CredGuild

CHEAT SHEET

CFA® Program Level I

Quantitative Methods Cheat Sheet

2027 curriculum · 11 readings · 133 formulas

credguild.com

Version v2027.3 · October 2026

ABOUT THIS EDITION

CredGuild · CFA® Program Level I · Quantitative Methods Cheat Sheet
2027 curriculum · Version v2027.3 · Generated October 6, 2026

SHARING

© 2026 CredGuild. You may share this file freely, complete and unchanged. Do not sell it, and do not republish it with the CredGuild name, links or watermark removed.

USING THIS BOOK

This sheet is a revision aid. Use it alongside the notes and the CFA Institute curriculum. The latest version is free at credguild.com.

FOUND A MISTAKE?

Email hello@credguild.com, or use "Report an issue with this question" under any question on credguild.com.

TRADEMARKS

CFA Institute does not endorse, promote, review, or warrant the accuracy or quality of the products or services offered by CredGuild. CFA®, Chartered Financial Analyst® and CFA Institute® are trademarks owned by CFA Institute.

Contents

Quantitative Methods **1**

1. Returns of Financial Assets and Instruments 1
2. Types of Financial Returns 2
3. Benchmarking Returns 4
4. The Time Value of Money in Finance 5
5. Statistical Characteristics of Asset Returns 7
6. Statistical Distributions for Financial Asset Prices and Returns 9
7. Estimation and Hypothesis Testing 10
8. The Return and Risk of a Financial Portfolio 12
9. Simulation of Financial Asset Prices and Returns 16
10. Applications of Simple Linear Regression in Finance 17
11. Introduction to Financial Data Science 19

Keep going at credguild.com **21**

The revision summary of the CredGuild notes for the 2027 CFA® Program Level I curriculum, Quantitative Methods: for each reading, the bottom line of every module, the exam shortcuts and the formulas. Use it in the last weeks before the exam, after you have studied the readings.

To learn a reading in full, open its notes at credguild.com. This sheet is free to print and share: credguild.com/cfa/level-1/cheat-sheet/

CFA® PROGRAM LEVEL I · CHEAT SHEET

Quantitative Methods

Exam weight 11–14% · 11 readings · 133 formulas

1 Returns of Financial Assets and Instruments

MODULE 1.1 · RATES OF RETURN

- Total (holding period) return is $(P_1 - P_0 + D)/P_0$, the price return plus the capital distribution return, with each component divided by the beginning price P_0.
- An ex post (actual) return uses prices and income already observed, while an ex ante (expected) return is an estimate for a future period.
- The geometric mean return $[(1 + R_1)(1 + R_2) \cdots (1 + R_n)]^{1/n} - 1$ is the measure for an annualized return or a compound annual growth rate; it cannot exceed the arithmetic mean, and the two are equal only when every period's return is identical.
- A holding period return is annualized with $(1 + HPR)^{365/\text{days}} - 1$, or $(1 + R_{\text{period}})^{c} - 1$ with c periods in a year; for a period shorter than a year this assumes the same return is earned in every remaining period.
- For a given positive stated annual rate, more frequent compounding gives a higher effective annual rate, $(1 + r_s/m)^m - 1$, and continuous compounding gives the highest, $e^r - 1$.
- The continuously compounded return is $r_{cc} = \ln(1 + HPR)$, which reduces to $\ln(P_1/P_0)$ only when no distributions are paid, and log returns are additive across periods.

EXAM SHORTCUTS

- The direction of an annualized return is known before calculating: a positive HPR over more than a year annualizes to a smaller figure and over less than a year to a larger one, while a negative HPR (between −100% and 0) becomes less negative over more than a year and more negative over less than a year.
- The geometric mean return cannot exceed the arithmetic mean return of the same returns, so an answer with a higher geometric mean can be ruled out.
- A log return is below the HPR for a gain and more negative than the HPR for a loss, which rules out answers on the wrong side of the HPR.

MODULE 1.2 · COMPONENTS AND MEASURES OF RETURN

- An interest rate can be read as a required rate of return, a discount rate or an opportunity cost of current consumption; a measure of risk is not one of these interpretations.
- Required return = real risk-free rate + expected inflation + risk premia, where the real risk-free rate plus expected inflation is the nominal risk-free rate, such as a short-term government bill yield.
- The real risk-free rate reflects only time preference and is linked to the nominal rate exactly by $(1 + r_{\text{nominal}}) = (1 + r_{\text{real}})(1 + \pi^e)$, or approximately by $r_{\text{real}} \approx r_{\text{nominal}} - \pi^e$.
- Gross return deducts trading commissions and other costs of generating the return but not management and administration fees; net return deducts both.
- With one tax rate on capital gains and another on distributions, the after-tax nominal return is price return $\times(1 - t_{cg})$ + distribution return $\times(1 - t_d)$.

- The leveraged return is $r_L = r_P + (V_D/V_E)(r_P - r_D)$, a gain or loss measured against the investor's own cash V_E.

EXAM SHORTCUTS

- The sign of $r_P - r_D$ shows whether borrowing raises or lowers a gain before any calculation: a gain is magnified only if r_P exceeds r_D, and a loss is always magnified.
- The real rate moves with the nominal risk-free rate and against expected inflation, so a falling bill yield with rising expected inflation means a lower real rate without computing it.

FORMULAS

Return over a single period	$\text{Price return} = \dfrac{P_1 - P_0}{P_0}$ $\text{Capital distribution return} = \dfrac{D}{P_0}$ $\text{Total (holding period) return} = \dfrac{P_1 - P_0 + D}{P_0} = \text{price return} + \text{capital distribution return}$
Averaging returns over several periods	$\bar{R}_{\text{arith}} = \dfrac{R_1 + R_2 + \cdots + R_n}{n}$ $R_{\text{geo}} = \left[(1+R_1)(1+R_2)\cdots(1+R_n)\right]^{1/n} - 1$
Annualizing a holding period return	$\text{Annualized return} = (1 + HPR)^{365/\text{days}} - 1$ $(1 + R_{\text{period}})^{c} - 1$
Compounding frequency and continuous compounding	$\text{Effective annual rate} = \left(1 + \dfrac{r_s}{m}\right)^{m} - 1$ $PV = \dfrac{FV}{(1 + r_s/m)^{m}}$ (one year ahead) $r_{cc} = \ln(1 + HPR)$ $HPR = e^{r_{cc}} - 1$
Building blocks of a required return	$\text{Required return} = \underbrace{\text{real risk-free rate} + \text{expected inflation}}_{\text{nominal risk-free rate}} + \text{risk premium(s)}$
Real return	$1 + r_{\text{real}} = \dfrac{1 + r_{\text{nominal}}}{1 + \pi}$ $\pi = \dfrac{1 + r_{\text{nominal}}}{1 + r_{\text{real}}} - 1$
Pretax vs. after-tax nominal return	$r_{\text{after-tax}} = \text{price return} \times (1 - t_{cg}) + \text{distribution return} \times (1 - t_d)$
Leveraged return	$r_L = \dfrac{r_P(V_E + V_D) - r_D V_D}{V_E} = r_P + \dfrac{V_D}{V_E}(r_P - r_D)$

2 Types of Financial Returns

MODULE 2.1 · FINANCIAL ASSETS, INSTRUMENTS, AND INDICATORS

- Total return is the capital appreciation (price) return plus the capital distribution return; distributions are realized when received, price gains only when the asset is sold, and unrealized gains still count in total return.

credguild.com

- Equity returns come mainly from price return, with variable and discretionary dividends, while a traditional bond's return comes mainly from its fixed coupon and price return is usually the smaller part.
- Interest rates, market indexes and foreign exchange rates are financial indicators: they influence returns but produce no cash flows of their own.
- A T-bill's holding period return is $(\text{Face value} - \text{Price})/\text{Price}$, annualized as $(1 + HPR)^{365/t} - 1$; when market interest rates rise, fixed-income prices fall, and longer-dated bonds fall more.
- The domestic currency return is $(1 + r_{\text{local}})(1 + r_{FX}) - 1$, where r_{FX} is the change in the value of the foreign currency and is negative if it depreciates.
- Volatility lowers the geometric (compound) return, so the arithmetic mean overstates multiperiod performance, most of all for the more volatile asset.

