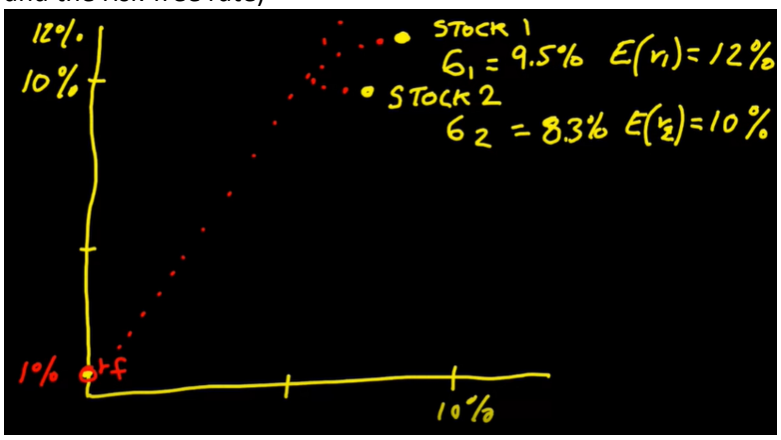
Week 4 Introduction to parametric models

12 July 2020 02:24

- model
  - signal
  - noise
- Gaussian model
  - standardisation (xi - mean) / std_dev
  - NORMSDIST(z) (cumulative prob)
    NORMSINV(prob)
  - central limit thm
  - Markowitz portfolio optimisation
    - stock i has expected return E(r_i); then expected return of portfolio E(rp) = wi * E(r_i) (sum_wi = 1)
    - x- axis: volatility; y-axis: return
      Want to find wi to maximise the Sharpe ratio (gradient of the line joining the pt and the risk-free rate)
    - 

- Linear regression (assuming X, Y gaussian)
  - standardising x and y
    => covariance = beta = R, alpha = 0 where best fit line is y = beta x + alpha, R^2 = correlation (correlation remains the same)
    - goal: minimise sigma((yi - yi_pred)^2) = sigma((yi - beta*xi -alpha)^2)
      => with standardisation, alpha = 0, so beta = sigma(xiyi)/sigma(xi^2) = covariance / 1 = covariance
      also, R = covariance / std(x)std(y) = covariance
  - Modelling error
    - Y = (beta * X + alpha) + error term
      with standardisation, phi(0,1) = phi(0, beta^2) + phi(0, stdev_error)
      => 1 = R^2 + stdev_error^2
    - A *Point* Forecast: $\hat{y}_i = \beta x_i$, and
      A *Probabilistic* Forecast: $\phi(\beta x_i, 1 - \beta^2)$.
    - can form a CI for the forecast
  - Information gain
    - differential entropy for CRV
    - H(Y) - H(Y|X) = H(phi(0,1)) - H(phi(0, stdev_error^2)
      signal + noise noise
    - Note H(X) = 2.05 + log(stdev) for gaussian
      So I(X;Y) = H(Y) - H(Y|X) = - log(stdev_error) = - log (sqrt(1-R^2))
    - percentage information gain = I(X;Y) / H(Y) = - log (sqrt(1-R^2))/2.05
  - 2 different kinds of forecast