Elements of Forecasting

in Business, Finance, Economics and Government

Francis X. Diebold
Department of Economics
University of Pennsylvania

"Solutions Manual"

Preface

This is quite a nonstandard “Solutions Manual,” but I use the term for lack of something more descriptively accurate. Many of the Problems and Complements don't ask questions, so they certainly don't have solutions; instead, they simply introduce concepts and ideas that, for one reason or another, didn't make it into the main text. Moreover, even for those Problems and Complements that *do* ask questions, the vast majority don't have explicit or unique solutions. Hence the “solutions manual” offers remarks, suggestions, hints, and occasionally, solutions. Most of the Problems and Complements are followed by brief remarks marked with asterisks, and in the (relatively rare) cases where there was nothing to say, I said nothing.

F.X.D.

Solutions

Chapter 1 Problems and Complements

1. (Forecasting in daily life: we are all forecasting, all the time)

   a. Sketch in detail three forecasts that you make routinely, and probably informally, in your daily life. What makes you believe that the forecast object is predictable? What factors might introduce error into your forecasts?

   b. What decisions are aided by your three forecasts? How might the degree of predictability of the forecast object affect your decisions?

   c. How might you measure the "goodness" of your three forecasts?

   d. For each of your forecasts, what is the value to you of a "good" as opposed to a "bad" forecast?

* Remarks, suggestions, hints, solutions: The idea behind all of these questions is to help the students realize that forecasts are of value only in so far as they help with decisions, so that forecasts and decisions are inextricably linked.

2. (Forecasting in business, finance, economics, and government) What sorts of forecasts would be useful in the following decision-making situations? Why? What sorts of data might you need to produce such forecasts?

   a. Shop-All-The-Time Network (SATTN) needs to schedule operators to receive incoming calls. The volume of calls varies depending on the time of day, the quality of the TV advertisement, and the price of the good being sold. SATTN must schedule staff to minimize the loss of sales (too few operators leads to long hold times, and people hang up if put on hold) while also considering the loss associated with hiring excess employees.

b. You're a U.S. investor holding a portfolio of Japanese, British, French and German stocks and government bonds. You're considering broadening your portfolio to include corporate stocks of Tambia, a developing economy with a risky emerging stock market. You're only willing to do so if the Tambian stocks produce higher portfolio returns sufficient to compensate you for the higher risk. There are rumors of an impending military coup, in which case your Tambian stocks would likely become worthless. There is also a chance of a major Tambian currency depreciation, in which case the dollar value of your Tambian stock returns would be greatly reduced.

c. You are an executive with Grainworld, a huge corporate farming conglomerate with grain sales both domestically and abroad. You have no control over the price of your grain, which is determined in the competitive market, but you must decide what to plant and how much, over the next two years. You are paid in foreign currency for all grain sold abroad, which you subsequently convert to dollars. Until now the government has bought all unsold grain to keep the price you receive stable, but the agricultural lobby is weakening, and you are concerned that the government subsidy may be reduced or eliminated in the next decade. Meanwhile, the price of fertilizer has risen because the government has restricted production of ammonium nitrate, a key ingredient in both fertilizer and terrorist bombs.

d. You run BUCO, a British utility supplying electricity to the London metropolitan area.

You need to decide how much capacity to have on line, and two conflicting goals must be resolved in order to make an appropriate decision. You obviously want to have enough capacity to meet average demand, but that's not enough, because demand is uneven throughout the year. In particular, demand skyrockets during summer heat waves -- which occur randomly -- as more and more people run their air conditioners constantly. If you don't have sufficient capacity to meet peak demand, you get bad press. On the other hand, if you have a large amount of excess capacity over most of the year, you also get bad press.

* Remarks, suggestions, hints, solutions: Each of the above scenarios is complex and realistic, with no clear cut answer. Instead, the idea is to get students thinking about and discussing relevant issues that run through the questions, such the forecast object, the forecast horizon, the loss function and whether it might be asymmetric, the fact that some risks can be hedged and hence need not contribute to forecast uncertainty, etc.

3. (The basic forecasting framework) True or false (explain your answers):

a. The underlying principles of time-series forecasting differ radically depending on the time series being forecast.

* Remarks, suggestions, hints, solutions: False - that is the beauty of the situation.

b. Ongoing improvements in forecasting methods will eventually enable perfect prediction.

* Remarks, suggestions, hints, solutions: False - the systems forecast in the areas that concern us are intrinsically stochastic and hence can never be perfectly forecast.

c. There is no way to learn from a forecast's historical performance whether and how it could be improved.

* Remarks, suggestions, hints, solutions: False. Indeed studying series of forecast errors can provide just such information. The key to forecast evaluation is that good forecasts shouldn't have forecastable forecast errors, so if the errors can be forecast then something is wrong.