EXAM SHORTCUTS

- A strengthening foreign currency puts the domestic return above the local return and a weakening one puts it below, so answers on the wrong side of the local return can be ruled out before compounding.

MODULE 2.2 · LONG-TERM RETURNS ON FINANCIAL ASSETS

- Under the risk-return tradeoff, expected long-term returns in a developed market rank short-term government bills < government bonds < corporate bonds < large-cap equities < small-cap equities, with bonds compared at similar maturities.
- The credit spread is the extra return corporate bonds offer over government bonds for their greater default risk, and longer-term fixed-income securities are expected to earn more than shorter-term ones because they face more inflation and interest rate risk.
- Systematic (macro) risk, such as inflation, interest rate and market risk, affects all investments to some degree and is not reduced by diversification; unsystematic (micro) risk, such as issuer and sector risk, is.
- The range, the highest return minus the lowest return, is the simplest measure of risk, and a wider range means more variability and more risk.
- A client who needs capital growth and tolerates risk tilts toward equities, a client who needs steady income with low risk tolerance tilts toward fixed-income securities, and a client who wants both holds a balanced mix.
- Each component of a portfolio's expected return is the weighted average of the asset-class components, $E(R_p) = \sum w_i E(R_i)$, and total return is the capital appreciation return plus the capital distribution return.

FORMULAS

Two sources of return	$\text{Total return} = \text{capital appreciation return} + \text{capital distribution return}$
Interest rates and the risk-free rate	$HPR = \dfrac{\text{Face value} - \text{Price}}{\text{Price}}$ $r_{\text{annual}} = (1 + HPR)^{365/t} - 1$
Foreign exchange rates	$1 + r_{\text{domestic}} = (1 + r_{\text{local}})(1 + r_{FX})$
The risk-return tradeoff	$\text{short-term government bills} < \text{government bonds} < \text{corporate bonds} < \text{large-cap equities} < \text{small-cap equities}$
Portfolio expected return	$E(R_p) = \sum_i w_i E(R_i)$ $\text{total return} = \text{capital appreciation return} + \text{capital distribution return}$

credguild.com

3 Benchmarking Returns

MODULE 3.1 · TIME-WEIGHTED AND MONEY-WEIGHTED RETURNS

- The money-weighted return is the IRR of the portfolio's cash flows, the rate that gives them an NPV of zero, with money going in and money coming out entered with opposite signs.
- The time-weighted return geometrically links the holding period returns of subperiods formed at each significant deposit or withdrawal, and is annualized with a root equal to the number of years, not the number of subperiods.
- The MWR depends on the timing and size of external cash flows: a deposit just before strong performance makes it higher than the TWR and a deposit just before poor performance makes it lower, while the TWR is unaffected.
- The MWR and the TWR are equal when there are no external cash flows during the period.
- The investment management industry prefers the TWR because managers usually do not control when clients add or withdraw money; the MWR suits a manager or investor who controls the cash flows.
- Firms claiming compliance with the voluntary GIPS must present time-weighted returns, except where the manager controls external cash flows and the portfolio is also closed-end, has a fixed life, calls committed capital or holds mainly illiquid investments.

EXAM SHORTCUTS

- With no deposits, withdrawals or paid-out income during the period, the MWR equals the TWR, so no IRR has to be computed.
- A deposit made just before the stronger subperiod puts the MWR above the TWR, and one made just before the weaker subperiod puts it below; withdrawals work the opposite way, so comparison questions need no IRR.
- If the subperiods are of equal length and each earns the same rate, the MWR and the TWR both equal that rate whatever the cash flows.

MODULE 3.2 · SECURITY MARKET INDEXES

- A price index uses constituent prices only and gives the price return; a total return index adds income, assumed reinvested, and its total return exceeds the price return when constituents pay distributions.
- A price-weighted index is the sum of prices divided by a divisor that is changed for splits, reverse splits, stock dividends and constituent changes, and its higher-priced stocks carry more weight.
- An equal-weighted index gives each constituent $1/N$ at each rebalancing, overweights small companies relative to their size and needs frequent, costly rebalancing as weights drift.
- A market-capitalization-weighted index weights each stock by its share of total market value, tracks aggregate investor wealth and rebalances automatically, but gives more weight to stocks whose prices have risen; a float-adjusted version uses market cap × float %.
- A fundamental-weighted index weights constituents by fundamentals such as earnings, dividends, cash flow, sales or book value and has a value tilt.
- Rebalancing resets weights to target after price moves and matters mainly for fundamental-weighted and equal-weighted indexes; reconstitution changes the constituents, with the divisor adjusted on that date.

EXAM SHORTCUTS

- For a period that starts at equal weights, an equal-weighted index return is the arithmetic mean of the constituent returns, so no share counts or divisor are needed.
- Only a price-weighted index changes its divisor for a stock split or stock dividend; equal-weighted, market-cap-weighted and fundamental-weighted indexes leave it unchanged.

FORMULAS

Money-weighted rate of return (MWR)

$$\sum_{t=0}^{T} \frac{CF_t}{(1+MWR)^t} = 0$$

Time-weighted rate of return (TWR)	$HPR_i = \frac{V_{end,i} - V_{begin,i} + D_i}{V_{begin,i}}$ $TWR_{annual} = [(1+HPR_1)(1+HPR_2)\cdots(1+HPR_n)]^{1/\text{years}} - 1$
What an index is and how its return is measured	$R_{\text{period}} = (1+R_1)(1+R_2)\cdots(1+R_n) - 1$ $V_n = V_0(1+R_1)(1+R_2)\cdots(1+R_n)$
Price-weighted index	$\text{Index} = \frac{\sum P_i}{D}$ $w_i = \frac{P_i}{\sum P_j}$
Market-capitalization-weighted (value-weighted) index	$w_i = \frac{P_i N_i}{\sum P_j N_j}$ $\text{Index} = \frac{\sum P_i N_i}{D}, \quad D = \frac{\text{base-period market cap}}{\text{base value}}$

4 The Time Value of Money in Finance

MODULE 4.1 · DISCOUNTED CASH FLOW VALUATION

- Every instrument here is valued as the present value of its expected cash flows, each discounted with $PV = FV/(1+r)^t$.
- A bond sells at a premium when its coupon rate is above its YTM, at par when they are equal and at a discount when the coupon rate is below the YTM.
- A perpetuity is worth PMT/r one period before its first payment, and preferred stock is valued the same way, D_p/r_p.
- In an amortizing loan each payment's interest is the rate times the opening balance, so the interest part falls and the principal part rises over the life of the loan.
- The Gordon growth model $V_0 = D_1/(r_e - g_c)$ uses next period's dividend and requires constant growth below the required return.
- In a multistage DDM the terminal value $P_n = D_{n+1}/(r_e - g_c)$ is discounted n periods, together with the dividend paid at n.

