Risk Modeling


DILBERT Scott Adams
DOGBERT THE FINANCIAL ADVISOR
YOU SHOULD INVEST ALL OF YOUR MONEY IN DISEASED LIVESTOCK.
scottadams@aol.com
www.dilbert.com
IT WOULD BE UNWISE TO INVEST IN JUST ONE SICK COW, BUT IF YOU AGGREGATE A BUNCH OF THEM TOGETHER, THE RISK GOES AWAY.
12-13-08 © 2008 Scott Adams, Inc./Dist. by UFS, Inc.
IT'S CALLED MATH.
SUDDENLY I FEEL ALL SAVVY.

Defining Risk as Standard Deviation

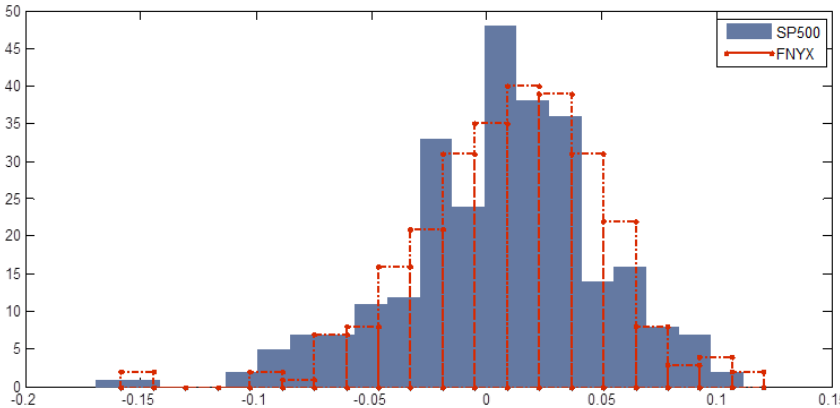
- That's what Harry Markowitz used.
- It has several important and useful properties:
  - Symmetric
  - Well-understood statistical properties.
  - Machinery exists for aggregation from asset to portfolio.

$$\sigma_P^2 = \mathbf{h}_P^T \cdot \mathbf{V} \cdot \mathbf{h}_P$$

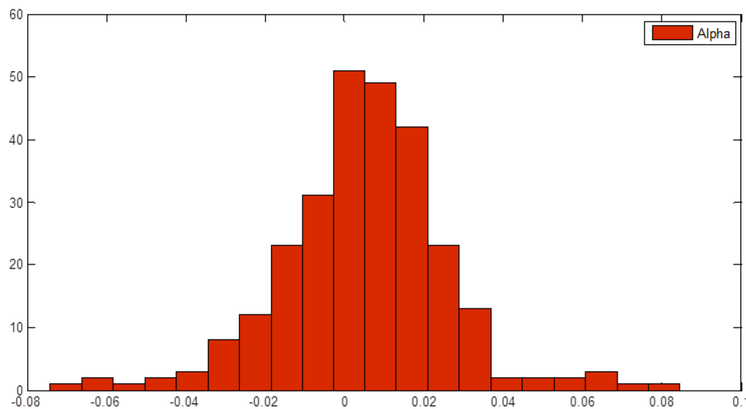
  - Predictable

Issues

- Non-normal distributions
  - Fat tails (kurtosis)
  - Skewness
- Other choices:
  - Semivariance
  - Shortfall probability
  - Value at Risk



- Example:
  - S&P500
  - Fidelity Contrafund
  - June 1989 – December 2011



Why Risk Models?

