

Time delta (δ)

Renzo Miguel Caballero Rosas

March 28, 2020

Note: We compute the following results using **initial_time.m**.

Copied from the paper (19/03/20):

We can estimate δ solving the problem

$$\delta \approx \arg \min_{\delta} \mathcal{L}_{\delta}(\boldsymbol{\theta}, \delta; V^{M,1}) = \arg \min_{\delta} \prod_{j=1}^M \rho_0(V_{j,t_0} | V_{j,t_{-\delta}}; \boldsymbol{\theta}, \delta).$$

Copied from MATLAB's slides (19/03/20):

We have that, for almost all days, at time $t = t_0$, $X(t_0) \neq p(t_0)$ and then $V(t_0) \neq 0$. However, by forecast construction, there should exist a time $t_{-\delta} < t_0$ such that $V(t_{-\delta}) = 0$.

We extrapolate linearly $p(t)$ so we can evaluate $p(t_{-\delta})$, and assume that $V(t_{-\delta}) = 0$. Then, for each day j , we have an initial transition $(V_{j,t_0} | V_{j,t_{-\delta}}; \boldsymbol{\theta}, \delta)$. We assume again that it is Beta and apply the same moment matching as for the rest of transitions. With our initial guess for $\boldsymbol{\theta}$, we can construct our initial guess for δ solving the problem

$$\delta \approx \arg \min_{\delta} \mathcal{L}_{\delta}(\boldsymbol{\theta}, \delta; V^{M,1}) = \arg \min_{\delta} \prod_{j=1}^M \rho_0(V_{j,t_0} | V_{j,t_{-\delta}}; \boldsymbol{\theta}, \delta).$$

We got $\delta \approx 80 \text{ min}$.

Step by step: (WLOG we can use $t_0 = 0$ and $t_{-\delta} = -\delta$)

We are imposing that $V_{j,t_{-\delta}} = 0$ for all $j \in \{1, \dots, M\}$, there $t_{-\delta} = t_0 - \delta$. Then, for each path j , we have that the first moment $m_1(t) = 0$ for all $t \in [t_{-\delta}, t_0]$.

The second moment is given by

$$\begin{cases} \frac{dm_2}{dt} &= -\underbrace{2\theta_t(1+\alpha)}_{k_1(t)} m_2 + \underbrace{2\theta_t\alpha}_{k_2(t)} \underbrace{p_t(1-p_t)}_{f(t)} \\ m_2(t_{-\delta}) &= 0. \end{cases} \quad (1)$$

Then, as $m_1(t_0) = 0$, the R.V. $V_{j,t_0} | (V_{j,t_{del}} = 0)$ can be extracted from

$$2 \times \text{Beta} \left(\frac{1 - m_2(t_0)}{2m_2(t_0)}, \frac{1 - m_2(t_0)}{2m_2(t_0)} \right) - 1,$$

which has mean and variance

$$\mu = 0, \quad \sigma^2 = m_2(t_0).$$

Step by step:

If we assume that $\theta_t = \theta_0$ in $[t_{-\delta}, t_0]$, the system (1) has solution

$$m_2(t) = e^{-k_1 t} \int_{t_{-\delta}}^t k_2 e^{k_1 s} f(s) ds.$$

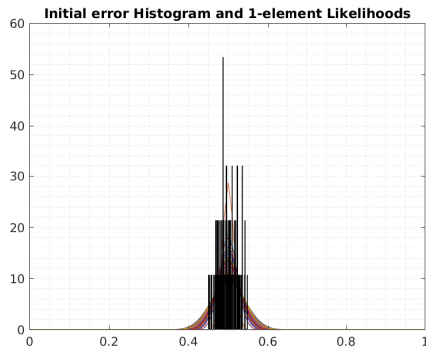
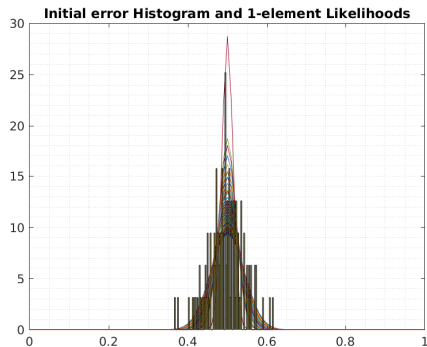
As $p_t(1 - p_t) \leq 1/4$ in $[t_{-\delta}, t_0]$, we have the bound

$$m_2(t) \leq \frac{\frac{k_2}{4} \int_{t_{-\delta}}^t e^{k_1 s} ds}{e^{k_1 t}} = \frac{k_2}{4k_1} \left(1 - \frac{e^{k_1 t_{-\delta}}}{e^{k_1 t}} \right) \leq \frac{1}{4} \left(1 - e^{k_1(t_{-\delta} - t)} \right).$$