EXAM SHORTCUTS

- Compare the coupon rate with the YTM before calculating a bond price: a coupon above the YTM means a premium, equal means par, below means a discount, so prices on the wrong side of par can be ruled out at once.
- A zero-coupon bond, a coupon bond and a level loan are all one TVM entry; a zero is the same entry with PMT = 0.
- A level perpetuity or a preferred share needs no worksheet: value = payment ÷ rate.

MODULE 4.2 · IMPLIED RETURNS AND CASH FLOW ADDITIVITY

- The implied return on a zero-coupon bond is $(FV/PV)^{1/N} - 1$; a coupon bond's YTM is found with the calculator (CPT I/Y).
- A bond's price and its YTM move in opposite directions.
- Rearranging the Gordon model gives the required return $r_e = D_1/P_0 + g_c$ and the implied growth rate $g_c = r_e - D_1/P_0$.
- Under cash flow additivity the PV of a stream equals the sum of the PVs of its pieces, so an uneven stream can be replicated with annuities and zero-coupon bonds.

- No arbitrage sets the implied forward rate $1y1y = (1+S_2)^2/(1+S_1) - 1$ and the forward exchange rate $\text{Forward}/\text{Spot} = (1+r_{\text{price}})^T/(1+r_{\text{base}})^T$.
- In a one-period binomial model the hedge ratio is $h = (c_u - c_d)/(S_u - S_d)$, the hedged portfolio earns the risk-free rate, and $c_0 = hS_0$ minus the PV of the hedged portfolio.

EXAM SHORTCUTS

- With an upward-sloping spot curve the implied forward rate lies above both spot rates, so a forward rate between the two spot rates can be ruled out.
- Check the direction of a forward exchange rate first: when the price currency has the higher interest rate the forward is above spot, and when it has the lower rate the forward is below spot.
- When two cash flow streams differ only by $+X$ earlier and $-X$ later, the stream with the earlier cash flow has the higher PV at any positive discount rate, with no NPV calculation needed.

FORMULAS

The time value of money	$FV = PV(1+r)^t$ $PV = \dfrac{FV}{(1+r)^t}$
Fixed-coupon bonds	$P = \sum_{t=1}^{N} \dfrac{C}{(1+YTM)^t} + \dfrac{F}{(1+YTM)^N}$
Perpetual bonds	$PV_{\text{perpetuity}} = \dfrac{PMT}{r}$
Amortizing bonds and annuities	$PV_{\text{annuity}} = PMT \times \dfrac{1-(1+r)^{-N}}{r}$ $PMT = \dfrac{r \times PV}{1-(1+r)^{-N}}$
Equity instruments	$V_{\text{preferred}} = \dfrac{D_p}{r_p}$ $V_0 = \dfrac{D_1}{r_e - g_c}$ $D_1 = D_0(1+g_c)$
Zero-coupon bond	$r = \left(\dfrac{FV}{PV}\right)^{1/N} - 1$
Equity: implied required return	$r_e = \dfrac{D_1}{P_0} + g_c = \text{dividend yield} + \text{constant growth rate}$
Equity: implied growth rate	$g_c = r_e - \dfrac{D_1}{P_0} = \text{required return} - \text{dividend yield}$
Forward interest rates	$(1+S_2)^2 = (1+S_1)(1+1y1y) \quad \Rightarrow \quad 1y1y = \dfrac{(1+S_2)^2}{1+S_1} - 1$ $(1+S_3)^3 = (1+S_1)(1+1y1y)(1+2y1y)$

credguild.com

Forward exchange rates	$\frac{\text{Forward}}{\text{Spot}} = \frac{(1+r_{\text{price}})^T}{(1+r_{\text{base}})^T}$ $\frac{\text{Forward}}{\text{Spot}} = \frac{1+r_{\text{price, period}}}{1+r_{\text{base, period}}}$
Option values: one-period binomial model	$h = \frac{c_u - c_d}{S_u - S_d}$

5 Statistical Characteristics of Asset Returns

MODULE 5.1 · MEASURES OF CENTRAL TENDENCY

- The arithmetic mean weights every observation equally and is sensitive to outliers, while the median, the middle value of sorted data (the average of the two middle values with an even count), is a robust estimator.
- The mode is the most frequent value; a dataset can have one mode, two (bimodal), several or none, and for continuous data the modal interval is reported.
- A trimmed mean drops a set percentage of the extreme observations, split equally between the two ends, and averages the rest; a winsorized mean replaces the extremes with percentile values and averages all n observations.
- When there are only a few outliers, a winsorized mean measures the center better than the arithmetic mean or a trimmed mean does.
- A quantile is a cut-off value with a given share of the observations at or below it, and the interquartile range $Q_3 - Q_1$ contains the central 50% of the data.
- Stationary data have a mean and variance that do not change over time; trending price levels are typically nonstationary, while their periodic percentage changes are much more likely to be stationary.

EXAM SHORTCUTS

- Changing only the largest or the smallest observation moves the mean but not the median, as long as that value keeps its rank.
- Quantile positions run in a fixed order (first decile, first quintile, first quartile, median), so an answer that puts a lower quantile above a higher one can be ruled out.
- A few huge positive values make the trimmed and winsorized means lower than the arithmetic mean, and a few huge negative values make them higher.

MODULE 5.2 · DISPERSION, SKEWNESS, AND KURTOSIS

- The range, maximum minus minimum, uses only the two extreme observations, while the mean absolute deviation $\sum|X_i - \bar{X}|/n$ averages the unsigned deviations from the mean.
- Sample variance divides the sum of squared deviations from the mean by $n - 1$ (Bessel's correction), which makes s^2 an unbiased estimator of σ^2; population variance divides by N.
- Taking the positive square root of the variance gives the standard deviation, which is in the same units as the data, while the variance is in squared units.
- Under the root of time law, with independent returns the variance over T periods is $T\sigma^2$ and the standard deviation is $\sigma\sqrt{T}$, although long-horizon variance is often lower in practice.
- A positively skewed distribution has a long right tail with mode < median < mean, a negatively skewed one has a long left tail with mean < median < mode, and a symmetrical one has all three equal.
- A normal distribution has kurtosis 3; excess kurtosis is kurtosis − 3, a leptokurtic distribution (above 3) has fatter tails, and more negative skew and greater excess kurtosis both mean greater risk.

EXAM SHORTCUTS

- In a skewed unimodal distribution the median lies between the mode and the mean, and the mean lies on the side of the long tail, so the direction of skew fixes the order of all three measures.

credguild.com

MODULE 5.3 · COVARIANCE, CORRELATION, AND ALTERNATIVE MEASURES OF DISPERSION

- Sample covariance is $\sum(X_i - \bar{X})(Y_i - \bar{Y})/(n-1)$; its sign shows the direction of co-movement, but its size depends on the units and it has no upper or lower bound.
- Correlation $\rho_{XY} = \text{Cov}_{XY}/(s_X s_Y)$ has no units, lies between −1 and +1, and $\rho = 0$ means no linear relationship.
- Correlation is not causation, outliers can create or destroy a correlation, and spurious correlation arises by chance or because both variables are related to a third variable.
- A scatter plot can reveal a nonlinear relationship that a correlation coefficient would miss.
- Semivariance uses only the outcomes below the mean, and the target downside deviation uses only the outcomes below a target B but keeps the full sample size minus one in the denominator.
- The coefficient of variation $CV = s/\bar{X}$ measures risk per unit of mean return, is unit-free, and a lower CV is better.