- Let's look at how we aggregate portfolio risk:

$$\sigma_P^2 = \mathbf{h}_P^T \cdot \mathbf{V} \cdot \mathbf{h} = \left[h_1\ h_2 \ldots h_N\right] \cdot \begin{bmatrix} \sigma_1^2\ \sigma_{12} \ldots \sigma_{1N} \\ \sigma_{12}\ \sigma_2^2 \ldots \\ \vdots \qquad \ddots \\ \sigma_{1N} \ldots \qquad \sigma_N^2 \end{bmatrix} \cdot \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_N \end{bmatrix}$$

$$= h_1^2 \cdot \sigma_1^2 + h_2^2 \cdot \sigma_2^2 + \ldots h_N^2 \cdot \sigma_N^2$$
$$+2h_1 \cdot h_2 \cdot \sigma_{12} + 2h_1 \cdot h_3 \cdot \sigma_{13} + \ldots + 2h_{N-1} \cdot h_N \cdot \sigma_{N-1,N}$$

- For N assets, we require $N(N+1)/2$ parameters. If N=1,000, we need to estimate 500,500 parameters.
- That is the challenge.

Historical Risk

- This is the most straight forward approach. Why doesn't it work?
- Observe N asset returns over T periods. We estimate:

$$\hat{\sigma}_{ij} = \left(\frac{1}{T-1}\right) \cdot \sum_{t=1}^{T} \left[r_i(t) - \bar{r}_i \right] \cdot \left[r_j(t) - \bar{r}_j \right]$$

- What are the problems with doing this?

Problems with historical risk

- Let's think about the number of parameters.
- We need to estimate $N(N+1)/2$ parameters, and we have NT observations. We require at a minimum, 2 observations per parameter estimate. (How can we estimate a variance from 1 number?)
- Hence, $NT \geq N(N+1)$, or $T > N$. This causes problems when we are looking at monthly returns for 1,000 assets.
  – Technical version: unless $T > N$, we will estimate a singular covariance matrix. What does that mean, mathematically? Intuitively?
- And even if we had 1,000 months of data (or more), we know that assets and markets change over time. We also regularly observe new assets. We don't have 1,000 months of data on Alphabet.

Single Factor Model

- Sharpe's Market Model*
- Decompose returns into market and residual components.
- Postulate that residual components are uncorrelated.

$$\mathbf{r} = \boldsymbol{\beta} \cdot r_{mkt} + \boldsymbol{\theta}$$

$$\mathbf{V} = \boldsymbol{\beta} \cdot \boldsymbol{\beta}^T \cdot \sigma_{mkt}^2 + \boldsymbol{\Delta}$$

$$\boldsymbol{\Delta} = \begin{bmatrix} \omega_1^2 & 0 \cdots & & 0 \\ 0 & \omega_2^2 & & \\ \vdots & & \ddots & \\ 0 & & & \omega_N^2 \end{bmatrix}$$

$$U = h^T \cdot \alpha - \lambda h^T V h + \text{constraints}$$

$$Corr\{r_n, r_m\} = \frac{\beta_n \cdot \beta_m \cdot \sigma_{mkt}^2}{\sqrt{\left(\beta_n^2 \cdot \sigma_{mkt}^2 + \omega_n^2\right) \cdot \left(\beta_m^2 \cdot \sigma_{mkt}^2 + \omega_m^2\right)}}$$

*Perspective: In his article, Sharpe stated that solving a 100-asset problem on an IBM 7090 computer required 33 minutes but this single factor model reduced that to 30 seconds. This model also allowed him to handle 2,000 assets where before he could handle only 249.

Single Factor Model

- Solves the number of parameters problem:
  - For N assets, it requires $2N+1$ parameters.
  - N betas, N residual risks, 1 market volatility.
- Problem with this model: it doesn't capture observed correlation structure in the market. Are ExxonMobil and Chevron correlated only through their market exposure?
- Advantage: simplicity makes this useful for back-of-the-envelope calculations.

Factor Models of Risk

- These are extensions of Sharpe's approach, designed to capture real market issues.

$$\mathbf{r} = \mathbf{X} \cdot \mathbf{b} + \mathbf{u}$$

- Separate returns into common factor and specific (idiosyncratic) pieces. We choose K factors, where $K << N$. The covariance matrix is then:

$$\mathbf{V} = \mathbf{X} \cdot \mathbf{F} \cdot \mathbf{X}^T + \mathbf{\Delta}$$

- The matrix **F** is the covariance matrix of the factors.

