pjb25

Portfolio Analysis with Random Portfolios

Patrick Burns
http://www.burns-stat.com

September 2006

Slide 1

pjb25 This was presented in London on 5 September 2006 at an event sponsored by UKSIP.

Patrick Burns, 07/09/2006

pjb1

Random Portfolios -- Outline

- Why
  - Performance measurement
  - Evaluating trading strategies
  - Setting constraint bounds
- How
  - Naïve ideas
  - Genetic algorithms

Slide 2

---

pjb1 This is a non-random selection of applications of random portfolios.

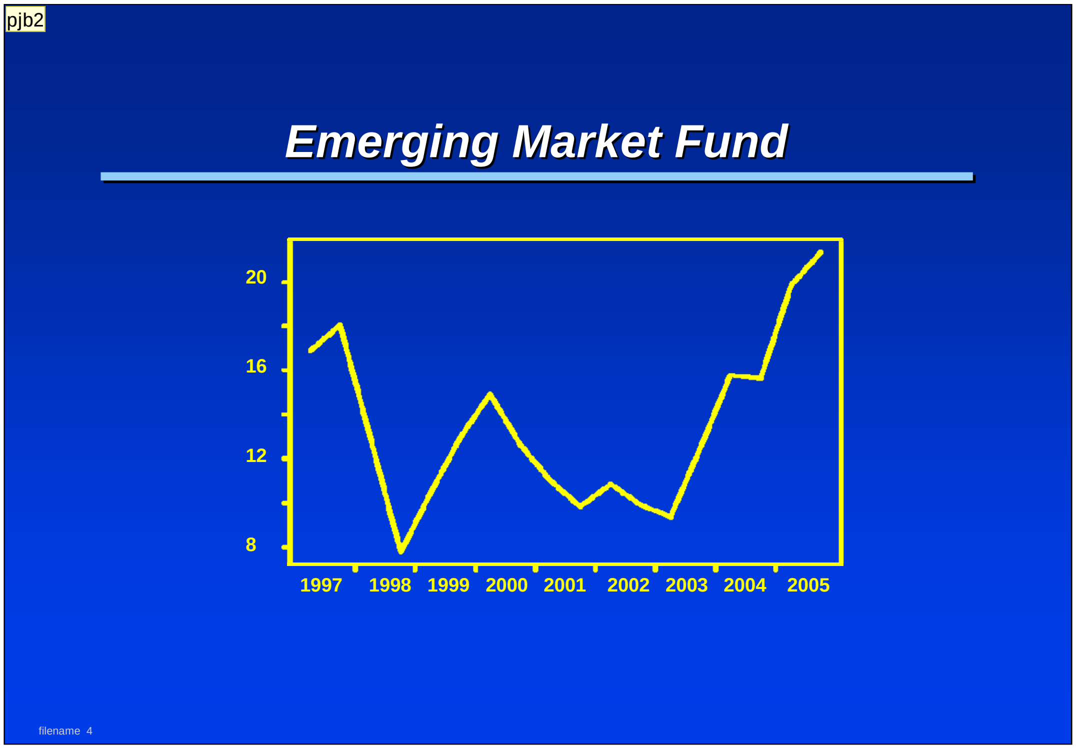
Random portfolios are a general and powerful technique. In my opinion, they should be in the toolbox of every quantitative analyst.

Patrick Burns, 06/09/2006

Why

Performance Measurement

filename 3


pjb2
Emerging Market Fund
20
16
12
8
1997
1998
1999
2000
2001
2002
2003
2004
2005
filename 4

Slide 4

pjb2 We are interested to know if this fund exhibits skill. Let's focus on the 2005 performance. This performance is good, but is it good enough that we can attribute skill to it with some amount of confidence.

Patrick Burns, 06/09/2006

pjb3

Perfect Performance Measurement

- Look at all possible portfolios that the manager might have held
- Take the return of each of these portfolios over the time period
- Compare actual return to the distribution from all possibilities
- Could be another measure instead of the return

Slide 5

pjb3 The "portfolios that the manager might have held" are not all portfolios comprising the fund manager's universe. We are ruling out portfolios that are too concentrated, too volatile, etcetera. We want to include only portfolios that meet the constraints that the fund is under, whether they be explicit or implicit constraints.
Patrick Burns, 07/09/2006

pjb4

The Catch

- The number of possible portfolios is finite, but astronomical
- Almost perfect is to use a random sample of all of the possible portfolios
- So, get random sample of portfolios that obey some given set of constraints

Slide 6

pjb4 The last line is precisely what I mean by 'random portfolio'. Consider the set of all portfolios that satisfy some number of specific constraints. We want a random sample from that set. The sampling pays no attention to whatever utility might be assigned to the various portfolios.

