Lecture 3

Fat tails:

Power laws distributions
Probability space (Kolmogorov)

1) \( p(x) \geq 0 \quad \forall x \in \Omega \)

2) \( p(\Omega) = 1 \)

3) For a set of disjoined events \( x_i \in \Omega \) with \( x_1 \cap x_2 = \emptyset \)
   \[
   p(x_1 \cup x_2 \cup x_3 \ldots) = \sum_i p(x_i)
   \]

Conditional probability

\[
p(A \mid B) = \frac{p(A \cap B)}{p(B)}
\]

Bayes Theorem

Two random events A and B are statistically independent if and only if

\[
p(A \cap B) = p(A)p(B)
\]

Chebyshev's inequality

\[
P\left( |X - m| > k\sigma \right) \leq \frac{1}{k^2}
\]

(defined variance)

The law of large numbers

(defined mean)
Central limit theorem

The sum of independent identically-distributed variables with finite variance will tend to be normally distributed.

\[ x_i(t) \text{ are N i.i.d. variables} \]

s.t. \[ \sigma_0^2 = \left< x_i^2 \right> - \left< x_i \right>^2 \text{ finite} \]

the sum \[ y = \sum_{i=1}^{N} x_i(t) \]

tends to the probability distribution

\[ p(y) \sim \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left( -\frac{(y - \langle y \rangle)^2}{2\sigma^2} \right) \]

if \( N \gg 1 \)

Any observable variable resulting from a sum of processes involving a number large enough of independent factors will be normally distributed.
Aggregate distribution for a sequence of numbers from 1 to 30k
Normal distributions must be everywhere....
...not quite everywhere!
absolute value of returns for S&P500
"How does the Eurodollar Interest Rate behave?"

T. Di Matteo and T. A.,
Income distribution in Australia

Fig. 1. – Complementary cumulative distributions for the Total annual income from all sources in Australia in the years 1993 – 1997.

Distribution of the Number of sexual partners (Sweden):

\[ p(k) \sim k^{-\alpha} \]

\[ \alpha \approx 3.4 \]

\[ p(k \leq 10) \sim 80\% \]

\[ p(k \geq 100) \sim 0.23\% \]

**FIG. 3.** The degree distribution of several real networks: (a) Internet at the router level. Data courtesy of Ramesh Govindan; (b) movie actor collaboration network. After Barabási and Albert 1999. Note that if TV series are included as well, which aggregate a large number of actors, an exponential cutoff emerges for large \( k \) (Amaral et al., 2000); (c) co-authorship network of high-energy physicists. After Newman (2001a, 2001b); (d) co-authorship network of neuroscientists. After Barabási et al. (2001).

more than 1600 over 700000 peoples

Power laws probabilities are VERY unusual
Central limit theorem
The sum of independent identically-distributed variables with finite variance will tend to be normally distributed accordingly to an ‘attractor distribution’.

Stable distributions
if a number of independent identically distributed random variables have a stable distribution, then a linear combination of these variables will have the same distribution.
A probability density is stable iff is of the form (Levy distribution):

\[
f(x, \alpha, \beta, c, \mu) = \int_{-\infty}^{+\infty} dk e^{ikx} \exp\left(ik\mu - ck|\alpha \left(1 - i\beta \frac{k}{|k|} \Phi(\alpha, k)\right)\right)
\]

\[0 < \alpha \leq 2 \quad -1 \leq \beta \leq 1 \quad c \geq 0\]

with \(\Phi(\alpha, k) = \tan\frac{\pi \alpha}{2}\) for \(\alpha \neq 1\)

or \(\Phi(\alpha, k) = \frac{2}{\pi} \log(|k|)\) for \(\alpha = 1\)

Asymptotic behavior for large \(x\)

\[
f(x, \alpha, \beta, c, \mu) \propto \frac{1}{|x|^{1+\alpha}}
\]

In practice...
How can we distinguish between fat tail distributions and normal-like distributions from a set of outcomes?

Power law distributions have extremely high kurtosis.