4. (Degrees of forecastability) Which of the following can be forecast perfectly? Which can not be forecast at all? Which are somewhere in between? Explain your answers, and be careful!

a. The direction of change tomorrow in a country's stock market;

* Remarks, suggestions, hints, solutions: Some would say imperfectly, some would say not at all.

b. The eventual lifetime sales of a newly-introduced automobile model;

* Remarks, suggestions, hints, solutions: Imperfectly.

c. The outcome of a coin flip;

* Remarks, suggestions, hints, solutions: Not at all, in the sense of guessing correctly more than fifty percent of the time (assuming a fair coin).

d. The date of the next full moon;

* Remarks, suggestions, hints, solutions: Perfectly.

e. The outcome of a (fair) lottery.

5. (Data on the web) A huge amount of data of all sorts are available on the web. Frumkin (2004) and Baumohl (2005) provide useful and concise introductions to the construction, accuracy and interpretation of a variety of economic and financial indicators, many of which are available on the web. Search the web for information on U.S. retail sales, U.K. stock prices, German GDP, and Japanese federal government expenditures. (The Resources for Economists page is a fine place to start: www.rfe.org) Using graphical methods, compare and contrast the

movements of each series and speculate about the relationships that may be present.

* Remarks, suggestions, hints, solutions: The idea is simply to get students to be aware of what data interests them and whether its available on the web.

6. (Univariate and multivariate forecasting models) In this book we consider both "univariate" and "multivariate" forecasting models. In a univariate model, 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.

In a multivariate model, 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.

a. Determine which of the following are examples of univariate or multivariate forecasting:

- Using a stock's price history to forecast its price over the next week;
- Using a stock's price history and volatility history to forecast its price over the next week;
- 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 models, consider a wine merchant seeking to forecast the price per case at which 1990 Chateau Latour, one of the greatest Bordeaux wines ever produced, will sell in the year 2015, at which time it will be fully mature.

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

* Remarks, suggestions, hints, solutions: Examine the prices from 1990 through the present and extrapolate in some "reasonable" way. Get the students to try to define "reasonable."

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

* Remarks, suggestions, hints, solutions: You might also use information in the prices of other similar wines, macroeconomic conditions, etc.

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

* Remarks, suggestions, hints, solutions: Multivariate approaches bring more information to bear on the forecasting problem, but at the cost of greater complexity. Get the students to expand on this tradeoff.

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

* Remarks, suggestions, hints, solutions: You decide!

Chapter 2 Problems and Complements

1. (Interpreting distributions and densities) The Sharpe Pencil Company has a strict quality control monitoring program. As part of that program, it has determined that the distribution of the amount of graphite in each batch of one hundred pencil leads produced is continuous and uniform between one and two grams. That is, f(y) = 1 for y in [1, 2], and zero otherwise, where y is the graphite content per batch of one hundred leads.

a. Is y a discrete or continuous random variable?

* Remarks, suggestions, hints, solutions: Continuous.

b. Is f(y) a probability distribution or a density?

* Remarks, suggestions, hints, solutions: Density.

c. What is the probability that y is between 1 and 2? Between 1 and 1.3? Exactly equal to 1.67?

* Remarks, suggestions, hints, solutions: 1.00, 0.30, 0.00.

d. For high-quality pencils, the desired graphite content per batch is 1.8 grams, with low variation across batches. With that in mind, discuss the nature of the density f(y).

* Remarks, suggestions, hints, solutions: f(y) is unfortunately centered at 1.5, not 1.8. Moreover, f(y) unfortunately shows rather high dispersion.

2. (Covariance and correlation) Suppose that the annual revenues of world's two top oil producers have a covariance of 1,735,492.

a. Based on the covariance, the claim is made that the revenues are "very strongly positively related." Evaluate the claim.

* Remarks, suggestions, hints, solutions: Can't tell - it depends on the units of measurement. Are they dollars, billions of dollars, or what?

b. Suppose instead that, again based on the covariance, the claim is made that the revenues are "positively related." Evaluate the claim.

* Remarks, suggestions, hints, solutions: True.

c. Suppose you learn that the revenues have a *correlation* of 0.93. In light of that new information, re-evaluate the claims in parts a and b above.

* Remarks, suggestions, hints, solutions: Indeed the revenues are unambiguously "very strongly positively related."

3. (Conditional expectations vs. linear projections) It is important to note the distinction between a conditional mean and a linear projection.

a. The conditional mean is not necessarily a linear function of the conditioning variable(s). In the Gaussian case, the conditional mean is a linear function of the conditioning variables, so it coincides with the linear projection. In non-Gaussian cases, however, linear projections are best viewed as approximations to generally non-linear conditional mean functions.

