Probability Concepts in Financial Engineering

1. Cumulative Distribution Function (CDF)

Definition:

The CDF of a random variable X, denoted F(x), gives the probability that X takes a value less than or equal to x.

F(x) = P(X <= x)

2. Probability Mass Function (PMF)

Definition:

Applicable to discrete random variables. The PMF p(x) gives the probability that X equals a specific value x.

Properties:

- p(x) >= 0 for all x.
- Sum of all p(x) = 1

p(x) = P(X = x)

3. Expected Value (Mean)

Definition:

The expected value E[X] of a discrete random variable X is the probability-weighted average of all possible values.

Interpretation:

Represents the long-run average outcome if the experiment is repeated many times.

E[X] = Sum x * p(x)

4. Variance

Definition:

The variance measures the spread or dispersion of a random variable around its mean.

Alternative Formula:

Var(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2

5. Binomial Distribution

Definition:

Describes the number of successes in n independent Bernoulli trials, each with success probability p.

Parameters:

- n: Number of trials
- p: Probability of success

Applications in Finance:

Used in risk modeling, option pricing, and evaluating fund manager performance.

P(X = k) = C(n, k) * p^k * (1 - p)^(n - k)

E[X] = n * p, Var(X) = n * p * (1 - p)

6. Poisson Distribution

Definition:

Models the number of events occurring in a fixed interval of time or space, with events occurring independently and at a constant average rate lambda.

Parameter:

- lambda: Average number of events

Applications in Finance:

Used in modeling default events, claims arrivals, and market jumps.

P(X = k) = (lambda^k * e^(-lambda)) / k!

E[X] = lambda, Var(X) = lambda

Conclusion

These probability tools form the mathematical backbone of many financial models, including:

- Pricing of fixed income securities
- Valuation of derivatives
- Risk management and portfolio theory

Understanding these concepts is essential for modeling uncertainty and making quantitative financial decisions.

7. Bayes' Theorem

Bayes' Theorem relates the conditional probability of two events A and B. It states that the probability of A given B is equal to the probability of A and B occurring together divided by the probability of B.

Key Points:

- Conditional Probability:

P(A|B) = P(A ? B) / P(B)

Alternative Expression:

The joint probability P(A ? B) can be rewritten using conditional probability as:

P(A ? B) = P(B|A) * P(A)

Partition of Sample Space:

If the event B can result from a set of mutually exclusive and exhaustive events A1, A2, ..., An, then the denominator P(B) can be expressed as:

P(B) = Sum P(B|Ai) * P(Ai)

Example:

Consider two six-sided dice. What is the probability that one die is greater than or equal to 4 given that the sum of the dice is greater than or equal to 8?

8. Continuous Random Variables and PDFs

For continuous random variables, we define a Probability Density Function (PDF) f(x), where probabilities are calculated using integrals.

Probability between a and b:

P(a <= X <= b) = ?[a to b] f(x) dx

9. Normal and Log-Normal Distributions

Normal Distribution:

Symmetric, bell-shaped curve. Widely used in finance for modeling returns.

PDF of Normal Distribution:

f(x) = (1 / (??(2?))) * exp(-(x - ?)^2 / (2?^2))

Log-Normal Distribution:

Used to model stock prices and other financial quantities that cannot be negative and are multiplicative in nature.

10. Conditional Expectation Identity

The expected value of a random variable X can be calculated using another variable Y by taking the expected value of the conditional expectation:

Key Idea:

The law of total expectation allows the decomposition of E[X] into:

E[X] = E[E[X | Y]]

11. Conditional Variance Identity

The variance of a random variable X can be decomposed into two parts:

1. The variance of the conditional expectation

2. The expected value of the conditional variance

This is known as the law of total variance:

$Var(X) = Var(E[X | Y]) + E[Var(X | Y)]$

Note:

Both E[X | Y] and Var(X | Y) are functions of Y and hence are themselves random variables.

Example:

A random sum of random variables is considered, where E[Sum] and Var(Sum) are evaluated using the above identities.

12. Indicator Functions

An indicator function I(A) takes the value 1 if event A occurs and 0 otherwise.

