

Columbia University

# **Portfolio Optimization through Information Theory**

Discrete Models (Prediction Markets)

Jan Vecer

Charles University

Frankfurt School of Finance & Management

9 October 2026

# Abstract

**Price can be viewed as information.** Positive assets represent probability distributions, and their normalized relative prices are likelihood ratios. Expected logarithmic returns, evaluated under the associated asset measure, are relative entropies.

This perspective connects pricing, portfolio choice, and hedging. Investor beliefs specify a desired distribution. Trading determines which distributions can be implemented. Portfolio choice then becomes a problem of matching probabilities or finding the closest attainable approximation.

We illustrate this general framework with binomial and negative-binomial models. They expose the roles of information flow, replication, and rebalancing, and allow a direct comparison of classical quadratic hedging with information-based criteria.

## Binary outcomes: beliefs, allocations, and odds

Two mutually exclusive outcomes: **Win** and **Loss**. The market supplies  $\mathbb{P}^Y$ ; the investor supplies beliefs  $\mathbb{P}^M$ .

| Outcome | Market $\mathbb{P}^Y$ | Belief / stake $\mathbb{P}^M$ | Gross odds | Payoff |
|---------|-----------------------|-------------------------------|------------|--------|
| Win     | 1/2                   | 2/3                           | 2          | 4/3    |
| Loss    | 1/2                   | 1/3                           | 2          | 2/3    |

Place  $\mathbb{P}^M(\omega)$  of initial wealth on outcome  $\omega$ . Each unit staked pays gross odds  $1/\mathbb{P}^Y(\omega)$  if that outcome occurs.

$$\frac{M(1, \omega)}{M(0)} = \underbrace{\mathbb{P}^M(\omega)}_{\text{stake fraction}} \cdot \underbrace{\frac{1}{\mathbb{P}^Y(\omega)}}_{\text{gross odds}} = \frac{\mathbb{P}^M(\omega)}{\mathbb{P}^Y(\omega)}.$$

Gross odds include the original stake. We will shortly show why allocating according to beliefs is log-optimal.

## Three outcomes: investing according to beliefs

The same construction works for **Win, Draw, and Loss**.

| Outcome | Market $\mathbb{P}^Y$ | Belief / stake $\mathbb{P}^M$ | Payoff $L_M$ |
|---------|-----------------------|-------------------------------|--------------|
| Win     | 0.50                  | 0.60                          | 1.20         |
| Draw    | 0.30                  | 0.30                          | 1.00         |
| Loss    | 0.20                  | 0.10                          | 0.50         |

Placing  $\mathbb{P}^M(\omega)$  on each outcome produces the portfolio  $M$ :

$$L_M(\omega) := \frac{M(1, \omega)}{M(0)} = \frac{\mathbb{P}^M(\omega)}{\mathbb{P}^Y(\omega)}, \quad \mathbb{E}^Y[L_M] = 1.$$

## Why is the investor's own probability the right allocation?

Vecer (2026), Analytical solution for Kelly's criterion for multiple outcomes.

## A positive payoff defines a probability

A payoff  $V(1) > 0$ , with price  $V(0) = \mathbb{E}^Y[V(1)]$ , induces

$$\mathbb{P}^V(\omega) = \mathbb{P}^Y(\omega) \frac{V(1, \omega)}{V(0)}, \quad \sum_{\omega} \mathbb{P}^V(\omega) = 1.$$

Equivalently, its normalized payoff is a likelihood ratio:

$$\frac{V(1)}{V(0)} = \frac{\mathbb{P}^V}{\mathbb{P}^Y}.$$

- A constant cash payoff corresponds to  $\mathbb{P}^V = \mathbb{P}^Y$ .
- The belief-based payoff  $\mathbb{P}^M/\mathbb{P}^Y$  corresponds to  $\mathbb{P}^V = \mathbb{P}^M$ .

**The payoff and its induced probability describe the same portfolio.**

For a strictly positive traded portfolio,  $\mathbb{P}^V$  is its numeraire measure.

## A one-dollar bet becomes a probability

Start with \$100. Bet \$1 on red at even money in an idealized fair game; keep \$99 in cash.

| Outcome | Market $\mathbb{P}^Y$ | Final wealth $V(1)$ | Induced $\mathbb{P}^V$ |
|---------|-----------------------|---------------------|------------------------|
| Red     | 0.50                  | \$101               | 0.505                  |
| Not red | 0.50                  | \$99                | 0.495                  |

$$\mathbb{P}^V(\text{red}) = \frac{1}{2} \frac{101}{100} = 0.505, \quad \mathbb{P}^V(\text{not red}) = \frac{1}{2} \frac{99}{100} = 0.495.$$

**The wealth pair (101, 99) and the probability pair (0.505, 0.495) describe the same position.**

Allocate \$50.50 to Red and \$49.50 to Not red at gross odds two: the same belief-based construction as before, now for a belief of 0.505.

# Relative entropy measures expected information

For two positive probability distributions  $\mathbb{P}$  and  $\mathbb{Q}$ , the **Kullback–Leibler divergence**, or **relative entropy**, is

$$D_{\text{KL}}(\mathbb{P} \parallel \mathbb{Q}) = \sum_{\omega} \mathbb{P}(\omega) \log \frac{\mathbb{P}(\omega)}{\mathbb{Q}(\omega)} = \mathbb{E}^{\mathbb{P}} \left[ \log \frac{\mathbb{P}}{\mathbb{Q}} \right].$$

- One outcome contributes log-likelihood evidence  $\log(\mathbb{P}(\omega)/\mathbb{Q}(\omega))$ .
- Averaging under  $\mathbb{P}$  gives the expected information favoring  $\mathbb{P}$  over  $\mathbb{Q}$ .
- $D_{\text{KL}}(\mathbb{P} \parallel \mathbb{Q}) \geq 0$ , with zero exactly when  $\mathbb{P} = \mathbb{Q}$ .

For  $\mathbb{P} = (2/3, 1/3)$  and  $\mathbb{Q} = (1/2, 1/2)$ ,

$$D_{\text{KL}}(\mathbb{P} \parallel \mathbb{Q}) = \frac{2}{3} \log \frac{4}{3} + \frac{1}{3} \log \frac{2}{3} = 0.056633.$$

Direction matters: generally  $D_{\text{KL}}(\mathbb{P} \parallel \mathbb{Q}) \neq D_{\text{KL}}(\mathbb{Q} \parallel \mathbb{P})$ . Cover–Thomas (2006), Ch. 2.

# Why maximize expected log wealth?

Repeat the same investment rule over independent, identically distributed periods. Let  $V(i) > 0$  be its gross wealth multiplier.

$$W(n) = W(0) \prod_{i=1}^n V(i), \quad \frac{1}{n} \log \frac{W(n)}{W(0)} = \frac{1}{n} \sum_{i=1}^n \log V(i).$$

If the model  $\mathbb{P}^M$  describes the outcomes and  $\mathbb{E}^M |\log V| < \infty$ , the law of large numbers gives

$$\frac{1}{n} \log \frac{W(n)}{W(0)} \longrightarrow \mathbb{E}^M [\log V] \quad \text{almost surely.}$$

**Maximizing expected log wealth maximizes this long-run growth rate under the model.**

Kelly (1956). This is geometric growth under the stated model, not the arithmetic expected return. Natural logarithms throughout.

