Critical phenomena in portfolio selection

Imre Kondor

Collegium Budapest and Eötvös University, Budapest

COST P10 Physics of Risk & Workshop on Complex System Science

Vilnius, Lithuania, 13-16 May, 2006

Contents

The subject of the talk lies at the crossroads of finance, statistical physics and computer science

I. Portfolio selection is very sensitive to sampling error, for a critical value of N/T the error actually diverges.

II. This is an example of an algorithmic phase transition

Contents

The subject of the talk lies at the crossroads of finance, statistical physics and computer science

I. Main message: Portfolio selection is very sensitive to sampling error, for a critical value of N/T the error actually diverges.

II. This is an example of an algorithmic phase transition

Contents

The subject of the talk lies at the crossroads of finance, statistical physics and computer science

I. Portfolio selection is very sensitive to sampling error, for a critical value of N/T the error actually diverges.

II. This is an example of an algorithmic phase transition

Coworkers

- **Szilárd Pafka** (ELTE PhD student → CIB Bank, →Paycom.net, California)
- **Gábor Nagy** (Debrecen University PhD student and CIB Bank, Budapest)

A portfolio …

… is a combination of assets or investment instruments.

- In this talk I will focus on equity portfolios.
- Rational portfolio selection seeks a tradeoff between risk and reward.

A portfolio …

… is a combination of assets or investment instruments.

- In this talk I will focus on equity portfolios.
- Rational portfolio selection seeks a tradeoff between risk and reward.

A portfolio …

… is a combination of assets or investment instruments.

- In this talk I will focus on equity portfolios.
- Rational portfolio selection seeks a tradeoff between risk and reward.

Risk and reward

- Financial reward can be measured in terms of the return of log return:

$$\frac{S_t - S_{t-1}}{S_{t-1}} \qquad \text{or} \qquad \ln\left(\frac{S_t}{S_{t-1}}\right)$$

- The characterization of risk is more controversial

The most obvious choice for a risk measure: Variance

- Variance is the average quadratic deviation from the average – a time honoured statistical tool
- Its use assumes that the probability distribution of the returns is sufficiently concentrated around the average, that there are no large fluctuations
- This is true in several instances, but we often encounter „fat tails", huge deviations with a non-negligible probability.

The most obvious choice for a risk measure: Variance

- Variance is the average quadratic deviation from the average – a time honoured statistical tool
- Its use assumes that the probability distribution of the returns is sufficiently concentrated around the average, that there are no large fluctuations
- This is true in several instances, but we often encounter „fat tails”, huge deviations with a non-negligible probability.

The most obvious choice for a risk measure: Variance

- Variance is the average quadratic deviation from the average – a time honoured statistical tool
- Its use assumes that the probability distribution of the returns is sufficiently concentrated around the average, that there are no large fluctuations
- This is true in several instances, but we often encounter „fat tails”, huge deviations with a non-negligible probability.

Alternative risk measures

- There are several alternative risk measures in use in the academic literature, practice, and regulation
- Value at risk (VaR): the best among the *p%* worst losses (not convex, punishes diversification)
- Mean absolute deviation (MAD): Algorithmics
- Coherent risk measures (promoted by academics):
  - Expected shortfall (ES): average loss beyond a high threshold
  - Maximal loss (ML): the single worst case

Alternative risk measures

- There are several alternative risk measures in use in the academic literature, practice, and regulation
- Value at risk (VaR): the best among the *p%* worst losses (not convex, punishes diversification)
- Mean absolute deviation (MAD): Algorithmics
- Coherent risk measures (promoted by academics):

    Expected shortfall (ES): average loss beyond a high threshold

    Maximal loss (ML): the single worst case

Alternative risk measures

- There are several alternative risk measures in use in the academic literature, practice, and regulation
- Value at risk (VaR): the best among the *p%* worst losses (not convex, punishes diversification)
- Mean absolute deviation (MAD): Algorithmics
- Coherent risk measures (promoted by academics):

  Expected shortfall (ES): average loss beyond a high threshold

  Maximal loss (ML): the single worst case

Alternative risk measures

- There are several alternative risk measures in use in the academic literature, practice, and regulation
- Value at risk (VaR): the best among the *p%* worst losses (not convex, punishes diversification)
- Mean absolute deviation (MAD): Algorithmics
- Coherent risk measures (promoted by academics):

  Expected shortfall (ES): average loss beyond a high threshold

  Maximal loss (ML): the single worst case

A portfolio is…

a weighted average of assets, with a set of weights w_i that add up to unity (the budget constraint).