* Remarks, suggestions, hints, solutions: This is one of the amazing and very convenient properties of the normal distribution.

b. The U.S. Congressional Budget Office (CBO) is helping the president to set tax policy. In particular, the president has asked for advice on where to set the average tax rate to maximize the tax revenue collected per taxpayer. For each of 23 countries the CBO has obtained data on the tax revenue collected per taxpayer and the average tax rate. Is tax revenue likely related to the tax rate? Is the relationship likely linear? (Hint: how much revenue would be collected at tax rates of zero or one hundred percent?) If not, is a linear regression nevertheless likely to produce a good approximation to the true relationship?

* Remarks, suggestions, hints, solutions: The relationship is not likely linear. Revenues would initially rise with the tax rate, but eventually decline as the rate nears 100 percent and people simply opt not to work, or to work but not report the income. (This is the famous "Laffer curve.") It appears unlikely that a linear approximation would be accurate.

4. (Conditional mean and variance) Given the regression model,

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t^2 + \beta_3 z_t + \varepsilon_t$$

$$\varepsilon_t \overset{iid}{\sim} (0, \sigma^2),$$

find the mean and variance of y_t conditional upon $x_t = x_t^*$ and $z_t = z_t^*$. Does the conditional mean adapt to the conditioning information? Does the conditional variance adapt to the conditioning information?

* Remarks, suggestions, hints, solutions: The conditional mean is

$$y_t = \beta_0 + \beta_1 x_t^* + \beta_2 x_t^{*2} + \beta_3 z_t^* + \varepsilon_t.$$

The conditional variance is simply σ^2.

5. (Scatter plots and regression lines) Draw qualitative scatter plots and regression lines for each of the following two-variable data sets, and state the R^2 in each case:

a. data set 1: y and x have correlation 1

b. data set 2: y and x have correlation -1

c. data set 3: y and x have correlation 0.

* Remarks, suggestions, hints, solutions: 1, 1, 0.

6. (Desired values of regression diagnostic statistics) For each of the diagnostic statistics listed below, indicate whether, other things the same, "bigger is better," "smaller is better," or neither. Explain your reasoning. (Hint: Be careful, think before you answer, and be sure to qualify your answers as appropriate.)

a. Coefficient

* Remarks, suggestions, hints, solutions: neither

b. Standard error

* Remarks, suggestions, hints, solutions: smaller is better

c. t statistic

* Remarks, suggestions, hints, solutions: bigger is better

d. Probability value of the t statistic

* Remarks, suggestions, hints, solutions: smaller is better

e. R-squared

* Remarks, suggestions, hints, solutions: bigger is better

f. Adjusted R-squared

* Remarks, suggestions, hints, solutions: bigger is better

g. Standard error of the regression

* Remarks, suggestions, hints, solutions: smaller is better

h. Sum of squared residuals

* Remarks, suggestions, hints, solutions: smaller is better

i. Log likelihood

* Remarks, suggestions, hints, solutions: bigger is better

j. Durbin-Watson statistic

* Remarks, suggestions, hints, solutions: neither -- should be near 2

k. Mean of the dependent variable

* Remarks, suggestions, hints, solutions: neither -- could be anything

l. Standard deviation of the dependent variable

* Remarks, suggestions, hints, solutions: neither -- could be anything

m. Akaike information criterion

* Remarks, suggestions, hints, solutions: smaller is better

n. Schwarz information criterion

* Remarks, suggestions, hints, solutions: smaller is better

o. F-statistic

* Remarks, suggestions, hints, solutions: bigger is better

p. Probability-value of the F-statistic

* Remarks, suggestions, hints, solutions: smaller is better

* Additional remarks: Many of the above answers need qualification. For example, the fact that, other things the same, a high R^2 is good in so far as it means that the regression has more explanatory power, does not mean that forecasting models should be selected on the basis of "high R^2."

7. (Mechanics of fitting a linear regression) On the book's web page you will find a second set of data on y, x and z, similar to, but different from, the data that underlie the analysis performed in

this chapter. Using the new data, repeat the analysis and discuss your results.

* Remarks, suggestions, hints, solutions: In my opinion, it's crucially important that students do this exercise, to get comfortable with the computing environment sooner rather than later.

8. (Regression with and without a constant term) Consider Figure 2, in which we showed a scatterplot of y vs. x with a fitted regression line superimposed.

a. In fitting that regression line, we included a constant term. How can you tell?

* Remarks, suggestions, hints, solutions: The fitted line does not pass through the origin.

b. Suppose that we had not included a constant term. How would the figure look?

* Remarks, suggestions, hints, solutions: The fitted line would pass through the origin.

c. We almost always include a constant term when estimating regressions. Why?

* Remarks, suggestions, hints, solutions: Except in very special circumstances, there is no reason to force lines through the origin.

d. When, if ever, might you explicitly want to exclude the constant term?