## Log optimality follows from two divergences

For any positive, unit-cost payoff  $V(1)/V(0) = \mathbb{P}^V/\mathbb{P}^Y$ ,

$$\mathbb{E}^M \log \frac{\mathbb{P}^V}{\mathbb{P}^Y} = D_{\text{KL}}(\mathbb{P}^M \parallel \mathbb{P}^Y) - D_{\text{KL}}(\mathbb{P}^M \parallel \mathbb{P}^V).$$

The first term does not depend on  $V$ .

If every outcome contract trades, we can choose  $\mathbb{P}^V = \mathbb{P}^M$  and make the second term zero. Hence

$$\arg \max_{\mathbb{P}^V} \mathbb{E}^M \log \frac{\mathbb{P}^V}{\mathbb{P}^Y} = \mathbb{P}^M, \quad \frac{M(1)}{M(0)} = \frac{\mathbb{P}^M}{\mathbb{P}^Y}.$$

The proof is the logarithmic identity

$$\log \frac{\mathbb{P}^V}{\mathbb{P}^Y} = \log \frac{\mathbb{P}^M}{\mathbb{P}^Y} - \log \frac{\mathbb{P}^M}{\mathbb{P}^V}.$$

**Starting from log-utility maximization leads to the likelihood ratio.**

# Performance equals information

For the ideal portfolio  $M$ , investor beliefs become its probability law:

$$\frac{M(1, \omega)}{M(0)} = \frac{\mathbb{P}^M(\omega)}{\mathbb{P}^Y(\omega)}.$$

$$\underbrace{\mathbb{E}^M \left[ \log \frac{M(1)}{M(0)} \right]}_{\text{performance}} = \mathbb{E}^M \left[ \log \frac{\mathbb{P}^M}{\mathbb{P}^Y} \right] = \underbrace{D_{\text{KL}}(\mathbb{P}^M \| \mathbb{P}^Y)}_{\text{information}}.$$

For our binary beliefs  $(2/3, 1/3)$  and market law  $(1/2, 1/2)$ ,

$$\mathbb{E}^M \log \frac{M(1)}{M(0)} = 0.056633, \quad e^{0.056633} \approx 1.0583.$$

**About 5.8% long-run geometric growth per period, if the model is correct.**

This is a geometric rate under  $\mathbb{P}^M$ , not the arithmetic expected return or a guaranteed one-period gain.

## Utility selects a law; log optimality reappears

Let utility  $U$ , beliefs  $\mathbb{P}^M$ , and market law  $\mathbb{P}^Y$  select a positive unit-cost payoff:

$$G_U \in \arg \max_{G>0, \mathbb{E}^Y[G]=1} \mathbb{E}^M[U(G)].$$

Its induced probability and likelihood payoff are

$$\mathbb{P}^U = \mathbb{P}^Y G_U, \quad G_U = \frac{\mathbb{P}^U}{\mathbb{P}^Y}.$$

For any other positive unit-cost payoff  $\mathbb{P}^V/\mathbb{P}^Y$ , the same KL identity gives

$$\mathbb{E}^U[\log G_U] - \mathbb{E}^U\left[\log \frac{\mathbb{P}^V}{\mathbb{P}^Y}\right] = D_{\text{KL}}(\mathbb{P}^U \parallel \mathbb{P}^V) \geq 0.$$

**The selected payoff is log-optimal under its own induced law  $\mathbb{P}^U$ .**

For log utility under  $\mathbb{P}^M$ ,  $\mathbb{P}^U = \mathbb{P}^M$ . Changing utility generally changes the associated probability.

## CRRA selects a geometric blend of probabilities

For constant relative risk aversion  $\gamma > 0$ , use  $U_\gamma(g) = g^{1-\gamma}/(1-\gamma)$  when  $\gamma \neq 1$ .

The full-menu first-order condition gives

$$\mathbb{P}^U(\omega) = \frac{\mathbb{P}^M(\omega)^{1/\gamma} \mathbb{P}^Y(\omega)^{1-1/\gamma}}{\sum_{\eta} \mathbb{P}^M(\eta)^{1/\gamma} \mathbb{P}^Y(\eta)^{1-1/\gamma}}.$$

In the binary example, with  $\gamma = 2$ ,

$$\mathbb{P}^U(\text{Win}) = \frac{\sqrt{2}}{1 + \sqrt{2}} = 2 - \sqrt{2} \approx 0.586.$$

The Win probability moves from the belief toward the market:

$$\mathbb{P}^Y(\text{Win}) = 0.500 < \mathbb{P}^U(\text{Win}) \approx 0.586 < \mathbb{P}^M(\text{Win}) \approx 0.667.$$

**Risk preferences select  $\mathbb{P}^U$ . The payoff  $\mathbb{P}^U/\mathbb{P}^Y$  is log-optimal under that law.**

At  $\gamma = 1$ , logarithmic utility gives  $\mathbb{P}^U = \mathbb{P}^M$ . The blend is convex in log probabilities for  $\gamma \geq 1$ .

## An incomplete prediction market

Suppose only the **Win contract and cash** trade. The contract's gross payoff is  $R_W = (2, 0, 0)$ .

$$V_w = 1 + w(R_W - 1), \quad \max_w \{0.6 \log(1+w) + 0.4 \log(1-w)\} \Rightarrow w^* = 0.6 - 0.4 = 0.2.$$

| Outcome | Market $\mathbb{P}^Y$ | Desired $\mathbb{P}^M$ | Attainable $\mathbb{P}^V$ |
|---------|-----------------------|------------------------|---------------------------|
| Win     | 0.50                  | 0.60                   | 0.60                      |
| Draw    | 0.30                  | 0.30                   | 0.24                      |
| Loss    | 0.20                  | 0.10                   | 0.16                      |

$$\underbrace{0.0400782}_{\text{ideal information}} = \underbrace{0.0201355}_{\text{attained log growth}} + \underbrace{0.0199427}_{\text{KL gap}}.$$

**Win is matched. The Draw/Loss split remains the market's 60:40.**

$V_{0.2} = (1.2, 0.8, 0.8)$  and  $\mathbb{P}^V = \mathbb{P}^Y V_{0.2}$ : for example, Draw has mass  $0.30 \times 0.80 = 0.24$ .

## A second traded asset: the binomial stock

Cash  $Y$  has gross return one. Now introduce a positive traded stock  $X$ , with two possible one-step returns:

$$R := \frac{X(1)}{X(0)} = \begin{cases} u = 3/2, & H \text{ (up)}, \\ d = 1/2, & T \text{ (down)}. \end{cases}$$

Its payoff induces the **stock-numeraire measure**

$$\mathbb{P}^X(\omega) = \mathbb{P}^Y(\omega)R(\omega), \quad R = \frac{\mathbb{P}^X}{\mathbb{P}^Y}.$$

Repeating the up/down rule gives a price process:  $X(n+1) = X(n)R(n+1)$ .