Patrick Burns, 06/09/2006

pjb5

Performance Measurement

- Generate random portfolios that satisfy fund constraints
- Find fraction of random portfolios that outperform the fund in the time period
- That fraction is the p-value of the statistical hypothesis test of no skill for that period

Slide 7

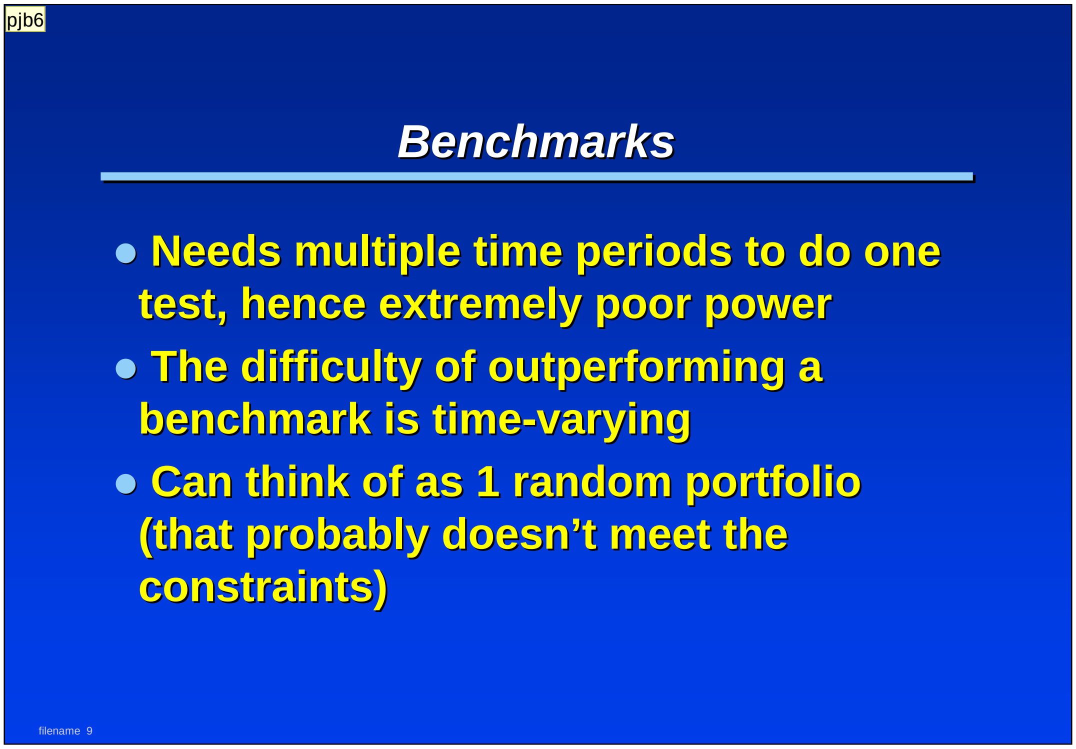
pjb5 Actually the p-value is a slight modification of that fraction -- see "Performance Measurement via Random Portfolios" for details.

How much evidence of skill you attribute to a small p-value depends on your personal taste and possibly to prior information that you have.

Patrick Burns, 07/09/2006

The Usual Suspects

- **Benchmark-relative performance**
- **Peer groups**


pjb6
Benchmarks
Needs multiple time periods to do one test, hence extremely poor power
The difficulty of outperforming a benchmark is time-varying
Can think of as 1 random portfolio (that probably doesn't meet the constraints)
filename 9

Slide 9

---

pjb6 An index like the FTSE 350 or the S&P 500 is not random in the mathematical sense, but it is random in the sense of having zero skill.
Patrick Burns, 06/09/2006

pjb7

Peer Groups

- Compare with "similar" funds
- Not clear what being the p^{th} percentile means
- What if no fund has skill?
- Can think of as "random" portfolios with unknown skill levels

filename 10

Slide 10

pjb7 There is a long list of problems with peer groups -- Ron Surz has written about that. The tension between wanting a lot of peers and wanting funds that do exactly the same thing as our target fund is just one of the problems.

If all funds have the same skill -- whether that value of skill is zero or not -- then the peer group ranking merely gives you the ranking of luck.

When we are thinking about peer groups as a random portfolio technique, we are drawing portfolios that roughly have the same constraints as our target fund. The problem is that we are not selecting randomly, but rather we are selecting portfolios from an unknown distribution of skill.

