HOMEWORK 4 – SAMPLE SOLUTIONS

IMPORTANT NOTE

These homework solutions show multiple approaches and some optional extensions for most of the questions in the assignment. You don't need to submit all this in your assignments; they're included here just to help you learn more – because remember, the main goal of the homework assignments, and of the entire course, is to help you learn as much as you can, and develop your analytics skills as much as possible!

Question 7.1

Describe a situation or problem from your job, everyday life, current events, etc., for which exponential smoothing would be appropriate. What data would you need? Would you expect the value of α (the first smoothing parameter) to be closer to 0 or 1, and why?

Here's one possible situation.

An automobile manufacturer might want to study the effect of preventive engine maintenance on gas mileage. Every car in the study is driven the same distance on the same indoor road course (to eliminate temperature effects) each day. For every car in the study, the car's fuel efficiency is recorded every week. Some cars receive preventive maintenance every 6 weeks, some every 10 weeks, and some every 14 weeks.

For each car, the manufacturer could build an exponential smoothing model, with weekly gas mileage as the values being studied. It would include cyclic effects (a cycle would be 6 weeks, 10 weeks, or 14 weeks, depending on the maintenance interval for the car), and the trend would help show how quickly gas mileage deteriorates over time.

I'm not a car expert, but I would expect that there wouldn't be too much variability in gas mileage from week to week for the same car (other than trend and cyclic effects). So, I'd expect the value of α to be closer to 1.

Question 7.2

Using the 20 years of daily high temperature data for Atlanta (July through October) from Question 6.2 (file `temps.txt`), build and use an exponential smoothing model to help make a judgment of whether the unofficial end of summer has gotten later over the 20 years. (Part of the point of this assignment is for you to think about how you might use exponential smoothing to answer this question. Feel free to combine it with other models if you'd like to. There's certainly more than one reasonable approach.)

Note: in R, you can use either `HoltWinters` (simpler to use) or the `smooth` package's `es` function (harder to use, but more general). If you use `es`, the Holt-Winters model uses `model="AAM"` in the function call (the first and second constants are used "A"dditively, and the third (seasonality) is used "M"ultiplicatively; the documentation doesn't make that clear).

Here's one possible solution. Please note that a good solution doesn't have to try all of the possibilities in the code; they're shown to help you learn, but they're not necessary.

The file `solution 7.2.R` shows how to use `HoltWinters` and run single, double, and triple exponential smoothing (including both additive and multiplicative seasonalities). In all cases, the final trend estimate seems to be just about zero, suggesting that the data don't show significant increases or decreases over the 20-year period.

To answer whether the unofficial end of summer has gotten later, we can look at the seasonal factors for every data point, and run a CUSUM analysis on them as in the previous homework (e.g., for every year, we find the day where a change is detected, and see if that date gets later over time). I won't repeat all of that analysis here, but for most values of C and T, it doesn't seem to show a change.