EXAM SHORTCUTS

- With a zero risk-free rate the CV is the inverse of the Sharpe ratio, so one can be read straight from the other.
- A higher target never lowers the target downside deviation, which rules out an answer in which raising the target reduces it.

FORMULAS

Population vs. sample	$\text{Sample mean: } \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ $\text{Population mean: } \mu = \frac{\sum_{i=1}^{N} X_i}{N}$
Range and mean absolute deviation	$\text{range} = \text{maximum value} - \text{minimum value}$ $\text{MAD} = \frac{\sum_{i=1}^{n} \lvert X_i - \bar{X} \rvert}{n}$
Skewness	$\text{skewness} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{s}\right)^3$
Kurtosis	$\text{kurtosis} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i - \bar{X}}{s}\right)^4$ $\text{excess kurtosis} = \text{kurtosis} - 3$
Covariance	$\text{Cov}_{XY} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{n-1} \quad \text{(sample)}$
Correlation	$\rho_{XY} = \frac{\text{Cov}_{XY}}{s_X s_Y}$ $\text{Cov}_{XY} = \rho_{XY}\, s_X\, s_Y$
Semivariance and target downside deviation	$\sigma^2_{\text{semi}} = \frac{\sum_{X_i < \bar{X}}(X_i - \bar{X})^2}{n}$ $s_{\text{target}} = \sqrt{\frac{\sum_{X_i < B}(X_i - B)^2}{n-1}}$
Coefficient of variation	$CV = \frac{s}{\bar{X}}$ $\left(\text{or } \frac{\sigma}{\mu}\right)$

credguild.com

6 Statistical Distributions for Financial Asset Prices and Returns

MODULE 6.1 · PROBABILITY DISTRIBUTIONS AND EXPECTED VALUES

- A discrete random variable has a probability mass function with positive probabilities at individual outcomes; a continuous variable has a density, zero probability at any single point and probabilities assigned to intervals.
- A probability function is valid only if every $p(x)$ lies between 0 and 1 and the probabilities sum to 1.
- The CDF is $F(x) = P(X \leq x)$, and $P(a < X \leq b) = F(b) - F(a)$.
- Expected value is $E(X) = \sum P(x_i)x_i$; variance is $\sum P(x_i)[x_i - E(X)]^2$ in squared units, and the standard deviation is its square root.
- The sign of the covariance shows the direction of co-movement, but its size depends on the units; correlation $\rho = \text{Cov}(X, Y)/(\sigma_X \sigma_Y)$ is unit-free and lies between −1 and +1.

EXAM SHORTCUTS

- When the outcomes listed are mutually exclusive and exhaustive and one probability is missing, it equals 1 minus the sum of the others.
- Variance can be found as $\sigma^2 = E(X^2) - [E(X)]^2$, which avoids computing each deviation.
- For a continuous variable, including or excluding the endpoints does not change the probability of an interval.

MODULE 6.2 · DISCRETE AND CONTINUOUS PROBABILITY DISTRIBUTIONS

- A binomial variable counts successes in n independent trials with a constant p: $E(X) = np$ and $\sigma^2 = np(1 - p)$; a single trial is a Bernoulli distribution.
- A Poisson variable counts events in an interval at a constant rate λ, and its mean and variance both equal λ.
- For a continuous uniform variable on $[a, b]$, $P(x_1 \leq X \leq x_2) = (x_2 - x_1)/(b - a)$.
- The normal distribution is symmetric, fully described by its mean and variance, and unbounded in both directions.
- If $\ln Y$ is normal, Y is lognormal: bounded below by zero and positively skewed, which suits asset prices.
- The logistic distribution is symmetric like the normal but has fatter tails (excess kurtosis +1.2).

EXAM SHORTCUTS

- Identify the distribution from the wording: successes in a fixed number of trials is binomial, a count of events in an interval is Poisson, random draws for simulation are continuous uniform, and asset prices are lognormal.
- Normal bands: about 68% of outcomes lie within ±1σ and 95% within ±2σ, so about 2.5% lie beyond 2σ on each side and about 13.5% lie between 1σ and 2σ on each side.
- For a Poisson variable, the probability of at least one event is $1 - e^{-\lambda}$.

MODULE 6.3 · CONDITIONAL EXPECTATIONS AND BAYESIAN UPDATING

- A joint probability is $P(A \text{ and } B)$, marginal probabilities are the row and column totals, and a conditional probability uses only the outcomes consistent with the condition.
- The conditional expectation is $E(X \mid S) = \sum P(x_i \mid S)x_i$, and the conditional variance is measured around $E(X \mid S)$.
- In a probability tree a path probability is the product of its branch probabilities.
- Bayes' formula is $P(E \mid I) = P(I \mid E)P(E)/P(I)$, where $P(I)$ adds the joint probabilities of the information with every state.
- $P(I \mid E)$ and $P(E \mid I)$ are generally different, and the posterior still depends on the prior.

EXAM SHORTCUTS

- For Bayes' formula, write one row per state with prior × likelihood, add the rows to get $P(I)$, and divide the row for the state of interest by that total.
- Check a conditional expectation answer: the conditional expectations weighted by their branch probabilities must give back the unconditional expectation.

FORMULAS

Cumulative distribution function (CDF)	$F(x) = P(X \leq x)$ $P(a < X \leq b) = F(b) - F(a)$
Unconditional expected value (mean)	$E(X) = \sum_{i=1}^{n} P(x_i)\, x_i$
Unconditional variance and standard deviation	$\sigma^2(X) = \sum_{i=1}^{n} P(x_i)\, [x_i - E(X)]^2$ $\sigma(X) = \sqrt{\sigma^2(X)}$
Unconditional covariance	$\text{Cov}(X, Y) = \sum_{i=1}^{n} P(x_i, y_i)\, [x_i - E(X)]\, [y_i - E(Y)]$
Discrete distributions	$E(X) = np$ $\sigma^2(X) = np(1-p)$ $P(k \text{ events in interval } t) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$ $\text{mean} = \text{variance} = \lambda \text{ (for one unit of time, } t = 1)$
Continuous distributions	$P(x_1 \leq X \leq x_2) = \frac{x_2 - x_1}{b - a}$ $\text{mean} = \frac{a+b}{2}$ $\sigma^2 = \frac{(b-a)^2}{12}$
Conditional mean, variance and covariance	$E(X \mid S) = \sum_{i} P(x_i \mid S)\, x_i$ $\text{Var}(X \mid S) = \sum_{i} P(x_i \mid S)\, [x_i - E(X \mid S)]^2 = E(X^2 \mid S) - [E(X \mid S)]^2$
Updating probabilities with Bayes' formula	$P(E \mid I) = \frac{P(I \mid E)}{P(I)} \times P(E) = \frac{P(I \mid E)\, P(E)}{P(I \mid E)P(E) + P(I \mid E^c)P(E^c)}$

7 Estimation and Hypothesis Testing

MODULE 7.1 · THE CENTRAL LIMIT THEOREM, CONFIDENCE INTERVALS, AND SAMPLING

- By the central limit theorem, the mean of samples of size $n \geq 30$ from any population with a finite variance is approximately normal with mean μ and variance σ^2/n.
- The sample mean's standard error is $\sigma/\sqrt{n}$, or $s/\sqrt{n}$ when σ is unknown.
- A confidence interval is the point estimate ± reliability factor × standard error, with a t reliability factor and $n - 1$ df when σ is unknown.
- A desirable estimator is unbiased, efficient, consistent and robust.
- Parametric VaR is $z_\alpha \times \sigma \times$ portfolio value with the whole α in one tail; it is the minimum loss at the stated probability, not the worst possible loss.

credguild.com

- Simple random, systematic, stratified and cluster sampling are probability methods; convenience and judgmental sampling are nonprobability methods.