The investor's belief is separate:  $\mathbb{P}^M(H) = 2/3$ .

**Starting from the stock dynamics, how do we recover  $\mathbb{P}^Y$  and  $\mathbb{P}^X$ ?**

# Classical pricing: recover probabilities from the dynamics

Let  $q_{\uparrow} = \mathbb{P}^Y(H)$  and  $p_{\uparrow} = \mathbb{P}^X(H)$ , with  $0 < d < 1 < u$ .

**Stock measured in cash is a martingale:**

$$\mathbb{E}^Y[R] = q_{\uparrow}u + (1 - q_{\uparrow})d = 1,$$

$$q_{\uparrow} = \frac{1 - d}{u - d} = \frac{1}{2}.$$

**Cash measured in stock is a martingale:**

$$\mathbb{E}^X[1/R] = \frac{p_{\uparrow}}{u} + \frac{1 - p_{\uparrow}}{d} = 1,$$

$$p_{\uparrow} = \frac{u(1 - d)}{u - d} = \frac{3}{4}.$$

Thus  $p_{\uparrow} = uq_{\uparrow}$  and  $1 - p_{\uparrow} = d(1 - q_{\uparrow})$ , recovering both likelihood ratios.

**Fundamental theorems of asset pricing:** no arbitrage gives existence of a martingale pricing law. Completeness gives uniqueness. More outcomes can leave several candidate laws.

$\mathbb{P}^M(H) = 2/3$  remains a belief input. Pricing recovers the numeraire laws, not the investor's forecast.

# Classical replication: the delta hedge

Given a target payoff  $V(1, \omega)$ , hold  $\Delta$  shares and  $B$  units of cash:

$$V(1, H) = \Delta \cdot X(1, H) + B,$$

$$V(1, T) = \Delta \cdot X(1, T) + B.$$

Subtract the two equations:

$$\Delta = \frac{V(1, H) - V(1, T)}{X(1, H) - X(1, T)}.$$

$$B = V(0) - \Delta \cdot X(0), \quad V(0) = \mathbb{E}^Y[V(1)].$$

**$\Delta$  is an absolute position, measured in shares.**

## Likelihood replication: matching probabilities

Divide the payoff equation by  $V(0)$  and multiply by  $\mathbb{P}^Y$ :

$$\mathbb{P}^V = w\mathbb{P}^X + (1 - w)\mathbb{P}^Y, \quad w = \frac{\Delta \cdot X(0)}{V(0)}.$$

For the target  $\mathbb{P}^V = \mathbb{P}^M$ , the up-state equation gives

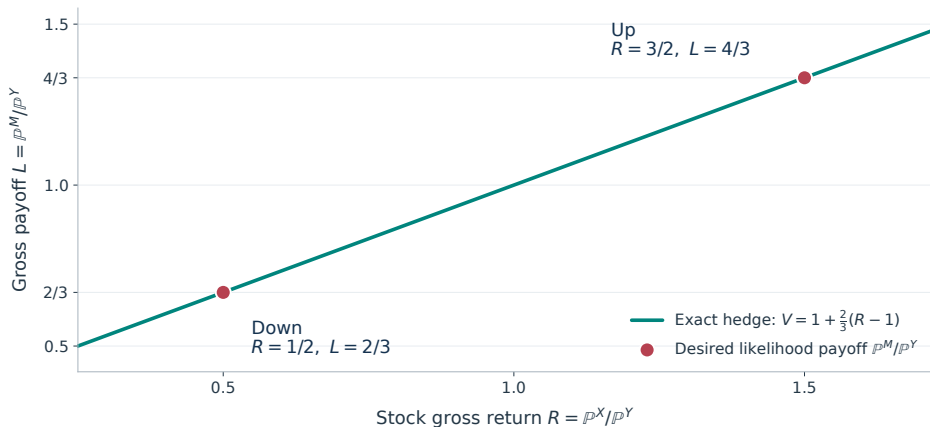
$$m_{\uparrow} := \mathbb{P}^M(H) = \frac{2}{3} = wp_{\uparrow} + (1 - w)q_{\uparrow}, \quad w = \frac{m_{\uparrow} - q_{\uparrow}}{p_{\uparrow} - q_{\uparrow}} = \frac{2}{3}.$$

$$\Delta = w \cdot \frac{V(0)}{X(0)}, \quad B = (1 - w)V(0).$$

**The two formulations give exactly the same hedge.**

$w$  is the initial stock weight. Write  $V_w := V(1)/V(0) = 1 + w(R - 1)$  for normalized wealth.

## Two outcomes, one exact hedge



The line  $V_w = 1 + w(R - 1)$  passes through both desired payoffs.

## Notation recap: targets and traded portfolios

| Symbol           | Meaning                                                                                               |
|------------------|-------------------------------------------------------------------------------------------------------|
| $\mathbb{P}^M$   | Investor beliefs.                                                                                     |
| $\mathbb{P}^Y$   | Market probabilities, also called the pricing law.                                                    |
| $\mathbb{P}^U$   | Target law selected by full-menu utility maximization; $\mathbb{P}^U = \mathbb{P}^M$ for log utility. |
| $\mathbb{P}^X$   | The traded stock's probability law, its stock-numeraire measure.                                      |
| $\mathbb{P}_w^V$ | The law of an attainable stock/cash portfolio.                                                        |

$$R = \frac{\mathbb{P}^X}{\mathbb{P}^Y}, \quad L_M = \frac{\mathbb{P}^M}{\mathbb{P}^Y}, \quad V_w = 1 + w(R - 1).$$

$$\mathbb{P}_w^V = w\mathbb{P}^X + (1 - w)\mathbb{P}^Y.$$

$w = \Delta \cdot X(0)/V(0)$  is the initial stock fraction. For  $0 \leq w \leq 1$ , this is a convex mixture.

# Binomial information accumulates linearly

After  $n$  independent periods, let  $K(n)$  count up moves. Then

$$\frac{X(n)}{X(0)} = u^{K(n)} d^{n-K(n)}, \quad L_M(n) = \left( \frac{m_{\uparrow}}{q_{\uparrow}} \right)^{K(n)} \left( \frac{1 - m_{\uparrow}}{1 - q_{\uparrow}} \right)^{n-K(n)}.$$

$$D_{\text{KL}}(\mathbb{P}^M(n) \parallel \mathbb{P}^Y(n)) = n \left[ m_{\uparrow} \log \frac{m_{\uparrow}}{q_{\uparrow}} + (1 - m_{\uparrow}) \log \frac{1 - m_{\uparrow}}{1 - q_{\uparrow}} \right].$$

In our example,

$$D_{\text{KL}}(\mathbb{P}^M(n) \parallel \mathbb{P}^Y(n)) = n \times 0.056633.$$

## What changes if we can trade only after the first up move completes a cycle?

The negative-binomial family will preserve linear information flow while exposing the effect of this trading restriction.

