Basic Considerations of Forecasting

Ka-fu WONG

University of Hong Kong

January 18, 2019

Contents

1. Basic considerations at the planning stage of any forecasting project

(1) The Decision Environment
(2) The Loss Function
(3) The Forecast Object
(4) The Forecast Horizon
(5) The Forecast Statement
(6) The Information Set
(7) The Forecast Method

Note:
(1) These considerations are *not independent* of one another.
(2) There may be *additional considerations* that we may need to address.

2. The Decision Environment

- Who will be using the forecast and for what purpose?

- (*Hypothetical*?) Examples
  - The Hong Kong Government has agreed to keep social security (under the Comprehensive Social Security Assistance Scheme, CSSA) increases in line with increases in the cost of consumption by our social security recipients as reflected by the Social Security Assistance Index of Prices (SSAIP). However, social security increases is better set before the increase in prices. Therefore, the Government may want to use the forecasted SSAIP inflation rate produced by a group of economists to set social security increases.
  - The Hong Kong Government may want to adjust its minimum wage periodically so that the minimum wage worker's living standard is maintained. In this case, we may want to adjust minimum wage according to the forecasted inflation rate.

- (*Hypothetical*?) Examples
  - The Chinese Government would like to keep its overall export competitiveness. Overall export competitiveness is related to the price of Chinese products to the US products, in real term. If overall export competitiveness is found declining, the government may need to implement some export promotion. Thus, the government is interested in forecasting the real exchange rate.
  - The Chinese Government may want to implement additional stimulus policies in order to maintain the economic growth close to its target if the economic growth is forecast to slow down.
  - The Bank of England is reviewing its monetary policy (whether to change its target interest rate). The Bank sets inflation targets. If inflation is going to be above certain level, it will raise interest. Otherwise, the Bank will keep the same interest rate. Thus, the Bank of England will be interested in forecasting the inflation rate.

3. The Loss Function

- Let D denote the decision that will depend partly or entirely on your forecast. That is,

$$D = D(f; x)$$

  where f is your forecast and x represents other stuff, possibly including other forecasts.

- For examples
  - For Hong Kong Government, D is the social security increase and f is the forecasted SSAIP inflation rate.
  - For Hong Kong Government, D is the new supply of public rental housing and f is the forecasted population growth.
  - For UK, D is the interest rate increase and f is the forecasted inflation rate.

- For Chinese Government, D is whether to increase export promotion expenditure and how much of the expenditure, and f is the forecasted real exchange rate.
- For People's Bank of China (PBOC), D is how much cash it should inject into the banking system to keep the short-term and medium term interbank lending rate in check, and f is the forecasted interbank lending rate.

The following report shows the important of forecast and the loss due to forecast.

> *While Australia's power demand has fallen 6 per cent since 2007, generating capacity [of CLP in Australia] has grown 13 per cent because plants take years to build and were based on outdated forecasts.*
>
> *Lower demand was caused by a rise in subsidised home solar panel installations and industrial plant closures linked to an appreciating Australian dollar.*
>
> —— South China Morning Post, February 28, 2014

- Let D^* denote the decision that would be optimal if there was *no forecast error*, i.e., if your forecast is perfectly accurate. D^* is the "*perfect foresight*" optimal decision.
- Then the cost of your forecast error will be the cost of making decision D instead of decision D^*.
- Naturally this cost can be derived from the consequence of a forecast error, say, the loss of profit. For example, the cost can be a loss of profit from investment based on the forecast.
- In certain decision-making settings this cost can be represented by a "loss function."

- Suppose that the decision depends on the forecast of a variable y. Let y^f denote your forecast of y. Suppose too, that the cost of making the decision $D(y^f)$ instead of the decision $D(y)[= D^*]$ depends only on the forecast error, $y - y^f$. That is,

$$L = L(e), \quad e = y - y^f$$

  where L is the cost or loss associated with the forecast error, e. $L(e)$ is called a loss function.

- The units of the loss function could be monetary units or utility units or ...

- Loss functions are particularly helpful is selecting an “optimal” forecast procedure.

Choose y^f to minimize $E(L)$

where $E(L)$ is the expected loss.

