Chapter 2

Universal Considerations

In Chapter 1 we sketched a variety of areas where forecasts are used routinely. Here we begin by highlighting, in no particular order, a number of considerations relevant for *any* forecasting task. We introduce the those considerations as *questions*.

1. (**Forecast Object**) What is the object that we want to forecast? Is it a time series, such as sales of a firm recorded over time, or an event, such as devaluation of a currency, or something else? Appropriate forecasting strategies depend on the nature of the object being forecast.

2. (**Information Set**) On what information will the forecast be based? In a time series environment, for example, are we forecasting one series, several, or thousands? And what is the quantity and quality of the data? Appropriate forecasting strategies depend on the information set, broadly interpreted to not only quantitative data but also expert opinion, judgment, and accumulated wisdom.

3. (**Model Uncertainty and Improvement**) Does our forecasting model match the true DGP? Of course not. One must never, ever, be so foolish as to be lulled into such a naive belief. All models are false: they are *intentional* abstractions of a much more complex reality. A model might be useful for certain purposes and poor for others. Models that once

worked well may stop working well. One must continually diagnose and assess both empirical performance and consistency with theory. The key is to work continuously toward model improvement.

4. (**Forecast Horizon**) What is the forecast horizon of interest, and what determines it? Are we interested, for example, in forecasting one month ahead, one year ahead, or ten years ahead (called **h-step-ahead forecasts**, in this case for $h = 1$, $h = 12$ and $h = 120$ months)? Appropriate forecasting strategies likely vary with the horizon.

5. (**Structural Change**)

   Are the approximations to reality that we use for forecasting (i.e., our models) stable over time? Generally not. Things can change for a variety of reasons, gradually or abruptly, with obviously important implications for forecasting. Hence we need methods of detecting and adapting to structural change.

6. (**Forecast Statement**) How will our forecasts be stated? If, for example, the object to be forecast is a time series, are we interested in a single "best guess" forecast, a "reasonable range" of possible future values that reflects the underlying uncertainty associated with the forecasting problem, or a full probability distribution of possible future values? What are the associated costs and benefits?

7. (**Forecast Presentation**)

   How best to *present* forecasts? Except in the simplest cases, like a single h-step-ahead point forecast, graphical methods are valuable, not only for forecast presentation but also for forecast construction and evaluation.

8. (**Decision Environment and Loss Function**) What is the **decision environment** in which the forecast will be used? In particular, what decision will the forecast guide? How do we quantify what we mean

by a "good" forecast, and in particular, the cost or loss associated with forecast errors of various signs and sizes?

9. (**Model Complexity and the Parsimony Principle**) What sorts of models, in terms of complexity, tend to do best for forecasting in business, finance, economics, and government? The phenomena that we model and forecast are often tremendously complex, but it does not necessarily follow that our forecasting models should be complex. Bigger forecasting models are not necessarily better, and indeed, all else equal, smaller models are generally preferable (the "parsimony principle").

10. (**Unobserved Components**) In the leading time case of time series, have we successfully modeled trend? Seasonality? Cycles? Some series have all such components, and some not. They are driven by very different factors, and each should be given serious attention.

2.1 The Forecast Object

There are many objects that we might want to forecast. In business and economics, the forecast object is typically one of three types: **event outcome**, **event timing**, or **time series**.

Event outcome forecasts are relevant to situations in which an event is certain to take place at a given time but the outcome is uncertain. For example, many people are interested in whether the current chairman of the Board of Governors of the U.S. Federal Reserve System will eventually be reappointed. The "event" is the reappointment decision; the decision will occur at the end of the term. The outcome of this decision is confirmation or denial of the reappointment.

Event timing forecasts are relevant when an event is certain to take place and the outcome is known, but the timing is uncertain. A classic example of an event timing forecast concerns business cycle turning points. There are two types of turning points: peaks and troughs. A peak occurs when the economy moves from expansion into recession, and a trough occurs when the economy moves from recession into expansion. If, for example, the economy is currently in an expansion, then there is no doubt that the next turning point will be a peak, but there is substantial uncertainty as to its *timing*. Will the peak occur this quarter, this year, or ten years from now?

Time series forecasts are relevant when the future value of a time series is of interest and must be projected. As we'll see, there are many ways to make such forecasts, but the basic forecasting setup doesn't change much. Based upon the history of the time series (and possibly a variety of other types of information as well, such as the histories of related time series, or subjective considerations), we want to project future values of the series. For example, we may have data on the number of Apple computers sold in Germany in each of the last 60 months, and we may want to use that data to forecast the number of Apple computers to be sold in Germany in each month of the next year.

