Statistics + Quantitative Finance

Wednesday, March 20th

About Me

Name: Carsen Sharkey

Year: Sophomore

Hometown: Westfield, New Jersey

Majors: Applied Mathematics and Statistics & Analytics

Minor: Data Science

Professional Interests: Quantitative Trading / Research

Brainteaser

How many trailing zeros does 1000! have?

- A trailing zero is added whenever a number is multiplied by 10
- Need to figure out how many factors of 2 and 5 there are (whichever is fewer)
- A factorial contains many more factors of 2 than factors of 5, so we just need to figure out how many factors of 5 there are.
- 1 out of every 5 numbers in the factorial is a factor of 5. However, factors of 25 contains two factors of 5, so they need to be counted twice, and so on for other powers of 5.
- So there are:
  - 1000 / 5 = 200 factors of 5
  - 1000 / 25 = 40 factors of 25
  - 1000 / 125 = 8 factors of 125
  - 1000 / 625 = 1.6 factors of 625 (which rounds down to just 1 factor, 625 itself)
- Answer: 200 + 40 + 8 + 1 = 249

Terminology

---

1) Random Variable:
   - A variable which represents the outcome of a trial, an experiment, or an event.
   - Example: You roll two dice and let X be the sum of the two dice. We say X is a random variable with possible values 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 12.

Terminology

---

2) Expected value (mean): The probability-weighted average of the possible outcomes of the random variable

$$E(X) = P(X_1)X_1 + P(X_2)X_2 + \ldots + P(X_n)X_n$$

3) Variance: The expected value of the squared deviation from the mean of a random variable

$$VAR(X) = E\{[X-E(X)]^2\}$$

4) Standard Deviation: The square root of variance

$$STDEV(X) = [VAR(X)]^{1/2}$$

Basic Example

The probability distribution of a company's sales is as follows:

Probability	Sales ($ millions)
0.20	50
0.30	40
0.50	30

What is the expected sales for the company?

E(Sales) = 0.20*50 + 0.30*40 + 0.50*30 = $37 million

Basic Example Part 2

Recall:

Probability	Sales ($ millions)
0.20	50
0.30	40
0.50	30

What is the variance of the company's sales?

Variance = $P(\$50)[\$50-E(Sales)]^2 + P(\$50)[\$50-E(Sales)]^2 + P(\$50)[\$50-E(Sales)]^2 = 0.20(\$50-\$37)^2 + 0.30(\$40-\$37)^2 + 0.50(\$30-\$37)^2 = \$61$ million

Basic Example Final Part

Recall:

Probability	Sales ($ millions)
0.20	50
0.30	40
0.50	30

What is the standard deviation of the company's sales?

Standard Deviation = sqrt($61 million) = $7.81 million

Application to Quantitative Finance

- Expected value and variance/standard deviation are used frequently in portfolio management

- The expected value implies an anticipated average value for an investment at some time in the future.

- The variance & standard deviation are measures of risk - a higher variance/standard deviation implies a higher risk on the investment

- The optimal portfolio has investments with relatively high expected values and relatively low standard deviations

Application to Quantitative Finance

---

1) The expected return on a portfolio is a weighted average of the expected returns on the component securities

$E(R) = w_1R_1 + w_2R_2 + \ldots + w_nR_n$ d

2) The variance of a portfolio's return is a function of the variance of the component assets as well as the covariance between each of them

For a 2-asset portfolio with assets A and B:

$Var(R) = w_A*Var(R_A) + w_B*Var(R_B) + 2*w_A*w_B*Cov(R_A, R_B)$

Application to Quantitative Finance

---

3) Covariance is a measure of the degree of co-movement between two random variables

- For example, there may be covariance between interest rates and inflation

Example Problem

You have a portfolio of two mutual funds, A and B, where 60% is invested in A and 40% is invested in B. The covariance between A and B is 0.0144.

Fund	A	B
Expected Return	15%	10%
Standard Deviation	20%	15%

What is the portfolio’s expected return and standard deviation of returns?

