PROBABILITY & STATISTICS

for Quantitative Finance

Core Mathematical Foundations

Distributions · Inference · Risk Measures

Probability Fundamentals

Core Concepts

Probability quantifies uncertainty in financial markets and asset returns.

Joint and Marginal Distributions:

Joint PDF: $f(x, y)$ with $\int \int f(x, y)\, dx\, dy = 1$

Marginal PDF: $f_X(x) = \int f(x, y)\, dy$

Conditional PDF: $f(y|x) = \frac{f(x,y)}{f_X(x)}$

Independence: $f(x, y) = f_X(x) f_Y(y)$

Multivariate Normal:

$$\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

Finance: Model correlated asset returns in portfolio.

Bayes' Theorem and Applications

Bayes' Theorem:

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

Finance Application:

Update market beliefs given new information:

Prior belief: $P(\text{Recession})$

New data arrives: Unemployment report

Posterior belief: $P(\text{Recession}|\text{Data}) = \frac{P(\text{Data}|\text{Recession}) \cdot P(\text{Recession})}{P(\text{Data})}$

Random Variables: Core Concepts

Definition

Random variables map outcomes to numerical values, modeling asset returns and prices.

Discrete Random Variables:

Probability Mass Function: $P(X = x_i) = p_i$ with $\sum_i p_i = 1$

Cumulative Distribution Function: $F(x) = P(X \leq x)$

Continuous Random Variables:

Probability Density Function: $P(a \leq X \leq b) = \int_a^b f(x)\,dx$

Normalization: $\int_{-\infty}^{\infty} f(x)\,dx = 1$

CDF: $F(x) = \int_{-\infty}^{x} f(t)\,dt$ and $f(x) = F'(x)$

Expected Value and Variance

Expected Value:

Discrete: $\mathbb{E}[X] = \sum_i x_i p_i$
Continuous: $\mathbb{E}[X] = \int_{-\infty}^{\infty} xf(x)\,dx$
Linearity: $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$ and $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$

Finance: Expected return on portfolio is weighted average of asset returns.

Variance:

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

Properties: $\mathrm{Var}(aX + b) = a^2\mathrm{Var}(X)$
Standard Deviation: $\sigma = \sqrt{\mathrm{Var}(X)}$

Finance: Volatility of asset returns, portfolio risk measurement.

Covariance and Correlation

Definition

Covariance and correlation measure the relationship between two random variables.

Covariance:

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Correlation:

$$\rho = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

Properties: $-1 \leq \rho \leq 1$

$\rho = 1$: Perfect positive correlation

$\rho = -1$: Perfect negative correlation

$\rho = 0$: Uncorrelated

Portfolio Variance and Diversification

Two-Asset Portfolio Variance:

$$\mathrm{Var}(w_1 R_1 + w_2 R_2) = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1 w_2 \rho\sigma_1\sigma_2$$

Key Insight:

Correlation drives diversification benefits.

If $\rho < 1$, portfolio variance is less than weighted average of individual variances.

If $\rho = -1$, can construct zero-variance portfolio.

Higher Moments:

Skewness: $\mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$ measures asymmetry, tail risk.

Kurtosis: $\mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^4\right]$ measures fat tails, extreme events.

Normal Distribution

Definition

The normal distribution is fundamental to financial modeling and option pricing.

Probability Density Function:

$$X \sim N(\mu, \sigma^2)$$
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Properties:

Mean: $\mathbb{E}[X] = \mu$ and Variance: $\mathrm{Var}(X) = \sigma^2$

Symmetric around mean

68-95-99.7 rule for standard deviations

Standard Normal: $Z \sim N(0, 1)$ with $Z = \frac{X-\mu}{\sigma}$

Lognormal Distribution and Stock Prices

Lognormal Distribution:

If $\ln(X) \sim N(\mu, \sigma^2)$, then $X \sim \text{Lognormal}(\mu, \sigma^2)$

$$\mathbb{E}[X] = e^{\mu+\sigma^2/2}$$

$$\text{Var}(X) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

Finance Application:

Stock prices follow lognormal distribution under Geometric Brownian Motion.

If log returns are normal: $\ln(S_t/S_0) \sim N(\mu, \sigma^2)$

Then stock price: $S_t \sim \text{Lognormal}$

Key property: Prices cannot go negative, which matches market reality.

Black-Scholes model assumes lognormal price distribution.

Key Distributions in Finance

Exponential Distribution:

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0$$

$$\mathbb{E}[X] = \frac{1}{\lambda}, \quad \mathrm{Var}(X) = \frac{1}{\lambda^2}$$

Memoryless property: $P(X > s + t | X > s) = P(X > t)$

Finance: Time between trades, default times, jump arrivals in Poisson processes.

Poisson Distribution:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

$$\mathbb{E}[X] = \mathrm{Var}(X) = \lambda$$

Finance: Number of defaults in credit portfolio, trade arrivals, jump events.