There are at least two reasons why time series forecasts are by far the most frequently encountered in practice. First, most business, economic and financial data are time series; thus, the general scenario of projecting the future of a series for which we have historical data arises constantly. Second, the technology for making and evaluating time-series forecasts is well-developed and the typical time series forecasting scenario is precise, so time series forecasts can be made and evaluated routinely. In contrast, the situations associated with event outcome and event timing forecasts arise less frequently and are often less amenable to quantitative treatment.

2.2 The Information Set

The quality of our forecasts is limited by the quality and quantity of information available when forecasts are made. Any forecast we produce is conditional upon the information used to produce it, whether explicitly or implicitly.

2.2.1 Univariate vs. Multivariate

The idea of an information set is fundamental to constructing good forecasts. In forecasting a series, y, using historical data from time 1 to time T, sometimes we use the **univariate information set**, which is the set of historical values of y up to and including the present,

$$\Omega_T = \{y_T,\ y_{T-1},\ ...,\ y_1\}.$$

In a univariate environment, then, a single variable is modeled and forecast solely on the basis of its own past. Univariate approaches to forecasting may seem simplistic, and in some situations they are, but they are tremendously important and worth studying for at least two reasons. First, although they are simple, they are not necessarily simplistic, and a large amount of accumulated experience suggests that they often perform admirably. Second, it's necessary to understand univariate forecasting models before tackling more complicated multivariate models.

Alternatively, sometimes we use the **multivariate information set**

$$\Omega_T = \{y_T,\ x_T,\ y_{T-1},\ x_{T-1},\ ...,\ y_1,\ x_1\},$$

where the x's are a set of additional variables potentially related to y. In a multivariate environment, a variable (or each member of a set of variables) is modeled on the basis of its own past, as well as the past of other variables, thereby accounting for and exploiting cross-variable interactions. Multivariate models have the *potential* to produce forecast improvements relative to univariate models, because they exploit more information to produce forecasts.

2.2.2 Expert Opinion and Judgment

Regardless of whether the information set is univariate or multivariate, it's always important to think hard about what information is available, what additional information could be collected or made available, the form of the information (e.g., quantitative or qualitative), and so on. A holistic view of an information involves far more than just the past history of one or a few quantitative variables; instead, it involves theoretical perspectives, expert judgment, contextual knowledge, and so on.

So you should take a broad view of what's meant by a "model." Try to incorporate views of experts and even non-experts. (Sometimes the alleged experts are not so expert, and the alleged non-experts are quite insightful.) Surveys, Bayesian priors and shrinkage, forecast combination, and prediction markets, which we'll discuss in due course, all attempt to do that.

2.2.3 Information Sets in Forecast Evaluation

The idea of an information set is also fundamental for evaluating forecasts: the basic principle of forecast evaluation is that a "good" forecast has corresponding errors that are unforecastable using using information available when the forecast is made. When evaluating a forecast, we're sometimes interested in whether the forecast could be improved by using a given set of information more efficiently, and we're sometimes interested in whether the forecast could be improved by using more information. Either way, the ideas of information and information sets play crucial roles in forecast evaluation.

2.3 Model Uncertainty and Improvement

One must never, ever, be so foolish as to be lulled into believing that one's model coincides with the true DGP. Indeed all models are false. Does that mean that models and modeling are somehow discredited, or worthless? Not at all, and their use continues to expand. The uninitiated are sometimes suspicious, because their lack of understanding of models and modeling leads them to have unreasonable expectations of models, but there's no other way forward. As George Box (1979) famously and correctly noted, "All models are false, but some are useful."

Related, a model might be useful for certain purposes and poor for others. Models that once worked well may stop working well. One must continually diagnose and assess both empirical performance and consistency with theory. That is, the key is to think continually about how to *improve* models. And always remember: *It takes a model to beat a model.*

2.4 The Forecast Horizon

2.4.1 h-Step-Ahead Forecasts

The forecast horizon is defined as the number of periods between today and the date of the forecast we make. For example, if we have annual data, and it's now year T, then a forecast of GDP for year $T+2$ has a forecast horizon of 2 steps. The meaning of a step depends on the frequency of observation of the data. For monthly data a step is one month, for quarterly data a step is one quarter (three months), and so forth. In general, we speak of an **h-step ahead forecast**, where the horizon h is at the discretion of the user.[1]

The horizon is important for at least two reasons. First, of course, the forecast changes with the forecast horizon. Second, the best forecasting model

[1]The choice of h depends on the decision that the forecast will guide. The nature of the decision environment typically dictates whether "short-term", "medium-term", or "long-term" forecasts are needed.

will often change with the forecasting horizon as well. All of our forecasting models are approximations to the underlying dynamic patterns in the series we forecast; there's no reason why the best approximation for one purpose (e.g., short-term forecasting) should be the same as the best approximation for another purpose (e.g., long-term forecasting).