Then, $\lim_{t_{-\delta} \rightarrow -\infty} m_2(t_0) \leq 1/4 \implies \sigma^2 \leq 1/4$.

Then, if we need more variance to model the initial transition, we can not use this model. Noticed that we assumed $\theta_t = \theta_0$ in $[t_{-\delta}, t_0]$, this is not true in general as $p_{t_{-\delta}}$ can be very near the boundaries ε and $1 - \varepsilon$. However, even with the assumption, to have limit variance is a problem as for some set of data the assumption can hold. Still, if the initial error is so large, we would not be able to model it correctly with limited variance.

Step by step: We remove points with $|V_{j,t_0}| > \eta = 0.1$



Here we can see all the likelihoods for each initial transition when $\delta = \text{*one day*}$, overlapped with histograms of the initial error. On the left, we use all the initial errors, and clearly, the distribution that fits that histogram needs more variance. On the right, we removed the points with more error, and we can see that we need $\delta < \text{*one day*}$.

Step by step: Explanation Problem-Solution

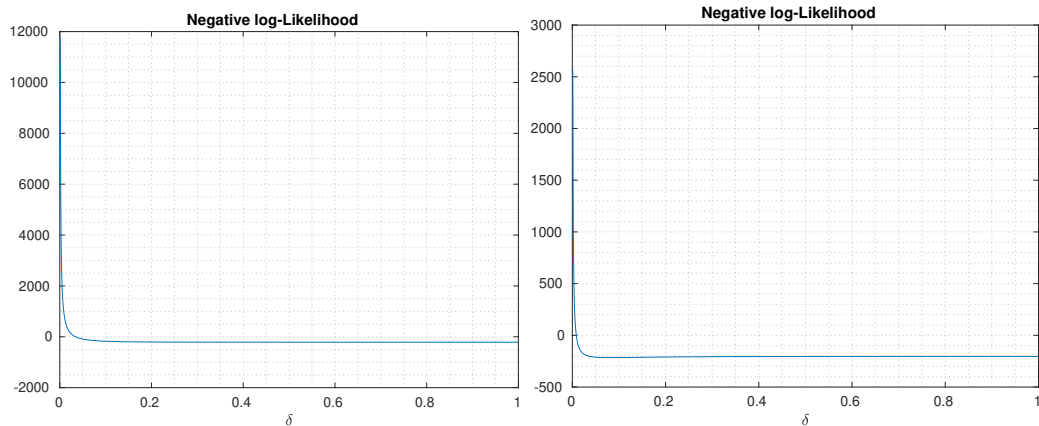
Basically, what happens is that the variance in the initial point grows with δ (the more in the past we *start*, the more variance we have at time $t = 0$ or t_0).

The problem is that, as $\delta \uparrow$, we have that $m_2(t_0) \uparrow$. At the same time, as $m_1(t) = 0$ (because we fix $V_{j,t-\delta} = 0$ for all $j \in \{1, \dots, M\}$), we can prove that $m_2(t_0) \leq 1/4$ for all possible $\delta > 0$.

Now, we have that $\xi_1 = \xi_2$ in the Beta distribution, and then $\sigma^2 = m_2(t_0)$ for the distribution of $V_{j,t-\delta}$. Then, the variance is bounded by $1/4$. This means that, if the set of errors at the initial time has a larger variance, we will not be able to reach the correct Beta distribution using the moments matching.

How we solve this problem? Two options: Change model to one with more possible variance, or remove the data with substantial error. **We remove all data s.t. $|V_{j,t_0}| > \eta$ for some $\eta > 0$.** To create this new dataset we use **new_batch_fixed_removed_samples.m**.

