Motivation

Example I: Petrobras returns
Histogram
Tail behavior

Example II: Automobile claims

Measuring variability

Measuring extremity
Historical facts
Computing kurtosis

Example III: Nasdaq returns

Example IV: Heavy vs thin

Final remarks

Beware the Long Tail[1]

Hedibert Freitas Lopes
Booth School of Business
University of Chicago

[1] Title borrowed from Rachel Ehrenberg's Science News article (November 5, 2011, www.sciencenews.org)

Outline

1. Motivation
2. Example I: Petrobras returns
   - Histogram
   - Tail behavior
3. Example II: Automobile claims
4. Measuring variability
5. Measuring extremity
   - Historical facts
   - Computing kurtosis
6. Example III: Nasdaq returns
7. Example IV: Heavy vs thin
8. Final remarks

Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

The Economist, Jan 22nd 2009

In Plato's cave

Mathematical models are a powerful way of predicting financial markets. But they are fallible.

> *...although the normal distribution closely matches the real world in the middle of the curve, where most of the gains or losses lie, it does not work well at the extreme edges, or "tails".*
>
> *Benoît Mandelbrot ... calculated that if the DJIA followed a normal distribution, it should have moved by more than 3.4% on 58 days between 1916 and 2003; in fact it did so 1,001 times. It should have moved by more than 4.5% on six days; it did so on 366. It should have moved by more than 7% only once in every 300,000 years; in the 20th century it did so 48 times.*

Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

DJIA returns: 1916-2003

Observed: more than 3.4% on **1,001** days
Expected: more than 3.4% on **58** days

Observed: more than 4.5% on **366** days
Expected: more than 4.5% on **6** days

Observed: more than 7% on **48** days
Expected: ?????

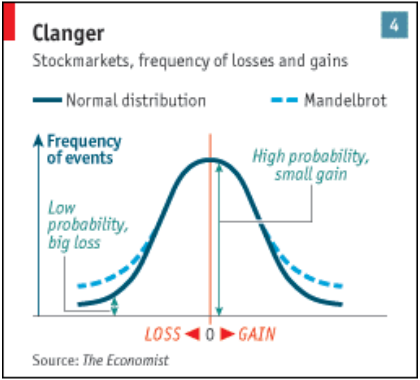
Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Miocene epoch

Observed: more than 7% on **48** days over **100 years**
Expected: more than 7% on **48** days over **14 million years!**

The apes arose and diversified during the Miocene epoch, becoming widespread in the Old World.

Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks



Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Petrobras[2]

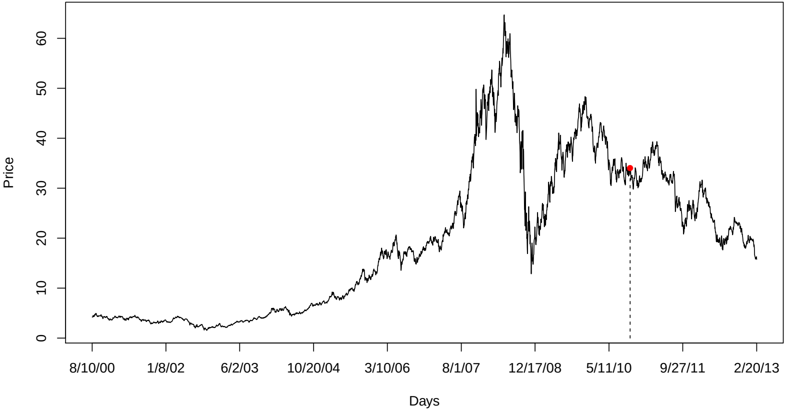
Let us take a closer look at Petrobras. Why?
Because I bought it at $34 in October 2010, that is why!

Date	Price
8/10/00	4.23
8/11/00	4.22
8/14/00	4.21
⋮	⋮
9/30/10	33.88
10/1/10	34.06
10/4/10	34.18
⋮	⋮
2/19/13	16.29
2/20/13	15.75

[2] Data source: Yahoo! Finance (finance.yahoo.com)

Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Prices



Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Returns

We define return as

$$R_t = 100(P_t/P_{t-1} - 1),$$

where P_t is the price at time t.