- The weights are not necessarily positive – short selling
- If there is no condition on the weights other than the budget constraint, then the domain over which the optimum is sought is unbounded

A portfolio is…

a weighted average of assets, with a set of weights w_i that add up to unity (the budget constraint).

- The weights are not necessarily positive – short selling
- If there is no condition on the weights other than the budget constraint, then the domain over which the optimum is sought is unbounded

A portfolio is…

a weighted average of assets, with a set of weights w_i that add up to unity (the budget constraint).

- The weights are not necessarily positive – short selling
- If there is no condition on the weights other than the budget constraint, then the domain over which the optimum is sought is unbounded

The variance of a portfolio

$$\sigma_P^2 = \sum_{i,j} w_i \sigma_{ij} w_j$$

- a quadratic form of the weights. The coefficients of this form are the elements of the covariance matrix that measures the co-movements between the various assets.

Markowitz' portfolio selection theory

Rational portfolio selection realizes the tradeoff between risk and reward by minimizing the risk functional over the weights, given the expected return, the budget constraint, and possibly other costraints.

How do we know the returns and the covariances?

- In principle, from observations on the market
- If the portfolio contains N assets, we need $O(N^2)$ data
- The input data come from T observations for N assets
- The estimation error is negligible as long as $NT >> N^2$, i.e. $N << T$
- In practice T is never longer than 4 years, i.e. $T \sim 1000$, whereas in a typical banking portfolio N is several hundreds or thousands.
- $N << T$ is therefore never fulfilled in practice.

How do we know the returns and the covariances?

- In principle, from observations on the market
- If the portfolio contains N assets, we need $O(N^2)$ data
- The input data come from T observations for N assets
- The estimation error is negligible as long as $NT >> N^2$, i.e. $N << T$
- In practice T is never longer than 4 years, i.e. T~1000, whereas in a typical banking portfolio N is several hundreds or thousands.
- $N << T$ is therefore never fulfilled in practice.

How do we know the returns and the covariances?

- In principle, from observations on the market
- If the portfolio contains N assets, we need $O(N^2)$ data
- The input data come from T observations for N assets
- The estimation error is negligible as long as $NT >> N^2$, i.e. $N << T$
- In practice T is never longer than 4 years, i.e. $T \sim 1000$, whereas in a typical banking portfolio N is several hundreds or thousands.
- $N << T$ is therefore never fulfilled in practice.

How do we know the returns and the covariances?

- In principle, from observations on the market
- If the portfolio contains N assets, we need $O(N^2)$ data
- The input data come from T observations for N assets
- The estimation error is negligible as long as $NT \gg N^2$, i.e. $N \ll T$
- In practice T is never longer than 4 years, i.e. $T \sim 1000$, whereas in a typical banking portfolio N is several hundreds or thousands.
- $N \ll T$ is therefore never fulfilled in practice.

How do we know the returns and the covariances?

- In principle, from observations on the market
- If the portfolio contains N assets, we need $O(N^2)$ data
- The input data come from T observations for N assets
- The estimation error is negligible as long as $NT >> N^2$, i.e. $N << T$
- In practice T is never longer than 4 years, i.e. $T \sim 1000$, whereas in a typical banking portfolio N is several hundreds or thousands.
- $N << T$ is therefore never fulfilled in practice.

How do we know the returns and the covariances?

- In principle, from observations on the market
- If the portfolio contains N assets, we need $O(N^2)$ data
- The input data come from T observations for N assets
- The estimation error is negligible as long as $NT \gg N^2$, i.e. $N \ll T$
- In practice T is never longer than 4 years, i.e. $T \sim 1000$, whereas in a typical banking portfolio N is several hundreds or thousands.
- $N \ll T$ is therefore never fulfilled in practice.

Information deficit

- Thus the Markowitz problem suffers from the „curse of dimensions”, or from information deficit
- The estimates will contain error and the resulting portfolios will be suboptimal
- How serious is this effect?
- How sensitive are the various risk measures to this kind of error?
- How can we reduce the error?