Choosing the Factors

- Art not science.
- Three general approaches:
  - Fundamental factors (MSCI/Barra, Northfield, Axioma)
    - Industries and investment themes.
  - Macroeconomic factors (Citigroup RAM, BIRR model)
    - Industrial productivity, inflation, interest rates, oil prices…
  - Statistical factors (Quantal, Northfield, APT(Sungard))
    - Use statistical factor modeling, principal components analysis…
- These three approaches involve different estimation issues, and exhibit different levels of effectiveness.
  - They are not mutually exclusive. Combinations are possible.

$r = X \cdot b + u$

Fundamental Models

Calculate X
Estimate b

- Typically about 60 factors for a major equity market.
- Calculate factor exposures, **X**, from fundamental data.
  - Industry membership.
  - Style exposures (e.g. value based on B/P)
- Run monthly cross-sectional GLS regressions to estimate factor returns.

$$\mathbf{b} = \left(\mathbf{X}^{\mathrm{T}} \cdot \boldsymbol{\Delta}^{-1} \cdot \mathbf{X}\right)^{-1} \cdot \mathbf{X}^{\mathrm{T}} \cdot \boldsymbol{\Delta}^{-1} \cdot \mathbf{r}$$

- Use *N* observations to estimate *K* factor returns.

Aside: Factor Portfolios

- Our extensive use of fundamental models makes it worth understanding factor portfolios in more detail.
- We estimate factor returns as:

$$\mathbf{b} = \left(\mathbf{X}^{\mathbf{T}} \cdot \mathbf{\Delta}^{-1} \cdot \mathbf{X}\right)^{-1} \cdot \mathbf{X}^{\mathbf{T}} \cdot \mathbf{\Delta}^{-1} \cdot \mathbf{r}$$

- This equation has the form:

$$\mathbf{b} = \mathbf{H}^{T} \cdot \mathbf{r}$$

- Each estimated factor return, b_j, is a weighted sum of asset returns. We can interpret those weights as the asset weights in a factor portfolio, or factor-mimicking portfolio.

Factor Portfolios

- The columns of **H** contain the portfolio weights, with one column for each factor.
- The GLS estimation approach guarantees that factor portfolio-*j* has:
  - Unit exposure to factor-*j*
  - Zero exposure to all other factors.
  - Minimum risk.
- Industry factor portfolios are typically fully invested, with long and short positions.
- Risk index factor portfolios are typically net zero investments, with positions in every stock.
- We will find factor portfolios quite useful later in this course.

e.g. Value $= \frac{B}{P}$
we standardize/ winsorize B/P ratios so exposures have mean 0 std 1

$r = X \cdot b + u$

Calculate b
Estimate X

Macroeconomic Models

- Typically about 9 factors for a major equity market.
- Calculate the change (or shock) in each macrovariable each month.
- Estimate exposure to such shocks, stock by stock, using time-series data.

$$r_n(t) = \sum_{k=1}^{K} b_k(t) \cdot X_{nk} + \varepsilon_n(t)$$

N separate regressions

- This approach requires *NK* parameter estimates.

Fundamental factor models require only K parameter estimates.

Statistical Models

- Start with only returns data.
- Use statistical analysis to determine number and identity of most important factors.
  - While this approach sounds completely objective, it involves many subjective decisions, e.g. choosing portfolios to build an initial historical covariance matrix.
- This approach implicitly assumes that factor exposures are constant over estimation period.
- Factors change from month to month.

Performance and Uses*

Barr Rosenberg — UC Berkeley Finance Prof

- Fundamental models
  - In general, best risk forecasting out-of-sample, but dependent on choosing the right factors.
  - Intuitive factors also useful for performance attribution and alpha forecasting.
- Macroeconomic models:
  - Poor at risk forecasting.
  - Direct macroeconomic connections can be useful for alpha forecasting.
- Statistical models:
  - Best in-sample forecasts. Will outperform fundamental models with poorly chosen factors.
  - Misses factors whose exposures change over time (especially momentum).
  - Not useful for performance attribution. Difficult to use for alpha forecasting.