## Another family with linear information flow

At each integer time  $n \geq 1$ , consider the **one-parameter negative-binomial family**

$$\mathbb{P}_a(Z(n) = k) = \binom{n+k-1}{k} (1-a)^n a^k, \quad k = 0, 1, \dots,$$

with shape  $n$  fixed by time and parameter  $0 < a < 1$ .

For beliefs  $\mathbb{P}_{a_M}$  and market probabilities  $\mathbb{P}_{a_Y}$ ,

$$D_{\text{KL}}(\mathbb{P}_{a_M}(n) \parallel \mathbb{P}_{a_Y}(n)) = n \left[ \log \frac{1-a_M}{1-a_Y} + \frac{a_M}{1-a_M} \log \frac{a_M}{a_Y} \right].$$

**At  $n = 1$  the law is geometric: down moves before the first up. Summing independent cycles gives  $Z(n)$ .**

Log-negative-binomial pricing precedent: Heston (1993).

# The geometric shock comes from the binary tree

One model period consists of  $G$  down moves followed by the first up move. The binary up probabilities remain  $1/2$ ,  $3/4$ , and  $2/3$ :

$$\mathbb{P}^Y(G = g) = \left(\frac{1}{2}\right)^{g+1}, \quad \mathbb{P}^X(G = g) = \frac{3}{4} \left(\frac{1}{4}\right)^g, \quad \mathbb{P}^M(G = g) = \frac{2}{3} \left(\frac{1}{3}\right)^g.$$

Thus the same  $u = 3/2$  and  $d = 1/2$  give

$$\boxed{\frac{X(1)}{X(0)} = R = u d^G = \frac{3}{2} \left(\frac{1}{2}\right)^G, \quad L_M = \frac{4}{3} \left(\frac{2}{3}\right)^G.}$$

After  $n$  completed shocks,

$$Z(n) = G(1) + \cdots + G(n), \quad \frac{X(n)}{X(0)} = u^n d^{Z(n)}.$$

$G$  counts down moves. The previous family's parameters are  $a_Y = 1/2$ ,  $a_X = 1/4$ , and  $a_M = 1/3$ .

## The same terminal law allows different trading rules

If trading is possible after *every binary step*, rebalancing to  $w = 2/3$  gives wealth multipliers

$$1 + \frac{2}{3}(u - 1) = \frac{4}{3}, \quad 1 + \frac{2}{3}(d - 1) = \frac{2}{3}.$$

The first up move completes the shock. Its duration is  $\tau = G + 1$  binary steps:

$$\frac{M(\tau)}{M(0)} = \frac{4}{3} \left( \frac{2}{3} \right)^G = L_M.$$

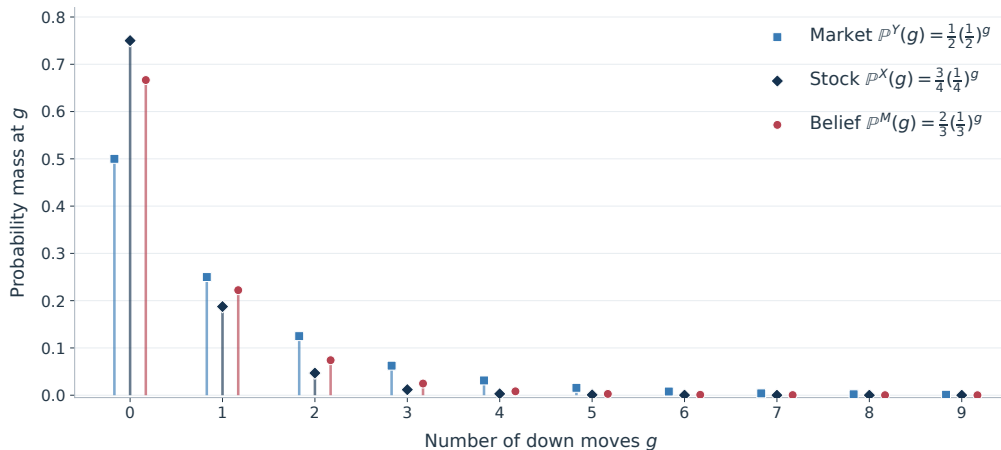
Wald's identity explains the larger information per completed shock:

$$\underbrace{0.0849495}_{\text{one geometric shock}} = \underbrace{\mathbb{E}^M[\tau]}_{3/2} \times \underbrace{0.0566330}_{\text{one binary step}}.$$

Expected total log gain = expected number of binary steps  $\times$  expected log gain per step.

**Our geometric market allows trading only between completed shocks.**

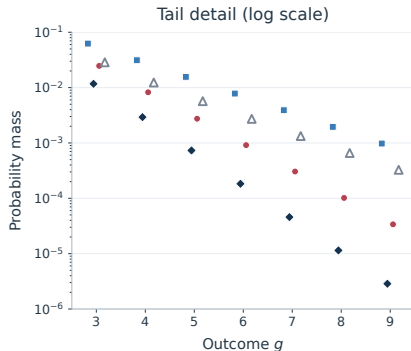
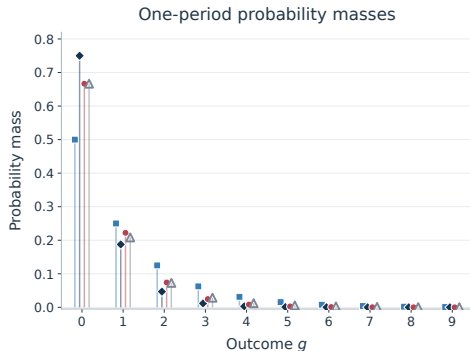
# Stock, cash, and the desired probability



$\mathbb{P}^M$  is the target distribution. Its desired payoff is  $\mathbb{P}^M/\mathbb{P}^Y$ .

# The binomial weight no longer matches every outcome

With stock and cash,  $V_w = 1 + w(R - 1)$  and  $\mathbb{P}_w^V = w\mathbb{P}^X + (1 - w)\mathbb{P}^Y$ .



■ Market  $\mathbb{P}^Y$     ◆ Stock  $\mathbb{P}^X$     ● Target  $\mathbb{P}^M$     △ Binomial mixture  $\mathbb{P}^V$  ( $w = 2/3$ )

$\mathbb{P}_{2/3}^V = \frac{2}{3}\mathbb{P}^X + \frac{1}{3}\mathbb{P}^Y$  matches the pooled tail, but not each outcome.

