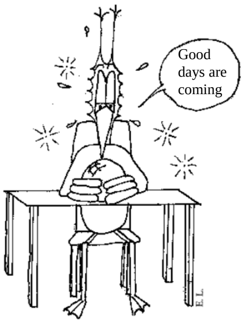
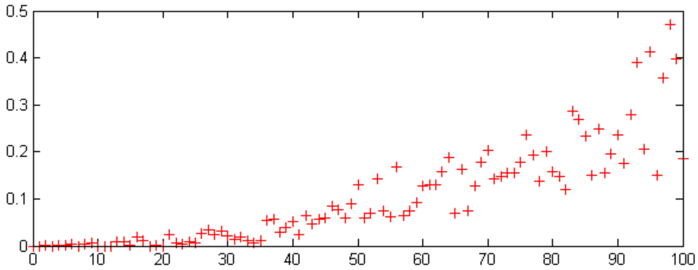
PERFORMANCE EVALUATION EXERCISES

FORECASTING

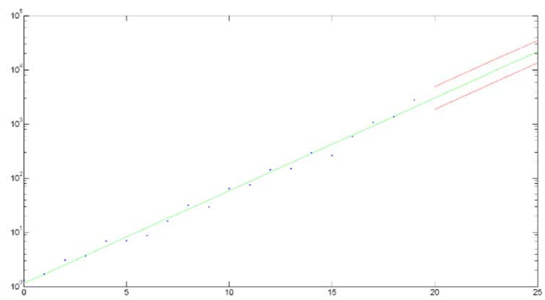
Jean-Yves Le Boudec, Spring 2021



1. The following data shows the amount of memory claimed by a server process, in percent of the total physical memory, as a function of times in seconds since last reboot. The server should be rebooted 10 seconds before the used memory reaches the threshold $\theta = 90\%$ (of the physical memory). Explain a method for deciding when to reboot.



2. We fit the log of virus expansion data using Laplace noise.



The model is

$$L_i = \ell + \alpha t_i + \varepsilon_i \text{ with } \varepsilon_i \sim \text{ iid Laplace}(\lambda)$$

where L_i is the logarithm of the ith value and t_i the time of measurement.

   (a) Write a linear program that you can use for estimating ℓ, α and λ.

   (b) When X is Laplace noise with parameter λ, for which value of η do we have $\mathbb{P}(|X| > \eta) = 0.05$?

   (c) We want to use the estimated model to predict the virus expansion at a time T. Give the formula for a 95%-prediction interval, assuming we can neglect the estimation uncertainty.

3. Δ_k is the differencing filter at lag k.

   (a) Is Δ_{16} stable ? Is it invertible ? If so, is the inverse stable ?

   (b) Say which is true

      i. ☐ Δ_{16} is a FIR filter

      ii. ☐ Δ_{16} is an AR filter

      iii. ☐ Both

      iv. ☐ None

   (c) Compute the $MA(\infty)$ and $AR(\infty)$ representations of Δ_{16}.

   (d) Let $F = \Delta_1 \Delta_{16}$ and $G = \Delta_{16}\Delta_1$. Give the operator- and the input-output-representations of F and G.

   (e) Is F stable ? Is it invertible ? If so, is the inverse stable ?

4. We want to forecast the temperature $T_1, T_2, ..$ where there is one measurement every hour. We want to use a differencing filter at lag 24. Let X_n be the differenced time series. Give the formulas to compute T from X and vice versa. We find that X_n looks almost iid with mean μ. We want to use this fact to give a point prediction for T_{n+5}, assuming we are at time n (where n is large). Give the formulas for this point prediction.

5. We have a times series Y_t. We computed the differenced time series $X_t = Y_t - Y_{t-1}$ and found that X_t can be modelled as an AR process:

$$X_t = 0.5X_{t-1} + \varepsilon_t \text{ with } \varepsilon_t \sim \text{ iid } N_{0,\sigma^2}$$

   for some value of σ that we have estimated.

   (a) Is this a valid ARIMA model ?

   (b) Compute a point forecast $\hat{X}_t(2)$

   (c) Compute a point forecast $\hat{Y}_t(2)$

   (d) Compute the first 3 terms of the impulse response of the filter $\varepsilon \mapsto Y$

   (e) Compute a prediction interval for Y_{t+2} done at time t.

   (f) Which of the following algorithms is a correct implementation of computing a prediction interval for Y_{t+2} done at time t using the bootstrap from residuals ?

   Algorithm A
   compute the time series $\varepsilon_s = X_s - 0.5X_{s-1}$ for $s = 3 : t$
   for $r = 1 : 999$ do
       draw e_s^r, $s = 3 : (t+2)$ with replacement from ε_s, $s = 3 : t$
       compute $X_{1:t}^r, Y_{1:t}^r$ and $\hat{Y}_{1:t}^r(2)$ using $X_s^r = 0.5X_{s-1}^r + e_s^r$, $Y_s^r = X_s^r + Y_{s-1}^r$
           and the formula you have found for $\hat{Y}_{1:t}^r(2)$
       $Y_{t+2}^r = e_{t+2}^r + 1.5e_{t+1}^r + \hat{Y}_t^r(2)$
   end do
   prediction interval is $[Y_{t+2}^{(25)};\ Y_{t+2}^{(975)}]$

Algorithm B
compute the time series $\varepsilon_s = X_s - 0.5X_{s-1}$ for $s = 3 : t$
for $r = 1 : 999$ do
    draw e_1^r, e_2^r with replacement from ε_s, $s = 3 : t$
    $Y_{t+2}^r = e_1^r + 1.5e_2^r + \hat{Y}_t(2)$
end do
prediction interval is $[Y_{t+2}^{(25)};\ Y_{t+2}^{(975)}]$

i. ☐ A
ii. ☐ B
iii. ☐ A and B
iv. ☐ None