Information deficit

- Thus the Markowitz problem suffers from the „curse of dimensions”, or from information deficit
- The estimates will contain error and the resulting portfolios will be suboptimal
- How serious is this effect?
- How sensitive are the various risk measures to this kind of error?
- How can we reduce the error?

Information deficit

- Thus the Markowitz problem suffers from the „curse of dimensions”, or from information deficit
- The estimates will contain error and the resulting portfolios will be suboptimal
- How serious is this effect?
- How sensitive are the various risk measures to this kind of error?
- How can we reduce the error?

Information deficit

- Thus the Markowitz problem suffers from the „curse of dimensions”, or from information deficit
- The estimates will contain error and the resulting portfolios will be suboptimal
- How serious is this effect?
- How sensitive are the various risk measures to this kind of error?
- How can we reduce the error?

Information deficit

- Thus the Markowitz problem suffers from the „curse of dimensions”, or from information deficit
- The estimates will contain error and the resulting portfolios will be suboptimal
- How serious is this effect?
- How sensitive are the various risk measures to this kind of error?
- How can we reduce the error?

Fighting the curse of dimensions

- Economists have been struggling with this problem for ages. Since the root of the problem is lack of sufficient information, the remedy is to inject external info into the estimate. This means imposing some structure on **σ**. This introduces bias, but beneficial effect of noise reduction may compensate for this.
- Examples:
  - single-index models (β's)
  - multi-index models
  - grouping by sectors
  - principal component analysis
  - Bayesian shrinkage estimators, etc.
  - Random matrix theory

All these help to various degrees. Most studies are based on empirical data

Our approach:

- To test the noise sensitivity of various risk measures we use simulated data
- The rationale behind this is that in order to be able to compare the sensitivity of various risk measures to noise, we better get rid of other sources of uncertainty, like non-stationarity. This can be achieved by using artificial data where we have total control over the underlying stochastic process.
- For simplicity, we use iid normal variables in the following.

Our approach:

- To test the noise sensitivity of various risk measures we use simulated data
- The rationale behind this is that in order to be able to compare the sensitivity of various risk measures to noise, we better get rid of other sources of uncertainty, like non-stationarity. This can be achieved by using artificial data where we have total control over the underlying stochastic process.
- For simplicity, we use iid normal variables in the following.

Our approach:

- To test the noise sensitivity of various risk measures we use simulated data
- The rationale behind this is that in order to be able to compare the sensitivity of various risk measures to noise, we better get rid of other sources of uncertainty, like non-stationarity. This can be achieved by using artificial data where we have total control over the underlying stochastic process.
- For simplicity, we use iid normal variables in the following.

- For such a simple underlying process the exact risk measure can be calculated.
- To construct the empirical risk measure, we generate long time series, and cut out segments of length T from them, as if making observations on the market.
- From these „observations" we construct the empirical risk measure and optimize our portfolio under it.
- The ratio q_0 of the empirical and the exact risk measure is a measure of the estimation error due to noise.

- For such simple underlying processes the exact risk measure can be calculated.
- To construct the empirical risk measure, we generate long time series, and cut out segments of length T from them, as if making observations on the market.
- From these „observations” we construct the empirical risk measure and optimize our portfolio under it.
- The ratio q_0 of the empirical and the exact risk measure is a measure of the estimation error due to noise.