The Expectation (Mean) of Random Variables

- Expectation definition
  - Discrete random variables

$$E(X) = \sum_i P_i X_i$$

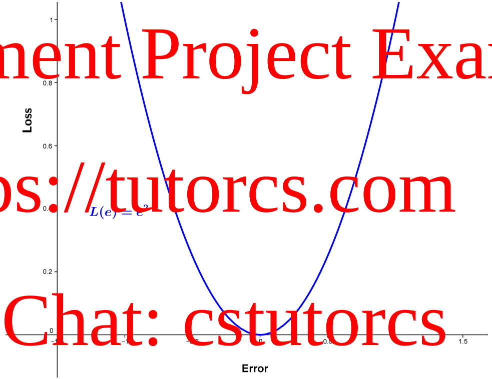
  - Continuous random variables

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

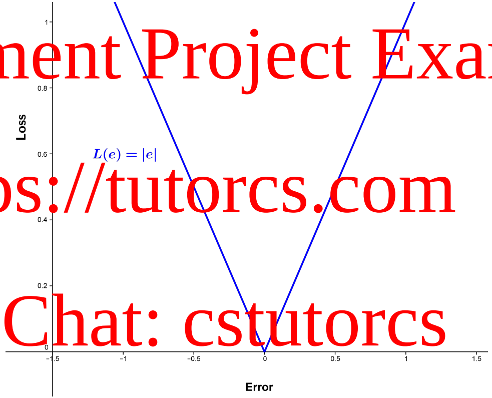
- Some useful results about expectation
  (1) $E(a) = a$
  (2) $E(bX) = bE(X)$
  (3) $E(a + bX) = a + bE(X)$
  (4) $E(aX + bY) = aE(X) + bE(Y)$

- The simplest and most commonly encountered loss functions are the quadratic and absolute loss functions:
  - Quadratic loss: $L = e^2$
  - Absolute loss: $L = \mid e \mid$
- These loss functions are *symmetric* loss functions, meaning that $L(e) = L(-e)$, i.e., the sign of the error doesn't matter, only its absolute value.
- The *difference* between the two is how the loss increases with the size of $\mid e \mid$:
  - the quadratic loss is increasing at an increasing rate with $\mid e \mid$.
  - the absolute loss is increasing proportionally with $\mid e \mid$.

The quadratic loss is increasing at an increasing rate with $| e |$.



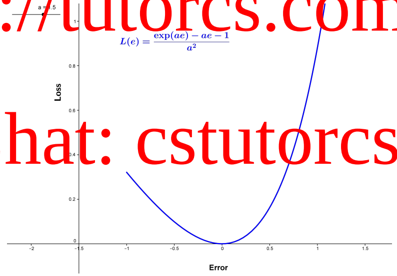
The absolute loss is increasing proportionally with $|\, e \,|$.



Loss needs not be symmetric in the errors. In some decision-making settings, other loss functions (e.g., asymmetric ones) would be more appropriate.

- For examples, the Chinese government concerns much more about the over-forecast of the export competitiveness than the under-forecast.

The Linex loss is asymmetric in e, i.e., $L(e) \neq L(-e)$



- Use quadratic loss if there is no strong reason for the alternative ones.
  - The choice of quadratic is consistent with the model selection criteria (usually done by least squares or its variation), as well as the approach of model estimation.
  - Model with quadratic loss is easier to estimate, computationally.

4. The Forecast Object

- Examples of forecast object
  - An event outcome (Timing of an event is known; the outcome of the event is unknown)
    - The winner of an election
    - The FOMC's fed funds rate target
  - An event timing (The outcome of the event is known; the timing of the event is unknown)
    - The beginning of the next recession
    - When the Hang Seng Index will next reach 40,000
  - A time series (We observe historical data on one or more economic variables and we want to forecast the future value(s) of these series.)
    - Real GDP during 2019.IV; the 2019 SSAIP

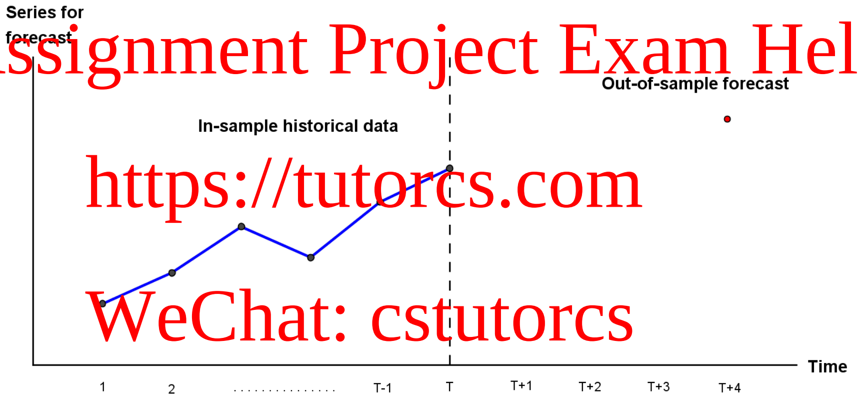
Time series are by far the most common forecast object. Forecasting time series will be our main concern.