Date	Price	Return
8/10/00	4.23	–
8/11/00	4.22	-0.2364066
8/14/00	4.21	-0.2369668
⋮	⋮	⋮
9/30/10	33.88	1.5283188
10/1/10	34.06	0.5312869
10/4/10	34.18	0.3523194
⋮	⋮	⋮
2/19/13	16.29	2.0676692
2/20/13	15.75	-3.3149171

Motivation

Example I: Petrobras returns
Histogram
Tail behavior

Example II: Automobile claims

Measuring variability

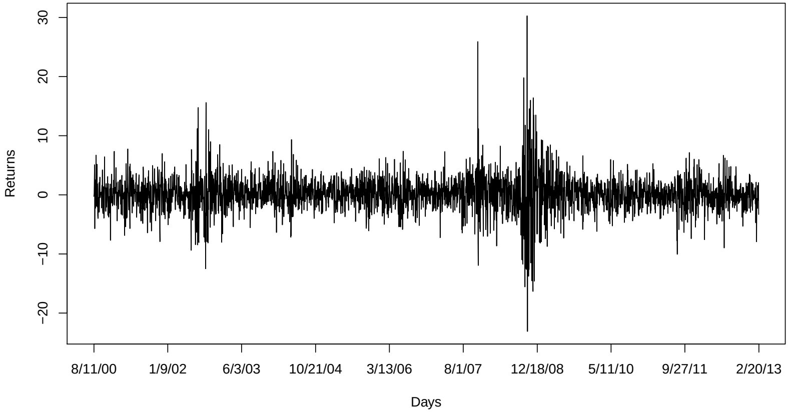
Measuring extremity
Historical facts
Computing kurtosis

Example III: Nasdaq returns

Example IV: Heavy vs thin

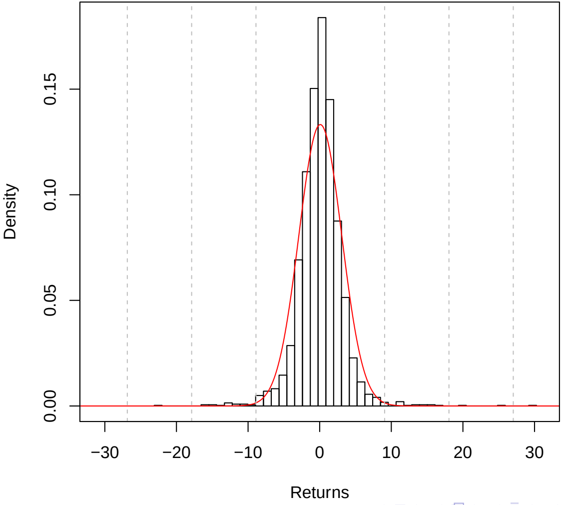
Final remarks

Returns



Empirical distribution of returns

Red curve: Fitted normal with data-driven mean and variance.
Dashed lines: 3,6,9 standard deviations from the mean.



Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Tail behavior

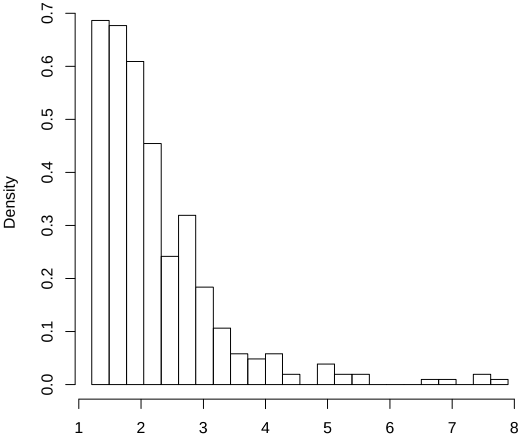
Number of observed and expected data points outside k standard deviations (out of 3145 daily returns).

