Exercises 8 · Probability, continued

Due Monday, April 11, 2016

Problem 1: sequential betting, revisited

(A) Consider a long-term investment in a risky asset where X_i is the return on your bet in period i. After T periods your wealth is given by the compound-interest formula

$$W_T = W_0 \cdot \prod_{i=1}^{T} (1 + X_i) \, .$$

Suppose that the X_i are independent, identically distributed samples from some probability distribution. Prove that the long-term growth rate of your investment is

$$R = e^{E[\log(1+X_i)]} - 1 \, ,$$

where E is expected value. By "long-term growth rate," I mean the equivalent rate R at which your investment would grow, assuming that you instead invested it in a riskless asset with interest rate R.

The implication of this result is that, to maximize the growth rate of an investment, one should attempt to maximize $E[\log(1 + X_i)]$.

Here are a couple of facts/hints to work with. Note that it may not be obvious how to cobble these together, but they are both relevant.

(1) Recall (or meet) the Law of Large Numbers. Suppose that $X_1, \ldots, X_T$ are independent, identically distributed random variables from some probability distribution. Then

$$\lim_{T \to \infty} T^{-1} \sum_{i=1}^{T} f(X_i) \longrightarrow E(f(X_i))$$

with probability 1. That is, long-run sample averages converge to their expected values.

(2) For a long-term investment with a constant per-period rate of return R, your wealth after T periods is

$$W_T = W_0 \prod_{i=1}^{T} (1 + R) \, .$$

Try matching the two cases (i.e. where your wealth after investing in the risky asset for T periods matches your wealth after investing in a fixed-rate asset for T periods) to solve for the implied growth rate R of the risky asset.

(B) Remember the scenario from a previous homework: you have W_0 dollars to your name, and that you are offered the following bet.

- With probability $p > 0.5$, you will win the bet.
- With probability $1 - p$, you will lose the bet.

You get to choose what fraction c of your total wealth W_0 to wager on the outcome of the bet, and to repeat the bet over and over again, 10,000 times. (Note that previously we had $p = 0.52$, but we'll leave it generic here.)

Previously you experimented with various choices of c. In light of the rule above: (1) What choice of c maximizes the long-term growth rate of your investment? And (2) what is the smallest value of c for which the long-term growth rate is negative (despite each individual round having a positive expected value)? After deriving your answer, go rerun the investment simulator a few times to convince yourself that your wealth really does grow quite fast under this "optimal" allocation.

Problem 2: Stocks and bonds

Let (X_{t1}, X_{t2}) denote the return in year t on the stock market (X_1) and on government bonds (X_2), respectively. Suppose that (X_1, X_2) are IID and follow a bivariate normal distribution with parameters:

$$\begin{aligned} \mu_1 = 0.065\,, &\quad \mu_2 = 0.015 \\ \sigma_1 = 0.20\,, &\quad \sigma_2 = 0.10 \\ \rho = -0.1\,. & \end{aligned}$$

Set up a Monte Carlo simulation that assess which of the following portfolio results in the highest long-term growth rate. (1) 50% stocks, 50% bonds; (2) 80% stocks, 20% bonds; or (3) 100% stocks. Assume that each year, you rebalance your portfolio to the target mix of assets (since otherwise it will drift randomly away from the target over time).

Remember from the previous problem that maximizing a portfolio's expected growth rate is equivalent to maximizing $E[\log(1 + X_i)]$ where X_i is the per-period return on your investment.

Describe your approach and your answer. Turn in your R code as a supplement to (but not a replacement of) this description. Make sure you print your R code in a fixed-width font.

Problem 3: a better model for soccer

Go read the article "One match to go!" by Spiegelhalter and Ng, linked from the course web page. They describe how they formulated their approach for predicting soccer matches. It is better than the simple approach we took in the course packet (though probably not as good as what actual bookies use).

Now go get data from this year's English Premiere League soccer season. For example, you can certainly find it here: `http://www.soccerstats.com/latest.asp?league=england`. Replicate their approach using this year's data. (You can use any combination of software tools you find helpful here, including Excel.) What is your estimated probability distribution of likely results for the match on April 17th between Leicester City and West Ham? How about the match between Manchester United and Aston Villa?