5. The Forecast Horizon

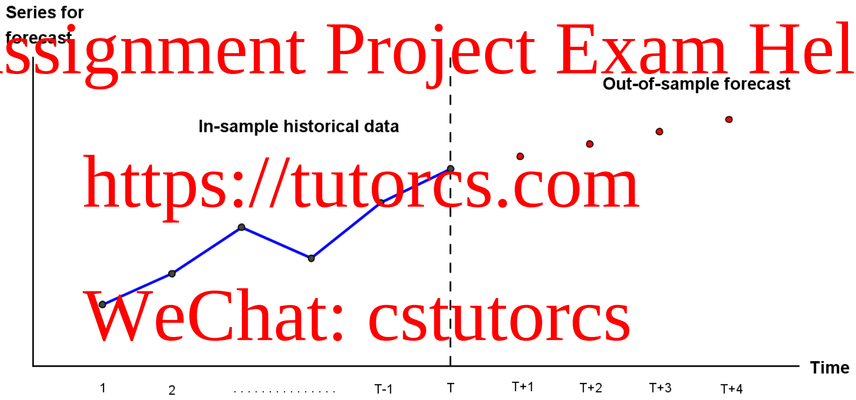
- Suppose that, standing at time T, we are interested in forecasting the value y_{T+h}, i.e., the value of y h periods into the future. We call this the h-step-ahead forecast of y and the length of the forecast horizon is h periods.

- Typical situations –
  - $h = 1$
  - $h = H$, for some $H > 1$
  - $h = 1, \ldots, H$ [H-step-ahead extrapolation forecasts]

Example: 4-Step-Ahead Point Forecast



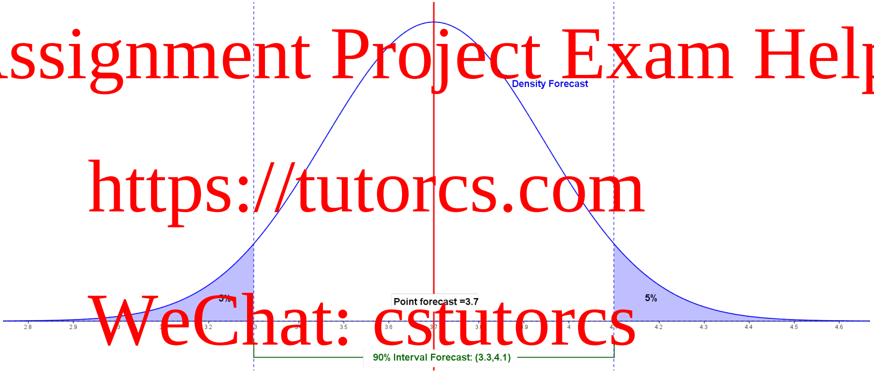
Example: 4-Step-Ahead Extrapolation Point Forecast



6. The Forecast Statement

- Suppose that our objective is to make a 1-step-ahead forecast of the time series y, i.e., we want to forecast y_{T+1} at time T. For example, we may want to forecast the 2019 SSAIP based on what we know in 2018. ($y = SSAIP$, $T = 2018$, $T + 1 = 2019$)
- The forecast can be stated in three forms:
  - a point forecast
  - a density forecast
  - an interval forecast

Example: Forecasting Hong Kong Real GDP Growth



Point Forecasts –

- The simplest forecast is a point forecast, which is a single number that represents our “best guess” of the forecast object: $y_{T,T+1}$
- For examples,
  - at the beginning of 2019 or the end of 2018, HKU reported the point forecast of the 2019 GDP growth;
  - at the beginning of 2019 or the end of 2018, the Government wanted a point forecast of the 2019 SSAIP.
- A point forecast is very concise and very simple, but
  - it *does not provide a sense of the uncertainty* that is bound to surround the forecast.
  - The decision-maker may want some sense of how precise your forecast is likely to be before making his/her decision.

Density Forecasts –

- A natural question that a decision-maker may raise is *how likely* is it that the actual value of the time series will turn out to be within a certain range of the point forecast?
  - E.g., what are the odds that the actual value will be more than 5% above/below the forecasted value? 10%? 25%?
- The way that we can approach this, which will be very natural when we use regression methods to generate our forecasts, is to
  - *think of the actual value of the time series as a random variable that is drawn from some probability distribution.*
- Providing an estimate of the probability distribution from which the actual value will be drawn is called a density forecast. *A density forecast is the most complete forecast statement.*

Example of Density Forecasts –

- Based upon the information available to us in 2018, suppose we think that the 2019 SSAIP will be drawn from an $N(\mu, \sigma^2)$ distribution, where the mean μ and variance σ^2 are unknown.

- Our forecasting task is to estimate these two parameters and, therefore, the precise distribution from which we think the 2019 SSAIP will be drawn.
  - Our estimate of μ is a reasonable choice for our point forecast of the 2019 SSAIP.