Standard deviations	Tail behavior: Observed	Tail behavior: Expected
1	688	998
2	140	143
3	41	9
4	22	0.2
5	9	0.002
6	4	6.2e-06
7	3	8.1e-09
8	2	4.2e-12
9	1	0
10	1	0

Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Another case: automobile claims

Automobile claims from 1988 to 2001 gathered from several European insurance companies, which are at least as large as 1.2 million Euros.



Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Measuring variability

Variance (or volatility) is not enough to understand, model, predict rare (tail, perhaps catastrophic) events.

How to identify heavy-tailness in the data?

How to model heavy-tail data?

Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Measure of extremity: kurtosis

Measures the degree to which exceptional values occur more frequently (high kurtosis) or less frequently (low kurtosis).

A reference distribution is the normal distribution, whose kurtosis is three.

High kurtosis results in exceptional values that are called fat tails".

Fat tails indicate a higher percentage of very low and very high returns than would be expected with a normal distribution.

Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Historical facts

KURTOSIS was used by Karl Pearson in 1905 in "Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson. A Rejoinder", *Biometrika*, 4, 169-212, in the phrase "the degree of kurtosis".

He introduced the terms leptokurtic, platykurtic and mesokurtic, writing in Biometrika (1905), 5. 173:

> *"Given two frequency distributions which have the same variability as measured by the standard deviation, they may be relatively more or less flat-topped than the normal curve. If more flat-topped I term them platykurtic, if less flat-topped leptokurtic, and if equally flat-topped mesokurtic".*

Motivation

Example I: Petrobras returns

Histogram

Tail behavior

Example II: Automobile claims

Measuring variability

Measuring extremity

Historical facts

Computing kurtosis

Example III: Nasdaq returns

Example IV: Heavy vs thin

Final remarks

In his "Errors of Routine Analysis" Biometrika, 19, (1927), p. 160 Student provided a mnemonic:

* In case any of my readers may be unfamiliar with the term "kurtosis" we may define mesokurtic as "having β_2 equal to 3," while platykurtic curves have $\beta_2 < 3$ and leptokurtic > 3. The important property which follows from this is that platykurtic curves have shorter "tails" than the

normal curve of error and leptokurtic longer "tails." I myself bear in mind the meaning of the words by the above *memoria technica*, where the first figure represents platypus, and the second kangaroos, noted for "lepping," though, perhaps, with equal reason they should be hares!

Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Computing kurtosis in `excel`

In fact, excel computes excess kurtosis (in excess to the the normal kurtosis, which equals 3):

$$\kappa = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s_x} \right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$$

where $\bar{x}$ and s_x are, respectively, the sample mean and the sample standard deviation of the data.

When n is large, it converges to

$$\kappa_n = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{x_i - \bar{x}}{s_x} \right)^4 - 3$$

Motivation

Example I: Petrobras returns
Histogram
Tail behavior

Example II: Automobile claims

Measuring variability

Measuring extremity
Historical facts
Computing kurtosis

Example III: Nasdaq returns

Example IV: Heavy vs thin

Final remarks

Petrobras ($n = 3145$)

$$\kappa = 9.761795 \quad \text{and} \quad \kappa_n = 9.736272$$

Automobile claims ($n = 371$)