They are especially useful for:

- Simplifying expected value calculations
- Counting the number of successful events in a random experiment

Example:

Used to model the number of eggs that hatch in a biological experiment.

I(A) = 1 if A occurs, 0 otherwise

13. Joint Cumulative Distribution Function (Joint CDF)

Defines the probability that multiple random variables X and Y are less than or equal to certain values x and y:

F(x, y) = P(X <= x, Y <= y)

14. Marginal CDF

The marginal CDF of a variable is derived from the joint CDF by fixing one variable and integrating or evaluating the other to infinity.

Example:

Marginal CDF of X:

F_X(x) = lim (y ? infinity) F(x, y)

15. Joint Probability Density Function (Joint PDF)

The joint PDF is the derivative of the joint CDF with respect to all variables:

f(x, y) = ?^2F(x, y) / (?x ?y)

16. Conditional CDF and PDF

These define the probability distribution of one variable given the value of another.

Conditional PDF of Y given X = x:

f(y | x) = f(x, y) / f_X(x)

17. Independence

Two random variables X and Y are independent if:

F(x, y) = F_X(x) * F_Y(y)

Or equivalently:

f(x, y) = f_X(x) * f_Y(y)

18. Mean Vector and Covariance Matrix

Mean Vector:

A vector of expected values for each variable:

? = [E[X], E[Y], ...]

Covariance Matrix:

Represents how two or more variables vary together.

If X = [X?, X?, ..., Xn]?, then:

Cov(X) = E[(X - ?)(X - ?)?]

19. Multivariate Normal Distribution

An n-dimensional vector X is said to be multivariate normal if every linear combination of its components is normally distributed.

It is defined by a mean vector ? and a covariance matrix Sum.

Probability Density Function (PDF):

f_X(x) = (1 / ((2?)^(n/2) * |Sum|^(1/2))) * exp(-0.5 * (x - ?)? Sum?^{1} (x - ?))

20. Standard Multivariate Normal

A special case of the multivariate normal distribution with mean vector 0 and identity covariance matrix I.

This implies that the variables are independent and identically distributed.

? = 0, Sum = I

21. Moment Generating Function (MGF)

The MGF of a multivariate normal distribution extends the univariate case:

It provides insights into the distribution's expected values and moments.

M_X(t) = exp(t?? + 0.5 * t?Sumt)

22. Marginal and Conditional Distributions

Marginal Distributions:

Each subset of the components of X is also normally distributed.

Conditional Distributions:

The conditional distribution of a subset of X given others is also multivariate normal.

23. Visualization and Interpretation

The relationships between variables in a multivariate normal can be visualized via ellipses or contours.

Knowing one variable provides information about the expected value and uncertainty (variance) of others, depending on the correlation structure encoded in Sum.

24. Martingales

Martingales are a class of stochastic processes that represent 'fair games' in probability theory.

A random process (X?) is a martingale if:

1. E[|X?|] < infinity for all n

2. E[X??? | ??] = X? for all n and m >= 0

Where ?? denotes the information available up to time n.

25. Information Filtration (??)

Information filtration represents the growing collection of information over time.

?? typically includes all prior observed values up to time n, i.e., {X?, X?, ..., X?}.

?? = ?(X?, X?, ..., X?)

26. Martingale Examples

Example 1: Random Walk

Let M? = ????? Y?, where Y? are IID with E[Y?] = 0.

Then M? is a martingale.

Example 2: Martingale Betting Strategy

A strategy where bets are doubled after each loss until a win.

The total winnings W? after n plays is a martingale.

27. Martingale Betting Winnings (W?)

In a Martingale betting scenario, the values of W? (total winnings after n bets) are:

- If the first win occurs on the nth bet: W? = 1

- If no win occurs in n bets: W? = -2??¹

This reflects the cumulative losses from all previous doubled bets.

To verify W? is a martingale, it must satisfy:

E[W??? | ??] = W?

This is shown using iterated expectations and the fairness of the betting setup.

E[W??? | ??] = W?

28. Martingale Cases in Coin Toss Game

In a coin toss betting game, two cases are considered to demonstrate that W? is a martingale:

Case 1: If W? = 1, the game stops because a win has occurred. The expected value of W??? given W? = 1 is also 1, confirming that W? is a martingale.

