# ***Beta and Regression***

## Administrative Items

**Getting help**

– See me Monday 3-5:30 or tomorrow after 2:30.
– Send me an e-mail with your question. (stine@wharton)
– Visit the StatLab/TAs, *particularly for help using the computer*.

**Problem set #5**

– Last problem set.
– Preparations for the project.

**Makeup Exam**

Next week. Be there or keep the *zero* you now have!

## Adjusting for Risk in Investments

**What's up with the stock market?**

**Dice assignment in Stat 101**

– Returns on several "simulated" investments.

| | Avg Annual Return | SD of Return | |
|---|---|---|---|
| White | 0% | 5% | "cash" |
| Green | 7.5% | 20% | "market" |
| Red | 71% | 130% | "internet" |

– Which of these investments worked out for your group?

– What about the mixed investment, called "Pink" which is half in Red and half in White?

| | Avg Annual Return | SD of Return | |
|---|---|---|---|
| Pink | 35.5% | 65% | "portfolio" |

**Effects of variation**

– Variability in returns is expensive!
– Start with an initial wealth $W_0$ of $100.
– Assume return is up 10% on one day, and down by 10% the next.
– Where do you end up after a few days?

$$S\,(1+.1)(1-.1) = S(1-.01) = S\,(0.99) \rightarrow \text{losing 1\% / two days}$$

or

losing 1/2 % every day!

## Risk-adjusted return

The risk-adjusted value of an investment with stochastic returns R (i.e., the returns change from day to day) is often calculated as (some will do this differently)

$$\text{Value} = E(R) - Var(R)/2$$

The second term is often called the "volatility drag" on the returns.

## Background (optional)

For the interested ones, here's an outline justifying the previous expression. Again, start with wealth $W_0$. If we denote the return on your investment on the $i^{th}$ day as $R_i$, then your wealth after *n* days is

$$W_n = W_0\,(1+R_1)(1+R_2)\,\ldots\,(1+R_n)$$

Rather than work with products, convert this to sums by taking logs.

$$\log W_n = \log W_0 + \Sigma \log(1+R_i)$$

and use the Taylor approximation $\log(1+x) \approx x - x^2/2$ to get

$$\log W_n \approx \log W_0 + \Sigma\,(R_i - R_i^2/2)$$

which has long run average

$$E \log W_n \approx \log W_0 + n\,(ER_i - Var(R_i)/2)$$

*So, if you want to maximize your expected wealth, you want to maximize*

$$E\,R_i - Var(R_i)/2$$

## Analysis of the dice assignment

– Returns on several "simulated" investments.

| | Avg Annual Return | SD of Return | Adjusted |
|---|---|---|---|
| White | 0% | 5% | 0 |
| Green | 7.5% | 20% | 5.5% |
| Red | 71% | 130% | –13.5% |

– What about the mixed investment which is half in Red and half in White?

| | Avg Annual Return | SD of Return | Adjusted |
|---|---|---|---|
| Pink | 35.5% | 65% | 14.4% |

### Portfolios

As in the dice example, one often combines individually unappealing investments (here white and red) to form an investment that is attractive. The trick is to decide how much of your money to invest.

Regression offers the key to figuring this out, because there is something very fishy about the previous experiment with dice that does not hold up in financial markets…

What's really artificial???

## Looking at Some Real Returns

### Correlations over time for several stocks

Note that all of the stock returns, even those from companies in very different industries, are positively correlated. (StockRet.jmp data file)

**Correlations**

| Variable | Sears | K-Mart | Penney | Mcdonalds | Citicorp |
|---|---|---|---|---|---|
| Sears | 1.000 | 0.603 | 0.631 | 0.584 | 0.525 |
| K-Mart | 0.603 | 1.000 | 0.682 | 0.429 | 0.357 |
| Penney | 0.631 | 0.682 | 1.000 | 0.533 | 0.419 |
| Mcdonalds | 0.584 | 0.429 | 0.533 | 1.000 | 0.491 |
| Citicorp | 0.525 | 0.357 | 0.419 | 0.491 | 1.000 |

### Artificial in the dice example?

Returns are close to independent over time, but there is something else artificial about the dice example.

