Exercises 9 · More Monte Carlo

Due Wednesday, May 6, 2015

Problem 1

Return to the data set in "peakdemand.csv" that you looked at a few weeks ago. Recall that in addition to the year, month, and day of the month for each observation, you also have the following variables:
PeakDemand: Actual peak demand for electricity in megawatts.
DailyTemp: Average temperature reading, in degrees Fahrenheit, over the course of the entire day.
Sat: coded 1 if that day was a Saturday, and 0 otherwise.
Sun: coded 1 if that day was a Sunday, and 0 otherwise.

The data set ends in 2009. Thus we will imagine ourselves back at the beginning of 2010 and trying to answer the following issue.

Duke Energy's current generating capacity is 5000 megawatts. If daily peak demand exceeds this price, they have to buy electricity on the spot market. This is expensive; they already had to do it twice during the 2006-2009 period. They want to know whether they should start working right now to expand capacity, or wait a year and start building new capacity at the beginning of 2011. It will take a year to build new capacity, so if they wait, and there are many days during calendar year 2011 where peak demand exceeds 5000 megawatts, they'll be on the hook for a large cash outlay.

Based on what we know up through 2009, what is the probability that there will be 7 or more days in calendar year 2011 where peak demand exceeds 5000 megawatts? (7 days is the pain point: 7 or more and they will build new capacity starting right now.) Build a statistical model to forecast peak demand (you should have already done this, but feel free to tweak it). Use Monte Carlo simulation to address the question, incorporating future sources of variation that are relevant to your model. In case you need information about temperature, the mean and standard deviation of temperatures over the last 50 years for every calendar day can be found in Temp50Year.csv. This also tells you which days will be Saturdays and Sundays over the period in question.

Problem 2

Continue the exercise from class on the value of diversification. To recap: you have $10000 to invest, and are contemplating allocating your wealth among five major asset classes, as represented by five ETF's:

- SPY, large-cap stocks
- TLT, long-term government bonds
- LQD, long-term corporate bonds
- DBC, commodities (oil, gold, metals, etc)
- VNQ, real estate

You will use Monte Carlo simulation combined with bootstrap resampling (using data from Yahoo finance from the beginning of 2008 through now) to compare the following three portfolio allocations:

1. 50% of your wealth in stocks (SPY), 50% in real estate (VNC)
2. 50% in commodities (DBC), 50% in government bonds (TLT)
3. 20% in each of the five asset classes.

Use these techniques to address the following two questions.

(A) Suppose you plan to take a buy-and-hold strategy for two weeks (10 trading days). That is, you do not rebalance your portfolio along the way to compensate for differing gains and losses in each of the assets. Estimate the 5% Value at Risk (VaR) for each of these portfolios over this horizon. Use enough Monte Carlo samples to get stable estimates.

(B) Now imagine a year's worth of trading in which you rebalance your portfolio every day to achieve the desired percentage allocation of your present wealth. The question is: which of these three portfolio-allocation strategies maximizes expected utility after a single year (250 days of compounded daily returns) of trading? Assume log utility, $u(w) = \log(w)$. Again, use enough Monte Carlo draws to get stable estimates.

To get enough Monte Carlo draws, you may need to leave your computer running awhile in the background. I would recommend checking your scripts first with relatively small Monte Carlo sample sizes, just to make sure they work as intended. Then increase the Monte Carlo sample size and take a break while your computer churns through the computations necessary to get your final answer.