2.4.2 $h - Step$ Ahead Path Forecasts

Let's distinguish between what we've called h-step-ahead forecasts and what we'll call ***h*-step-ahead path forecasts**, sometimes also called ***h*-step-ahead extrapolation forecasts**. In h-step-ahead forecasts, the horizon is always fixed at the same value, h. For example, every month we might make a 4-month-ahead forecast. Alternatively, in path forecasts, the horizon includes all steps from 1-step-ahead to h-steps-ahead. There's nothing particularly deep or difficult about the distinction, but it's useful to make it, and we'll use it subsequently.

Suppose, for example, that you observe a series from some initial time 1 to some final time T, and you plan to forecast the series.[2] We illustrate the difference between h-step-ahead and h-step-ahead path forecasts in Figures 2.1 and 2.2. In Figure 2.1 we show a 4-step-ahead point forecast, and in Figure 2.2 we show a 4-step-ahead path point forecast. The path forecast is nothing more than a set consisting of 1-, 2-, 3-, and 4-step-ahead forecasts.

2.4.3 Nowcasting and Backcasting

Quite apart from making informative and useful guesses about the future (forecasting), often we're interested in the present ("nowcasting") – which is also subject to lots of uncertainty – or even the past ("backcasting"). Many

[2] For a sample of data on a series y, we'll typically write $\{y_t\}_{t=1}^T$. This notation means "we observe the series y from some beginning time $t = 1$ to some ending time $t = T$".

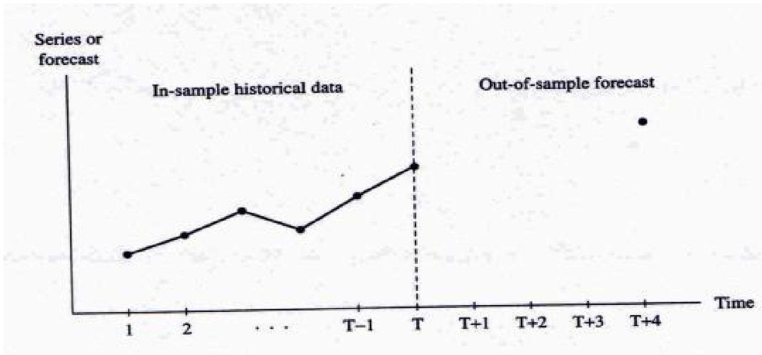

Figure 2.1: 4-Step-Ahead Point Forecast

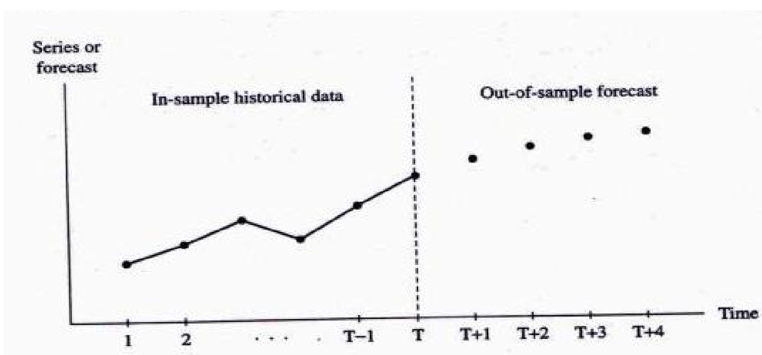

Figure 2.2: 4-Step-Ahead Path Forecast

of our models and methods will be relevant there as well.[3]

2.5 Structural Change

In time series, we rely on the future being like the present/past in terms of dynamic relationships (e.g., the next twenty years vs. the last twenty years). But that's not always true. Structural change can be gradual or abrupt.

In cross sections, we rely on fitted relationships being relevant for new cases from the original population, and often even for new populations. But again, that's not always true. For example, the effect of class size on test scores may differ for 10-year olds in California vs. 6-year olds in Maine.

Structural change can affect any or all parameters of a model, and the breaks can be large or small.

Structural change is a type of non-linearity; indeed abrupt structural change is often handled with dummy variable models, and gradual structural change is often handled with smoothly-time-varying parameter models.

2.6 The Forecast Statement

2.6.1 Time Series

When we make a forecast, we must decide if the forecast will be (1) a single number (a "best guess"), (2) a range of numbers, into which the future value can be expected to fall a certain percentage of the time, or (3) an entire probability distribution for the future value. In short, we need to decide upon the forecast type.