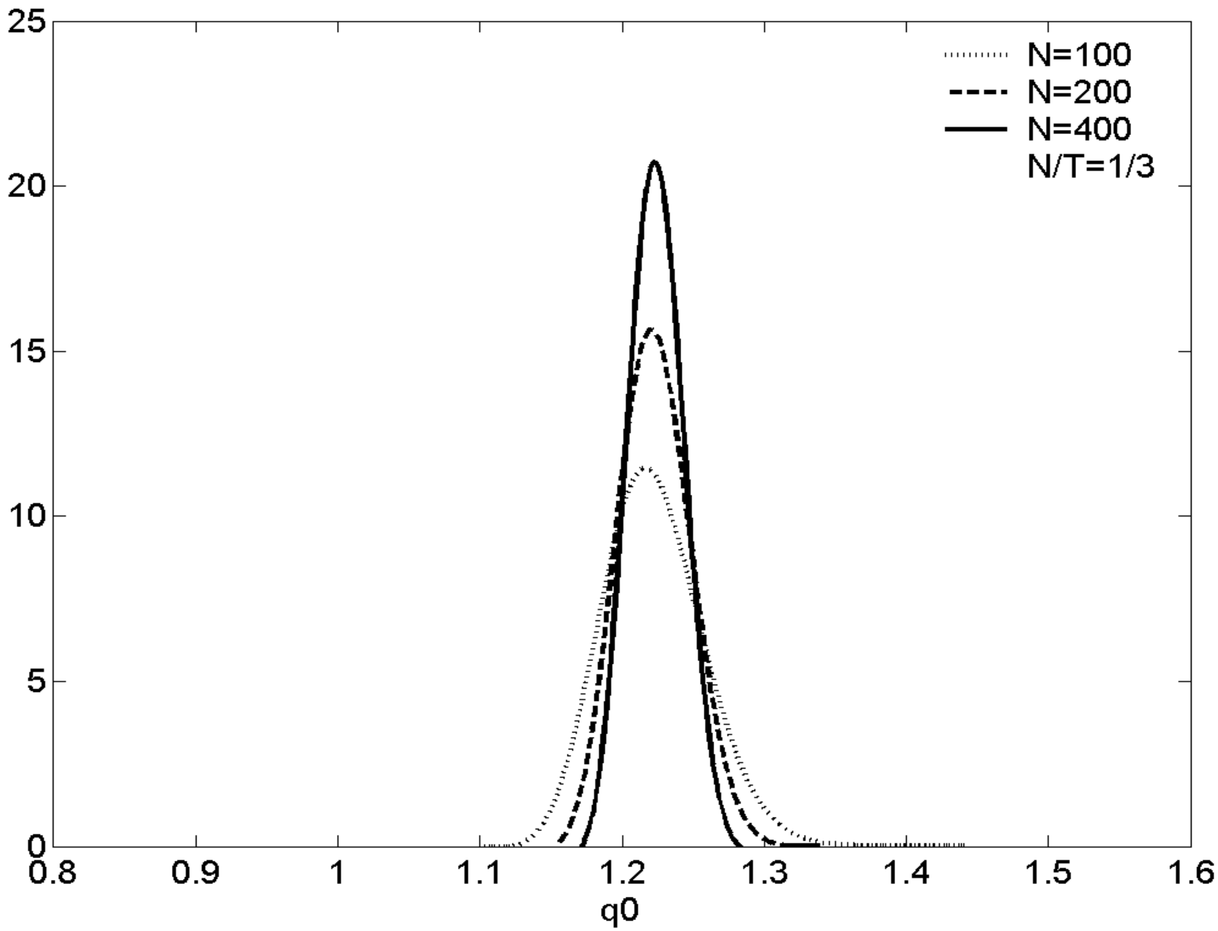
- For such simple underlying processes the exact risk measure can be calculated.
- To construct the empirical risk measure, we generate long time series, and cut out segments of length T from them, as if making observations on the market.
- From these „observations” we construct the empirical risk measure and optimize our portfolio under it.
- The ratio q_o of the empirical and the exact risk measure is a measure of the estimation error due to noise.

- For such simple underlying processes the exact risk measure can be calculated.
- To construct the empirical risk measure, we generate long time series, and cut out segments of length T from them, as if making observations on the market.
- From these „observations" we construct the empirical risk measure and optimize our portfolio under it.
- The ratio q_0 of the empirical and the exact risk measure is a measure of the estimation error due to noise.

