Multivariate Distribution

Motivation

- Understanding dependence between risks is crucial for risk aggregation and management.
- Correlation is the most common measure, but has important limitations.
- More general concepts are needed for complex portfolios.

Dependence via Covariance

Definition (Covariance)

The covariance between random variables X and Y is:

$$\mathrm{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$$

The correlation coefficient is:

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} = \frac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

where $\rho(X, Y) \in [-1, 1]$.

Positive and Negative Dependence

- **Positive dependence:** $\text{Cov}(X, Y) > 0$
  - When X is above its mean, Y tends to be above its mean
  - Variables move in the same direction on average
- **Negative dependence:** $\text{Cov}(X, Y) < 0$
  - When X is above its mean, Y tends to be below its mean
  - Variables move in opposite directions on average
- **Uncorrelated:** $\text{Cov}(X, Y) = 0$ (necessary but not sufficient)

Financial Market Examples

Example (Positive Dependence)

Stock Returns in Same Sector:

- Apple (AAPL) and Microsoft (MSFT) returns
- Both tech companies affected by similar factors
- During tech boom: both rise; during tech crash: both fall
- Correlation $\rho \approx 0.6 - 0.8$

Example (Negative Dependence)

Stock and Bond Returns:

- S&P 500 returns vs. 10-year Treasury bond returns
- When stocks fall (flight to quality), bonds rise
- When stocks rise (risk-on), bonds may fall
- Correlation $\rho \approx -0.2$ to -0.5 (varies over time)

Example (Currency Pairs)

USD/EUR vs. EUR/GBP: Often negatively correlated due to USD strength affecting both pairs oppositely.

Independence

- X and Y are independent if
  $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$ for all x, y.
- Equivalently, X and Y are independent iff

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$$

  for all measurable f, g.

Independence and correlation

- Independence $\implies$ zero correlation, but not vice versa.
- (Counter)example: $X \sim N(0,1)$, $Y = X^2$.
- Here, $\text{Cov}(X, Y) = 0$, hence X and Y are uncorrelated.
- But X and Y are clearly dependent since

$$\mathbb{P}(X > 1, Y > 1) \neq \mathbb{P}(X > 1)\mathbb{P}(Y > 1).$$

Example

- Let $Y = aX + Z$ for constants a and independent r.v. X and Z.
- Then $\text{Cov}(X, Y) = a\,\text{Var}(X)$.
- If $a > 0$, positive dependence;
- If $a < 0$, negative dependence.
- If $a = 0$, uncorrelated.

Problem setup

- Dataset: Historical daily stock prices for two companies (e.g., Apple and Microsoft)

$$\{X_i : i = 0, \dots, n\}, \quad \{Y_i : i = 0, \dots, n\}$$

- Compute daily returns:

$$x_i = \frac{X_i - X_{i-1}}{X_{i-1}}, \quad y_i = \frac{Y_i - Y_{i-1}}{Y_{i-1}}$$

- Goal: Estimate correlation $\rho(x_n, y_n)$.

Averaging using past m days

- $\mathrm{cov}_n := \mathrm{Cov}(x_n, y_n) = E[x_n y_n] - E[x_n]E[y_n]$
- Assume $\mathbb{E}x_n = \mathbb{E}y_n = 0$ (demeaned returns)
- $\mathrm{cov}_n = E[x_n y_n]$
- Sample mean estimation:

$$\mathrm{cov}_n = \frac{1}{m}\sum_{i=1}^{m} x_{n-i} y_{n-i}$$

- Sample correlation estimation:

$$\mathrm{var}_{x,n} = \frac{1}{m}\sum_{i=1}^{m} x_{n-i}^2, \quad \mathrm{var}_{y,n} = \frac{1}{m}\sum_{i=1}^{m} y_{n-i}^2$$

$$\mathrm{cor}_n = \frac{\mathrm{cov}_n}{\sqrt{\mathrm{var}_{x,n}\mathrm{var}_{y,n}}}$$

- Weighted average is similar.

EWMA

- Exponentially Weighted Moving Average (EWMA) assigns exponentially decreasing weights to past observations.
- Formula:

$$\mathrm{cov}_n = \lambda \mathrm{cov}_{n-1} + (1-\lambda) x_{n-1} y_{n-1}$$

where $\lambda \in (0, 1)$ is the decay factor.

Example - EWMA

- Suppose that $\lambda = 0.95$
- and that the estimate of the correlation between two variables X and Y on day $n-1$ is 0.6.
- Suppose that the estimate of the volatilities for X and Y on day $n-1$ are 1% and 2%, respectively.
- Suppose that the percentage changes in X and Y on day $n-1$ are 0.5% and 2.5%, respectively.
- (Q) Find the correlation estimate using EWMA on day n.

Example - EWMA (continued)

- $\text{cov}_{n-1} = 0.6 \times 0.01 \times 0.02 = 0.00012$
- $\sigma^2_{x,n} = 0.95 \times 0.01^2 + 0.05 \times 0.005^2 = 0.00009625$
- $\sigma^2_{y,n} = 0.95 \times 0.02^2 + 0.05 \times 0.025^2 = 0.00041125$
- $\text{cov}_n = 0.95 \times 0.00012 + 0.05 \times 0.005 \times 0.025 = 0.00012025$
- The new volatility of X is $\sqrt{0.00009625} = 0.981\%$, and the new volatility of Y is $\sqrt{0.00041125} = 2.028\%$.
- The new correlation between X and Y is:

$$\frac{0.00012025}{0.00981 \times 0.02028} = 0.6044$$

GARCH

- Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models are used to estimate time-varying volatility.
- GARCH(1,1) model:

$$\rho_n = \frac{\mathrm{cov}_n}{\sigma_{x,n}\sigma_{y,n}}$$

where

$$\mathrm{cov}_n = \omega_1 + \alpha_1 x_{n-1} y_{n-1} + \beta_1 \mathrm{cov}_{n-1}$$
$$\sigma_{x,n}^2 = \omega_2 + \alpha_2 x_{n-1}^2 + \beta_2 \sigma_{x,n-1}^2$$
$$\sigma_{y,n}^2 = \omega_3 + \alpha_3 y_{n-1}^2 + \beta_3 \sigma_{y,n-1}^2$$

- Parameters $\omega_i, \alpha_i, \beta_i$ are estimated from historical data.