## Investing in a Market of One

### Problem

How much "pink" should you buy?

### Investing in one security

What share $\gamma$ of your wealth should you invest in a risky asset, holding the rest in cash?

## Goal with one asset

One objective is to maximize your long-term wealth. The average gain with share $\gamma$ is

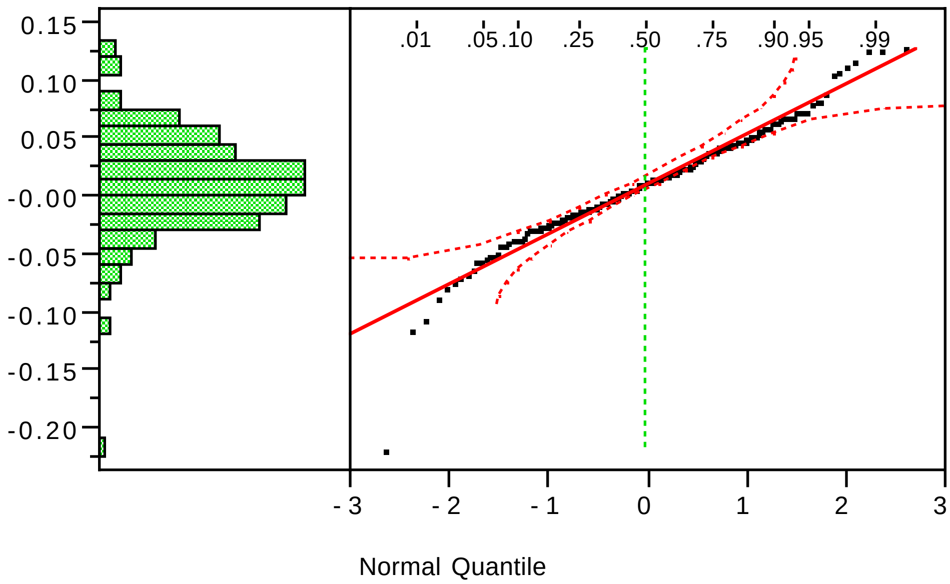
$$E(\gamma R) - Var(\gamma R)/2 = \gamma E(R) - \gamma^2 Var(R)/2$$

which we can maximize as a function of the share, finding the maximum long-term value is obtained with

$$\gamma = E(R)/Var(R)$$

## Example for investing in the S&P 500

Using the monthly value-weighted returns over the 19 year period 1976-1994, we get this summary for the S&P 500




with mean 0.0116 and variance .0019. This return is not adjusted for inflation, however, and removing the risk-free return .0059 gives an inflation adjusted return of .0057. Divided by the variance, you get a net share of

$$\gamma = E(R)/Var(R) = .0057/.0019 = 3$$

Yup, it would have been nice to have been leveraged in the market. (Things are not so impressive, though, if you adjust for the cost of having to borrow that money that you just used to invest further in the market.)

# Making a Small Portfolio

### Problem

We start with initial wealth, say \$1 at time 0. We can invest in *two* securities, with returns R1 and R2. Portfolios with the dice looked pretty good, but those were independent. How much of my current wealth should I put into each investment?

In other words, for the dice, why put half in red and half in white? And how do you decide this when red and white are not independent?

### Special trick

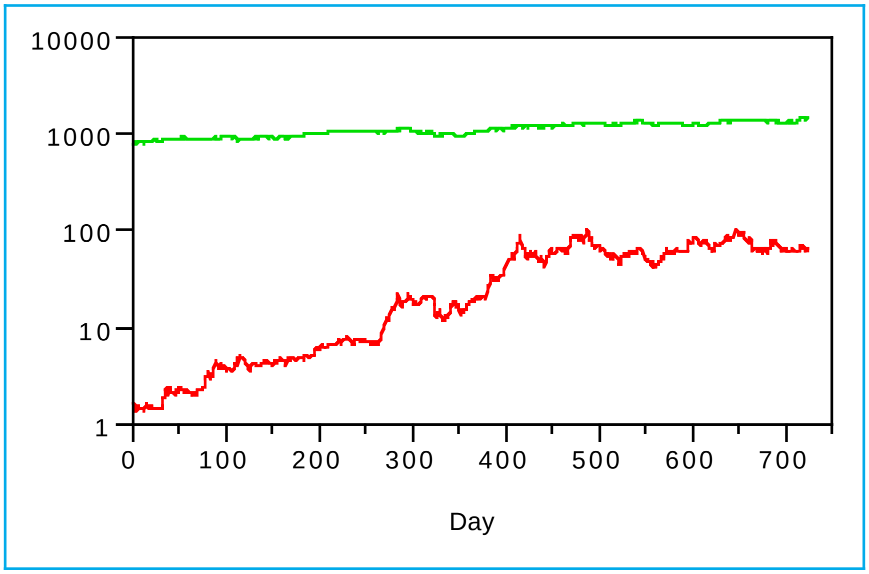
The investment shares $\gamma$ and are simply