Student's t and Other Distributions

Student's t-Distribution:

Heavier tails than normal distribution, controlled by degrees of freedom ν.

As $\nu \to \infty$, converges to normal distribution.

Finance: Better fit for asset returns with fat tails and extreme events. Used in risk models when normal distribution underestimates tail risk.

Binomial Distribution:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$\mathbb{E}[X] = np, \quad \mathrm{Var}(X) = np(1 - p)$$

Finance: Binomial option pricing model, number of defaults in portfolio.

Limit Theorems

Foundation of Statistical Inference

Limit theorems justify using normal approximations in many financial applications.

Law of Large Numbers:

Sample mean converges to expected value:

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{n\to\infty} \mathbb{E}[X]$$

Finance: Long-run average returns approach expected return.

Central Limit Theorem:

For independent X_i with $\mathbb{E}[X_i] = \mu$ and $\text{Var}(X_i) = \sigma^2$:

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} N(0,1)$$

Applications of Central Limit Theorem

Portfolio Returns:

Sum of many independent returns approximately normal, even if individual returns are not normal.

Justifies using normal distribution for portfolio analysis.

Risk Management:

Allows construction of confidence intervals for portfolio returns.

Statistical Testing:

Enables hypothesis tests on sample means and portfolio performance metrics.

Limitation:

CLT requires independence and finite variance. Financial returns often exhibit dependence and fat tails, so normal approximation may fail during crises.

Statistical Inference: Estimation

Definition

Statistical inference uses sample data to make conclusions about population parameters.

Point Estimation:

Sample Mean: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ estimates μ

Sample Variance: $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ estimates σ^2

Confidence Intervals:

For mean with known variance:

$$\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

For mean with unknown variance:

$$\bar{X} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$$

Hypothesis Testing

Framework:

Null Hypothesis H_0: Default assumption to test against

Alternative H_1: What we want to show evidence for

Test Statistic: Measures how far data deviates from H_0

p-value: Probability of observing data at least as extreme if H_0 is true

Decision Rule:

If $p < \alpha$: Reject H_0 (significant result)

If $p \geq \alpha$: Fail to reject H_0 (insufficient evidence)

Finance Application:

Test if trading strategy alpha is significantly different from zero.

H_0: $\alpha = 0$ (no skill) vs H_1: $\alpha \neq 0$ (generates alpha)

Regression Analysis

Definition

Regression models the relationship between dependent and independent variables.

Simple Linear Regression:

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

Ordinary Least Squares estimates:

$$\hat{\beta}_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

R^2: Proportion of variance explained by model

Finance Application - CAPM:

Multiple Regression and Factor Models

Multiple Regression:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon$$

Matrix form: $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$

OLS solution: $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$

Fama-French Three-Factor Model:

$$R_i - r_f = \alpha + \beta_1(R_m - r_f) + \beta_2 \text{SMB} + \beta_3 \text{HML} + \epsilon$$

SMB: Small minus big (size factor)

HML: High minus low (value factor)

Factor loadings explain cross-sectional variation in returns.

Time Series Analysis

Definition

Time series models capture temporal dependence in financial data.

Stationarity:

Weak stationarity requires constant mean, constant variance, and autocovariance depending only on lag.

Asset returns are typically stationary, but prices are not.

Autoregressive Model AR(1):

$$X_t = \phi X_{t-1} + \epsilon_t$$

Stationarity requires $|\phi| < 1$

Finance: Models mean reversion in returns, interest rates, volatility.

GARCH Models

Volatility Clustering:

Asset returns exhibit periods of high and low volatility.

Variance is not constant over time.

GARCH(1,1) Model:

$$r_t = \mu + \epsilon_t, \quad \epsilon_t = \sigma_t z_t, \quad z_t \sim N(0,1)$$
$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2$$

Conditional variance depends on past squared errors and past variance.

Finance: Forecasting volatility for option pricing, risk management, portfolio allocation.

Captures time-varying volatility better than constant variance models.

Value at Risk

Definition

VaR measures the maximum expected loss over a time horizon at a given confidence level.

Mathematical Definition:

$$P(\text{Loss} > \text{VaR}_\alpha) = 1 - \alpha$$

95% VaR: Maximum loss exceeded 5% of the time

Parametric VaR:

Assume returns are normal: $R \sim N(\mu, \sigma^2)$

$$\text{VaR}_\alpha = -(\mu + \sigma z_\alpha)$$

where z_α is the α-quantile of standard normal distribution.

Example: For 95% VaR, use $z_{0.05} = -1.645$

Conditional VaR and Other Risk Measures

Conditional VaR (Expected Shortfall):

$$\text{CVaR}_\alpha = \mathbb{E}[\text{Loss}|\text{Loss} > \text{VaR}_\alpha]$$

Average loss in the worst $(1 - \alpha)$

Better than VaR because it accounts for tail risk severity.

Sharpe Ratio:

$$\text{SR} = \frac{\mathbb{E}[R_p] - r_f}{\sigma_p}$$

Risk-adjusted return per unit of volatility. Higher is better.