Step by step:

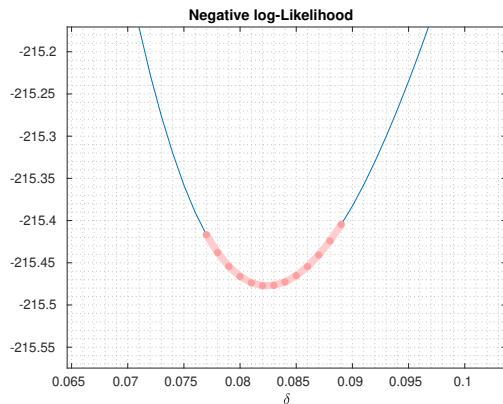


On the left, negative log-likelihood using all the data. On the right, using only the ***corrected data*** with $\eta = 0.1$. In the next slide we will see a ***zoom*** of the right plot.

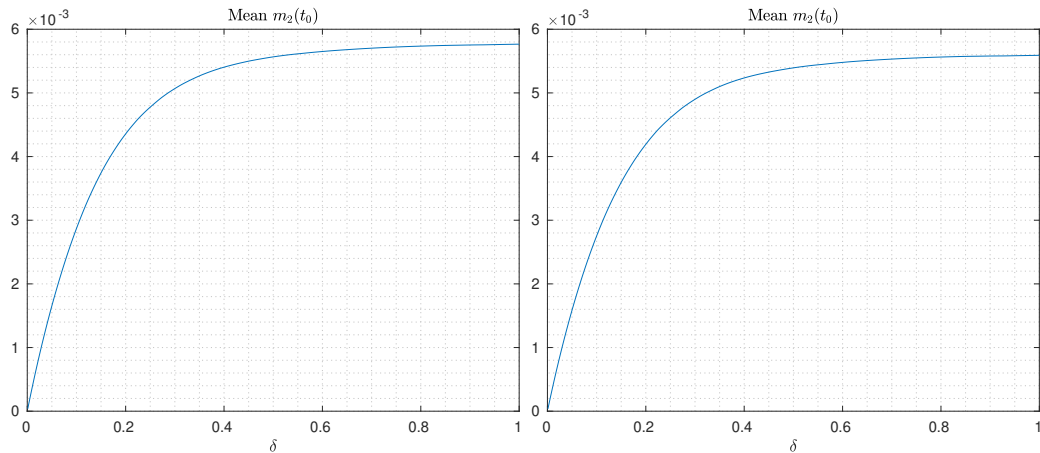
Step by step:

We can observe that the minimum, for $\eta = 0.1$, is in $\delta \approx 0.083 \approx 120 \text{ min} = 2 \text{ h}$.

If we choose (let say...) $\eta = 0.08$, we obtain that $\delta \approx 90 \text{ min}$.

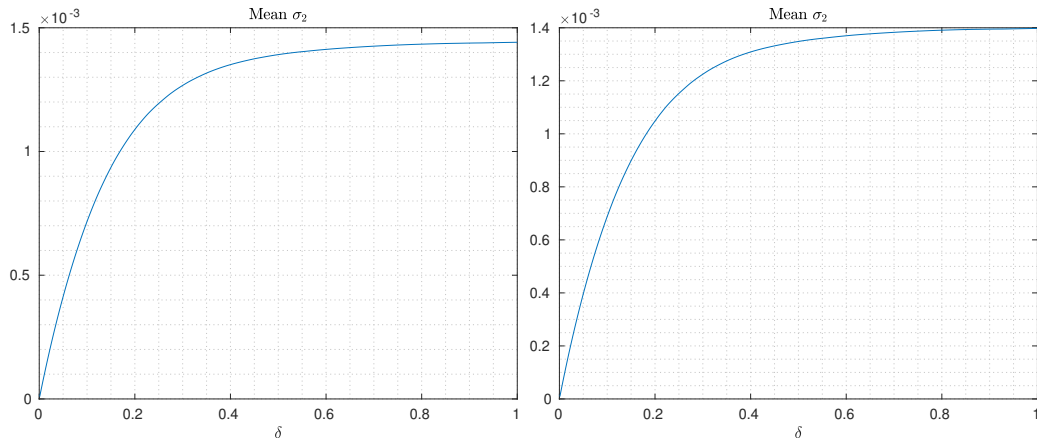


Step by step: Averaged $m_2(t_0)$ over all the the used samples



On the left, all initial points. On the right, only the ***corrected data*** with $\eta = 0.1$.