  - Why settled at point forecast? *With an estimate of μ and σ^2, we have an estimate of the entire distribution* from which the 2019 SSAIP will be drawn and are in a position to answer questions of the form:

    * What is the estimated probability that the actual value of the 2019 SSAIP will deviate from the point forecast by more that x-percent?

Density Forecasts –

- Note that a density forecast is more informative than a point forecast. However, there are a number of reasons why these are not necessarily preferred to simple point forecasts and *why*,
  - *In practice, point forecasts are much more common than density forecasts*.

  (1) Additional assumptions: Generating a sound density forecasts requires *more assumptions* about how the time series is determined. If these assumptions are wrong, not only will they lead to misleading density forecasts, but the implied point forecast may be worse than the point forecast that would have been generated under a weaker set of assumptions.

  (2) Computing power: Density forecasts may require substantially *more computational* effort than is necessary to make a sound point forecast.

Density Forecasts –

- In some decision-making environments, there is no benefit to having a density forecast.
  - For instance, the Government plans to set the social security increase equal to the point forecast of the SSAIP inflation rate. There decision does not depend on the degree of uncertainty about the actual SSAIP.
- Can you think of a decision-making environment in which a density forecast might be more valuable than a point forecast?
  - Inflation forecast!! Why?
  - A central bank's decision to raise interest rate based on inflation forecast affects the whole economy!!

Interval Forecasts –

- An typical interval forecast might have the following form
  - The 90% interval forecast for the 2019 SSAIP is 140 $\pm$ 5 or, equivalently, [135,145].
- The interpretation of this forecast interval is that
  - the point forecast of the SSAIP is 140 and
  - 90% of the time (under current conditions) this procedure provides an interval that will contain next period's SSAIP.
- We can, in principle, create an interval forecast using any x-percent as confidence, $0 < x < 100$.
- Note that the width of the interval forecast will increase with x.

Interval Forecasts –

- Interval forecasts provide more information than point forecasts but less information than the density forecasts, from which they are implicitly or explicitly derived.
- That is, they provide the decision-maker with some sense of the magnitude of the uncertainty surrounding the forecast (in contrast to a point forecast) but in a more concise way than density forecasts provide.
- They are subject to the same limitations that pertain to density forecasts (i.e., sensitivity to the assumptions required to construct these forecasts and potentially high computational costs).

7. The Information Set

- *What information will we use as input?*
- This decision will be made *jointly with the choice of the forecast method*.
- If we use simple time series methods (i.e., the unobserved components approach to modeling and forecasting time series), the information set will simply be the current and past values of the series we are interested in forecasting. That is, if we are trying to forecast y_{T+h} at time T, then our information set will be $y_T, y_{T-1}, \ldots, y_1$, where y_1 is the first observation of y that is available to us.

- If we use a multiple regression method in which we begin by modeling y_T as a function of not only its own past values, but also the current and past values of certain other variables, say, $x_1, x_2, \ldots, x_k$, then the information set will be:

$$y_T, y_{T-1}, \ldots, y_1$$
$$x_{1,T}, x_{1,T-1}, \ldots, x_{1,1}$$

$$x_{k,T}, x_{k,T-1}, \ldots, x_{k,1}$$

  - For example, in the Bank of England forecast of inflation rate, many other variables are available: interest rates of various horizon, GDP growth, foreign trade, etc.
- As opposed to simple time series approach, we *need to decide what additional variable to include in the information set*. Among the considerations that will be at work in making this decision:
  - data availability
  - data relevance
  - econometric efficiency

8. Forecast Method

- The selection of a forecast method, i.e., the statistical or econometric model that will generate the forecasts, will be made
  - partly in the context of the other considerations that we have discussed,
  - partly based on the experiences of others who have been involved in applied forecasting exercises, and
  - partly based on the theory underlying the various approaches that are available.
- In others words –
  - What are you trying to accomplish?
  - What has worked well for others who have engaged in similar projects?
  - What do statistical and econometric theory suggest ought to work well in your forecast environment?

- Although the economy is very complex, experience suggests that *unless you need to simultaneously generate a large number of interrelated time series, relatively simple methods tend to produce better forecasts than more complicated methods*.
- This is not to say that we think we do a very good job at forecasting economic variables or that more complicated methods will not end up being developed that work substantially better than the simple methods that are currently widely used.
- Instead, it is saying that given the current art and science of economic forecasting, the parsimony and *KISS* principles lead us to prefer to work with relatively simple models.

KISS: Keep It Straight and Simple

9. An important assumption in preparing forecast

- In the past centuries, the sun rises from the East. Tomorrow, the sun will rise from the East.
- Out of 10 black Fridays, 9 of them had stock market crash. Tomorrow is a black Friday, the stock market will crash with 90% of chance.