$$\gamma_1 = E(R_1)/Var(R_1) \text{ and } \gamma_2 = E(R_2)/Var(R_2)$$

when the investments are uncorrelated. So, how can we make investments uncorrelated? Use regression…

### Two investments

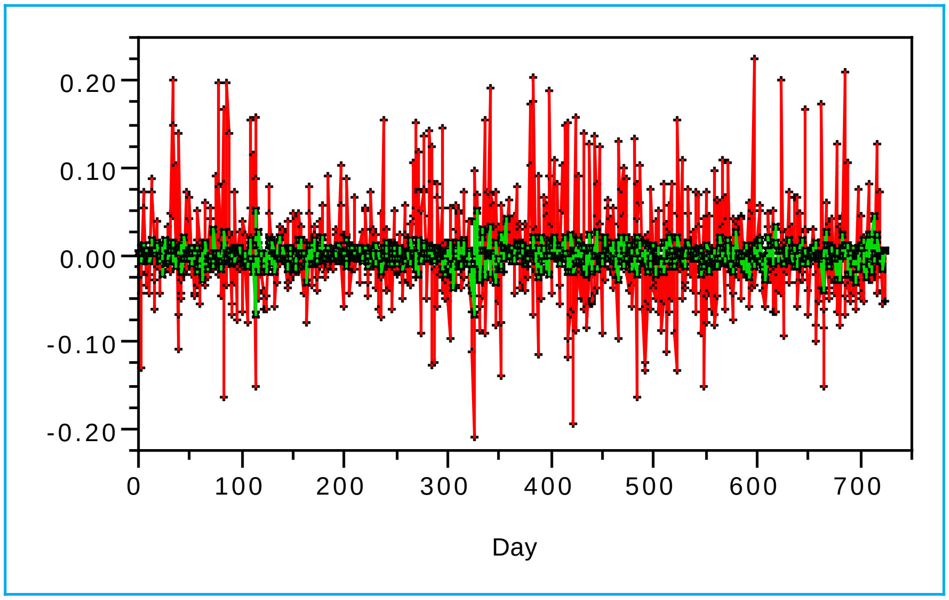
For an example, consider investing either in Amazon or the S&P 500. Here's a plot of their value (thanks to Yahoo) for the last couple of years on a log scale




with Amazon in red and the S&P in green at the top (very flat).

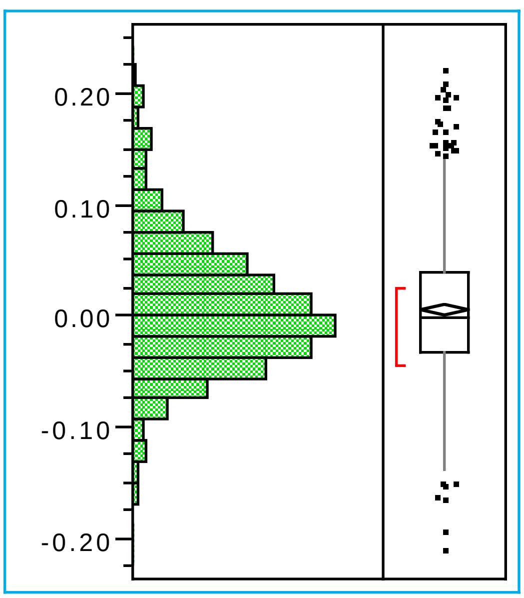
## Returns

The next plot shows the daily returns for the same two. Now you can start to see some of the volatility in the returns for Amazon