$r = X \cdot b + u$

$E\{r\} = X \cdot E\{b\}$

*See Gregory Connor, "The Three Types of Factor Models: A Comparison of their Explanatory Power." *Financial Analysts Journal*, May-June 1995, pp. 42-46.

Covariance Matrix Estimation

- For fundamental and macroeconomic models (and even to some extent for statistical models), we still need to estimate the covariance matrix, given the factor return history.
- We also need to estimate the specific (idiosyncratic) risk matrix.

An Example

- Want best forecast of portfolio risk over the next several months. (Horizon will depend on use.)
- Given that risk varies over time, we would like to overweight more recent observations.
- At the same time, we have the parameter estimation challenge: we want $T>>K$. Lowering the weight on historical observations effectively lowers T.
- Here is one approach.

US Equity Covariance Matrix

- Historical monthly data back to 1973. Between 65 and 70 factors.
- Step 0: Determine the factors.
- Step 1: Use exponential smoothing to estimate factor covariance matrix*:

$$F_{ij}(T+1)=\frac{\sum_{t=1}^{T}\left[b_i(t)-\overline{b}_i\right]\cdot\left[b_j(t)-\overline{b}_j\right]\cdot Exp\{-\gamma\cdot(T-t)\}}{\sum_{t=1}^{T}Exp\{-\gamma\cdot(T-t)\}}$$

*There are many different approaches to estimating a covariance matrix from a series of returns. This is just one example.

Step 2: Specific Risk Model

- Our monthly factor return estimations also estimate specific returns:

$$\mathbf{r} = \mathbf{X} \cdot \mathbf{b} + \mathbf{u}$$

- We can use a similar approach:

$$Var\left\{u_n\left(T+1\right)\right\} = \frac{\sum_{t=1}^{T}\left[u_n\left(t\right)-\bar{u}_n\right]^2 \cdot Exp\left\{-\gamma' \cdot \left(T-t\right)\right\}}{\sum_{t=1}^{T} Exp\left\{-\gamma' \cdot \left(T-t\right)\right\}}$$

A Useful Property

- We have occasion to invert the covariance matrix, i.e. $\mathbf{V}^{-1} \cdot \boldsymbol{\alpha}, \quad \mathbf{V}^{-1} \cdot \mathbf{e}$
- This is typically an operation of order N^3.
- If the covariance matrix has the factor form, then: $V = XFX^T + \Delta$

$$\mathbf{V}^{-1} = \boldsymbol{\Delta}^{-1} - \boldsymbol{\Delta}^{-1} \cdot \mathbf{X} \cdot \left\{ \mathbf{X}^T \cdot \boldsymbol{\Delta}^{-1} \cdot \mathbf{X} + \mathbf{F}^{-1} \right\}^{-1} \cdot \mathbf{X}^T \cdot \boldsymbol{\Delta}^{-1}$$

- This is sometimes known as the Woodbury identity, after Max Woodbury.

Testing Risk Forecasts

- Given risk forecast $\sigma(t)$ at time t, and return $r(t)$ between t and $t+\Delta t$, how can we test whether $\sigma(t)$ is a good risk forecast?

$t \quad t+\Delta t \quad t+2\Delta t \quad \ldots \quad t+T$

- Bias test:
  - Convert returns to standardized outcomes:

$$x(t) = \frac{r(t)}{\sigma(t)}$$

If $\sigma(t) = \sigma$ (i.e. indep of t)

Compare σ to $\mathrm{Std}\{r(t) \mid t, t+T\}$

  - The bias statistic is the sample standard deviation of these outcomes:

$$bias = StDev\left\{x(t) \mid t = 1, \ldots T\right\}$$

Is $\dfrac{\mathrm{Std}\{r(t) \mid t, T+t\}}{\sigma} = 1$?