More precisely, we must decide if the forecast will be (1) a **point forecast**, (2) an **interval forecast**, or (3) a **density forecast**. A point forecast is a single number. For example, one possible point forecast of the growth rate of

[3]For an example of nowcasting see the ADS Index at FRB Philadelphia, and for an example of backcasting see GDPplus, also at FRB Philadelphia.

the total number of web pages over the next year might be +23.3%; likewise, a point forecast of the growth rate of U.S. real GDP over the next year might be +1.3%. Point forecasts are made routinely in numerous applications, and the methods used to construct them vary in difficulty from simple to sophisticated. The defining characteristic of a point forecast is simply that it is a single number.

A good point forecast provides a simple and easily-digested guide to the future of a time series. However, random and unpredictable "shocks" affect all of the series that we forecast. As a result of such shocks, we expect nonzero forecast errors, even from very good forecasts. Thus, we may want to know the degree of confidence we have in a particular point forecast. Stated differently, we may want to know how much uncertainty is associated with a particular point forecast. The uncertainty surrounding point forecasts suggests the usefulness of an interval forecast.

An interval forecast is not a single number; rather, it is a range of values in which we expect the realized value of the series to fall with some (prespecified) probability.[4] Continuing with our examples, a 90% interval forecast for the growth rate of web pages might be the interval [11.3%, 35.3%] (23.3% ± 12%). That is, the forecast states that with probability 90% the future growth rate of web pages will be in the interval [11.3%, 35.3%]. Similarly, a 90% interval forecast for the growth rate of U.S. real GDP might be [-2.3%, 4.3%] (1.3% ± 3%); that is, the forecast states that with probability 90% the future growth rate of U.S. real GDP will be in the interval [-2.3%, 4.3%].

A number of remarks are in order regarding interval forecasts. First, the length (size) of the intervals conveys information regarding forecast uncertainty. The GDP growth rate interval is much shorter then the web page

[4] An interval forecast is very similar to the more general idea of a *confidence interval* that you studied in statistics. An interval forecast is simply a confidence interval for the true (but unknown) future value of a series, computed using a sample of historical data. We'll say that $[a, b]$ is a $100(1-\alpha)\%$ interval forecast if the probability of the future value being less than a is $\alpha/2$ and the probability of the future value being greater than b is also $\alpha/2$.

growth rate interval; this reflects the fact that there is less uncertainty associated with the real GDP growth rate forecast than the web page growth rate forecast. Second, interval forecasts convey more information than point forecasts: given an interval forecast, you can construct a point forecast by using the midpoint of the interval.[5] Conversely, given only a point forecast, there is no way to infer an interval forecast.

Finally, we consider density forecasts. A density forecast gives the entire density (or probability distribution) of the future value of the series of interest. For example, the density forecast of future web page growth might be normally distributed with a mean of 23.3% and a standard deviation of 7.32%. Likewise, the density forecast of future real GDP growth might be normally distributed with a mean of 1.3% and a standard deviation of 1.83%. As with interval forecasts, density forecasts convey more information than point forecasts. Density forecasts also convey more information than interval forecasts, because given a density, interval forecasts at any desired confidence level are readily constructed. For example, if the future value of a series x is distributed as $N(\mu, \sigma^2)$, then a 95% interval forecast of x is $\mu \pm 1.96\sigma$, a 90% interval forecast of x is $\mu \pm 1.64\sigma$, and so forth. Continuing with our example, the relationships between density, interval, and point forecasts are made clear in Figures 2.3 (web page growth) and 2.4 (GDP growth).

To recap, there are three time series forecast types: point, interval, and density. Density forecasts convey more information than interval forecasts, which in turn convey more information than point forecasts. This may seem to suggest that density forecasts are always the preferred forecast, that density forecasts are the most commonly used forecasts in practice, and that we should focus most of our attention in this book on density forecasts.

In fact, the opposite is true. Point forecasts are the most commonly used

[5] An interval forecast doesn't *have* to be symmetric around the point forecast, so that we wouldn't necessarily infer a point forecast as the midpoint of the interval forecast, but in many cases such a procedure is appropriate.