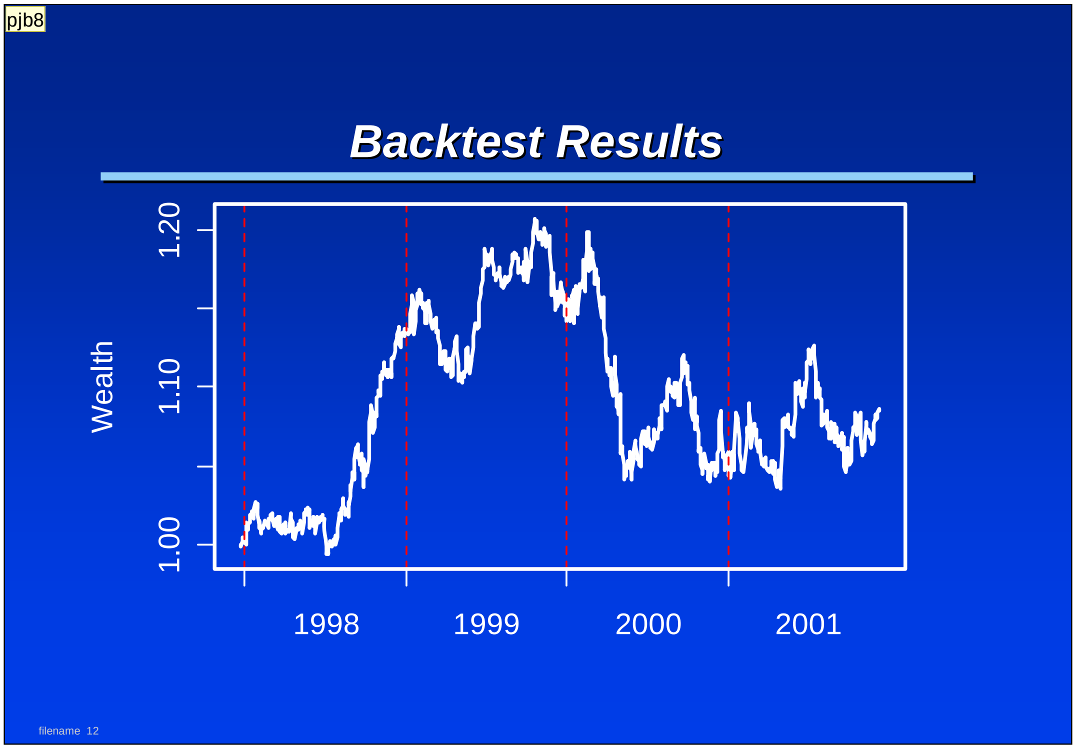
In order to know what the peer group ranking really means, we need to know what the distribution of skill is. But we are doing the peer group in the first place in order to learn that. It's a dog chasing his tail.

Patrick Burns, 07/09/2006

Why

Evaluating Trading Strategies

filename 11


pjb8
Backtest Results
Wealth
1.00
1.10
1.20
1998
1999
2000
2001
filename 12

Slide 12

pjb8 This is a long-short portfolio that is dollar neutral.

In the absence of random portfolios, there is not an especially good way to assess the quality of this (or any) backtest.

Patrick Burns, 06/09/2006

Backtest

- Start with a specific portfolio
- Optimise the trade every day based on predictions and constraints

filename 13

pjb11

Backtest Random Portfolios

- Generate 100 random paths
- Each path starts with the same initial portfolio
- Each path trades randomly each day, obeying the constraints
- Might possibly add constraints based on the optimised trading

filename 14

Slide 14

pjb11 We are creating 100 backtests that mimic the real backtest in everything except that these backtests have exactly zero skill.

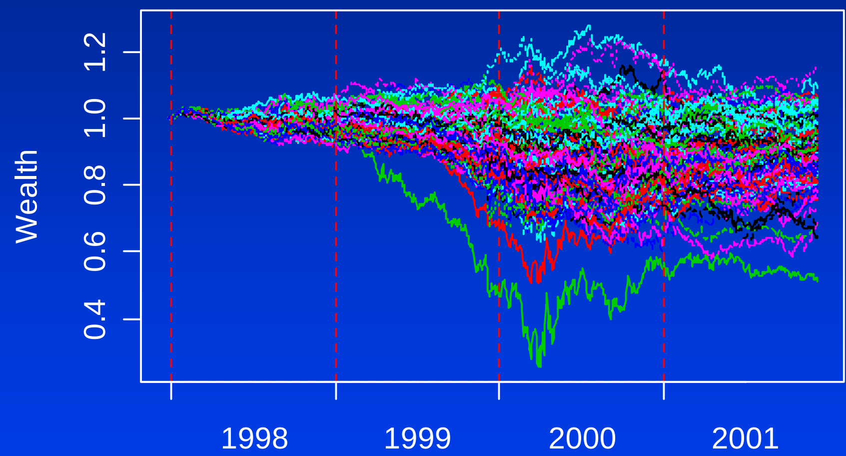
I'll give two examples of constraints that the random paths might be made to obey which are not constraints on the original strategy, but there could be others as well.