Case 2: If W? = -2??^{1}, the next value W??? can be either 1 (if the next toss is a win) or -2??^{1} + 1 (if the toss is a loss). The expected value in this case is also equal to W?, reinforcing the martingale property.

E[W??? | ??] = W?

29. Polya's Urn Model as a Martingale

Polya's Urn is an example of a martingale where balls are drawn from an urn containing red and green balls. After each draw, the color of the drawn ball is added back to the urn, which changes the composition. The ratio of red balls to total balls forms a martingale.

The martingale property holds because the expected ratio of red balls to total balls after each draw is equal to the current ratio.

E[R??? / T??? | ??] = R? / T?

30. Brownian Motion

Brownian motion (or Wiener process) is a key stochastic process in finance, especially in the Black-Scholes framework. It is defined as a random process (x?) with parameters ? (drift) and ? (volatility) that satisfies the following conditions:

1. Increments (x?? - x??, x?? - x??, ...) are mutually independent.

2. For s > 0, the increment (x??? - x?) follows a normal distribution with mean ?s and variance ?^2s.

3. The function x? is continuous over time.

31. Historical Context

Brownian motion was introduced into finance by Louis Bachelier in 1900 to model stock price movements. It was later formalized in the 1920s by Norbert Wiener, who introduced the Wiener process as the mathematical foundation.

32. Standard Brownian Motion

When the drift ? = 0 and volatility ? = 1, the process is called standard Brownian motion, denoted as W?. It represents a Brownian motion with no drift and unit volatility, which is often used as a benchmark in financial models.

W? = B? with ? = 0 and ? = 1

33. Independent Increments Property

The increments of Brownian motion are independent over non-overlapping intervals. This means that the future increments of the process are independent of the information available up to the current time.

E[(x??? - x?) | F?] = 0 for s > 0

34. Sample Paths

Sample paths of Brownian motion are continuous but exhibit jagged behavior. This means the process is continuous in time but exhibits random fluctuations, making it a powerful model for financial data.

35. Information Filtration

At any time t, the information available is denoted by F?. The increments of Brownian motion are independent of the information available at time t, reflecting the 'memoryless' nature of the process.

F? = ?(x?, x?, ..., x?)

36. Geometric Brownian Motion (GBM)

Geometric Brownian Motion (GBM) is a key stochastic process used in finance, particularly for modeling stock prices and other financial assets. It is defined as:

X? = e^{(? - ?^2/2)t + ?W?}

Where:

- W? is standard Brownian motion
- ? is the drift (expected return)
- ? is the volatility (standard deviation)

X? = e^{(? - ?^2/2)t + ?W?}

37. Properties of Brownian Motion

The process W? (standard Brownian motion) is normally distributed with mean 0 and variance t. This gives us the key feature that increments are normally distributed and independent over time.

W? ~ N(0, t)

38. Simulation of GBM

GBM can be simulated at various time points by using random variables, making it particularly useful for modeling asset prices and other processes that evolve over time. The process can be discretized and simulated for discrete time steps.

X? = X??? * exp[(? - ?^2/2)?t + ???t * Z?]

39. Expected Value of GBM

The expected value of X??? given X? is derived using the moment-generating function of a normal distribution. This leads to the result that the expected growth rate of GBM is equal to ?.

E[X??? | X?] = X? * e^{?s}

40. Path Characteristics of GBM

The paths of GBM are continuous but jagged. This means that while the process is continuous in time, it exhibits random fluctuations without jumps, which is a crucial feature in financial modeling.

41. Key Properties of GBM

The following are important properties of Geometric Brownian Motion:

- Ratios of GBM values at different times are independent.
- If X? > 0, then X??? > 0 for all s > 0.
- The distribution of returns depends only on the time increment (s), not on the current value (X?).

42. Application: GBM in Black-Scholes Model

GBM serves as the underlying model for the Black-Scholes option pricing formula, which is used for pricing European call and put options. The assumption of GBM helps in modeling the continuous price evolution of financial assets.