Conditional Probability

Definition (Conditional Probability)

For events A and B with $P(B) > 0$, the conditional probability is:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

- Represents the probability of A occurring given that B has occurred
- $P(\cdot|B)$ is a probability measure on the reduced sample space
- Key properties: $P(A|B) \geq 0$, $P(\Omega|B) = 1$, additivity

Conditional Distribution

Definition (Conditional CDF)

For random variables X and Y, the conditional CDF of Y given $X = x$ is:

$$F_{Y|X}(y|x) = P(Y \leq y|X = x)$$

- For continuous variables, we use conditional density:

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

  where $f_X(x) > 0$
- Conditional distribution captures how Y behaves when we know X

Conditional Expectation

Definition (Conditional Expectation)

The conditional expectation of continuous Y given $X = x$ is:

$$E[Y|X = x] = \int y \, f_{Y|X}(y|x) \, dy$$

For discrete Y:

$$E[Y|X = x] = \sum_{y} y \, P(Y = y|X = x)$$

More generally, $E[Y|X]$ is a random variable that equals $E[Y|X = x]$ when $X = x$.

Example: Discrete Case

Let X and Y be discrete random variables with the following joint distribution:

x	y	$P(X = x, Y = y)$
0	0	0.1
0	1	0.2
1	0	0.3
1	1	0.4

- Conditional Distribution of Y given $X = 0$?
- Conditional Expectation $\mathbb{E}[Y \mid X = 0]$?

Example (continued)

We compute:

$$P(Y = y \mid X = 0) = \frac{P(X = 0, Y = y)}{P(X = 0)}$$

Then:

$$P(Y = 0 \mid X = 0) = \frac{0.1}{0.3} = \frac{1}{3}, \quad P(Y = 1 \mid X = 0) = \frac{0.2}{0.3} = \frac{2}{3}$$

$$\mathbb{E}[Y \mid X = 0] = \sum_{y} y \cdot P(Y = y \mid X = 0) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}$$

Properties of Conditional Expectation

- **Law of Total Expectation:** $E[Y] = E[E[Y|X]]$
- **Linearity:** $E[aY + bZ|X] = aE[Y|X] + bE[Z|X]$
- **Independence:** If $Y \perp X$, then $E[Y|X] = E[Y]$
- **Measurability:** $E[g(X)Y|X] = g(X)E[Y|X]$
- **Tower Property:** $E[E[Y|X, Z]|X] = E[Y|X]$

Financial Example: Conditional Expectation

Example (Stock Return Prediction)

Let S_t be stock price and M_t be market return. Model:

$$R_{t+1} = \alpha + \beta M_t + \varepsilon_{t+1}$$

where $\varepsilon_{t+1} \perp M_t$ with $E[\varepsilon_{t+1}] = 0$.

Conditional expectation:

$$E[R_{t+1} | M_t] = \alpha + \beta M_t$$

Interpretation:

- Given today's market return M_t, expected stock return is $\alpha + \beta M_t$
- $\beta > 1$: stock is more volatile than market (high beta)
- $\beta < 1$: stock is less volatile than market (low beta)
- This is the foundation of the Capital Asset Pricing Model (CAPM)

Multivariate Normal Distribution

- A random vector $\mathbf{X} = (X_1, X_2, \ldots, X_k)^T$ is said to follow a multivariate normal distribution if every linear combination of its components is normally distributed.
- Notation: $\mathbf{X} \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where:
  - $\boldsymbol{\mu} \in \mathbb{R}^k$ is the mean vector, i.e., $\boldsymbol{\mu} = E[\mathbf{X}]$
  - $\boldsymbol{\Sigma} \in \mathbb{R}^{k \times k}$ is the covariance matrix, i.e., $\boldsymbol{\Sigma}_{ij} = \mathrm{Cov}(X_i, X_j)$

Properties of Multivariate Normal Distribution

- If $\mathbf{X} \sim \mathcal{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then for any vector $\mathbf{a} \in \mathbb{R}^k$:
  - The linear combination $Y = \mathbf{a}^T\mathbf{X}$ is normally distributed:

$$Y \sim \mathcal{N}(\mathbf{a}^T\boldsymbol{\mu}, \mathbf{a}^T\boldsymbol{\Sigma}\mathbf{a})$$

- Marginal distributions are also normal:
  - The marginal distribution of any subset of components of $\mathbf{X}$ is multivariate normal.

Application: One factor model

- One factor model:

$$U_i = a_i F + \sqrt{1 - a_i^2} Z_i$$

  where $F, Z_i \sim \mathcal{N}(0, 1)$ independent.
- F is the common factor, Z_i are idiosyncratic noises.
- It models a multivariate normal $(U_1, U_2, \dots, U_N) \sim \mathcal{N}(0, \Sigma)$ with

$$\Sigma_{ii} = 1, \Sigma_{ij} = a_i a_j, i \neq j$$

- To determine the model, we only need to estimate a_i's, instead of Σ_{ij}'s.