# The same probability mismatch, three criteria

We want  $\mathbb{P}^M$ , but can attain only  $\mathbb{P}_w^V = w\mathbb{P}^X + (1-w)\mathbb{P}^Y$ . Write  $e_g = \mathbb{P}_w^V(g) - \mathbb{P}^M(g)$ .

| Criterion        | Loss to minimize                                                                                                             |
|------------------|------------------------------------------------------------------------------------------------------------------------------|
| Exact KL (Kelly) | $D_{\text{KL}}(\mathbb{P}^M \parallel \mathbb{P}_w^V) = \sum_g \mathbb{P}_g^M \log \frac{\mathbb{P}_g^M}{\mathbb{P}_w^V(g)}$ |
| KL2              | $\frac{1}{2} \sum_g \frac{e_g^2}{\mathbb{P}_g^M} = \frac{1}{2} \chi^2(\mathbb{P}_w^V \parallel \mathbb{P}^M)$                |
| Quadratic hedge  | $\sum_g \frac{e_g^2}{\mathbb{P}_g^Y} = \mathbb{E}^Y[(V_w - L_M)^2]$                                                          |

For convex  $f$  with  $f(1) = 0$ , an  $f$ -**divergence** is  $D_f(\mathbb{P} \parallel \mathbb{Q}) = \sum_g \mathbb{Q}_g f(\mathbb{P}_g / \mathbb{Q}_g)$ .

With  $(\mathbb{P}, \mathbb{Q}) = (\mathbb{P}_w^V, \mathbb{P}^M)$ ,  $f(t) = -\log t$  gives exact KL and  $f(t) = (t-1)^2/2$  gives KL2. The quadratic hedge instead uses the third law  $\mathbb{P}^Y$  in its denominator.

## Minimal information loss

The attainable distribution is  $\mathbb{P}_w^V = w\mathbb{P}^X + (1 - w)\mathbb{P}^Y$ .

$$\min_w D_{\text{KL}}(\mathbb{P}^M \parallel \mathbb{P}_w^V) \iff \max_w \mathbb{E}^M \log[1 + w(R - 1)].$$

The interior first-order condition is

$$\mathbb{E}^M \left[ \frac{R - 1}{1 + w(R - 1)} \right] = 0.$$

Strict concavity gives a unique optimum in this example:

$$w_{\text{KL}} \approx 0.745.$$

**The penalty is the expected log growth lost through imperfect matching.**

The domain requires  $V_w(g) > 0$  in every state here:  $-2 < w \leq 1$ . Negative wealth cannot induce a probability law.

## Quadratic hedging gives an analytical solution

The target  $L_M = \mathbb{P}^M/\mathbb{P}^Y$  and portfolio  $V_w = 1 + w(R - 1)$  both cost one. Their pricing-measure error has mean zero, so

$$\min_w \mathbb{E}^Y[(L_M - V_w)^2] = \min_w \text{Var}^Y(L_M - V_w).$$

The linear first-order equation gives the regression coefficient

$$w_Q = \frac{\text{Cov}^Y(L_M, R)}{\text{Var}^Y(R)} = \frac{\mathbb{E}^M[R] - 1}{\text{Var}^Y(R)} = \frac{1/5}{2/7} = \frac{7}{10}.$$

**Analytical tractability explains the appeal of quadratic and mean–variance methods.**

Unconstrained solutions use linear equations and explicit weights. The criterion alone allows negative terminal wealth.

Here exact KL instead requires solving a nonlinear first-order equation numerically.

Föllmer–Sondermann (1986), pricing-measure hedging; Schweizer (1995), physical-measure hedging.

# The precise link to mean–variance optimization

For any unit-cost payoff  $V$  and  $L_M = \mathbb{P}^M / \mathbb{P}^Y$ ,

$$\mathbb{E}^Y[(L_M - V)^2] = \underbrace{\mathbb{E}^Y[L_M^2]}_{\text{constant}} + 1 - 2\mathbb{E}^M[V] + \text{Var}^Y(V).$$

Therefore quadratic hedging is exactly equivalent to

$$\max_V \left\{ \mathbb{E}^M[V] - \frac{1}{2} \text{Var}^Y(V) \right\}.$$

This has mean–variance form: the target supplies the expected-payoff term.

Ordinary Markowitz under investor beliefs uses

$$\max_V \left\{ \mathbb{E}^M[V] - \frac{\gamma}{2} \text{Var}^M(V) \right\}.$$

**The connection is exact, with a different probability law for variance.**

## KL2 is a chi-square approximation

For a small relative mismatch  $\delta = (V_w - L_M)/L_M$ ,

$$D_{\text{KL}}(\mathbb{P}^M \parallel \mathbb{P}_w^V) = -\mathbb{E}^M \log(1 + \delta) \approx \frac{1}{2} \mathbb{E}^M[\delta^2].$$

The linear term vanishes because  $\mathbb{E}^M[\delta] = \mathbb{E}^Y[V_w - L_M] = 0$ . Thus the chi-square criterion introduced earlier is the second-order approximation:

$$\text{KL}_2(w) = \frac{1}{2} \chi^2(\mathbb{P}_w^V \parallel \mathbb{P}^M) = \frac{1}{2} \mathbb{E}^Y \left[ \frac{(V_w - L_M)^2}{L_M} \right].$$

Here

$$\text{KL}_2(w) = \frac{24}{65} w^2 - \frac{3}{5} w + \frac{1}{4}, \quad w_{\text{KL}_2} = \frac{13}{16} = 0.8125.$$

**The chi-square denominator is the target  $\mathbb{P}^M$ . The approximation is local.**

## The quadratic hedge beats KL2 in this example

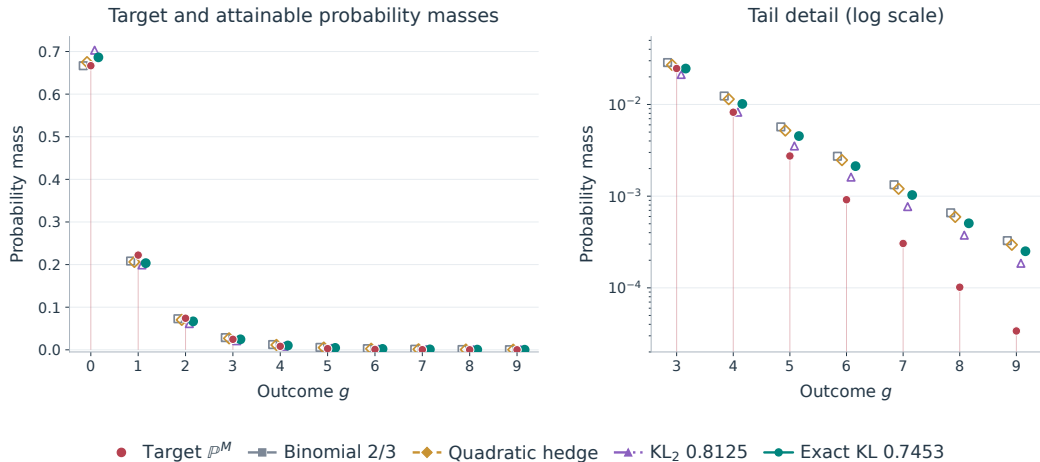
| Portfolio         | Stock weight    | Log growth       | Exact KL gap     |
|-------------------|-----------------|------------------|------------------|
| Binomial weight*  | 0.666667        | 0.0802308        | 0.0047187        |
| KL2 hedge         | 0.812500        | 0.0804213        | 0.0045282        |
| Quadratic hedge   | 0.700000        | 0.0810414        | 0.0039081        |
| <b>KL-optimal</b> | <b>0.745331</b> | <b>0.0814623</b> | <b>0.0034872</b> |

$$\underbrace{\mathbb{E}^M \log V_w}_{\text{attained log growth}} + \underbrace{D_{\text{KL}}(\mathbb{P}^M \| \mathbb{P}_w^V)}_{\text{gap}} = \underbrace{D_{\text{KL}}(\mathbb{P}^M \| \mathbb{P}^Y)}_{0.0849495}.$$

**The quadratic hedge minimizes ordinary squared hedging error by construction.**

\*Applied to the geometric stock. Values are per model period and rounded. KL2 minimizes its own approximation, not exact KL.