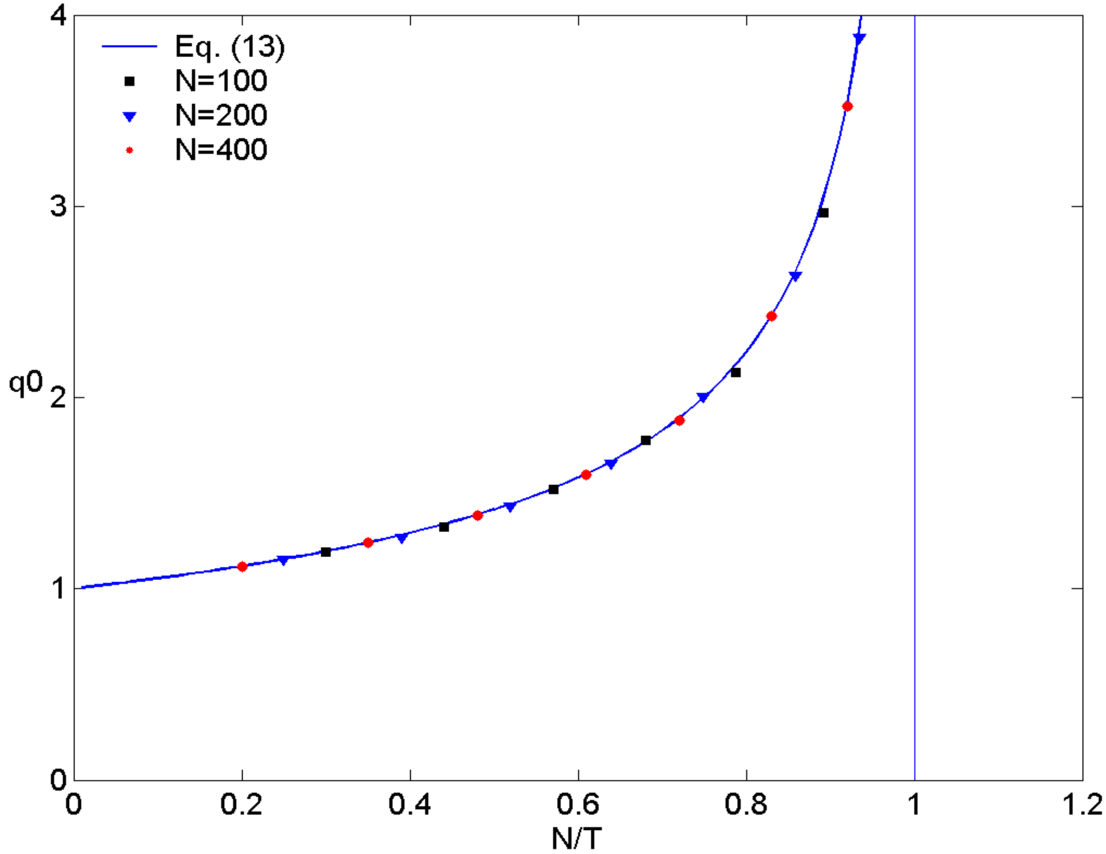
Sample to sample fluctuations

- The relative error q_0 of the optimal portfolio is a random variable, fluctuating from sample to sample.
- The weights of the optimal portfolio also fluctuate.

The distribution of *qo* over the samples



The expectation value of *qo* as a function of *N/T*



The critical point: $N/T = 1$

- As N approaches T, the relative error is increasing and diverges at the critical point $N=T$.
- The expectation value of the error can be shown to be:

$$q_0 = \frac{1}{\sqrt{1 - \frac{N}{T}}}$$

- This formula was first published by Sz. Pafka and I.K.
- The variance of the distribution of q_0 diverges even more strongly, with an exponent -3/4.

Fluctuation of weights

- The weights wildly fluctuate within a given sample as well as from sample to sample.
- The optimization hardly determines the weights even far from the critical point!
- The standard deviation of the weights relative to their exact average value also diverges at the critical point

If short selling is banned

If the weights are constrained to be positive, the instability will manifest itself by more and more weights becoming zero – the portfolio spontaneously reduces its size!

Explanation: the solution would like to run away, the constraints prevent it from doing so, therefore it will stick to the walls.

Similar effects are observed if we impose any other linear constraints, like bounds on sectors, etc.

It is clear, that in these cases the solution is determined more by the constraints than the objective function.

If short selling is banned

- If the weights are constrained to be positive, the instability will manifest itself by more and more weights becoming zero – the portfolio spontaneously reduces its size!
- Explanation: the solution would like to run away, the constraints prevent it from doing so, therefore it will stick to the walls.
- Similar effects are observed if we impose any other linear constraints, like bounds on sectors, etc.
- It is clear, that in these cases the solution is determined more by the constraints than the objective function.

If short selling is banned

If the weights are constrained to be positive, the instability will manifest itself by more and more weights becoming zero – the portfolio spontaneously reduces its size!

Explanation: the solution would like to run away, the constraints prevent it from doing so, therefore it will stick to the walls.

Similar effects are observed if we impose any other linear constraints, like bounds on sectors, etc.

It is clear, that in these cases the solution is determined more by the constraints than the objective function.