  - If the bias>1, we have under-predicted risk, and vice versa.

Testing Risk Forecasts

- Statistical significance: Remember that for normally distributed random numbers:

underpredicted risk: bias > 1

$$SE\{\sigma\} = \frac{\sigma}{\sqrt{2T}}$$

$$\text{bias} > 1 + 2\frac{\sigma}{\sqrt{2T}}$$

- The bias test estimates whether we are accurate on average. We can apply it to total, residual, common factor, and specific risk.
- We can use further tests to see if our forecasts are above average when realized risk is above average, etc.

Total and Active Risk

- Given the covariance matrix, we can estimate total and active risk:

$V = XFX^T + \Delta$

$x_p = X \cdot h_p$

$$\sigma_P^2 = \mathbf{h}_P^T \cdot \mathbf{V} \cdot \mathbf{h}_P = \mathbf{x}_P^T \cdot \mathbf{F} \cdot \mathbf{x}_P + \mathbf{h}_P^T \cdot \mathbf{\Delta} \cdot \mathbf{h}_P$$

$$\psi_P^2 = \mathbf{h}_{PA}^T \cdot \mathbf{V} \cdot \mathbf{h}_{PA} = \mathbf{x}_{PA}^T \cdot \mathbf{F} \cdot \mathbf{x}_{PA} + \mathbf{h}_{PA}^T \cdot \mathbf{\Delta} \cdot \mathbf{h}_{PA}$$

Active Risk: 1

- We can also estimate the correlation of returns from two portfolios:

$$Corr\{r_A, r_B\} = \frac{\mathbf{h}_A^T \cdot \mathbf{V} \cdot \mathbf{h}_B}{\sigma_A \cdot \sigma_B}$$

Active Risk

r_P = portfolio return

r_B = benchmark return

δ_P = active return $= r_P - r_B$

ψ_P = active risk $= \mathrm{Std}\{r_P - r_B\}$

Beta

- We often think of estimating betas using time-series regressions.

$$r_P(t) = \alpha + \beta \cdot r_B(t) + \varepsilon(t)$$

$$\beta = \frac{Cov\{r_P, r_B\}}{Var\{r_B\}}$$

Beta

- In that regression view, we estimate the covariance and variance using the observed time-series.
- But given our covariance matrix, which is a forward-looking forecast, we can (hopefully) develop a better beta forecast:

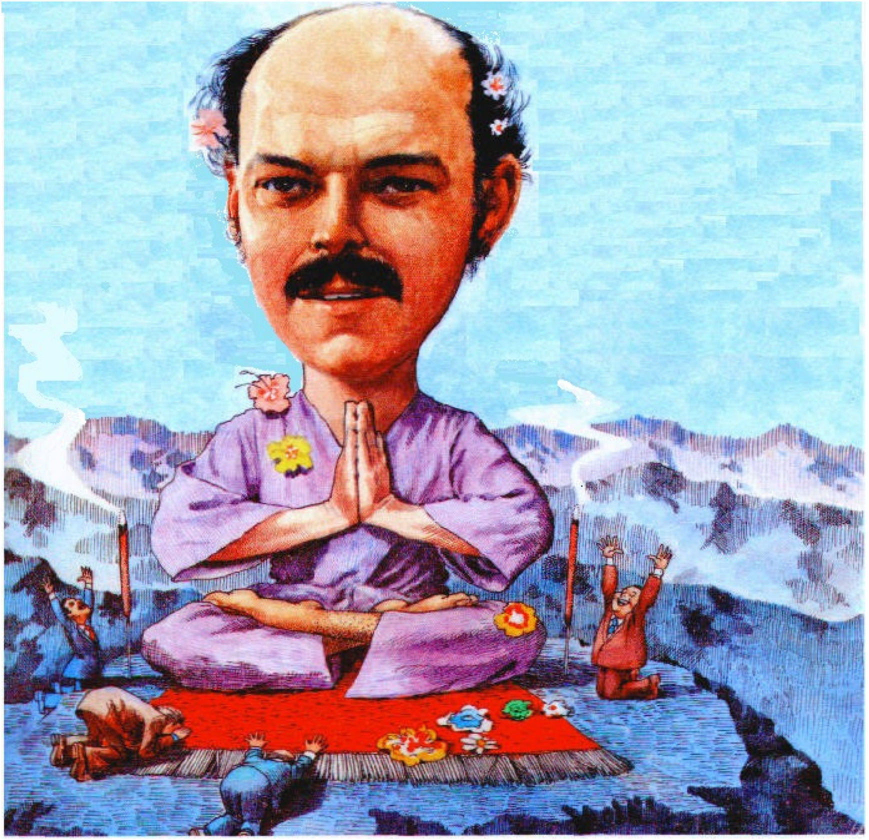
$$\beta_P = \frac{\mathbf{h}_P^T \cdot \mathbf{V} \cdot \mathbf{h}_B}{\sigma_B^2} = \mathbf{h}_P^T \cdot \boldsymbol{\beta}, \quad \boldsymbol{\beta} = \frac{\mathbf{V} \cdot \mathbf{h}_B}{\sigma_B^2}$$