$$\kappa = 8.303171 \quad \text{and} \quad \kappa_n = 8.115396$$

Motivation

Example I: Petrobras returns

Histogram

Tail behavior

Example II: Automobile claims

Measuring variability

Measuring extremity

Historical facts

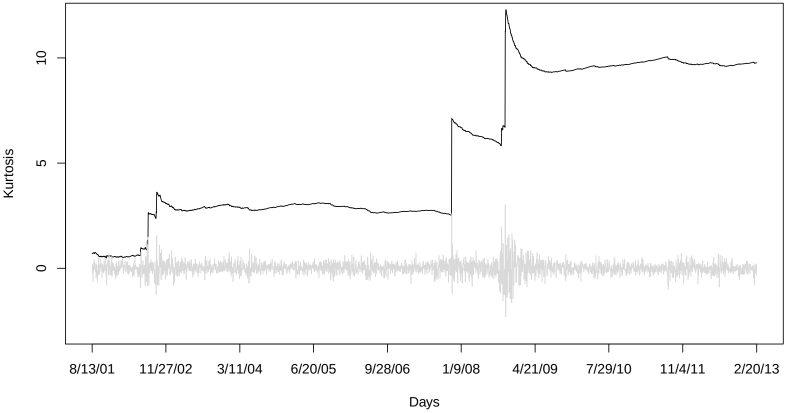
Computing kurtosis

Example III: Nasdaq returns

Example IV: Heavy vs thin

Final remarks

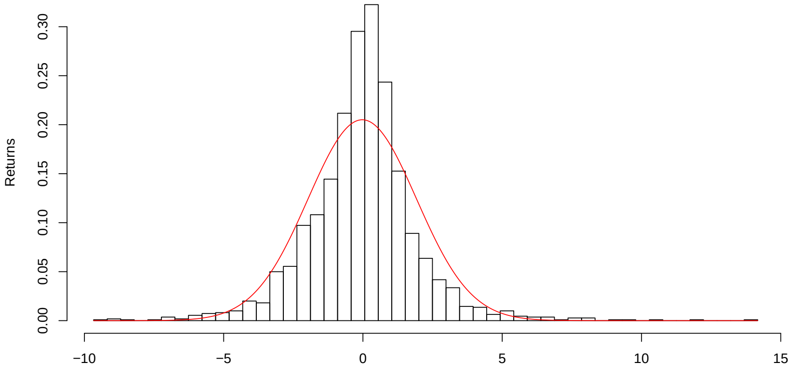
Petrobras kurtosis



Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Nasdaq returns

From 1/1/2000 to 12/312008 (2262 obs.)
Standard deviation = 1.945645
Excess kurtosis = 4.687585



Motivation

Example I: Petrobras returns

Histogram

Tail behavior

Example II: Automobile claims

Measuring variability

Measuring extremity

Historical facts

Computing kurtosis

Example III: Nasdaq returns

Example IV: Heavy vs thin

Final remarks

Tail behavior

	Empirical		Normal model	
Extreme	Prob.	Years	Prob.	Years
4.386	98.0%	0.2	98.83%	0.3
5.526	99.0%	0.4	99.78%	2.0
10.231	99.9%	4.0	100.00%	59000

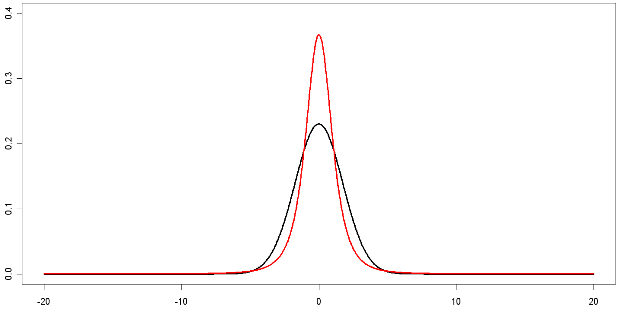
Prob. = Probability of the right tail

Years = expected number of years until rare event.

Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Synthetic data

Same mean $(= 0)$
Same variance $(= 3)$
Same skewness $(= 0)$
Different kurtosis



black curve: kurtosis $= 0.054$ (thin-tail distribution)
red curve: kurtosis $= 65.18$ (fat-tail distribution)

Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Tail behavior

Percentage of observations below threshold

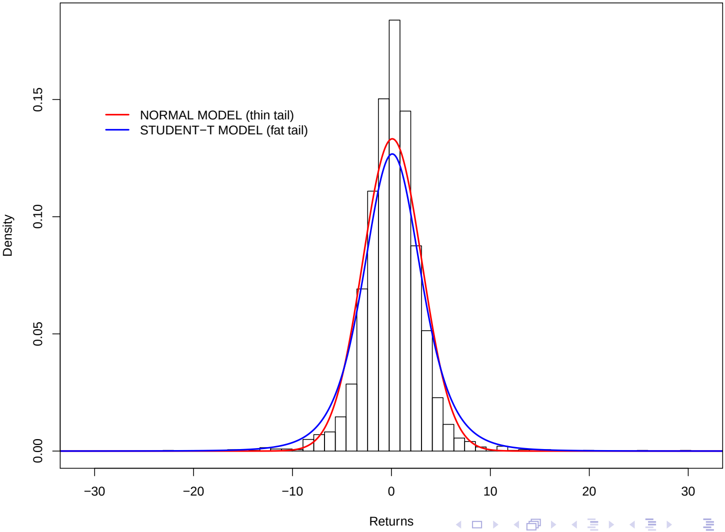
threshold	fat-tail	thin-tail
-6	0.4636	0.0571
-5	0.7696	0.3571
-4	1.4004	1.6004
-3	2.8834	5.1393
-2	6.9663	11.8255
-1	19.5501	19.4970

Fat-tail: On average, once every 1/0.004636=216 days (about one year) one data point would fall below 6 standard deviations.

Thin-tail: On average, once every 1/0.000571=1750 days (about seven years) one data point would fall below 6 standard deviations.

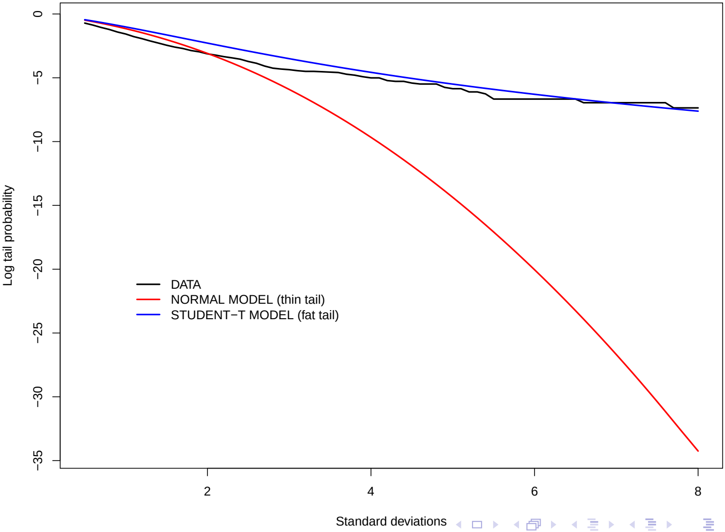
Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Revisiting Petrobras



Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Log tail probabilities



Motivation

Example I: Petrobras returns

Histogram

Tail behavior

Example II: Automobile claims

Measuring variability

Measuring extremity

Historical facts

Computing kurtosis

Example III: Nasdaq returns

Example IV: Heavy vs thin

Final remarks

Tail percentage

Standard Deviations	Data	Model	
		Student-t	Normal
4	0.667727	1.032342	0.006334
5	0.286169	0.410472	0.000057
6	0.127186	0.184614	0.000000
7	0.095390	0.091675	0.000000
8	0.063593	0.049291	0.000000
9	0.031797	0.028268	0.000000

Motivation
Example I: Petrobras returns
Histogram
Tail behavior
Example II: Automobile claims
Measuring variability
Measuring extremity
Historical facts
Computing kurtosis
Example III: Nasdaq returns
Example IV: Heavy vs thin
Final remarks

Final remarks

Lesson:
Studying volatilities is important, but studying kurtosis is too.

Modeling the tails:

- Extreme value theory
  - Extreme floods, large insurance losses, equity risks
  - Embrechts *et al.* (1997) *Modelling extremal events for insurance and finance*
  - Novak (2011) *Extreme value methods with applications to finance*
- Time-varying volatility modeling
- Multivariate modeling