* Remarks, suggestions, hints, solutions: If, for example, an economic "production function" were truly linear, then it should pass through the origin. (No inputs, no outputs.)

9. (Interpreting coefficients and variables) Let $y_t = \beta_0 + \beta_1 x_t + \beta_2 z_t + \varepsilon_t$, where y_t is the number of hot dogs sold at an amusement park on a given day, x_t is the number of admission tickets sold that day, z_t is the daily maximum temperature, and ε_t is a random error.

a. State whether each of y_t, x_t, z_t, β_0, β_1 and β_2 is a coefficient or a variable.

* Remarks, suggestions, hints, solutions: variable, variable, variable, coefficient, coefficient, coefficient

b. Determine the units of β_0, β_1, and β_2, and describe the physical meaning of each.

* Remarks, suggestions, hints, solutions: Units are hot dogs. The coefficients measure the responsiveness (formally the partial derivative) of hot dog sales to the various variables.

c. What does the sign of a coefficient tell you about its corresponding variable affects the number of hot dogs sold? What are your expectations for the signs of the various coefficients (negative, zero, positive or unsure)?

* Remarks, suggestions, hints, solutions: Sign tells whether the relationship is positive or inverse. Sign on admissions is surely expected to be positive. I don't have strong feelings about the sign of the temperature coefficient; that is, I'm not sure whether people eat more or fewer hot dogs when it's hot. Maybe the coefficient is zero.

d. Is it sensible to entertain the possibility of a non-zero intercept (i.e., $\beta_0 \neq 0$)? $\beta_0 > 0$? $\beta_0 < 0$?

* Remarks, suggestions, hints, solutions: Taken rigidly, it's probably not sensible to allow a non-zero intercept. (Presumably hot dog sales must be zero if admissions are zero.) But more generally, if we view this linear model a merely a linear approximation to a potentially non-linear relationship, the intercept may well be non-zero (of either sign).

10. (Nonlinear least squares) The least squares estimator discussed in this chapter is often called "ordinary" least squares. The adjective "ordinary" distinguishes the ordinary least squares estimator from fancier estimators, such as the nonlinear least squares estimator. When we estimate by nonlinear least squares, we use a computer to find the minimum of the sum of squared residual function directly, using numerical methods. For the simple regression model discussed in this chapter, ordinary and nonlinear least squares produce the same result, and ordinary least squares is simpler to implement, so we prefer ordinary least squares. As we will see, however,

some intrinsically nonlinear forecasting models can't be estimated using ordinary least squares but can be estimated using nonlinear least squares. We use nonlinear least squares in such cases.

For each of the models below, determine whether ordinary least squares may be used for estimation (perhaps after transforming the data).

a. $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$

b. $y_t = \beta_0 e^{\beta_1 x_t} \varepsilon_t$

c. $y_t = \beta_0 + e^{\beta_1 x_t} + \varepsilon_t$.

* Remarks, suggestions, hints, solutions: OLS is fine for a, fine for b after taking logs, and no good for c.

11. (Regression semantics) Regression analysis is so important, and used so often by so many people, that a variety of associated terms have evolved over the years, all of which are the same for our purposes. You may encounter them in your reading, so it's important to be aware of them. Some examples:

a. Ordinary least squares, least squares, OLS, LS.

b. y, left-hand-side variable, regressand, dependent variable, endogenous variable

c. x's, right-hand-side variables, regressors, independent variables, exogenous variables, predictors

d. probability value, prob-value, p-value, marginal significance level

e. Schwarz criterion, Schwarz information criterion, SIC, Bayes information criterion, BIC

* Remarks, suggestions, hints, solutions: Students are often confused by statistical/econometric jargon, particularly the many redundant or nearly-redundant terms. This complement presents

some commonly-used synonyms, which many students don't initially recognize as such.

Chapter 3 Problems and Complements

1. (Data and forecast timing conventions) Suppose that, in a particular monthly data set, time t=10 corresponds to September 1960.

   a. Name the month and year of each of the following times: t+5, t+10, t+12, t+60.

   b. Suppose that a series of interest follows the simple process $y_t = y_{t-1} + 1$, for $t = 1, 2, 3, \ldots$, meaning that each successive month's value is one higher than the previous month's. Suppose that $y_0 = 0$, and suppose that at present t=10. Calculate the forecasts $y_{t+5,t}$, $y_{t+10,t}$, $y_{t+12,t}$, $y_{t+60,t}$, where, for example, $y_{t+5,t}$ denotes a forecast made at time t for future time t+5, assuming that t=10 at present.

* Remarks, suggestions, hints, solutions: t+5 is February 1961, and so on. $y_{t+5,t} = y_{15,10} = 15$, and so on.

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

* Remarks, suggestions, hints, solutions: d satisfies the criteria, immediately by inspection.