EXAM SHORTCUTS

- Quadrupling the sample size halves the standard error.
- Reliability factors: two-tailed 90%, 95% and 99% use 1.65, 1.96 and 2.58; one-tailed 10%, 5% and 1% (VaR) use 1.28, 1.65 and 2.33.
- A higher confidence level widens an interval and a larger sample narrows it, which rules out answer choices that move the wrong way.

MODULE 7.2 · HYPOTHESIS TESTS OF THE POPULATION MEAN

- H_0 always contains the equality and is usually the claim the researcher hopes to reject; H_a is the claim the researcher seeks evidence for.
- The test statistic is (sample mean − μ_0) ÷ standard error, following a t-distribution with $n - 1$ df when s is used.
- H_0 is rejected when the statistic falls beyond the critical value; otherwise the decision is to fail to reject, and the null is never accepted.
- A Type I error rejects a true null with probability α; a Type II error fails to reject a false null with probability β; power is $1 - \beta$.
- With the sample size unchanged, a lower α makes a Type II error more likely and reduces power.

EXAM SHORTCUTS

- The hypothesis that contains the equality sign is H_0.
- A two-tailed 10% test and a one-tailed 5% test share the critical value 1.65.
- For a two-tailed test, H_0 is rejected when the sample mean falls outside $\mu_0 \pm$ critical value × standard error, so a confidence interval answers the test directly.

MODULE 7.3 · OTHER PARAMETRIC HYPOTHESIS TESTS

- Independent samples with equal variances use the pooled difference-in-means t-test with $n_1 + n_2 - 2$ df; the same subjects measured twice use the paired comparisons test with $n - 1$ df.
- A single variance is tested with $\chi^2 = (n-1)s^2/\sigma_0^2$ and $n - 1$ df; the chi-square distribution is right-skewed and bounded below by zero.
- Two variances are tested with $F = s_1^2/s_2^2$ and $n_1 - 1$, $n_2 - 1$ df.
- A correlation is tested with $t = r\sqrt{n-2}/\sqrt{1-r^2}$ and $n - 2$ df; its absolute value rises with $|r|$ and with n.
- The p-value is the smallest significance level at which H_0 can be rejected.
- p-hacking raises the frequency of Type I errors.

EXAM SHORTCUTS

- Choose the test from what is tested: a mean uses t when the population variance is unknown (z when it is known, or as a large-sample approximation), a test that a correlation is zero uses t when both variables are normally distributed, one variance uses chi-square, and two variances use F.
- With the larger variance in the numerator of the F-statistic, only the upper critical value is needed.
- When a p-value is given, reject H_0 if it is below α; no critical value is needed.

MODULE 7.4 · NONPARAMETRIC HYPOTHESIS TESTS

- A nonparametric test is used for ranks or categories, when the assumptions of a parametric test do not hold, or when the question is not about a parameter.
- The Spearman rank correlation is $r_s = 1 - 6\sum d_i^2/[n(n^2-1)]$, tested in large samples ($n > 30$) with the same t-statistic as a Pearson correlation.
- In a chi-square test of independence the expected count is row total × column total ÷ total, $\chi^2 = \sum(O-E)^2/E$ with $(r-1)(c-1)$ df, and only the right tail is used.
- A standardized residual is $(O-E)/\sqrt{E}$, and cells beyond ±2 are flagged by the rule of thumb.

EXAM SHORTCUTS

- Rank data point to the Spearman test; category counts point to a chi-square test on a contingency table.
- Contingency table df are $(r-1)(c-1)$, counting only category rows and columns and leaving out the totals.

FORMULAS

The central limit theorem (CLT)	$\bar{X} \dot{\sim} N\left(\mu, \frac{\sigma^2}{n}\right)$ for large n (rule of thumb: $n \geq 30$)
Standard error and sampling error	$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$ (σ known) $s_{\bar{X}} = \frac{s}{\sqrt{n}}$ (σ unknown, the usual case)
Confidence intervals	$\text{point estimate} \pm \underbrace{(\text{reliability factor} \times \text{standard error})}_{\text{margin of error}}$
Value at risk (VaR) — a one-tailed interval	$\text{VaR}_\alpha = z_\alpha \times \sigma \times \text{portfolio value}$
Test statistic for a mean	$\text{test statistic} = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$ (use $\sigma/\sqrt{n}$ if σ is known)
Spearman rank correlation test	$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$
Test of independence using a contingency table	$E_{i,j} = \frac{(\text{total for row } i) \times (\text{total for column } j)}{\text{total number of observations}}$ $\chi^2 = \sum_i \sum_j \frac{(O_{i,j} - E_{i,j})^2}{E_{i,j}}$ $\text{df} = (r-1)(c-1)$

8 The Return and Risk of a Financial Portfolio

MODULE 8.1 · PORTFOLIO EXPECTED RETURN

- A portfolio's expected return is the weighted average $E(R_P) = \sum w_i E(R_i)$, with each weight equal to the asset's share of total portfolio value and the weights summing to 1.
- A portfolio's historical return for one period is $\sum w_{i,t-1} R_{i,t}$, using the weights at the start of the period.
- The cumulative (compound) return $\prod(1 + R_{P,t}) - 1$ and the geometric average include compounding, while the sum and the arithmetic average of the period returns ignore it.
- Without rebalancing, market moves shift the weights (portfolio drift), so each period's return must use the drifted weights from start-of-period market values.
- Style drift is a portfolio moving away from its stated strategy, either through portfolio drift or through changes in the holdings themselves, such as a small-cap stock growing into the mid-cap category.
- Exam convention: drift raises the weights of high-return securities in an equal-weighted portfolio and enlarges already large allocations in a portfolio that is not equally weighted, so it behaves like a momentum strategy with possibly higher concentration risk; current practice: in any unrebalanced portfolio the outperformers gain weight, and whether concentration or risk rises depends on which assets outperform.

EXAM SHORTCUTS

- For a portfolio that is not rebalanced, the cumulative return equals the initial-weight average of each asset's cumulative return, so the cumulative figure needs no period-by-period weights.

MODULE 8.2 · PORTFOLIO RISK MEASURES

- Two-asset portfolio variance is $\sigma_P^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\text{Cov}_{1,2}$, and the portfolio standard deviation is its square root.
- Covariance can be written $\text{Cov}_{1,2} = \rho_{1,2}\sigma_1\sigma_2$, which exam calculations use; an expected covariance, however, cannot be found by multiplying separate expected values of the three inputs.
- With positive weights, a correlation of +1 gives a portfolio standard deviation equal to the weighted average of the standard deviations (no diversification benefit), and any lower correlation gives less, with a larger benefit as ρ falls.
- With $\rho = -1$, $\sigma_P = |w_1\sigma_1 - w_2\sigma_2|$, which is zero only when $w_1\sigma_1 = w_2\sigma_2$.
- A negative covariance makes the cross term negative, so the portfolio is much less risky than the weighted average of the standard deviations suggests.

EXAM SHORTCUTS

- For two assets with positive weights, the portfolio standard deviation equals the weighted average of the two standard deviations only when $\rho = +1$ and is below it for any lower correlation, so with $\rho < +1$ an answer at or above the weighted average can be ruled out.

