

Econ 803: Homework#2

immediate

Due: Mar 27, 11:59pm

1 Filtering an artificial time series(50pts)

This question requires you to generate an artificial time series and then apply several smoothers to extract the signal component.

1. (10pts)Using Stata to construct an artificial data that meets the following properties.

- (a) Set 100 observations
- (b) The series follows the equation

$$\begin{cases} y_t = t + \varepsilon_t & t \neq 20, t \neq 70 \\ y_{20} = 40 & t = 20 \\ y_{70} = 35 & t = 70 \end{cases}$$

where ε_t is the random term that follows normal distribution with a mean of 0 and a standard deviation of 2

- (c) You may rename or label the series where appropriate.
2. After generating the series, please answer the following questions
 - (a) (2pts)Plot the trend and the raw series on one graph.
 - (b) (8pts)Apply the following smoothers and generate the smoothed series.

- Smoothers: Span-3 median smoother; Hanning smoother; Hanning and Span-3 median compound smoother; Span-7 median smoother;
- (c) (5pts) Evaluate the performance of four smoothers by plotting the residuals between the raw data and the smoothed values. Report the graph and briefly show your eyeball findings.
- (d) (5pts) Now evaluate the performance of four smoother through the distributions of the residuals. Report the graphs and briefly show your findings.
- (e) (5pts) Now evaluate the performance of four smoothers using the mean and standard deviations of the residuals. Show the descriptive statistics of the residuals and make an educated guess on which filter performs better.
- (f) (15pts) Now suppose that the acceptable percentage of deviation from raw data should not exceed 2%, please re-evaluate the performance of each smoother. Hint:
- In the first step, you might need to generate the percentage deviation from raw data using the formula:

$$\text{percent of deviation} = \frac{\text{residuals}}{\text{raw value}}$$

- Then, you might need to count the number of observations whose absolute value of percent of deviation is less than 0.02. Command hint: `count if abs(percent of deviation)<0.02`
- Next, you generate a local variable that record the count. Command hint: `local numerator "r(N)"`
- And finally, you calculate the share of observations that fall in the ± 0.02 acceptance band. Command hint: `display 100*'numerator'/100`

2 Forecasting the Consumption Expenditures(50pts)

Please go to <https://fred.stlouisfed.org/> and download the Personal Consumption Expenditures(Not Seasonally Adjusted) from Jan 1,1980 to Oct 1, 2020. The series code is **NA000349Q**.

Please answer the following questions:

1. (2pts)Convert the date variable from the 01jan1980 format to 1980q1 format; And generate the log of consumption. Hint:

```
. format observation_date %td  
  
. gen date_quarter=qofd(observation_date) /*Notes: observation_date is your time variable */  
  
. format date_quarter %tq  
  
. tsset date_quarter, quarterly
```

2. (3pts)Plot the time series of log of consumption. (4pts)Briefly describe what types of signal the series contains. (6pts)Based on the types of signal, select the most appropriate filter to project 11 years using the full sample, that is, the projection date is on 2020Q3.
3. (5pts)Plot both the log of consumption and the smoothed value after 2009Q4. Report your graph
4. (30pts)A standard way of evaluating the reliability of a forecast model is **backtesting**. The basic idea is to apply the model to an earlier period and compare the projections of the model with the actual outcome. With this in mind, please do the followings,
 - (a) (15pts)**Train the model(filter) you selected in 2 using data prior to 2010Q1, then use that estimate to project the remaining quarters.** (Hint: Count the number of quarters from 2010Q1 to 2020Q3)
 - (b) (3pts)**Generate the forecast errors between the forecast values and the actuals.**
 - (c) (12pts)Plot the **log of consumption, forecast values, and the forecast errors** on one graph between 2000Q1 and 2020Q3; Evaluate the predictive power of the model for

short-run and long-run forecast, respectively. (Hint: In the graph, you may need to put **forecast errors** on the other y axis.)

3 Bonus question: When will COVID-19 end?(10pts)

When is this finally going to end? That's the question on many minds after a year of living through the COVID-19 pandemic.

But public health experts say we do have an answer, and you're not going to like it: COVID-19 is never going to end. It now seems poised to become an endemic disease—one that is always a part of our environment, no matter what we do.

The World Health Organization declared COVID-19 a pandemic on March 11, 2020. A year later, the virus has infected 118 million people worldwide and killed over 2.6 million, including more than 530,000 Americans, according to data compiled by Johns Hopkins University.

At the same time, several effective COVID vaccines were developed at an unprecedented pace and have already been administered to nearly 330 million people worldwide, the vaccination and containment measures will eventually get the pandemic under control, potentially turning COVID into another disease we simply learn to live with.

The past month has seen a significantly drop in the new cases across the World. Now let's make a forecast on the "date when COVID is eventually eradicated"

Please find the dataset "covid19 cases.xlsx" on Blackboard→course material→homework→HW#2. This dataset tracks the daily new cases from 1/22/2020 to 3/11/2021 for each country in the World. Please select a country of interest and make forecast using appropriate filters.