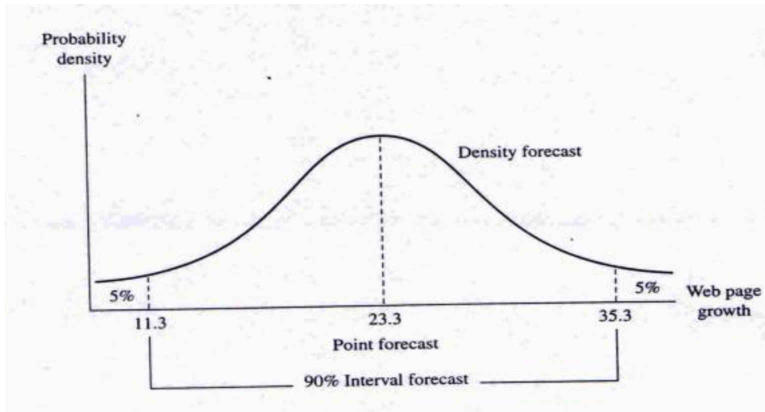

Figure 2.3: Web Page Growth Forecasts

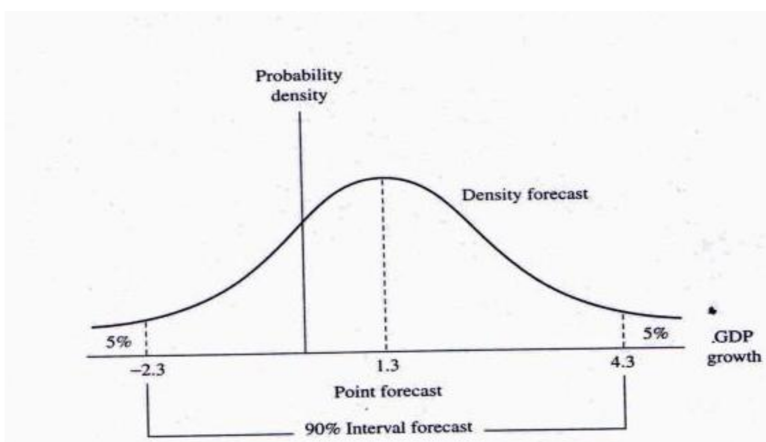

Figure 2.4: GDP Growth Forecasts

forecasts in practice, interval forecasts are a rather distant second, and density forecasts are rarely made. There are at least two reasons. First, the construction of interval and density forecasts requires either (a) additional and possibly incorrect assumptions relative to those required for construction of point forecasts, or (b) advanced and computer-intensive methods involving, for example, extensive simulation. Second, point forecasts are often easier to understand and act upon than interval or density forecasts. That is, the extra information provided by interval and density forecasts is not necessarily an advantage when information processing is costly.

2.6.2 Events

Thus far we have focused exclusively on types of time series forecasts, because time series are so prevalent and important in numerous fields. It is worth mentioning another forecast type of particular relevance to event outcome and event timing forecasting, the **probability forecast**. To understand the idea of a probability forecast, consider forecasting which of two politicians, X or Y, will win an election. (This is an event-outcome forecasting situation.) If our calculations tell us that the odds favor X, we might issue the forecast simply as "X will win." This is roughly analogous to the time series point forecasts discussed earlier, in the sense that we're not reporting any measure of the uncertainty associated with out forecast. Alternatively, we could report the probabilities associated with each of the possible outcomes; for example, "X will win with probability .6, and Y will win with probability .4." This is roughly analogous to the time series interval or density forecasts discussed earlier, in the sense that it explicitly quantifies the uncertainty associated with the future event with a probability distribution.

Event outcome and timing forecasts, although not as common as time series forecasts, do nevertheless arise in certain important situations and are often stated as probabilities. For example, when a bank assesses the probability of default on a new loan or a macroeconomist assesses the probability that a business cycle turning point will occur in the next six months, the banker or macroeconomist will often use a probability forecast.

2.6.3 Probability Forecasts as Point and/or Density Forecasts

2.7 Forecast Presentation

2.7.1 Graphics for Forecasts

2.7.2 Graphics for Forecast Evaluation

2.8 The Decision Environment and Loss Function

Forecasts are not made in a vacuum. The key to generating good and useful forecasts, which we will stress now and throughout, is recognizing that forecasts are made to guide decisions. The link between forecasts and decisions sounds obvious – and it is – but it's worth thinking about in some depth. Forecasts are made in a wide variety of situations, but in every case forecasts are of value because they aid in decision making. Quite simply, good forecasts help to produce good decisions. Recognition and awareness of the decision making environment is the key to effective design, use and evaluation of forecasting models.

2.8.1 Loss Functions

Let y denote a series and $\hat{y}$ its forecast. The corresponding **forecast error**, e, is the difference between the realization and the previously-made forecast,

$$e = y - \hat{y}.$$

We consider loss functions of the form $L(e)$. This means that the loss associated with a forecast depends only on the size of the forecast error. We might require the loss function $L(e)$ to satisfy three conditions:

1. $L(0) = 0$. That is, no loss is incurred when the forecast error is zero. (A zero forecast error, after all, corresponds to a perfect forecast!)

2. $L(e)$ is continuous. That is, nearly-identical forecast errors should produce nearly-identical losses.

3. $L(e)$ is increasing on each side of the origin. That is, the bigger the absolute value of the error, the bigger the loss.

Apart from these three requirements, we impose no restrictions on the form of the loss function.