If the amount that the strategy trades varies, then the random paths could be forced to trade essentially the same amount at each point as the real strategy trades.

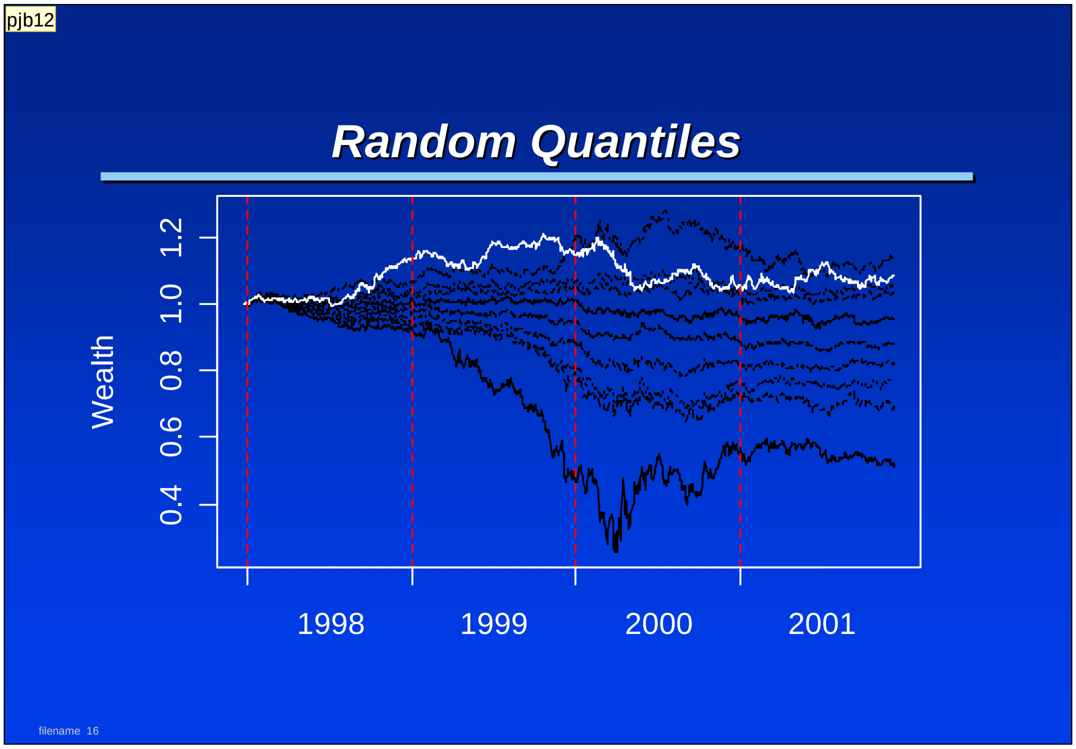
The random paths could be forced to close the same number of positions at each time point as the real strategy happens to close.

Patrick Burns, 07/09/2006

The Random Paths



filename 15


pjb12
Random Quantiles
Wealth
0.4
0.6
0.8
1.0
1.2
1998
1999
2000
2001
filename 16