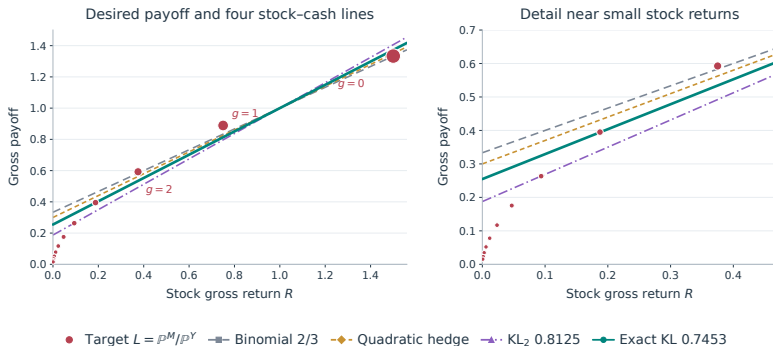
# Different criteria select different attainable probabilities



We want  $\mathbb{P}^M$ , but can implement only  $\mathbb{P}_w^V = w\mathbb{P}^X + (1 - w)\mathbb{P}^Y$ .

# A nonlinear payoff and linear portfolios

A nonlinear target  $L_M(R)$ , approximated by lines  $V_w(R) = 1 + w(R - 1)$ .



KL2 gives more relative weight to small-target crash states through the factor  $1/L_M$ . This favors a lower cash intercept  $1 - w$ , hence more stock here (0.8125 versus 0.700).

These are different regressions. Exact KL also changes the loss to a log ratio, averaged under  $\mathbb{P}^M$ .

# The gap is the value of unavailable rebalancing

The same terminal stock law supports different attainable payoffs.

| Trading opportunities                           | Expected log growth |
|-------------------------------------------------|---------------------|
| Rebalance to 2/3 after each binary step         | 0.0849495           |
| Trade only before and after the completed shock | 0.0814623           |

$$0.0849495 - 0.0814623 = 0.0034872.$$

## The information loss has a concrete trading interpretation.

Both figures are attained growth rates, so their difference is pricing-law invariant. Trading every binary step uniquely fixes  $\mathbb{P}^Y(G = g) = 2^{-(g+1)}$ .

# Hedging and portfolio choice: where does the target come from?

## Hedging a given nonnegative claim $H$

$$H(0) = \mathbb{E}^Y[H] > 0, \quad \mathbb{P}^H(\omega) = \mathbb{P}^Y(\omega) \frac{H(\omega)}{H(0)}.$$

The claim fixes the target probability once the pricing law is chosen.  
The remaining problem is how closely trading can reproduce it.

## Choosing an investment portfolio

- The investor supplies beliefs and risk preferences.
- Markowitz summarizes return beliefs by means and covariances.
- Utility selects the target  $\mathbb{P}^U$ . For log utility,  $\mathbb{P}^U = \mathbb{P}^M$ .

**Utility chooses  $\mathbb{P}^U$ ; its ideal likelihood payoff  $\mathbb{P}^U/\mathbb{P}^Y$  is log-optimal under  $\mathbb{P}^U$ .**

## A concrete change of pricing probabilities

Cash and one stock impose only two pricing equations on countably many outcome probabilities. Many laws therefore price the same assets.

Keep  $Q'_g = P_g^Y$  for  $g \geq 3$ , but set  $Q'_0 = 17/32$ ,  $Q'_1 = 5/32$ ,  $Q'_2 = 3/16$ . Both laws satisfy

$$\sum_g Q'_g = 1, \quad \mathbb{E}^{Q'}[R] = 1.$$

| Stock weight     | Original $Q$ | Alternative $Q'$ |
|------------------|--------------|------------------|
| Quadratic hedge  | 0.700        | 0.641            |
| Exact KL optimum | 0.745        | 0.745            |

Beliefs  $P^M$  and the stock payoff  $R$  stay fixed. The desired ideal payoff  $P^M/Q$  changes.

# The exact portfolio choice survives a change of pricing law

Hold beliefs  $\mathbb{P}^M$ , traded return  $R$ , and feasible portfolios  $V_w$  fixed.

Any equivalent pricing law  $\mathbb{Q}$  satisfying  $\mathbb{E}^{\mathbb{Q}}[R] = 1$  gives

$$\mathbb{P}_{\mathbb{Q}}^{V_w} = \mathbb{Q} V_w,$$

$$D_{\text{KL}}(\mathbb{P}^M \| \mathbb{P}_{\mathbb{Q}}^{V_w}) = D_{\text{KL}}(\mathbb{P}^M \| \mathbb{Q}) - \mathbb{E}^M \log V_w.$$

| Changes with $\mathbb{Q}$            | Stays the same                                           |
|--------------------------------------|----------------------------------------------------------|
| Ideal benchmark and absolute KL gaps | Exact KL-optimal weight                                  |
| Quadratic and KL2 criteria           | Attained expected log growth<br>Ranking by exact KL gaps |

**The decision depends on beliefs and traded payoffs.**

# The information interpretation of hedging

- A target probability specifies a desired likelihood payoff.
- Exact replication matches both the payoff and its probability.
- Incomplete trading permits only a restricted family of probabilities.
- KL minimization retains as much expected log growth as the available assets allow.

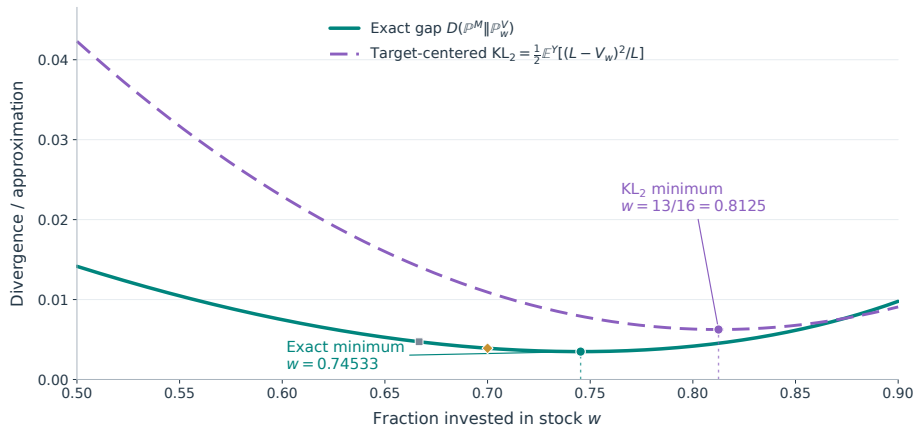
$$\text{attainable log growth} = \text{ideal information} - \text{KL gap}.$$

The decomposition uses the selected pricing law. The optimal attainable growth is invariant to that choice.