The **quadratic loss** function is tremendously important in practice. First, it's often an arguably-reasonable approximation to realistic loss structures. Second, it's mathematically convenient: It is usually easy to compute, because quadratic objectives have linear first-order conditions.[6]

Quadratic loss is given by

$$L(e) = e^2 ,$$

and we graph it as a function of the forecast error in Figure 2.5. Because of the squaring associated with the quadratic loss function, it is symmetric around the origin, and in addition, it increases at an increasing rate on each side of the origin, so that large errors are penalized much more severely than small ones.

Another important symmetric loss function is **absolute loss**, or **absolute error loss**, given by

$$L(e) = |e|.$$

Like quadratic loss, absolute loss is increasing on each side of the origin,

[6] In contrast, optimal forecasting under **asymmetric loss** is rather involved, and the tools for doing so are still under development. See, for example, Christoffersen and Diebold, 1997.

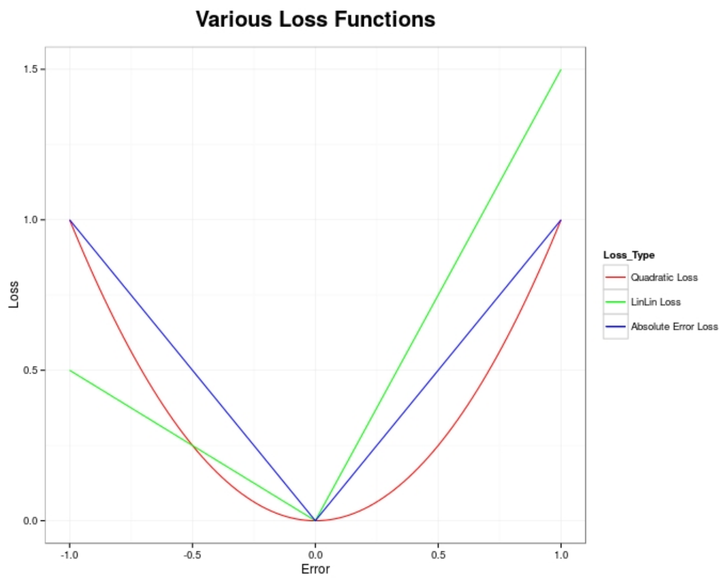

Figure 2.5: Quadratic, Absolute, and Linlin Loss Functions

but loss increases at a constant (linear) rate with the size of the error. We illustrate absolute loss in Figure 2.5.

In certain contexts, symmetric loss functions may not be an adequate distillation of the forecast / decision environment, as would be the case, for example, if negative forecast errors were for some reason generally less costly than positive errors. An important asymmetric loss function is **"linlin loss"** (linear on each side of the origin, with generally different slopes), given by

$$L(e) = \begin{cases} a|e|, \text{ if } e > 0 \\ \\ b|e|, \text{ if } e \leq 0. \end{cases}$$

We show asymmetric linlin loss in Figure 2.5.

2.8.2 Optimal Forecasts with Respect to a Loss Function

Much of this book is about how to produce **optimal forecasts**. What precisely do we mean by an optimal forecast? That's where the loss function comes in: the optimal forecast is the forecast with smallest conditionally expected loss.

$$\hat{y}(x)^* = argmin_{\hat{y}(x)} \int \int L(y - \hat{y}(x))\, f(y,x)\; dy\, dx.$$

Here are some key results:

- Under quadratic loss, the optimal forecast is the conditional mean. That is,

  $$\hat{y}(x)^* = E(y|x).$$

  Note that x could be lagged y.

- Under absolute loss, the optimal forecast is the conditional median,

  $$\hat{y}(x)^* = Q_{.50 \cdot 100\%}(y|x),$$

  where $Q_{d \cdot 100\%}(\cdot)$ denotes the d-percent conditional quantile function.

- Under lin-lin loss, the optimal forecast is the conditional $d \cdot 100\%$ quantile, where

  $$d = \frac{b}{a+b} = \frac{1}{1+a/b}.$$

  That is,

  $$\hat{y}(x)^* = Q_{d \cdot 100\%}(y|x).$$

Quite generally under asymmetric $L(e)$ loss (e.g., linlin), optimal forecasts are biased, whereas the conditional mean forecast is unbiased.[7] Bias is optimal under asymmetric loss because we can gain on average by pushing the forecasts in the direction such that we make relatively few errors of the more costly sign.

[7] A forecast is unbiased if its error has zero mean.