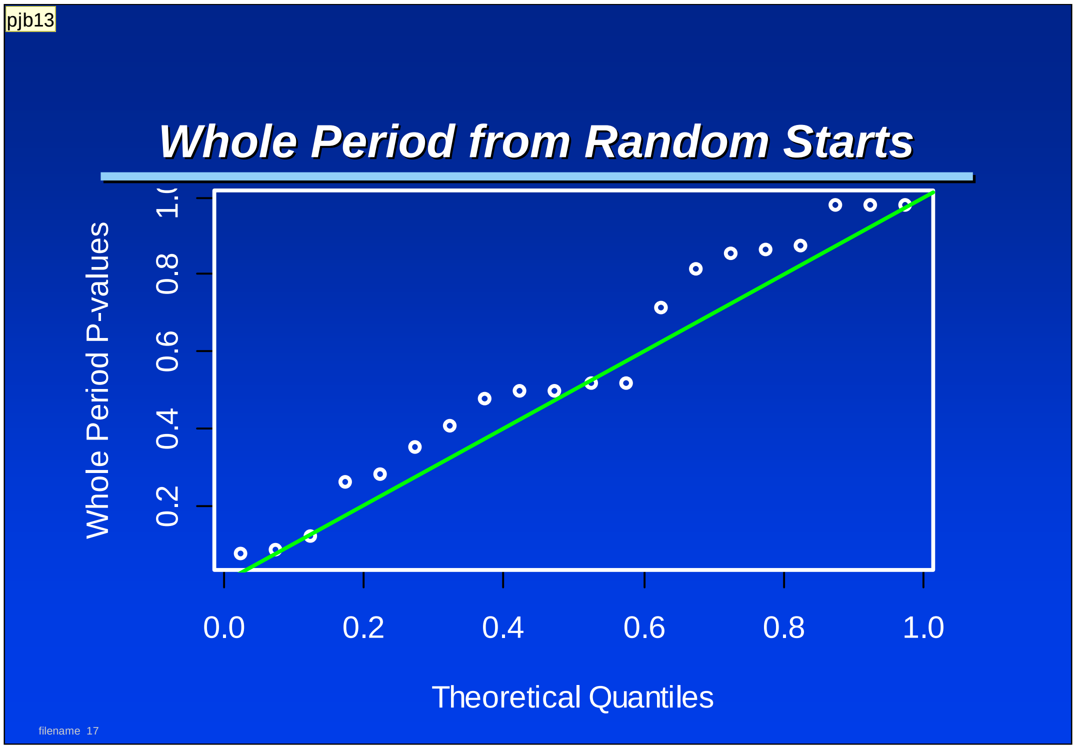
Slide 16

pjb12 The white line is the original strategy, the black lines are selected quantiles of the random paths -- including the minimum and maximum.

The strategy does well in the last half of 1998. It stays better than the best of the random paths through almost all of 1999. When the dotcom bubble burst, then the strategy burst as well.

In the end it winds up being better than all but 2 of the random paths. Thus even though it is up only 8% in 4 years, that turns out to be quite good -- very surprising.

Patrick Burns, 07/09/2006


pjb13
Whole Period from Random Starts
Whole Period P-values
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
Theoretical Quantiles
filename 17

Slide 17

pjb13 We don't care about our strategy working well from one particular initial portfolio, we want it to work no matter where we start.

So let's redo the whole process again from 20 different initial portfolios that were randomly selected. From each of the 20 experiments only retain the whole period p-value (basically the fraction of random paths that outperform).

If there is zero skill, then getting the p-values is precisely the same as getting a random sample of size 20 from the Uniform(0, 1). In which case the points in the plot will fall close to the green line. If there is skill, then the p-values will, in general, be small. If there is negative skill, the p-valules will tend to be close to 1.

Patrick Burns, 06/09/2006

Why

Setting Constraint Bounds

filename 18

FTSE Example

- FTSE 350 Data
- 10,000 portfolios generated for each set of constraints
- Returns: 2006 Jan 01 – 2006 June 01
- Long-only
- 90 – 100 assets in portfolio
- Nested set of linear constraints

Linear Constraints

- Large cap versus Mid cap
  - 10% - 30%   70% - 90%
  - 13% - 27%   73% - 87%
  - 17% - 23%   77% - 83%
- High yield versus Low yield
  - 50% - 70%   30% - 50%
  - 53% - 67%   33% - 47%
  - 57% - 63%   37% - 43%