Discussion, book, and contact details

# Finite mismatch separates the two criteria

## OPTIONAL DETAIL



**The local approximation changes both the loss and the minimizing weight.**

# This call pays in only one outcome

## OPTIONAL DETAIL

The strike-one call pays  $C = (R - 1)^+$ , where  $R_g = \frac{3}{2}(\frac{1}{2})^g$ .

| Outcome    | Stock return $R_g$ | Call payoff $C_g$ | $\mathbb{P}^Y(G = g)$ |
|------------|--------------------|-------------------|-----------------------|
| $g = 0$    | 1.50               | 0.50              | 1/2                   |
| $g = 1$    | 0.75               | 0                 | 1/4                   |
| $g = 2$    | 0.375              | 0                 | 1/8                   |
| $g \geq 3$ | $R_g \leq 0.1875$  | 0                 | tail mass 1/8         |

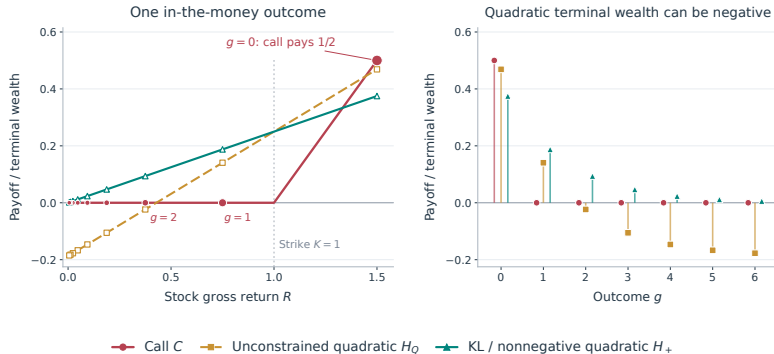
$$C = \frac{1}{2} \mathbf{1}_{\{G=0\}}, \quad C(0) = \mathbb{E}^Y[C] = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}.$$

**Only  $g = 0$  puts the stock above the strike. Every other call payoff is zero.**

# Quadratic fitting can produce negative terminal wealth

## OPTIONAL DETAIL

$$H_Q = \frac{7}{16}R - \frac{3}{16}, \quad \mathbb{E}^Y[(C - H_Q)^2] = \frac{1}{128}. \quad H_+ = \frac{1}{4}R > 0.$$



All portfolios cost  $1/4$ . The unconstrained quadratic portfolio's terminal wealth is negative for  $g \geq 2$ . The stock position is  $\Delta_Q = 7/[16X(0)]$ .

# KL approximation and the wealth constraint

## OPTIONAL DETAIL

At the same budget  $c = C(0) = 1/4$ , let  $H_w = c[1 + w(R - 1)]$ . The call's probability is  $\mathbb{P}^C = \mathbb{P}^Y C/c = \mathbf{1}_{\{G=0\}}$ .

$$\min_{H_w \geq 0} D_{\text{KL}}(\mathbb{P}^C \| \mathbb{P}^{H_w}) = \min_{-2 < w \leq 1} \left\{ -\log \left[ \frac{1}{2} \left( 1 + \frac{w}{2} \right) \right] \right\}.$$

$$w_{\text{KL}} = 1, \quad H_{\text{KL}} = \frac{1}{4}R, \quad D_{\text{KL}}(\mathbb{P}^C \| \mathbb{P}^{H_{\text{KL}}}) = \log \frac{4}{3}.$$

Negative wealth is outside the probability domain. Infeasible portfolios receive  $+\infty$  **in the extended optimization objective**.

Quadratic fitting with the same constraint also selects  $H = R/4$ , with squared error  $1/56$  versus the unconstrained  $1/128$ .

**Quadratic loss alone does not enforce this wealth constraint.**

Zero wealth gives infinite KL where the target is positive (here,  $g = 0$ ).

# The binomial model pools the tail outcomes

## OPTIONAL DETAIL

Retain only  $\mathcal{C} = \{G = 0, G \geq 1\}$  and coarsen the payoffs:

$$\bar{R} = \mathbb{E}^Y[R \mid \mathcal{C}], \quad \bar{L}_M = \mathbb{E}^Y[L_M \mid \mathcal{C}].$$

| Event      | $\mathbb{P}^Y$ | $\mathbb{P}^X$ | $\mathbb{P}^M$ | $\bar{R}$ | $\bar{L}_M$ |
|------------|----------------|----------------|----------------|-----------|-------------|
| $G = 0$    | 1/2            | 3/4            | 2/3            | 3/2       | 4/3         |
| $G \geq 1$ | 1/2            | 1/4            | 1/3            | 1/2       | 2/3         |

$$\underbrace{D_{\text{KL}}(\mathbb{P}_{\mathcal{C}}^M \parallel \mathbb{P}_{\mathcal{C}}^Y)}_{0.056633} \leq \underbrace{D_{\text{KL}}(\mathbb{P}_G^M \parallel \mathbb{P}_G^Y)}_{0.084950}.$$

Finer outcomes carry more information. Cover–Thomas (2006), Ch. 2: data-processing inequality.

# The information discarded by coarsening

## OPTIONAL DETAIL

The KL chain rule gives

$$\begin{aligned} D_{\text{KL}}(\mathbb{P}_G^M \| \mathbb{P}_G^Y) &= D_{\text{KL}}(\mathbb{P}_C^M \| \mathbb{P}_C^Y) \\ &\quad + \sum_{c \in \mathcal{C}} \mathbb{P}^M(c) D_{\text{KL}}(\mathbb{P}^M(G | c) \| \mathbb{P}^Y(G | c)). \end{aligned}$$

The event  $G = 0$  has no remaining uncertainty. On  $G \geq 1$ , geometric memorylessness applies.

$$D_{\text{fine}} = D_{\text{coarse}} + \frac{1}{3} D_{\text{fine}}.$$

Thus,

$$0.0849495 = 0.0566330 + 0.0283165.$$

KL chain rule: Cover–Thomas (2006), Ch. 2.