Step by step: Averaged $\sigma^2(t_0)$ over all the used samples

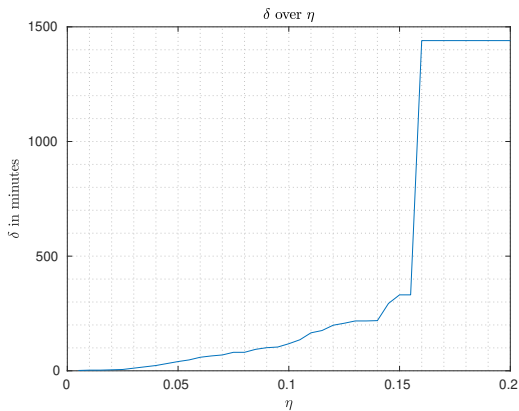
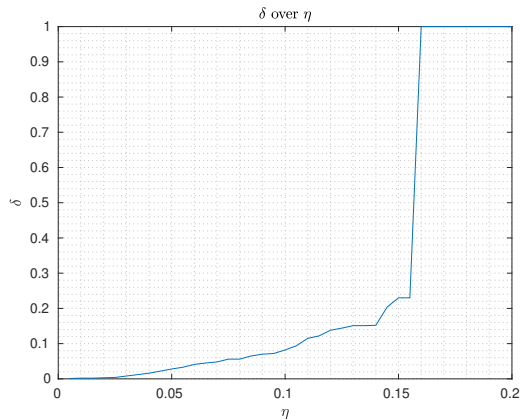


On the left, all initial points. On the right, only the ***corrected data*** with $\eta = 0.1$.

Some remarks:

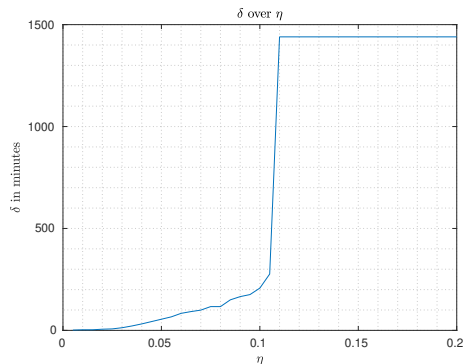
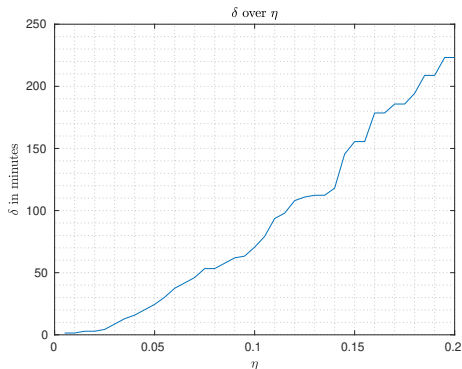
- ▶ We showed plots with $\eta = 0.1$. Smaller values also work, but we want to modify the data as few as possible. **Can we plot δ w.r.t. η ? See next slide.**
- ▶ $1/4$ is a very large upper bound, as we assumed $p_t(1 - p(t)) = 1/4$ for all time. This is the reason why the plots show values of variance smaller by orders of magnitude.
- ▶ By construction of our model, if $m_1(t_0) = 0$, then $\xi_1 = \xi_2$. And then, $\sigma^2 = m_2(t_0)$.
- ▶ We bounded the variance using that $m_1(t) = 0$. With no this assumption, maybe we can reach any value of variance.

δ over η :



We choose $\eta = 0.15$ which leads to $\delta = 330 \text{ min} = 5.5 \text{ h}$. However, this plot is highly dependent on α and θ_0 . Then, our selection of δ may change before and after finding the optimal parameters.

δ over η :



On the left, we use the initial (θ_0, α) ($\alpha\theta_0 = 0.1$) and we get $\delta = 220$ min.

On the right, we use the optimal ones. As the optimal product $\alpha\theta_0$ ($\alpha\theta_0 = 0.08$) is smaller than the initial one, we need a larger δ to match the variance of the initial error, or to remove some data. If we remove up to $\eta = 0.1$, we get $\delta = 220$ min.