pjb14

Linear Constraints

- **5 Sectors**
  - **10% - 30%**
  - **13% - 27%**
  - **17% - 23%**

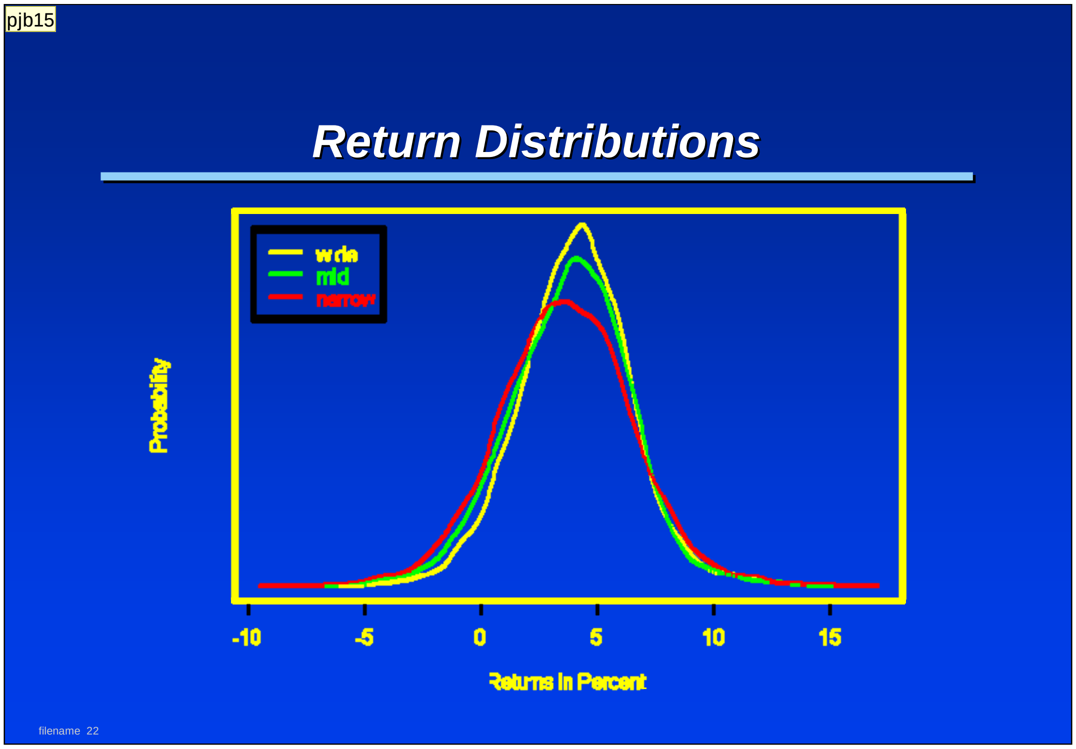
filename 21

Slide 21

pjb14 The key thing in this example is that we have a nested set of bounds for the linear constraints. The yellow is the loosest bounds, and red is the tightest bounds.

The purpose of constraints is to keep the portfolio from doing something stupid. So we should expect the red distribution to be less dispersed than the yellow distribution.

Patrick Burns, 06/09/2006

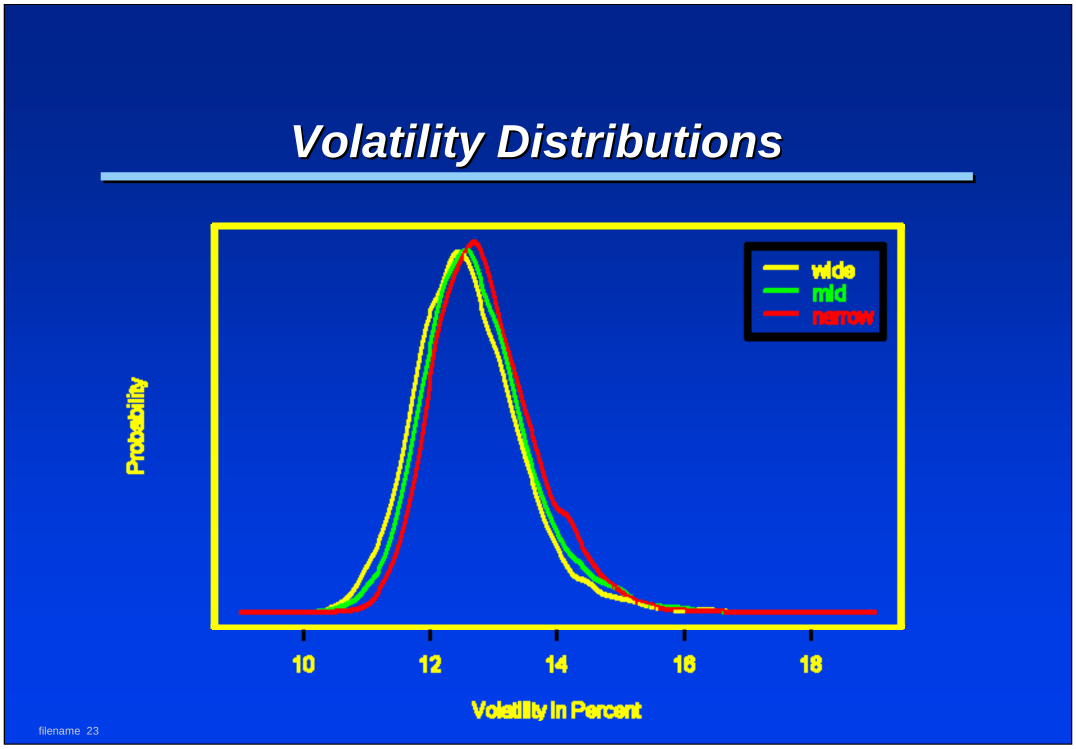

pjb15
Return Distributions
wide
mid
narrow
Probability
-10
-5
0
5
10
15
Returns in Percent
filename 22