$\sigma_B^2 = h_B^T V h_B$

- These were Barr's better betas.

Barr Rosenberg

Former Berkeley professor of finance



Residual Risk

$$\boxed{\omega_P^2 = \sigma_P^2 - \beta_P^2 \cdot \sigma_B^2 = \psi_P^2 - \beta_{PA}^2 \cdot \sigma_B^2}$$

- The residual covariance matrix:

$$\mathbf{VR} \equiv \mathbf{V} - \boldsymbol{\beta} \cdot \boldsymbol{\beta}^T \cdot \sigma_B^2$$

- Hence:

$$\omega_P^2 = \mathbf{h}_P^T \cdot \mathbf{VR} \cdot \mathbf{h}_P = \mathbf{h}_{PA}^T \cdot \mathbf{VR} \cdot \mathbf{h}_{PA}$$

$$\Rightarrow \mathbf{h}_{PA}^T \cdot \mathbf{V} \cdot \mathbf{h}_{PA} \quad \text{if } \beta_P = 1$$

Marginal Contributions

- It is often useful to know at the margin the significant contributors to total, residual, and active risk:

$$\frac{\partial \sigma_P}{\partial \mathbf{h}_P^T} = \frac{\mathbf{V} \cdot \mathbf{h}_P}{\sigma_P} = \mathbf{MCTR}$$

$$\frac{\partial \omega_P}{\partial \mathbf{h}_{PA}^T} = \frac{\mathbf{VR} \cdot \mathbf{h}_{PA}}{\omega_P} = \mathbf{MCRR}$$

$$\frac{\partial \psi_P}{\partial \mathbf{h}_{PA}^T} = \frac{\mathbf{V} \cdot \mathbf{h}_{PA}}{\psi_P} = \mathbf{MCAR}$$

what is VR?

marginal contribution to total risk

$$\sigma_P^2 = h_P^T V h_P$$

$$\frac{\partial \sigma_P}{\partial h_P^T} = \frac{V \cdot h_P}{\sigma_P}$$

- Why is this useful? Remember first order portfolio construction conditions:

$$\boldsymbol{\alpha} = 2\lambda \mathbf{V} \cdot \mathbf{h}_{PA} = 2\lambda \psi_P \cdot \mathbf{MCAR}$$

- At optimality, alphas are proportional to marginal contributions. Note that we can calculate marginal contributions with only the portfolio and covariance matrix.
- In principle, we can also use marginal contributions to estimate the change in risk as we make small changes in positions. With today's fast computing, this isn't as useful as in the past.

Fundamental Risk Factors

Barra Models:

Industries 30 – 60 of these

Styles / "Risk Indices"

- Value (Book/Price)
- Size
- Momentum
- Volatility
- Dividend Yield
- Earnings Yield
- ⋮

~ 10 of these