Statistics Review

Professor Lasse H. Pedersen

Outline

- Statistics review
  - Random variables
  - Expected value, variance, standard deviation, covariance, correlation
- Properties of linear combinations of random variables, such as portfolio returns
  - Return on portfolio
  - Expected return on portfolio
  - Variance and standard deviation of portfolio
  - Risk-reduction and diversification

Random Variables

- Definition: a value representing an outcome of an uncertain event
- The outcome may be
  - Discrete:
    - “scenario analysis”
    - Example: economic environments becomes (recession, normal, boom)
    - Example: flip of a coin (heads, tails)
  - Continuous:
    - Example: the weight of a baby
    - Example: the return of a stock
Distribution

- The likelihood of each possible event
- Example of a discrete outcome
  - Fair coin: 50% head, 50% tail
  - Rigged coin: 60% head, 40% tail
- For continuous outcome
  - Normal distribution

Normal Distribution (Bell Curve)

Expected Value

- The expected value is the average outcome if the event was repeated infinitely often.
- It is probability-weighted average of the possible outcomes.
- Suppose the return $R_i$ on an asset $i$ is equal to $R_i(s)$ with probability $p(s)$ for $s=1,\ldots,S$.
- Then the expected return is:

$$E(R_i) = \sum_{s=1}^{S} R_i(s) p(s)$$
Variance and Standard Deviation

- The **variance** is the average squared deviation from the expected value:
  \[ \sigma_i^2 = E[(R_i(s) - E(R_i))^2] \]
  \[ = \sum_{s=1}^{S} (R_i(s) - E(R_i))^2 p(s) \]
- The **standard deviation** (SD) is the square root of the variance:
  \[ \sigma_i = \sqrt{\sigma_i^2} \]

Covariance

- The **covariance** between two random variables is the average of the products of their deviations from the mean:
  \[ \text{cov}(R_i, R_j) = E[(R_i - E(R_i))(R_j - E(R_j))] \]
  \[ = \sum_{s=1}^{S} (R_i(s) - E(R_i))(R_j(s) - E(R_j)) p(s) \]
- The covariance is
  - **Positive** if it tends to be that if the one random variable is unusually high (or low) then the other is also high (or low)
  - **Negative** if the one variable tends to be high when the other is low, and vice versa

Correlation

- The **correlation** is the covariance between two random variables, divided by their standard deviations:
  \[ \rho_{ij} = \frac{\text{cov}(R_i, R_j)}{\sigma_i \sigma_j} \]
- The correlation is scaled so that:
  \[ -1 \leq \rho_{ij} \leq 1 \]
Estimating Mean, Variance, and Covariance from Historical Data

- Use the "sample counterpart" of the definition:

\[
\hat{\sigma}^2 = \frac{1}{T-1} \sum_{t=1}^{T} [R_t(t) - \hat{\mu}(R_t)]^2
\]

\[
cov(R_i, R_j) = \frac{1}{T-1} \sum_{t=1}^{T} [R_i(t) - \hat{\mu}(R_i)][R_j(t) - \hat{\mu}(R_j)]
\]

Portfolio

- A combination of N assets, with returns \(R_1, \ldots, R_N\).
- Portfolio \(p\), with portfolio weights \(\omega_1, \ldots, \omega_N\):
  - \(\omega_i\) is percentage of wealth invested in asset \(i\):
    \[
    \omega_i = \frac{\text{\$ value of stock } i\text{'s position}}{\text{total } \$ \text{ value of portfolio}}
    \]
  - Portfolio weights sum to one: \(\omega_1 + \cdots + \omega_N = 1\).
  - A negative weight indicates a short position.

Portfolio Return and Portfolio Expected Return

- The return on the portfolio is:
  \[
  R_p = \sum_{i=1}^{N} \omega_i R_i
  \]
- The expected return on the portfolio is:
  \[
  E(R_p) = \sum_{i=1}^{N} \omega_i \hat{\mu}(R_i)
  \]
Portfolio Variance and SD

- With 2 securities (N=2), the portfolio variance is:
  \[ \sigma_p^2 = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2 \omega_1 \omega_2 \rho_{12} \sigma_1 \sigma_2 \]

- In general, the portfolio variance is:
  \[ \sigma_p^2 = \sum_{i=1}^{N} \omega_i^2 \sigma_i^2 + 2 \sum_{i=1}^{N} \sum_{j>i}^{N} \omega_i \omega_j \rho_{ij} \sigma_i \sigma_j \]

- The standard deviation of the portfolio is:
  \[ \sigma_p = \sqrt{\sigma_p^2} \]