2.8.3 State-Dependent Loss

In some situations, the $L(e)$ form of the loss function is too restrictive. Although loss will always be of the form $L(y,\ \hat{y})$, there's no reason why y and $\hat{y}$ should *necessarily* enter as $y - \hat{y}$. In predicting financial asset returns, for example, interest sometimes focuses on direction of change. A **direction-of-change forecast** takes one of two values – up or down. The loss function associated with a direction of change forecast might be:[8]

$$L(y,\ \hat{y}) = \begin{cases} 0, \text{ if } sign(\Delta y) = sign(\Delta \hat{y}) \\ 1, \text{ if } sign(\Delta y) \neq sign(\Delta \hat{y}). \end{cases}$$

With this loss function, if you predict the direction of change correctly, you incur no loss; but if your prediction is wrong, you're penalized.

This is one example of a **state-dependent loss function**, meaning that loss actually depends on the state of the world (y), as opposed to just depending on e. This may sometimes make sense; the cost of a given error may be higher or lower, for example, in different states of the world as indexed by y.

Under direction-of-change loss, the optimal forecast is the conditional mode. That is,

$$\hat{y}(x)^* = Mode(y|x).$$

2.9 Model Complexity and the Parsimony Principle

It's crucial to tailor forecasting tools to forecasting tasks, and doing so is partly a matter of judgment. Typically the specifics of the situation (e.g., decision environment, forecast object, forecast statement, forecast horizon, information set, etc.) will indicate the desirability of a specific method or

[8]The operator "Δ" means "change." Thus Δy_t is the change in y from period $t-1$ to period t, or $y_t - y_{t-1}$.

modeling strategy. Moreover, as we'll see, formal statistical criteria exist to guide model selection within certain classes of models.

We've stressed that a variety of forecasting applications use a small set of common tools and models. You might guess that those models are tremendously complex, because of the obvious complexity of the real-world phenomena that we seek to forecast. Fortunately, such is not the case. In fact, decades of professional experience suggest just the opposite – simple, parsimonious models tend to be best for out-of-sample forecasting in business, finance, and economics. Hence, the **parsimony principle**: other things the same, simple models are usually preferable to complex models.

There are a number of reasons why smaller, simpler models are often more attractive than larger, more complicated ones. First, by virtue of their parsimony, we can estimate the parameters of simpler models more precisely. Second, because simpler models are more easily interpreted, understood and scrutinized, anomalous behavior is more easily spotted. Third, it's easier to communicate an intuitive feel for the behavior of simple models, which makes them more useful in the decision-making process. Finally, enforcing simplicity lessens the scope for "data mining" – tailoring a model to maximize its fit to historical data. Data mining often results in models that fit historical data beautifully (by construction) but perform miserably in out-of-sample forecasting, because it tailors models in part to the *idiosyncrasies* of historical data, which have no relationship to unrealized future data.

Finally, note that simple models should not be confused with naive models. All of this is well-formalized in the **KISS principle** (appropriately modified for forecasting): "Keep it Sophisticatedly Simple." We'll attempt to do so throughout.

2.10 Unobserved Components

Trend, seasonal, cycle, noise. Deterministic vs. stochastic trend and seasonality.

$$y_t = T_t + S_t + C_t + \varepsilon_t.$$

Or maybe

$$y_t = T_t \times S_t \times C_t \times \varepsilon_t,$$

but of course that's just

$$lny_t = lnT_t + lnS_t + lnC_t + ln\varepsilon_t.$$

2.11 Concluding Remarks

This chapter obviously deals with broad issues of general relevance. For the most part, it avoids detailed discussion of specific modeling or forecasting techniques. The rest of the book drills down more deeply.

2.12 Exercises, Problems and Complements

1. Properties of loss functions.

   State whether the following potential loss functions meet the criteria introduced in the text, and if so, whether they are symmetric or asymmetric:

   a. $L(e) = e^2 + e$

   b. $L(e) = e^4 + 2e^2$

   c. $L(e) = 3e^2 + 1$

   d. $L(e) = \begin{cases} \sqrt{e} \text{ if } e > 0 \\ \\ |e| \text{ if } e \leq 0. \end{cases}.$

2. Relationships among point, interval and density forecasts.

   For each of the following density forecasts, how might you infer "good" point and ninety percent interval forecasts? Conversely, if you started with your point and interval forecasts, could you infer "good" density forecasts?

   Be sure to defend your definition of "good."

   a. Future y is distributed as $N(10, 2)$.

   b. $P(y) = \begin{cases} \frac{y-5}{25} \text{ if } 5 < y < 10 \\ \\ -\frac{y-15}{25} \text{ if } 10 < y < 15 \\ \\ 0 \text{ otherwise.} \end{cases}$

3. Forecasting at short through long horizons.

Consider the claim, "The distant future is harder to forecast than the near future." Is it sometimes true? Usually true? Always true? Why or why not? Discuss in detail. Be sure to define "harder."