If short selling is banned

If the weights are constrained to be positive, the instability will manifest itself by more and more weights becoming zero – the portfolio spontaneously reduces its size!

Explanation: the solution would like to run away, the constraints prevent it from doing so, therefore it will stick to the walls.

Similar effects are observed if we impose any other linear constraints, like bounds on sectors, etc.

It is clear, that in these cases the solution is determined more by the constraints than the objective function.

If the variables are not iid

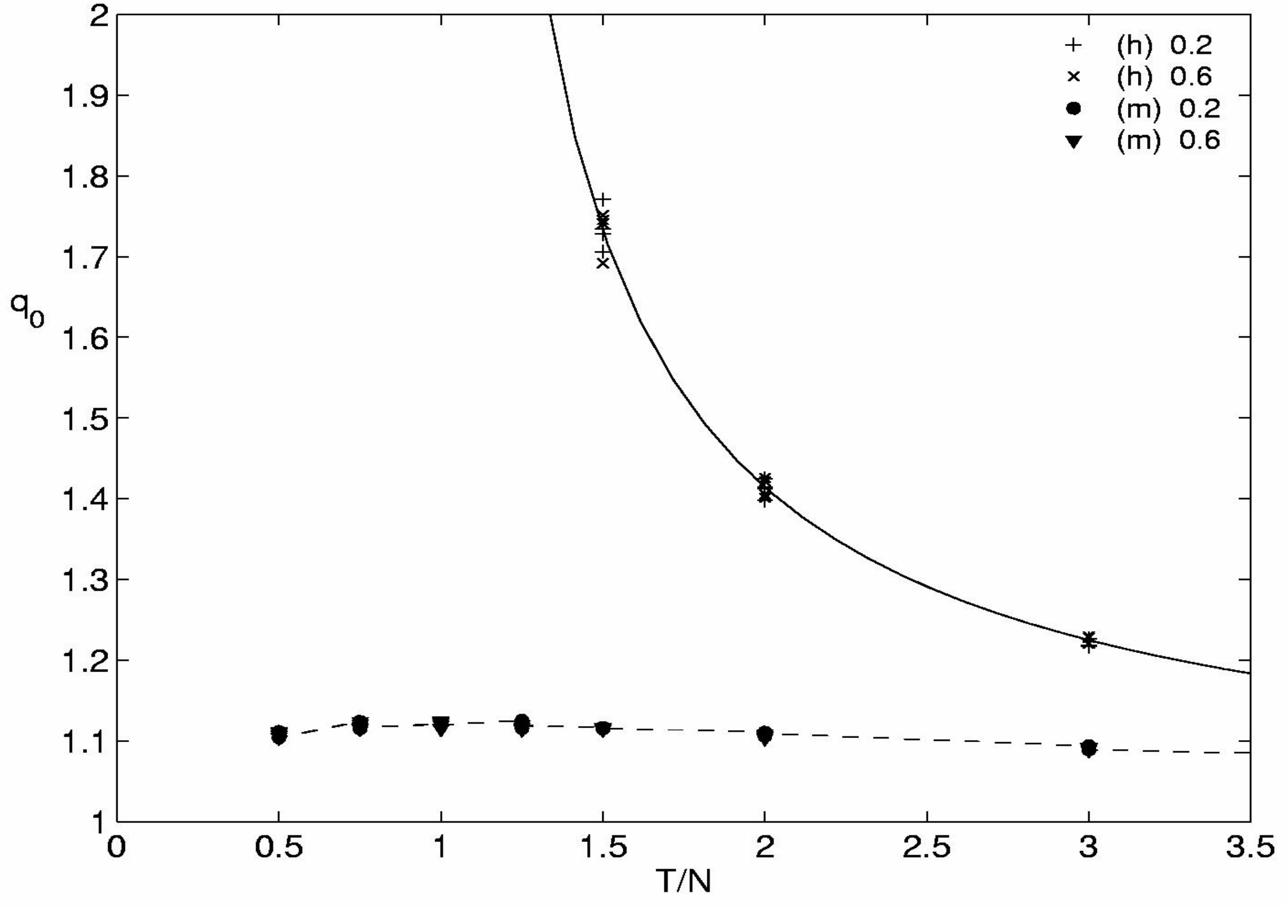
Experimenting with various market models (one-factor, market plus sectors, positive and negative covariances, etc.) shows that the main conclusion does not change.