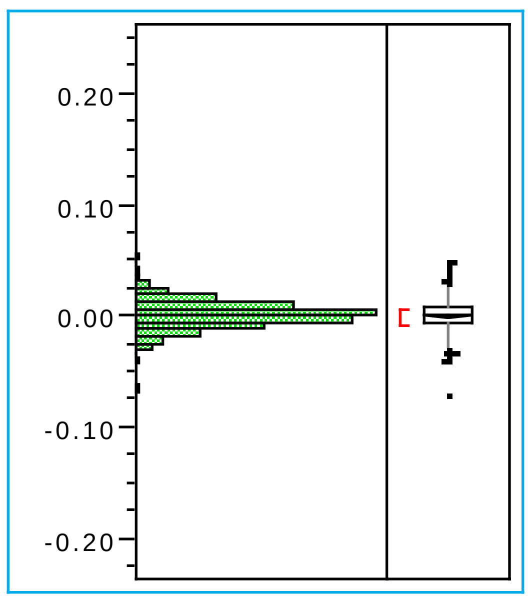



Some days Amazon was up 20%, some days down 20%. Here are the associated summary statistics (neither adjusted for inflation)

Amazon




S&P500

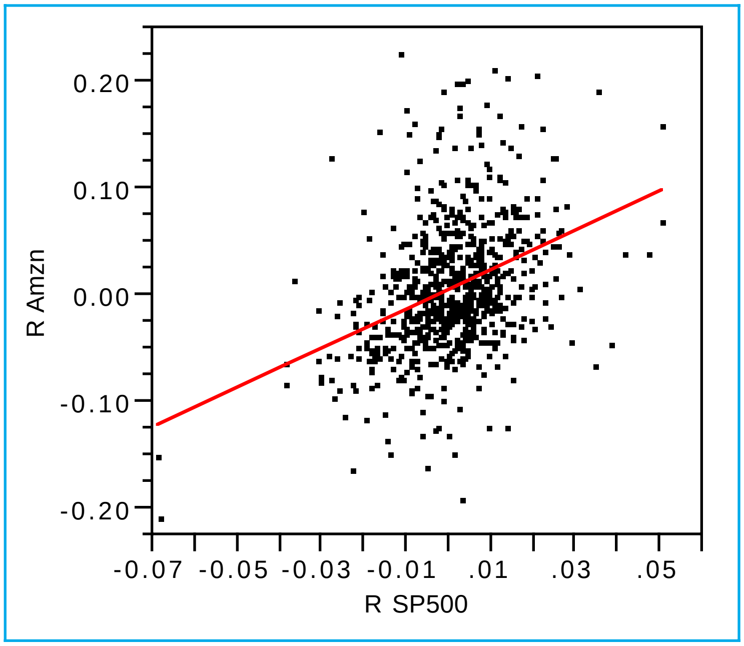



| | Amazon | | S&P500 |
|---|---|---|---|
| Mean | 0.0068 | Mean | 0.00090 |
| Variance | 0.0038 | Variance | 0.00015 |

However, these two are correlated, as seen in the next plot, so we cannot set our investment shares based on these summaries alone.

## Regression

Regression allows us to easily construct two investments. Here's a plot of Amazon on the S&P




and the associated bivariate (i.e., single predictor) regression results:

R Amzn = 0.00518 + 1.8332 R SP500

| | |
|---|---|
| RSquare | 0.138 |
| Root Mean Square Error | 0.057 |
| Mean of Response | 0.007 |
| Observations (or Sum Wgts) | 725 |

**Parameter Estimates**

| Term | Estimate | Std Error | t Ratio | Prob>\|t\| |
|---|---|---|---|---|
| Intercept | 0.0052 | 0.0021 | 2.44 | 0.0151 |
| R SP500 | 1.8332 | 0.1707 | 10.74 | <.0001 |

Residuals are not correlated with the predictor

Synthetic investments

# Next Class on Wednesday

**Simple regression in finance**

The term “ beta” in finance refers to a regression coefficient. On Weds, we’ll take a look at what makes this coefficient so interesting.