MODULE 8.3 · CORRELATION AND DIVERSIFICATION BENEFITS

- The diversification benefit grows as correlation falls: two-asset combinations lie on a straight line at $\rho = +1$, bow to the left when correlation is below +1, and include one zero-risk mix at $\rho = -1$.
- With N assets the variance-covariance matrix holds N variances and $(N^2 - N)/2$ unique covariances.
- Expected portfolio variance from scenarios is $\sum_m p_m[R_{P,m} - E(R_P)]^2$, with scenario probabilities that sum to 1.
- An equally weighted portfolio of N assets has variance $\bar{\sigma}^2/N + [(N-1)/N]\,\overline{\text{Cov}}$, which approaches the average covariance as N grows and is approximated for a large portfolio by $\bar{\rho}\,\bar{\sigma}^2$.
- Unsystematic risk is removed by diversification and is not rewarded with higher expected return; systematic risk is not removed and is rewarded.
- Roughly 20 to 30 well-chosen assets capture most of the diversification benefit, and beyond about 30 assets each new holding adds close to none.

EXAM SHORTCUTS

- The expected portfolio return from scenarios equals $\sum w_i E(R_i)$, the assets' expected returns weighted by portfolio weight, so it can be found without the scenario portfolio returns; the variance still needs them.

MODULE 8.4 · THE MINIMUM-VARIANCE PORTFOLIO AND THE EFFICIENT FRONTIER

- The minimum-variance frontier holds the lowest-risk portfolio for each level of expected return, and the global minimum-variance portfolio is its point of lowest risk.
- The efficient frontier is the upper part of the minimum-variance frontier from the global minimum-variance portfolio upward; points to the left of the frontier are unattainable and attainable points below the efficient frontier are inefficient.
- For two risky assets, $w_A = (\sigma_B^2 - \text{Cov}_{A,B})/(\sigma_A^2 + \sigma_B^2 - 2\text{Cov}_{A,B})$ and $w_B = 1 - w_A$ give the minimum-variance portfolio.
- The optimal risky (tangency) portfolio is where a line from the risk-free rate touches the efficient frontier, the efficient portfolio with the highest Sharpe ratio $(E(R_P) - R_f)/\sigma_P$.
- Combining the risk-free asset with portfolio P gives $E(R_C) = w_P E(R_P) + (1 - w_P)R_f$ and $\sigma_C = w_P\sigma_P$, with $w_P = (r_{\text{target}} - R_f)/(E(R_P) - R_f)$ for a target return and a weight above 100% meaning borrowing at the risk-free rate.
- Return-constrained optimization minimizes risk for a given return, risk-constrained optimization maximizes expected return for a given risk, and risk-adjusted return optimization seeks the highest Sharpe ratio.

EXAM SHORTCUTS

- Expected returns do not enter the two-asset minimum-variance weights, so any expected returns given in such a question can be set aside.

MODULE 8.5 · RISK AVERSION

- Optimal portfolio theory assumes rational, risk-averse investors who maximize utility, a single-period horizon, no taxes or transaction costs, and borrowing and lending at a constant risk-free rate.
- A risk-seeking investor has $A < 0$ and downward-sloping indifference curves, a risk-neutral investor has $A = 0$ and flat curves, and a risk-averse investor has $A > 0$ and upward-sloping curves.
- Utility is $U = E(r) - \frac{1}{2}A\sigma^2$ with the variance in decimals, so a higher coefficient of risk aversion A means a bigger penalty for risk.
- Risk tolerance, the inverse of risk aversion, can change with age, wealth and economic conditions.
- A risk-averse investor's indifference curves slope upward and are convex, curves higher and to the left give higher utility, and a more risk-averse investor has steeper curves.

MODULE 8.6 · CAPITAL ALLOCATION LINE AND CAPITAL MARKET LINE

- Under its assumptions (a riskless risk-free asset, a risky portfolio with return and risk above the risk-free rate, only these two assets, and unlimited borrowing and lending at the risk-free rate), the CAL is $E(R_C) = R_f + [(E(R_P) - R_f)/\sigma_P]\sigma_C$, with intercept R_f and slope equal to the risky portfolio's Sharpe ratio.
- Points on the CAL between the intercept and the risky portfolio involve lending at the risk-free rate, and points beyond the risky portfolio involve borrowing.
- An investor's optimal portfolio is where the CAL touches the highest attainable indifference curve; a more risk-averse investor chooses a lower-risk point on the same line.
- The CML is the CAL whose risky portfolio is the market portfolio, which holds all risky assets weighted by market value and is mean-variance efficient.
- Under the mutual fund theorem, every investor can obtain an efficient portfolio by combining the risk-free asset with the market portfolio.
- A CAL allows heterogeneous expectations, so each investor can have a different CAL, while the CML assumes homogeneous expectations and is the same for all investors.

EXAM SHORTCUTS

- Every portfolio on a CAL with a positive allocation to the risky portfolio has the Sharpe ratio of that risky portfolio, however much is lent or borrowed, so its Sharpe ratio can be read from the risky portfolio directly.

MODULE 8.7 · THE CAPITAL ASSET PRICING MODEL

- Because unsystematic risk can be diversified away at no cost, the market rewards investors only for bearing systematic risk.
- The CAPM is a single-factor model, $E(R_i) = R_f + \beta_i[E(R_M) - R_f]$, that assumes homogeneous expectations, utility-maximizing investors holding efficient portfolios, price takers and markets in equilibrium.
- Beta is $\mathrm{Cov}(R_i, R_M)/\sigma_M^2 = \rho_{i,M}\sigma_i/\sigma_M$, and the beta of the market portfolio is 1.
- Exam convention: a beta above 1 indicates an expected return above the market's and a beta below 1 one below it; current practice: this ordering holds only when the market risk premium is positive.
- The SML plots expected return against beta with intercept R_f and slope equal to the market risk premium and applies to any asset, while the CML plots against total risk and applies only to efficient portfolios.
- Portfolio theory is limited by non-normal returns, unstable correlations, imperfectly efficient markets, sensitivity to estimation error and a future that may differ from the past, and it ignores transaction costs and taxes.

FORMULAS

Expected return of a portfolio	$E(R_P) = \sum_{i=1}^{N} w_i\, E(R_i)$ $\sum_{i=1}^{N} w_i = 1$ $R_{P,t} = \sum_{i=1}^{N} w_{i,t-1}\, R_{i,t}$
Portfolio variance	$\sigma_P^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j\, \mathrm{Cov}(R_i, R_j)$ $\sigma_P^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\, \mathrm{Cov}_{1,2}$ $\sigma_P = \sqrt{\sigma_P^2}$
Covariance and correlation	$\mathrm{Cov}_{1,2} = \rho_{1,2}\,\sigma_1\,\sigma_2 \quad\Longrightarrow\quad \sigma_P^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho_{1,2}\sigma_1\sigma_2$
More assets: the variance–covariance matrix	$\text{number of unique covariances} = \dfrac{N^2 - N}{2}$ $\sigma_P^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \rho_{i,j}\sigma_i\sigma_j$ $\rho_{i,i} = 1$
Large portfolios: variance approaches average covariance	$\sigma_P^2 = \dfrac{\bar{\sigma}^2}{N} + \dfrac{N-1}{N}\,\overline{\mathrm{Cov}}$ $\sigma_P^2 \approx \bar{\rho}\,\bar{\sigma}^2$
Minimum-variance portfolio of two risky assets	$w_A = \dfrac{\sigma_B^2 - \mathrm{Cov}_{A,B}}{\sigma_A^2 + \sigma_B^2 - 2\,\mathrm{Cov}_{A,B}}$ $w_B = 1 - w_A$
Adding a risk-free asset	$\text{Sharpe ratio} = \dfrac{E(R_P) - R_f}{\sigma_P}$ $E(R_C) = w_P E(R_P) + (1 - w_P)R_f$ $\sigma_C = w_P\,\sigma_P$ $w_P = \dfrac{r_{\text{target}} - R_f}{E(R_P) - R_f}$
Utility function	$U = E(r) - \frac{1}{2}A\sigma^2$
The capital allocation line (CAL)	$E(R_C) = R_f + \left[\dfrac{E(R_P) - R_f}{\sigma_P}\right]\sigma_C$
The capital market line (CML)	$E(R_P) = R_f + \left[\dfrac{E(R_M) - R_f}{\sigma_M}\right]\sigma_P$
Return-generating models and the CAPM	$E(R_i) = R_f + \beta_i\,[E(R_M) - R_f]$