Solution

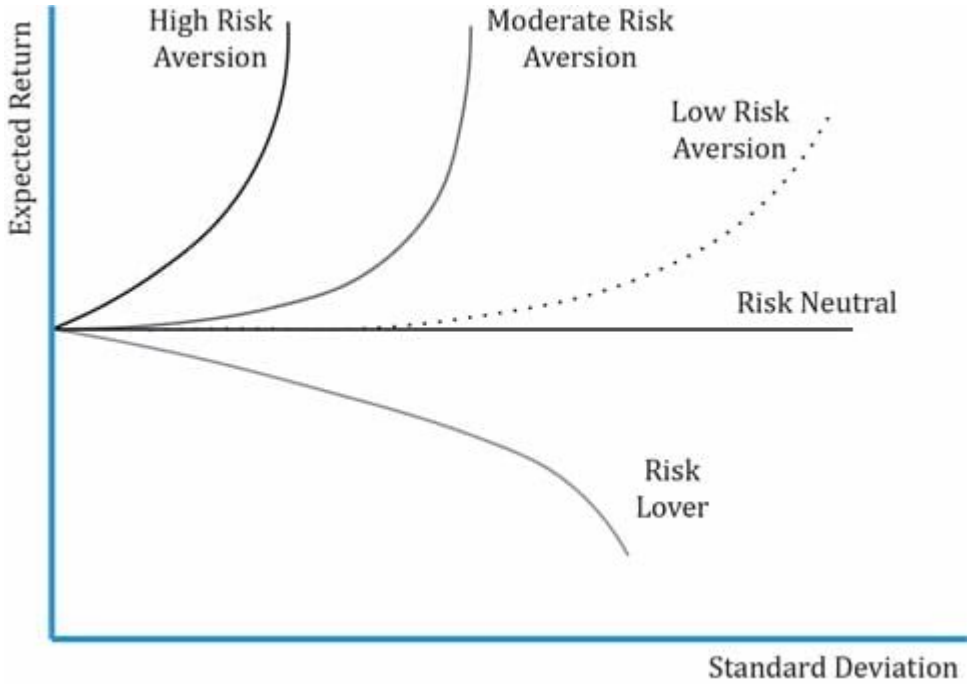
$E(R) = 0.6(15\%) + 0.4(10\%) = 13\%$

$Variance(R) = 0.6^2(0.2)^2 + 0.4^2(0.15)^2 + 2*0.6*0.4*0.0144$

$= 0.024912$

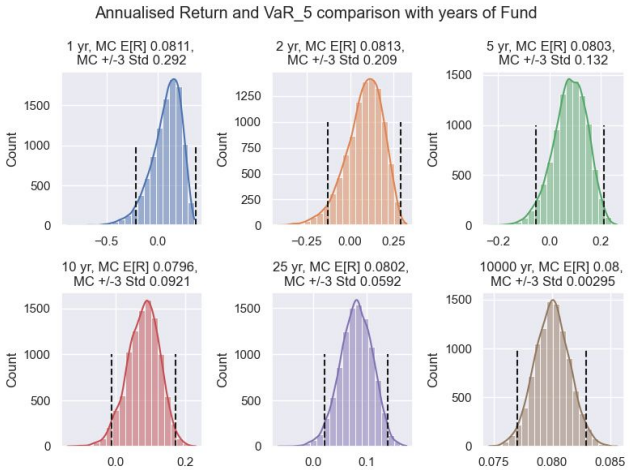
$Standard\ Deviation(R) = 0.1578 = 15.78\%$

Portfolio Management Curve



Probability Distributions

- A probability distribution is a mathematical function that assigns probabilities to various outcomes
- For example, we can assign a probability to the outcome of a certain stock increasing in value or decreasing in value
- Investors use probability distributions to calculate returns on assets and potential catastrophic events so that they can understand and then hedge their risk
- Applications in risk management, option pricing, and portfolio optimization, which are all prominent areas of quantitative finance
- Types:
  - Discrete
  - Binomial
  - Normal (Gaussian): Most commonly used, many real-life phenomena follow normal distributions
  - Poisson



Example Interview Question

“You flip a weighted coin that comes up Heads 40% of the time and Tails 60%. If you flip this coin 5 times, what is the probability that you see at least 3 Tails, rounded to the nearest percent? Please input your answer as a decimal rounded to the nearest hundredth.”

X = The number of tails

$X \sim \text{Binomial}(5, 0.6)$

(5 = # of trials; 0.6 = probability of success)

$$P(X \geq 3) = \binom{5}{3}(.6)^3(.4)^2 + \binom{5}{4}(.6)^4(.4)^1 + \binom{5}{5}(.6)^5(.4)^0$$