43. Vectors and Their Fundamental Concepts

Vectors are fundamental in linear algebra. They are collections of real numbers that can be arranged as row vectors or column vectors. The notation R? indicates a vector with n components, representing a point or direction in an n-dimensional space.

44. Examples of Vectors

Vectors can be represented in 2D space (R?) with components corresponding to the x and y axes. For instance, the vector $v = (3, 4)$ can represent a point in the 2D plane with coordinates (3, 4).

v = (x?, x?)

45. Linear Combinations

A linear combination of vectors involves multiplying them by real numbers (scalars) and adding them together. For vectors v?, v?, ..., v? and scalars a?, a?, ..., a?, the linear combination is defined as:

a?v? + a?v? + ... + a?v?.

c = a?v? + a?v? + ... + a?v?

46. Linear Independence

Two vectors are linearly independent if one cannot be written as a scalar multiple of the other. This means that no vector in the set is a linear combination of the others.

a?v? + a?v? = 0, if and only if a? = a? = 0

47. Basis

A basis for a vector space is a set of linearly independent vectors that span the entire space. For R?, a basis consists of exactly n elements, and every vector in the space can be expressed as a linear combination of the basis vectors.

v = c?v? + c?v? + ... + c?v?

48. Standard Basis

The standard basis in R? consists of vectors with one component equal to one and all others equal to zero. These vectors serve as building blocks to represent any vector in R?.

e? = (1, 0, 0, ..., 0), e? = (0, 1, 0, ..., 0), ..., e? = (0, 0, ..., 1)

49. Vector Length and Norms

In this section, we discuss the concept of vector length using two commonly used norms: the l? norm and the l? norm. The length of a vector is important in many areas, including distance measurement and optimization in finance.

50. Vector Length Calculation

For a vector in R^2 with components (4, 3), the length is calculated using the l? norm formula:

Length = ?(4^2 + 3^2) = 5.

Length = ?(x?^2 + x?^2 + ... + x?^2)

51. Properties of Length

The length of a vector has several important properties:

- Length is always >= 0.
- A vector with length 0 is the zero vector.
- Scaling a vector by a factor (positive or negative) scales its length by the absolute value of that factor.
- The triangle inequality states that the length of the sum of two vectors is less than or equal to the sum of their lengths.

52. l? Norm (Euclidean Distance)

The l? norm is the standard way of calculating the length of a vector. It is commonly used for measuring Euclidean distance in various applications, including in finance for portfolio optimization and risk management.

l? Norm = ?(x?^2 + x?^2 + ... + x?^2)

53. l? Norm (Manhattan Distance)

The l? norm measures the distance between two points by summing the absolute differences of their coordinates, resulting in a Manhattan or taxicab distance. For example, for points (4, 3) and (1, 2), the distance would be:

|4 - 1| + |3 - 2| = 3 + 1 = 4.

l? Norm = |x? - y?| + |x? - y?| + ... + |x? - y?|

54. Inner Product (Dot Product)

The dot product of two vectors relates to the cosine of the angle between them. It is calculated by multiplying corresponding components of the vectors and summing the results:

v · w = v?w? + v?w? + ... + v?w?.

v · w = v?w? + v?w? + ... + v?w?

55. Angle Calculation between Vectors

The cosine of the angle ? between two vectors v and w can be calculated using the formula:

cos(?) = (v · w) / (||v|| ||w||), where ||v|| and ||w|| are the lengths of vectors v and w, respectively.

cos(?) = (v · w) / (||v|| ||w||)

56. Matrices and Operations

This section introduces matrices and their operations. Matrices are fundamental in linear algebra and are widely used in finance for portfolio optimization, risk management, and pricing models.

57. Definition of a Matrix

A matrix is a rectangular array of real numbers, defined by its number of rows (m) and columns (n). For example, a matrix with 2 rows and 3 columns is called a 2 by 3 matrix.

58. Matrix Elements

Matrix elements are indexed by their row and column positions. For example, the element in the first row and second column of matrix A is denoted as A??.

A?? = element at row i, column j

59. Types of Vectors

There are two types of vectors: row vectors (1 row, multiple columns) and column vectors (1 column, multiple rows). These vectors are used to represent quantities in matrix equations.