# Target-relative quadratic error and the near-cash limit

## OPTIONAL DETAIL

For a positive target  $L$  and an equal-cost approximation  $V$ ,

$$D_{\text{KL}}(\mathbb{P}^M \parallel \mathbb{P}^V) \approx \frac{1}{2} \mathbb{E}^Y \left[ \frac{(V - L)^2}{L} \right].$$

This gives the weighted regression coefficient

$$w_{\text{KL}2} = \frac{\mathbb{E}^Y[(R - 1)(L - 1)/L]}{\mathbb{E}^Y[(R - 1)^2/L]}.$$

If  $L = 1 + \varepsilon h$  and  $V = 1 + \varepsilon a$ , with suitable regularity,

$$D_{\text{KL}}(\mathbb{P}^M \parallel \mathbb{P}^V) = \frac{\varepsilon^2}{2} \mathbb{E}^Y[(h - a)^2] + o(\varepsilon^2).$$

**Ordinary quadratic hedging appears in the additional near-cash limit.**

# The closest pricing law determines attainable growth

## OPTIONAL DETAIL

Let  $\mathcal{Q} = \{Q > 0 : \sum_g Q_g = 1, \mathbb{E}^Q[R] = 1\}$  include *all* pricing laws for cash and stock. In this example,

$$\max_w \mathbb{E}^M \log V_w = \min_{Q \in \mathcal{Q}} D_{\text{KL}}(\mathbb{P}^M \| Q) = 0.0814623.$$

At the interior optimum,  $V^* = 1 + w_{\text{KL}}(R - 1)$  and

$$Q_g^* = \frac{\mathbb{P}_g^M}{V^*}, \quad \sum_g Q_g^* = 1, \quad \mathbb{E}^{Q^*}[R] = 1, \quad \boxed{\frac{\mathbb{P}^M}{Q^*} = V^*}.$$

The target relative to  $Q^*$  is attainable. The market remains incomplete for other claims.

Goll–Rüschendorf (2001), §6.2: logarithmic utility and  $f(t) = -\log t$ . Explicitly,  $D_{-\log}(Q \| \mathbb{P}^M) = D_{\text{KL}}(\mathbb{P}^M \| Q)$ .

# Why the optimal reciprocal defines a pricing law

## OPTIONAL DETAIL

The interior first-order condition and  $V^* = 1 + w^*(R - 1)$  give

$$\mathbb{E}^M \frac{R - 1}{V^*} = 0, \quad \mathbb{E}^M \frac{1}{V^*} = 1 - w^* \mathbb{E}^M \frac{R - 1}{V^*} = 1.$$

Thus  $\mathbb{Q}^* = \mathbb{P}^M / V^*$  is a probability and  $\mathbb{E}^{\mathbb{Q}^*}[R] = 1$ . For any  $\mathbb{Q} \in \mathcal{Q}$  and any positive unit-cost stock/cash portfolio  $V$ ,

$$D_{\text{KL}}(\mathbb{P}^M \| \mathbb{Q}) = \mathbb{E}^M \log V + D_{\text{KL}}(\mathbb{P}^M \| \mathbb{Q} V) \geq \mathbb{E}^M \log V.$$

At  $\mathbb{Q}^*$  and  $V^*$ , the last divergence is zero, proving equality.

**This is a dual optimization over pricing laws, with beliefs fixed.**

# The dual law and the numeraire portfolio

## OPTIONAL DETAIL

| Outcome $g$             | 0      | 1      | 2      | 3      |
|-------------------------|--------|--------|--------|--------|
| Original $\mathbb{P}^Y$ | 0.5000 | 0.2500 | 0.1250 | 0.0625 |
| Optimal $\mathbb{Q}^*$  | 0.4857 | 0.2731 | 0.1387 | 0.0626 |

For every admissible stock/cash portfolio,

$$\mathbb{E}^M \left[ \frac{V_w}{V^*} \right] = \mathbb{E}^{\mathbb{Q}^*} [V_w] = 1.$$

Thus  $V^*$  is the **numeraire portfolio** within this feasible set.

$\mathbb{Q}^*$  prices the terminal stock, but does not price the intermediate binary steps. Allowing rebalancing imposes those additional pricing constraints.

Long (1990); Karatzas–Kardaras (2007).

# What does ordinary Markowitz give here?

## OPTIONAL DETAIL

With both moments under  $\mathbb{P}^M$  and risk coefficient  $\gamma = 1$ ,

$$\text{Var}^M(R) = \frac{54}{275}, \quad w_{\text{MV}} = \frac{\mathbb{E}^M[R] - 1}{\text{Var}^M(R)} = \frac{55}{54} \approx 1.01852.$$

Its terminal wealth is

$$V_{\text{MV}} = 1 + \frac{55}{54}(R - 1) = \frac{55R - 1}{54},$$

which is negative precisely when  $G \geq 7$ , with probability

$$\mathbb{P}^M(G \geq 7) = \left(\frac{1}{3}\right)^7 = \frac{1}{2187} \approx 0.0457\%.$$

**Log wealth is undefined on these states: the unconstrained portfolio is inadmissible for log utility.**

Positive wealth requires  $-2 < w \leq 1$ . With that constraint, this Markowitz objective selects  $w = 1$ .

# Selected references: likelihoods and information

## OPTIONAL DETAIL

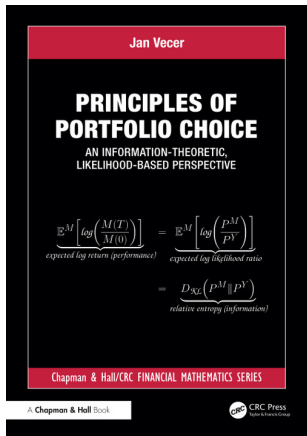
- Kelly, J. L. (1956). A new interpretation of information rate. *Bell System Technical Journal* 35, 917–926.
- Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*, 2nd ed. Wiley, Chapter 2.
- Geman, H., El Karoui, N. and Rochet, J.-C. (1995). Changes of numeraire, changes of probability measure and option pricing. *Journal of Applied Probability* 32, 443–458.
- Heston, S. L. (1993). Invisible parameters in option prices. *Journal of Finance* 48, 933–947.
- Vecer, J. (2026). Analytical solution for Kelly's criterion for multiple outcomes. *Quantitative Finance*. doi:10.1080/14697688.2026.2629952.

# Selected references: hedging and the dual portfolio

## OPTIONAL DETAIL

- Markowitz, H. (1952). Portfolio selection. *Journal of Finance* 7, 77–91.
- Föllmer, H. and Sondermann, D. (1986). Hedging of non-redundant contingent claims. In *Contributions to Mathematical Economics*, 205–223.
- Schweizer, M. (1995). Variance-optimal hedging in discrete time. *Mathematics of Operations Research* 20, 1–32.
- Long, J. B., Jr. (1990). The numeraire portfolio. *Journal of Financial Economics* 26, 29–69.
- Karatzas, I. and Kardaras, C. (2007). The numéraire portfolio in semimartingale financial models. *Finance and Stochastics* 11, 447–493.
- Goll, T. and Rüschendorf, L. (2001). Minimax and minimal distance martingale measures and their relationship to portfolio optimization. *Finance and Stochastics* 5, 557–581.

# Further reading and staying in touch



## Principles of Portfolio Choice

An Information-Theoretic,  
Likelihood-Based Perspective

Jan Vecer

[janvecer.com](http://janvecer.com)



Website and research



Connect on LinkedIn