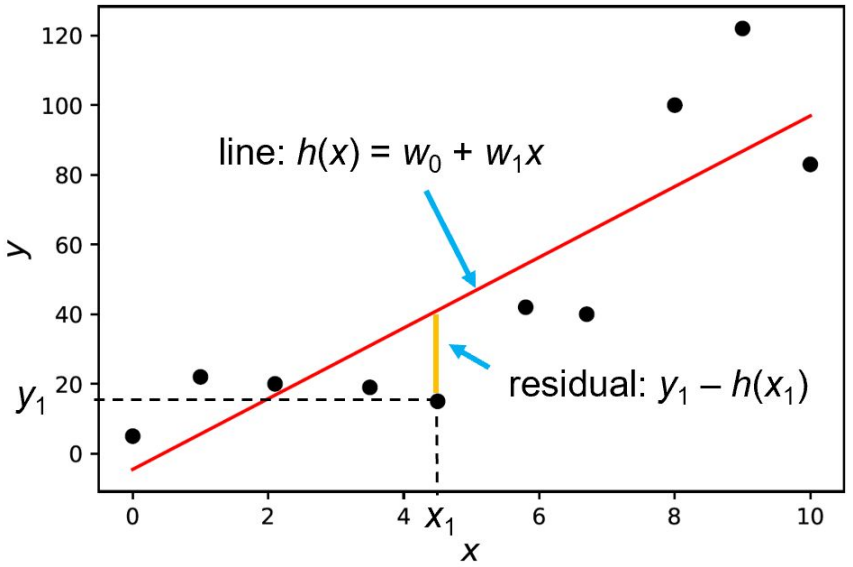
$$= \boxed{0.68}$$

Linear Regression

- Statistical method for determining how one or more independent variables and a dependent variable are related
  - Helps trading professionals to spot patterns, trends, and correlations that might have gotten lost in the immense sea of market-related information and data
- Potential uses of linear regression:
  - Predict future stock prices or asset returns based on historical data and other relevant factors
  - Understand the impact of various market factors on the price of financial assets
  - Estimate the risk and return characteristics of investment portfolios
  - Analyze the relationship between financial variables and macroeconomic indicators
- Overall, regression analysis helps financial analysts and researchers to better understand and quantify the relationships between different financial variables, which can inform investment decisions, risk management strategies, and financial modeling

Line of Best Fit

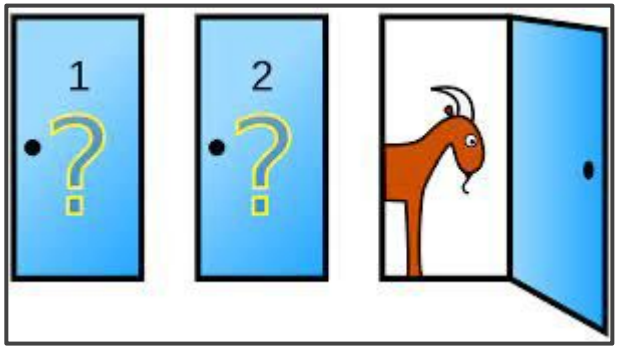
- The fitted values are the predictions
- The residuals - the vertical distances from the data points to the line - help assess the quality of the fit

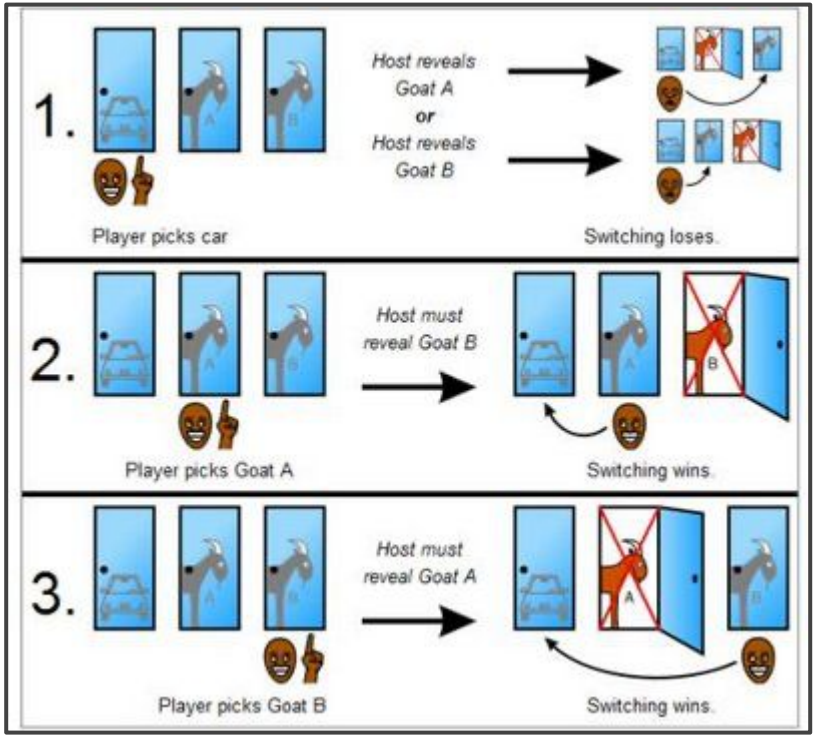


Hypothesis Testing

- At its core, science is about making falsifiable hypotheses about the world, testing them experimentally, then using the experiment outcomes to refute or refine the hypotheses
- The scientific method is an integral part of quantitative finance; it provides a framework we can use to identify and analyze trading signals or anomalies
- We set up a null hypothesis and test whether there is sufficient evidence to reject the null hypothesis
- The p-value is the probability that random data can explain your results
- The p-value is what determines whether or not you can reject the null hypothesis (show whether or not there is enough data to support a particular relationship existing)

Brainteaser - Monty Hall Problem





Attendance Tracker for QFA Members

CLE Credit