Problem Set 2

Due: Wednesday January 22, 4pm

Data Exercises

1. An Excel data file `realgdpgrowth.xlsx` is posted on Canvas. It contains quarterly observations from 1947q2 to 2024q3 for most categories of U.S. GDP, calculated as quarterly percentage changes at annual rates in real GDP and its component categories. To answer the following questions, it should be sufficient to use the "import data wizard" and the `mean`, `std`, and `quantile` commands (e.g., `quantile(gdp, [0.2 0.7])`).

   (a) The series `pce_nondurables` is personal consumption expenditures on nondurable goods. What was the percentage change (annual rate) in 2024q3?
   4.6%

   (b) Assuming growth rates are independent across quarters, so using no other information other than from the basic summary commands (such as `mean`, `std`, `quantile`, etc.), what is our point forecast for future values of `pce_nondurables`?
   Without specification, we think the loss function is quadratic. Under quadratic loss, the optimal point forecast is the unconditional mean, which is 2.5742%

   (c) Using the normal rule, give a 90% forecast interval.
   To use the normal rule, I need to have the mean, standard deviation, and Z score. We have $Z_{0.05} = 1.645$, $\mu = 2.5742$, **and** $\sigma = 3.6914$. **Thus, two quantiles are** $\mu + \sigma \cdot Z_{\alpha/2} = 8.6466$, **and** $\mu - \sigma \cdot Z_{\alpha/2} = -3.4982$. **The 90% forecast interval using normal rule is** $[-3.4982, 8.6466]$

   (d) Using empirical percentiles, give a 90% forecast interval.
   [-3.30,8.10]

2. Repeat question 1 for the series `pce_durables` (personal consumption expenditures on durable goods).

   (a) What was the percentage change (annual rate) in 2024q3?
   7.6%

   (b) What is our point forecast for future values of `pce_durables`?
   Similarly, under the quadratic loss function, the optimal point forecast is the unconditional mean, which is 6.3558.

(c) Using the normal rule, give a 90% forecast interval.
We have $Z_{0.05} = 1.645$, $\mu = 6.3558$, and $\sigma = 3.6914$. The 90% forecast interval using normal rule is $[-20.6480, 33.3596]$.

(d) Using empirical percentiles, give a 90% forecast interval.
[-15.10,27.40]

(e) What are the important differences between your answers to questions 1 and 2? Why are there such large differences for these two different components of personal consumption?
There are several differences: Firstly, the mean for nondurable goods is 2.57%, and the mean for durable goods is 6.36%. That means the consumption of durable goods, on average, was changing more rapidly between quarters. Secondly, the forecast interval of durable goods is much wider than the forecast interval of nondurable goods. That means the consumption of nondurable goods is volatile over time. Sometimes there is a huge increase in nondurable goods and sometimes a big decrease between quarters. The difference in the two means likely arises because nondurable goods include essential goods such as food and water, which are typically consumed in fixed amounts over time. In contrast, durable goods often consist of items like electronic devices, where advancing technology drives increased demand for newer and better products. The difference in quantiles between durable and nondurable goods can be attributed to the volatility in nondurable goods consumption. This volatility stems from the unstable demand for many nondurable goods, which may fluctuate with changes in price, seasonality, and the economic environment.

Theoretical Questions

3. Suppose you are an intern at a company. For a meeting, you are responsible for bringing photocopies of the agenda (a one-page sheet, cost is $0.05 each) for everyone attending the meeting. You are told that it is important that everyone has a copy. You do not know exactly how many people will be at the meeting, so you are uncertain about how many copies to print. Describe the loss problem. Is the loss function symmetric?
In this loss problem, you must decide how many photocopies of the agenda to prepare for a meeting, carefully balancing the losses from overprinting and underprinting. The loss function is asymmetric: overprinting incurs a minor loss of $0.05 per extra copy. In contrast, underprinting could severely disrupt the meeting's efficiency, causing significant financial losses for the firm. Your boss may dock your pay or even fire you, resulting in personal losses of hundreds or thousands of dollars.

4. Imagine a similar situation, but instead you are told to bring copies of Hamilton's *Time Series Analysis*, one per person, which you need to purchase at a cost of $150 per copy. You are told that it would be okay if participants shared copies. Describe

the loss problem. Is the loss function similar to question 3, or how is it different?
In this new loss problem, you must decide how many copies of books to prepare for participants. Unlike Question 3, the loss function here is different. Overpreparing even one extra book results in a significant loss of $150, far exceeding the minor loss of $0.05 in the previous scenario. Conversely, underpreparing is less critical since participants can share copies with others. While the loss function in this case is also asymmetric, it is different to 3. The higher loss is incurred when overestimating the required number of books rather than underestimating.

Question 5 and 6 are written in another file.

5. In this exercise, you will verify that the median is the best point forecast under absolute loss. Suppose that Y is a random variable with the cdf $F(y) = y^2$ and pdf $f(y) = 2y$ both on the interval $[0, 1]$.

   (a) Calculate the expectation of $|Y - c|$,

   $$E|Y - c| \equiv \int_0^1 |y - c| f(y)\, dy,$$

   where c is a constant.

   (b) Find c which minimizes $E|Y - c|$.

   (c) Find the median of F.

   (d) What is the optimal (best) point forecast for Y under absolute loss? How is it related to your answers to (b) and (c)?

6. Suppose that X is distributed uniformly on $[0, 1]$, i.e., $\text{Prob}(X \leq c) = c$ for $c \in [0, 1]$.

   (a) Find the point optimal forecast of X given the loss function

   $$L(e) = \begin{cases} 3e & e \geq 0, \\ -e & e < 0, \end{cases} \qquad e = X - \hat{X}.$$

   (b) Compare $L(e)$ and your result in (a) with the absolute loss and the point optimal forecast under the absolute loss. Discuss.