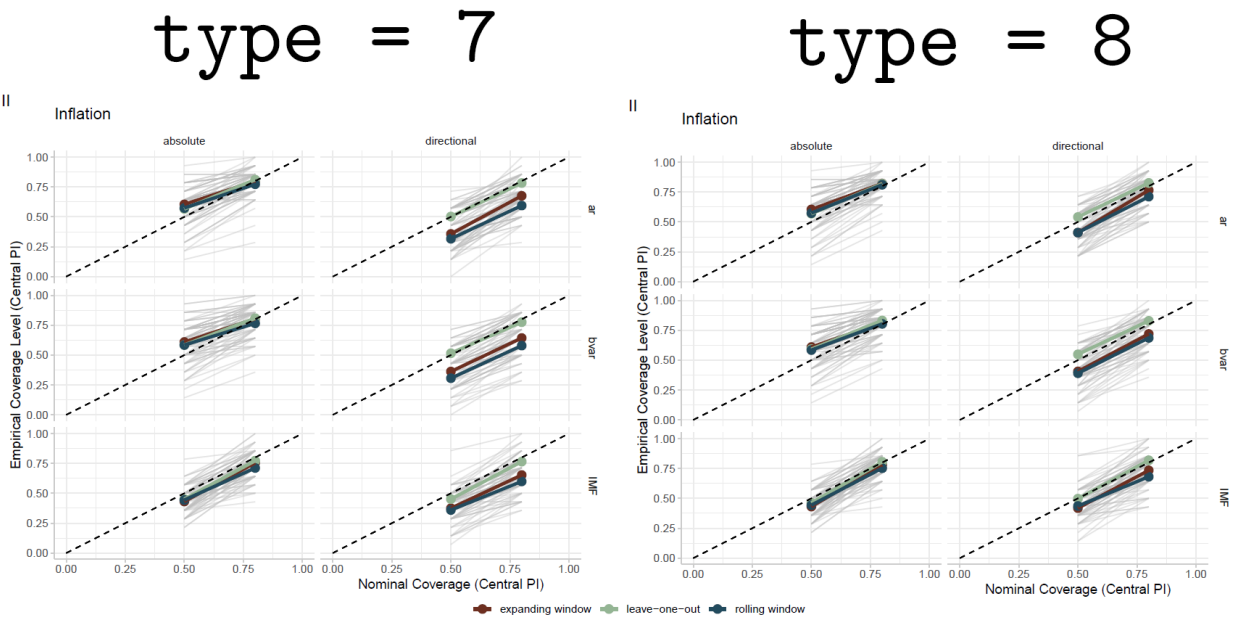


Figure 2: Changes in central interval coverage - Inflation.