60. Identity Matrix

The identity matrix is a square matrix with 1s on the diagonal and 0s elsewhere. It acts as the multiplicative identity in matrix multiplication, meaning any matrix multiplied by the identity matrix remains unchanged.

I = [1 0 0; 0 1 0; 0 0 1]

61. Transpose Operation

The transpose of a matrix is obtained by switching its rows and columns. For example, the transpose of a column vector becomes a row vector.

A? = transpose of matrix A

62. Matrix Multiplication

Matrix multiplication requires that the number of columns in the first matrix equals the number of rows in the second matrix. The resulting matrix's dimensions are determined by the outer dimensions of the two matrices.

$A\ (m \times n) \times B\ (n \times p) = C\ (m \times p)$

63. Inner Product

The inner product of two vectors can be expressed using matrix multiplication. For example, if v is a row vector and w is a column vector, the inner product is given by v · w = v?w.

v · w = v?w

64. Connection to Linear Functions

Matrices and vectors are often used to express linear functions. In the next modules, we will explore how matrices and vectors relate to linear transformations and systems of equations.

65. Linear Functions and Properties

This section introduces the concept of linear functions, focusing on their properties and relationships with vectors and matrices. Linear functions are essential in finance for modeling transformations and optimizing portfolios.

66. Definition of a Linear Function

A function is considered linear if it satisfies the property that for any vectors x and y, and real numbers ? and ?, the function evaluated at the linear combination (?x + ?y) equals the same linear combination of the function evaluations: f(?x + ?y) = ?f(x) + ?f(y).

f(?x + ?y) = ?f(x) + ?f(y)

67. Linear Functions and Matrices

A linear function can be represented as the multiplication of a vector by a matrix, denoted as $f(x) = Ax$, where A is the matrix and x is the input vector.

$f(x) = Ax$

68. Rank of a Matrix

The rank of a matrix indicates the number of linearly independent columns (column rank) and rows (row rank). For any matrix, the row rank and column rank are equal.

Rank(A) = Number of linearly independent rows or columns

69. Range of a Matrix

The range of a matrix describes the set of all possible output vectors that can be generated by multiplying the matrix by different input vectors.

Range(A) = {Ax | x ? R^n}

70. Invertibility of a Matrix

A square matrix is invertible if its rank equals the number of its rows (or columns), meaning all rows and columns are linearly independent. If a matrix is invertible, it has a unique inverse.

$A?^{1}$ = Inverse of matrix A, if Rank(A) = n

71. Linear Optimization through Hedging

This section covers the key concepts of linear optimization, focusing on a hedging example in financial markets. Linear optimization is crucial in finance for portfolio optimization and risk management.

72. Types of Optimization Problems

The module introduces primal and dual optimization problems and their relationship. Primal problems involve minimizing or maximizing a linear objective subject to constraints, while dual problems provide bounds on the primal solution.

73. Hedging Problem

The hedging problem involves managing multiple assets with uncertain future prices across different states. The objective is to minimize the cost of hedging while ensuring that the portfolio meets the required payoffs.

74. State Representation

Prices can be represented in two ways: by state (where all asset prices in a state are listed) or by asset (where the price of a specific asset across states is shown). This dual representation helps in constructing the optimization problem.

75. Matrix Representation

A matrix is used to represent asset prices, with rows indicating states and columns indicating assets. The matrix representation simplifies the analysis of price structures in different states.

76. Obligation and Portfolio

An obligation vector depends on the state, indicating the required payoff in each state. A portfolio is then chosen to hedge this obligation, meaning the portfolio?s value in each state should meet or exceed the obligation.

77. Linear Optimization Problem

The goal of the linear optimization problem is to minimize the cost of the portfolio while ensuring that the payoff in each state meets or exceeds the obligation. This forms a linear program with objective functions and constraints.

78. Importance of Price Matrix Rank

The rank of the price matrix determines whether all payoffs can be generated by the portfolio. This leads to concepts of complete and incomplete markets, where a complete market allows all payoffs to be generated, while an incomplete market restricts certain payoffs.