credguild.com

Beta	$\beta_i = \frac{\mathrm{Cov}(R_i, R_M)}{\sigma_M^2} = \frac{\rho_{i,M}\,\sigma_i\,\sigma_M}{\sigma_M^2} = \frac{\rho_{i,M}\,\sigma_i}{\sigma_M}$

9 Simulation of Financial Asset Prices and Returns

MODULE 9.1 · HISTORICAL SIMULATION

- Every simulation sets parameters, generates scenarios, evaluates them and compiles the results; historical simulation, bootstrapping and Monte Carlo simulation differ mainly in where the scenarios come from.
- Historical simulation applies actual past changes in prices, returns or risk factors to the current portfolio positions and assumes that past data represent the future.
- Historical simulation is nonparametric and captures skewness, excess kurtosis, fat tails and volatility clustering, but it can only produce outcomes from the lookback period and is weak when there are regime changes or unprecedented events.
- Missing historical data are filled with a proxy, a substitute instrument or a weighted combination of instruments with data available.
- VaR is the minimum loss over a specified period in the worst 5% (or 1%) of outcomes; the 95% historical VaR is the fifth percentile of the simulated P&L distribution, and daily VaR usually assumes a mean return of zero.
- Because a bond's duration shrinks as it approaches maturity, historical simulation for bonds applies past yield changes for the same remaining maturity to today's cash flows instead of using the bond's own past prices.

EXAM SHORTCUTS

- With N historical scenarios, the 95% VaR lies between the $(0.05N)$th-worst P&L and the next-worst one, so only those two outcomes need to be located (either one or a weighted average of the two is acceptable).

MODULE 9.2 · BOOTSTRAP RESAMPLING

- Bootstrapping is a nonparametric method that uses the sample at hand as a proxy for the population and builds new datasets by drawing from it at random with replacement, so each of n observations has a $1/n$ chance on every draw.
- In bootstrapping the same observation can be drawn several times within one scenario and can appear in many samples or in none.
- Historical simulation replays the past dataset as if it were the population, while bootstrapping builds new combinations of past observations, which reduces the risk of overfitting to one historical sequence.
- Bootstrapping can give the distribution of a European call's payoff at maturity under historical returns, which is not by itself the option's current no-arbitrage price, and can estimate the volatility of target date funds.
- Bootstrapping needs minimal assumptions and works with small samples, but it carries any bias in the original sample into every resample, cannot generate outcomes outside the observed range and is not suited to autocorrelated time series.

MODULE 9.3 · MONTE CARLO SIMULATION

- Monte Carlo simulation draws scenarios from probability distributions specified by the analyst and runs thousands of trials to build a simulated frequency distribution, with more trials giving more precise estimates.
- Monte Carlo simulation is used to estimate portfolio performance and to value securities without a closed-form pricing formula, such as Asian-style options, mortgage-backed securities and convertible bonds.
- Exam convention: a Monte Carlo value for a European call should match the Black-Scholes-Merton value, getting closer as more trials are run; current practice: this holds only when the simulated prices are lognormal with constant volatility and a risk-neutral expected return, with the average payoff discounted at the risk-free rate.
- The Cholesky decomposition maps independent standard normal draws into correlated normal variables using an assumed correlation; it does not estimate correlations or covariances.

credguild.com

- Monte Carlo simulation is flexible, forward-looking and handles nonnormal distributions and correlations, but it is complex, model-dependent, computationally intensive and data-hungry, and it gives a statistical estimate without the cause-and-effect insight of analytic methods.
- Historical simulation assumes past data represent the future, bootstrapping assumes the sample represents the population, and Monte Carlo simulation assumes the process generating outcomes can be modeled statistically.

EXAM SHORTCUTS

- Identify the method from the source of the scenarios: actual past changes mean historical simulation, resampling with replacement from the observed sample means bootstrapping, and random draws from an assumed distribution mean Monte Carlo simulation.

FORMULAS

Multivariate simulation and the Cholesky decomposition

$$X_1 = \sigma_1 Z_1$$
$$X_2 = \sigma_2 \left(\rho Z_1 + \sqrt{1-\rho^2}\, Z_2\right)$$

10 Applications of Simple Linear Regression in Finance

MODULE 10.1 · LINEAR REGRESSION BASICS

- Simple linear regression explains the variation of a dependent variable Y, $\sum(Y_i - \bar{Y})^2$, with the variation of a single independent variable X.
- Ordinary least squares chooses the line that minimizes the sum of squared errors, $\text{SSE} = \sum(Y_i - \hat{Y}_i)^2$.
- The OLS estimates are $\hat{b}_1 = \text{Cov}(X, Y)/\sigma_X^2$ and $\hat{b}_0 = \bar{Y} - \hat{b}_1\bar{X}$, so the fitted line passes through the point of means.
- The slope is the expected change in Y for a one-unit rise in X, and the intercept is the predicted value of Y when X equals zero.
- The residual $Y_i - \hat{Y}_i$ is the gap between the actual and the fitted value; exam convention treats the residual, error term and disturbance term as the same thing, while current practice treats the residual as the sample estimate of the unobservable error term.
- Time series data follow one variable over time, cross-sectional data cover many subjects at one point in time, and panel data repeat cross-sectional observations over time.

EXAM SHORTCUTS

- The regression line passes through $(\bar{X}, \bar{Y})$, so once the slope is known the intercept is $\bar{Y} - \hat{b}_1\bar{X}$ with no further sums.

MODULE 10.2 · ANALYSIS OF VARIANCE (ANOVA) AND GOODNESS OF FIT

- Simple linear regression assumes a linear relationship, homoskedastic residuals, residuals uncorrelated with one another, and normally distributed residuals.
- Residual variance that widens or narrows with X or over time is heteroskedasticity, and residuals correlated with one another, such as a recurring seasonal pattern, show autocorrelation.
- ANOVA splits total variation into explained and unexplained parts, $\text{SST} = \text{SSR} + \text{SSE}$.
- In a simple regression $\text{MSE} = \text{SSE}/(n-2)$, $\text{SEE} = \sqrt{\text{MSE}}$, and $R^2 = \text{SSR}/\text{SST} = 1 - \text{SSE}/\text{SST}$, which equals r^2.
- $F = \text{MSR}/\text{MSE}$ with 1 and $n-2$ degrees of freedom is a one-tailed test, and with one independent variable it tests $H_0 : b_1 = 0$, the same hypothesis as the slope's t-test.
- Regressing Y on an indicator variable gives a slope equal to the difference in the mean of Y between periods with and without the condition, and an intercept equal to the mean when the indicator is 0.