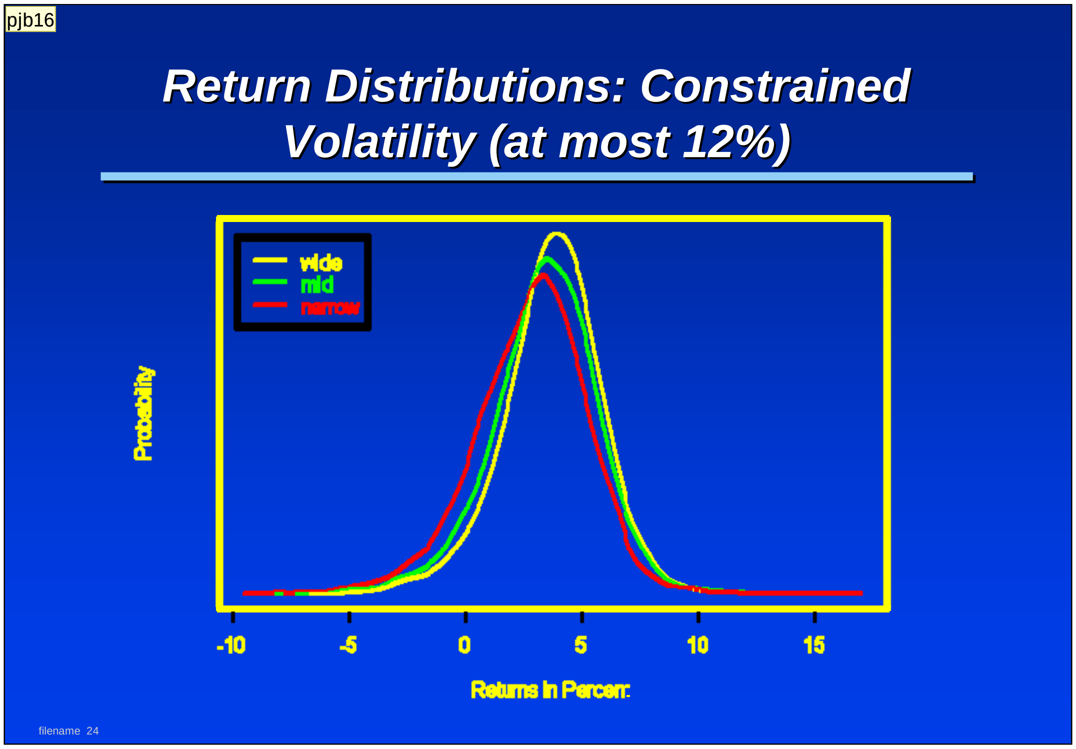
Slide 22

pjb15 The distributions in this example exhibit the opposite behavior to what we would expect -- the more we constrain, the wider the distribution.

One suggestion from the audience was that perhaps the more constrained these portfolios, the more concentrated they are forced to be.

Patrick Burns, 06/09/2006


Volatility Distributions
wide
mid
narrow
Probability
10
12
14
16
18
Volatility in Percent
filename 23


pjb16
Return Distributions: Constrained Volatility (at most 12%)
wide
mid
narrow
Probability
-10
-5
0
5
10
15
Returns in Percent
filename 24

Slide 24

pjb16 With the volatility constraint added, the spreads of the distribution are more similar. However, there is the worse feature that the more constrained portfolios have significantly smaller returns.

I have no idea if this is a very pathological case that I happened upon, or if this is more common than we would like to suppose. I don't see a way of doing this type of analysis without random portfolios.
Patrick Burns, 07/09/2006

How

Naïve Ideas

filename 25

pjb17

Convex Polytope Methods

- If the shoe fits, probably good
- Generally not applicable
- Non-linear constraints
  - tracking error constraints
  - variance constraints
- Integer constraints
  - number of names traded
  - number of names in the portfolio

Slide 26

pjb17 If all constraints are linear, then the feasible region is a convex polytope. Algorithms are available for random samples from a convex polytope.
Patrick Burns, 06/09/2006

Rejection Method

- Generate some portfolio, accept it if it meets all of the constraints
- Exactly uniform if original generation is
- For many practical problems the waiting time could be years

How

Genetic Algorithms

pjb18

Outline of Genetic Algorithms

Used for optimisation

Have a population of solutions

The population is improved through time via random mechanisms

Might be called the Hollywood Algorithm

- **Sex**
- **Violence**

Slide 29

pjb18 Genetic algorithms as discussed here are for optimization -- we'll get to the connection with random portfolios later.