Monte Carlo Simulation

Definition

Monte Carlo uses random sampling to compute expected values and price derivatives.

Basic Algorithm:

1. Generate N random paths for underlying asset price
2. Calculate payoff for each simulated path
3. Average payoffs and discount to present value

$$\hat{V} = e^{-rT} \frac{1}{N} \sum_{i=1}^{N} \text{Payoff}_i$$

Monte Carlo: Finance Applications

Option Pricing:

Price path-dependent options like Asian options, barrier options, lookback options.

Simulate stock price paths under risk-neutral measure using GBM.

Portfolio Risk:

Generate correlated asset returns using multivariate normal or copulas.

Compute VaR and CVaR from simulated portfolio returns.

Variance Reduction:

Antithetic variates: Use Z and $-Z$ to reduce variance

Control variates: Leverage known prices of similar securities

Importance sampling: Sample more frequently in relevant regions

Maximum Likelihood Estimation

Definition

MLE finds parameter values that maximize the probability of observing the data.

Likelihood Function:

For data $x_1, \ldots, x_n$ and parameter θ:

$$L(\theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

MLE Estimator:

$$\hat{\theta}_{MLE} = \arg\max_{\theta} \ell(\theta)$$

Solve: $\frac{\partial \ell(\theta)}{\partial \theta} = 0$

MLE in Finance

Example: Normal Distribution

Given returns $r_1, \ldots, r_n$ assumed $\sim N(\mu, \sigma^2)$:

$$\hat{\mu}_{MLE} = \frac{1}{n}\sum_{i=1}^{n} r_i = \bar{r}$$

$$\hat{\sigma}^2_{MLE} = \frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^2$$

Finance Applications:

Calibrate GARCH models, fit distributions to returns, estimate factor loadings.

Bootstrap Methods

Definition

Bootstrap resamples from observed data to estimate sampling distributions without parametric assumptions.

Basic Bootstrap Algorithm:

1. Draw n observations with replacement from original sample
2. Compute statistic of interest $\hat{\theta}^*$
3. Repeat steps 1-2 for B times (typically $B = 1000$ or more)
4. Use empirical distribution of $\{\hat{\theta}_1^*, \ldots, \hat{\theta}_B^*\}$ for inference

Bootstrap Confidence Interval:

Percentile method: Use $\alpha/2$ and $1 - \alpha/2$ quantiles of bootstrap distribution.

Bootstrap in Finance

Applications:

Sharpe ratio confidence intervals: Non-normal returns make analytical CI unreliable

VaR estimation: Bootstrap historical returns to estimate tail quantiles

Strategy backtesting: Resample returns to assess statistical significance

Block Bootstrap:

For time series with dependence, resample blocks of consecutive observations.

Preserves autocorrelation structure in returns.

Caution:

Bootstrap assumes sample is representative of population.

May fail during regime changes or structural breaks in financial markets.

Copulas and Dependence

Definition

Copulas model dependence structures separately from marginal distributions.

Sklar's Theorem:

Any joint distribution $F(x, y)$ can be written as:

$$F(x, y) = C(F_X(x), F_Y(y))$$

where C is a copula function on $[0, 1]^2$.

Key Insight:

Marginals F_X, F_Y and dependence structure C are modeled separately.

Allows flexible modeling: Normal margins with non-normal dependence.

Common Copulas in Finance

Gaussian Copula:

$$C(u, v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$$

where Φ_ρ is bivariate normal CDF with correlation ρ.

Used extensively in CDO pricing (pre-2008 crisis).

t-Copula:

Similar to Gaussian but with heavier tails and tail dependence.

Better captures joint extreme events in financial markets.

Clayton Copula:

Exhibits lower tail dependence: Assets more correlated in crashes.

Useful for modeling downside risk in portfolios.

Extreme Value Theory

Definition

EVT provides statistical tools to model rare but severe losses in financial markets.

Generalized Pareto Distribution:

Models exceedances over high threshold u:

$$F(x) = 1 - \left(1 + \xi \frac{x - u}{\beta}\right)^{-1/\xi}$$

where ξ is tail index, β is scale parameter.

Finance Applications:

Tail risk measurement: Estimate probability of extreme losses

Stress testing: Model losses beyond historical worst cases

Better CVaR estimates in fat-tailed distributions

Key Takeaways

Probability Foundations:

Joint distributions, Bayes' theorem, and multivariate models underpin portfolio theory.

Statistical Distributions:

Normal and lognormal for prices; t-distribution for fat tails; GARCH for volatility.

Risk Measures:

VaR and CVaR quantify downside risk; Sharpe ratio for risk-adjusted performance.

Advanced Methods:

Monte Carlo for derivatives pricing; MLE for parameter estimation; copulas for dependence; EVT for tail risk.

Critical Understanding:

Models are approximations. Markets exhibit fat tails, regime changes, and contagion effects that standard models may not fully capture.