EXAM SHORTCUTS

- With one independent variable, $F = t^2$ for the test of $H_0 : b_1 = 0$ against $H_a : b_1 \neq 0$, so at the same significance level the F-test and the two-tailed slope t-test reach the same decision and only one needs computing.
- When only R^2 and SSR are given, SST $=$ SSR$/R^2$ and SSE $=$ SST $-$ SSR, which lead straight to the MSE and the SEE.
- In a simple regression $r = \pm\sqrt{R^2}$, taking the sign of the slope.

MODULE 10.3 · PREDICTED VALUES AND FUNCTIONAL FORMS OF REGRESSION

- A predicted value is $\hat{Y} = \hat{b}_0 + \hat{b}_1 X_p$, using the intercept and the sign of each term.
- The prediction interval for Y is $\hat{Y} \pm t_c s_f$ with $n - 2$ degrees of freedom, where s_f is the standard error of the forecast.
- The forecast is more precise when the SEE is smaller, the sample is larger, X_p is closer to $\bar{X}$ and the independent variable has a greater variance.
- A log-lin model ($\ln Y$ on X) has a slope giving the relative change in Y for an absolute change in X, a lin-log model the absolute change in Y for a relative change in X, and a log-log model an elasticity.
- Exam convention: a higher R^2, a higher F-statistic and a lower SEE identify the better-fitting functional form; current practice: R^2 and SEE are directly comparable only when the dependent variable is the same.

EXAM SHORTCUTS

- The standard error of the forecast is always larger than the SEE, so a prediction interval narrower than $\hat{Y} \pm t_c \times$ SEE can be ruled out.

MODULE 10.4 · ESTIMATING CAPM VALUES WITH SIMPLE LINEAR REGRESSION

- The CAPM is estimated by regressing the asset's excess return $R_i - R_f$ on the market risk premium $R_M - R_f$; the slope is the asset's beta and the intercept is zero if the CAPM holds exactly.
- A security market index proxies the market return, short-term government debt yields proxy the risk-free rate, and at least 30 observations is the usual minimum.
- Beta is tested with $t = (\hat{b}_1 - \text{hypothesized value})/s_{\hat{b}_1}$, and its confidence interval is the slope $\pm$ critical value $\times$ the slope's standard error.
- The expected return is $E(R_i) = R_f + \hat{\beta}[E(R_M) - R_f]$, and a change in inputs gives $\Delta E(R_i) = \Delta R_f + \hat{\beta}\,\Delta[E(R_M) - R_f]$.
- The R^2 of the CAPM regression is the share of the asset's excess-return variation that is systematic, and $1 - R^2$ is the unsystematic share.
- Longer samples give more precise estimates but may include periods with a different beta, and more frequent data can be noisy for thinly traded assets.

EXAM SHORTCUTS

- A two-tailed 95% confidence interval for beta that excludes 1 means $H_0 : \beta = 1$ is rejected at the 5% level, and one that includes 1 means it is not, so either result gives the other.

FORMULAS

The model and the regression line

$$Y_i = b_0 + b_1 X_i + \varepsilon_i \quad \text{(population model)}$$

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i \quad \text{(estimated regression line, the line of best fit)}$$

Least squares criterion	$\text{SSE} = \sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$ $\hat{b}_1 = \dfrac{\text{Cov}(X,Y)}{\sigma_X^2}$ $\hat{b}_0 = \bar{Y} - \hat{b}_1\bar{X}$
Analysis of variance (ANOVA)	$\text{SST} = \text{SSR} + \text{SSE}$
Goodness-of-fit measures	$\text{SEE} = \sqrt{\text{MSE}} = \sqrt{\dfrac{\text{SSE}}{n-2}}$ $R^2 = \dfrac{\text{SSR}}{\text{SST}} = 1 - \dfrac{\text{SSE}}{\text{SST}}$
F-test	$F = \dfrac{\text{MSR}}{\text{MSE}}$ $df = 1\ (\text{numerator}),\ n-2\ (\text{denominator})$
t-test of a regression coefficient	$t = \dfrac{\hat{b}_1 - b_1}{s_{\hat{b}_1}}$ $df = n-2$
Predicted values	$\hat{Y} = \hat{b}_0 + \hat{b}_1 X_p$
Confidence (prediction) interval for a predicted value	$\hat{Y} \pm t_c\, s_f$ $df = n-2$ $s_f^2 = \text{SEE}^2\left[1 + \dfrac{1}{n} + \dfrac{(X_p - \bar{X})^2}{(n-1)s_X^2}\right]$
The CAPM as a one-factor regression	$E(R_i) = R_f + \beta_i\,[E(R_M) - R_f]$ $Z_{i,t} = b_0 + b_1 Z_{M,t} + \varepsilon_t$ $Z_{i,t} = R_{i,t} - R_{f,t}, \quad Z_{M,t} = R_{M,t} - R_{f,t}$

11 Introduction to Financial Data Science

- Financial data science applies quantitative and qualitative analysis to gain insight into a specific financial question, with main uses in evaluating investment opportunities, optimizing portfolios and mitigating risk.
- Big data can be structured (rows and columns), semi-structured (tagged, such as HTML code) or unstructured (social media posts, voice recordings, pictures, sensor output), and comes from traditional sources or from alternative sources: individuals, business processes and sensors.
- The characteristics of big data are volume, velocity (low latency means real time), variety of structures and, for financial data especially, veracity, meaning whether the data are reliable and credible.
- Data science processes data through capture, curation (ensuring data quality), storage, search and transfer, and a data lake keeps structured, semi-structured and unstructured data in their raw format.
- Data mining searches large datasets for patterns to predict outputs, while machine learning, a subcategory of AI, learns from input data without assumptions about their distribution and without human help.

credguild.com

- Supervised learning uses labeled inputs and outputs, unsupervised learning finds structure in unlabeled inputs, and reinforcement learning learns by trial and error from rewards and penalties.
- Overfitting comes from a model that is too complex and treats noise as a true pattern; underfitting comes from one that is not complex enough and treats true parameters as noise.
- The training dataset is the largest, about 60%–80% of the data, with validation and test sets splitting the rest equally; observations are normally assigned at random, but time series data keep their chronological order through a time-based split or rolling-window validation.

EXAM SHORTCUTS

- Classify a data source by where it comes from, not by its structure: records from individuals, business processes and sensors are alternative data even when they are structured.
- Match the type of machine learning to what the algorithm is given: labeled inputs and outputs mean supervised learning, unlabeled inputs alone mean unsupervised learning, and rewards and penalties for its own actions mean reinforcement learning.

credguild.com

Keep going at credguild.com

The sheet is for revision. These are free on the site and do what a PDF cannot:

- **Diagnostic test:** 36 questions weighted like the exam, a score in each topic and the topics to start with. credguild.com/cfa/level-1/diagnostic/
- **Practice questions:** 398 free questions with full explanations, scored as you answer. A free account answers any question in the bank, up to 50 new questions a day. credguild.com/cfa/level-1/practice/
- **Notes:** every reading of the 2027 curriculum, with worked examples and exam traps, on the site and as a PDF. credguild.com/cfa/level-1/notes/
- **Study plan:** pick your exam window and weekly hours and get a week-by-week schedule. credguild.com/cfa/level-1/plan/
- **Mock exam:** 180 questions in two sessions, on the 2027 topic weights; the first one is free with an account. credguild.com/cfa/level-1/mock/
- **Cheat sheet and formula sheet:** every reading's bottom line and every formula, by topic. credguild.com/cfa/level-1/cheat-sheet/

Found a mistake? Email hello@credguild.com.

credguild.com