Most optimization algorithms have a single solution at each point in time, and that solution is improved as the algorithm progresses. In contrast genetic algorithms have a population of solutions at each point in time.

Genetic algorithms require a way of combining solutions together, and a way of killing off solutions so that there is not a population explosion.

Patrick Burns, 06/09/2006

pjb19

The "Standard" Genetic Algorithm

- **Uses binary vectors**

  Parent A: 0010110

  Parent B: 1011010

  Child: 0011010

- **Binary strings are spliced**
- **Mutations may occur**
- **A new generation replaces the old**

filename 30

Slide 30

pjb19 The parameter vector is represented as a (long) binary string. The main genetic operation is "crossover" in which a random place in the string is picked (for each mating) and the child gets the bits from the start to the crossover from one parent and the rest of the string from the other parent. Mutations can then occur.

A new generation is built up. At some point the old generation is discarded, and the child generation becomes the parent generation.

Patrick Burns, 07/09/2006

pjb20

Problems with the Standard

- Natural parameter space probably not binary
- Inefficient search of the binary parameter space
- The best solution is often thrown away

filename 31

Slide 31

pjb20 It is hard to overemphasize just how bad the standard genetic algorithm is.

Not many problems naturally have a binary parameter space. When you restructure the parameters, it is hard to get the genetics to make sense. Getting the genetics right is, oddly enough, quite important in a genetic algorithm.

Even if the parameter space is binary, the standard algorithm does a terribly inefficient search of the space.

The generational scheme means that the best solution found so far can easily be lost.
Patrick Burns, 07/09/2006

pjb21

A Modified Genetic Algorithm

- Keep original parameter space
- Two parents marry and produce children
- Best two of parents plus children survive
- If a child survives, do simulated annealing

Slide 32

pjb21 This scheme of improving the population means that it will converge. Typically it will converge to a non-optimal point, but you are guaranteed of convergence.

Combining a simulated annealing type operation is extremely useful. Genetic algorithms are good at globally searching the parameter space (if the genetics are right); simulated annealing is very good at searching locally. The two approaches together are more powerful than either alone.

Patrick Burns, 07/09/2006

pjb22

A Wedding in Cherokee County

Parents		Raw Twins		Twins	
A .2	D .2	A .2	D .4	A .222	D .364
D .4	G .3	D .2	E .1	D .222	E .091
E .1	P .3	P .3	G .3	P .333	G .273
Z .3	Z .2	Z .2	Z .3	Z .222	Z .273

pjb22 Portfolios are really a collection of asset names plus the number of shares or contracts or whatever that are held. In this example, we simplify by supposing that portfolios are described by weights that sum to 1.

The two parents have 2 assets in common: D and Z.
All children will be forced to have D and Z as well. In this example the weights for D and Z happen to land the same way in the children as the parents, but that is a chance event.

Assets that are in only one parent are randomly assigned to one or the other twin. The raw twins have weights that no longer sum to 1, but just rescale the weights so that they do.

The twins will be evaluated, and compared to the parents -- the best are the ones that survive.
Patrick Burns, 07/09/2006

pjb23

Back to Random Portfolios

- **Cast search for a random portfolio as an optimisation**
- **Objective is zero if all constraints are met**
- **Objective is positive by an amount dependent on how much the constraints are violated**
- **Optimiser stops when it gets to zero**

filename 34

Slide 34

pjb23 The objective function is identically zero over the set of portfolios that satisfy the constraints. It is strictly positive elsewhere.

Patrick Burns, 07/09/2006

pjb24

Random Portfolio Generation

- Fill population with totally random portfolios
- Run genetic optimisation until some portfolio meets all constraints
- Gives one random portfolio
- Start over again with completely new totally random portfolios

Slide 35

pjb24 Once a portfolio with objective zero has been found, the other portfolios in that population need to be discarded since they will be correlated with the portfolio we are saving. It is important to start fresh when searching for a second portfolio.

Patrick Burns, 07/09/2006

More Information

http://www.burns-stat.com

- **Random portfolios page**
- **Working papers**
- **Software**