4. "Real" forecasts vs. "goal" or "advocacy" forecasts.

   Many things that seem like forecasts are not at all real forecasts. Every politician forecasts that she will win the election. Should you take such forecasts seriously? Every lawyer forecasts that his client will win. Should you take such forecasts seriously? Simultaneously, hidden away from the public, serious, scientifically disinterested forecasts are routinely made and used successfully in numerous endeavors. The problem is that the public routinely sees the former (e.g., from television pundits) and rarely sees the latter.

5. Univariate and multivariate information sets.

   a. Which of the following modeling situations involve univariate information sets? Multivariate?

      i. Using a stock's price history to forecast its price over the next week;

      ii. Using a stock's price history and volatility history to forecast its price over the next week;

      iii. Using a stock's price history and volatility history to forecast its price and volatility over the next week.

   b. Keeping in mind the distinction between univariate and multivariate information sets, consider a wine merchant seeking to forecast the price per case at which a fine vintage of Chateau Latour, one of the greatest Bordeaux wines, will sell when it is thirty years old, at which time it will be fully mature.

i. What sorts of univariate forecasting approaches can you imagine that might be relevant?

ii. What sorts of multivariate forecasting approaches can you imagine that might be relevant? What other variables might be used to predict the Latour price?

iii. What are the comparative costs and benefits of the univariate and multivariate approaches to forecasting the Latour price?

iv. Would you adopt a univariate or multivariate approach to forecasting the Latour price? Why?

6. Assessing forecasting situations.

   For each of the following scenarios, discuss the decision environment, the nature of the object to be forecast, the forecast type, the forecast horizon, the loss function, the information set, and what sorts of simple or complex forecasting approaches you might entertain.

   a. You work for Airborne Analytics, a highly specialized mutual fund investing exclusively in airline stocks. The stocks held by the fund are chosen based on your recommendations. You learn that a newly rich oil-producing country has requested bids on a huge contract to deliver thirty state-of-the-art fighter planes, but that only two companies submitted bids. The stock of the successful bidder is likely to rise.

   b. You work for the Office of Management and Budget in Washington DC and must forecast tax revenues for the upcoming fiscal year. You work for a president who wants to maintain funding for his pilot social programs, and high revenue forecasts ensure that the programs keep their funding. However, if the forecast is too high, and the president runs a large deficit at the end of the year, he will be seen as fiscally irresponsible, which will lessen his probability of reelection.

Furthermore, your forecast will be scrutinized by the more conservative members of Congress; if they find fault with your procedures, they might have fiscal grounds to undermine the President's planned budget.

c. You work for D&D, a major Los Angeles advertising firm, and you must create an ad for a client's product. The ad must be targeted toward teenagers, because they constitute the primary market for the product. You must (somehow) find out what kids currently think is "cool," incorporate that information into your ad, and make your client's product attractive to the new generation. If your hunch is right, your firm basks in glory, and you can expect multiple future clients from this one advertisement. If you miss, however, and the kids don't respond to the ad, then your client's sales fall and the client may reduce or even close its account with you.

7. Box vs. Wiener on Models and Modeling.

   We earlier mentioned George Box's memorable view that "All models are false, but some are useful." Norbert Wiener, an equally important applied mathematician on whose work much of this book builds, had a different and also-memorable view, asserting that "The best material model of a cat is another, or preferably the same, cat."[9] What did Wiener mean? What is your view?

8. Forecasting as an ongoing process in organizations.

   We could add another very important item to this chapter's list of considerations basic to successful forecasting – forecasting in organizations is an ongoing process of building, using, evaluating, and improving forecasting models. Provide a concrete example of a forecasting model used in business, finance, economics or government, and discuss ways in which

[9]Attributed by Wikiquote to Wiener and Rosenblueth's *Philosophy of Science*, 1945.

each of the following questions might be resolved prior to, during, or after its construction.

a. Are the data "dirty"? For example, are there "**ragged edges**" (different starting and ending dates of different series)? Are there **missing observations**? Are there aberrant observations, called **outliers**, perhaps due to **measurement error**?

   Are the data stored in a format that inhibits computerized analysis?

b. Has software been written for importing the data in an ongoing forecasting operation?

c. Who will build and maintain the model?

d. Are sufficient resources available (time, money, staff) to facilitate model building, use, evaluation, and improvement on a routine and ongoing basis?

e. How much time remains before the first forecast must be produced?

f. How many series must be forecast, and how often must ongoing forecasts be produced?

g. What level of data **aggregation** or **disaggregation** is desirable?

h. To whom does the forecaster or forecasting group report, and how will the forecasts be communicated?

i. How might you conduct a "forecasting audit"?