Overwhelmingly positive correlations tend to enhance the instability, negative ones decrease it, but they do not change the power of the divergence, only its prefactor

If the variables are not iid

- Experimenting with various market models (one-factor, market plus sectors, positive and negative covariances, etc.) shows that the main conclusion does not change.
- Overwhelmingly positive correlations tend to enhance the instability, negative ones decrease it, but they do not change the power of the divergence, only its prefactor

After filtering the noise is much reduced, and we can even penetrate into the region below the critical point $T<N$



Similar studies under mean absolute deviation, expected shortfall and maximal loss

- Lead to similar conclusions, except that the effect of estimation error is even more serious
- In addition, no convincing filtering methods exist for these measures
- In the case of coherent measures the existence of a solution becomes a probabilistic issue, depending on the sample
- Calculation of this probability leads to some intriguing problems in random geometry

Similar studies under mean absolute deviation, expected shortfall and maximal loss

- Lead to similar conclusions, except that the effect of estimation error is even more serious
- In addition, no convincing filtering methods exist for these measures
- In the case of coherent measures the existence of a solution becomes a probabilistic issue, depending on the sample
- Calculation of this probability leads to some intriguing problems in random geometry

Similar studies under mean absolute deviation, expected shortfall and maximal loss

- Lead to similar conclusions, except that the effect of estimation error is even more serious
- In addition, no convincing filtering methods exist for these measures
- In the case of coherent measures the existence of a solution becomes a probabilistic issue, depending on the sample
- Calculation of this probability leads to some intriguing problems in random geometry

Similar studies under mean absolute deviation, expected shortfall and maximal loss

- Lead to similar conclusions, except that the effect of estimation error is even more serious
- In addition, no convincing filtering methods exist for these measures
- In the case of coherent measures the existence of a solution becomes a probabilistic issue, depending on the sample
- Calculation of this probability leads to some intriguing problems in random geometry

A wider context

- Hard computational problems (combinatorial optimization, random assignment, graph partitioning, satisfiability): the length of the algorithm grows exponentially with the size of the problem. These are practically untractable.
- Their difficulty may depend on some internal parameter (e.g. the density of constraints in a satisfiability problem)
- Recently it has been observed that the difficulty does not change gradually with the variations of this parameter, but there is a critical value where the problem becomes hard abruptly
- This critical point is preceded by a number of critical phenomena

A wider context

- Hard computational problems (combinatorial optimization, random assignment, graph partitioning, and satisfiability problems): the length of the algorithm grows exponentially with the size of the problem. These are practically untractable.
- Their difficulty may depend on some internal parameter (e.g. the density of constraints in a satisfiability problem)
- Recently it has been observed that the difficulty does not change gradually with the variations of this parameter, but there is a critical value where the problem becomes hard abruptly
- This critical point is preceded by a number of critical phenomena

A wider context

- Hard computational problems (combinatorial optimization, random assignment, graph partitioning, and satisfiability problems): the length of the algorithm grows exponentially with the size of the problem. These are practically untractable.
- Their difficulty may depend on some internal parameter (e.g. the density of constraints in a satisfiability problem)
- Recently it has been observed that the difficulty does not change gradually with the variations of this parameter, but there is a critical value where the problem becomes hard abruptly
- This critical point is preceded by a number of critical phenomena

A wider context

- Hard computational problems (combinatorial optimization, random assignment, graph partitioning, and satisfiability problems): the length of the algorithm grows exponentially with the size of the problem. These are practically untractable.
- Their difficulty may depend on some internal parameter (e.g. the density of constraints in a satisfiability problem)
- Recently it has been observed that the difficulty does not change gradually with the variations of this parameter, but there is a critical value where the problem becomes hard abruptly
- This critical point is preceded by a number of critical phenomena

- The critical phenomena we observe in portfolio selection are analogous to these, they represent a new „random Gaussian” universality class within this family, where a number of modes go soft in rapid succession, as one approaches the critical point.
- Filtering corresponds to discarding these soft modes.

- The critical phenomena we observe in portfolio selection are analogous to these, they represent a new „random Gaussian" universality class within this family, where a number of modes go soft in rapid succesion, as one approaches the critical point.
- Filtering corresponds to discarding these soft modes.

Similar examples from everyday life, and …