If we have a finite measurement of a set of outcomes of a given process we can spot deviation from the normal distribution by looking at the kurtosis and comparing it with the one expected from a normal distribution with same mean and standard deviation (excess kurtosis $\mu_4/\sigma^4-3$).
A more reliable test is to look at the inverse cumulate distribution $P_>(x)$. If the distribution has power law tails, a log-log plot will reveal a linear decreasing trend at large fluctuations.

The “rank-frequency” log-log plot is a very convenient to plot the cumulate distribution especially when the dataset size is not large. In this plot one first sorts the $n$ observed values in ascending order, and then plot them against the vector $[1,(n-1)/n,(n-2)/n,\ldots,1/n]$. 
Estimating the tail

Let $x_1, \ldots, x_n$ a sequence of $n$ observation (iid). The **Extreme Value Theory** tells us that regardless the underlying distribution function (under some little restrictive conditions) the probability to find $M_n = \max\{x_1, \ldots, x_n\}$ larger than $x$ must follow one of the three distributions:

$G(x) = e^{-e^{-x}}$

$G(x) = \begin{cases} 
0 & x \leq 0 \\
\frac{1}{\alpha} x^{-1/\alpha} & x > 0, \alpha > 0 \\
0 & x < 0, \alpha < 0 
\end{cases}$

$G(x) = \begin{cases} 
0 & x \geq 0 \\
\frac{1}{\alpha} (-x)^{-1/\alpha} & x < 0, \alpha < 0 
\end{cases}$

(Gumbel) (Fréchet) (Weibull)
For ‘fat-tailed’ data only the Frèche distribution is relevant i.e the probability that $M_n > x$ converges towards

\[
G(x) = \begin{cases} 
0 & x \leq 0 \\
 e^{-x^{-1/\alpha}} & x > 0 \quad \alpha > 0 
\end{cases}
\]

Essentially the theorem says that the inverse cumulate distribution tends to $G(x)$ in the region of large $x$. In such region $G(x)$ behave as a power law with tail exponent $\alpha$

**Hill estimator for the tail index**

\[
p(x) \sim \frac{\alpha - 1}{x_{\min}} \left( \frac{x}{x_{\min}} \right)^{-\alpha}
\]

\[
\alpha^* \sim 1 + \frac{n}{\sum_{i=1}^{n} \ln \frac{x_i}{x_{\min}}}
\]
Where do power laws come from?
Origin of large fluctuations

1) Collective phenomena such as avalanches

Power law tails in the avalanches’ sizes: No characteristic size
Origin of large fluctuations

2) Complex interactions (scale free networks)

Distance 1 from a Manhattan bank board

shareholder network for NYSE

shareholder network for Nasdaq

Productivity distributions

In complex systems, power laws are often seen as signatures of hierarchy.
Origin of large fluctuations

3) Irrational human behaviors
**Simon’s, Price’s model** *(cumulative advantage)*

Simon’s idea to explain power laws in wealth distributions: The rate at which a person increases his wealth is proportional to the wealth that he already has. Rich get richer.

Additional material and exercises
Maximum likelihood derivation of Hill estimator

\[ p(x) \sim \frac{\alpha - 1}{x_{\text{min}}} \left( \frac{x}{x_{\text{min}}} \right)^{-\alpha} \]

\[ p\{x\} | \alpha) = \prod_{i=1}^{n} p(x_i) = \frac{\alpha - 1}{x_{\text{min}}} \prod_{i=1}^{n} \left( \frac{x_i}{x_{\text{min}}} \right)^{-\alpha} \]

\[ p(\alpha | \{x\}) = p(\{x\} | \alpha) \frac{p(\alpha)}{p(\{x\})} = p(\{x\} | \alpha) \]

\[ \ln p\{x\} | \alpha) = n \ln \frac{\alpha - 1}{x_{\text{min}}} - \alpha \sum_{i=1}^{n} \ln \left( \frac{x_i}{x_{\text{min}}} \right) \]

\[ \frac{\partial}{\partial \alpha} \ln p\{x\} | \alpha) = 0 \]

\[ \alpha^* \sim 1 + \frac{n}{\sum_{i=1}^{n} \ln \frac{x_i}{x_{\text{min}}}} \]
Exercise

Download the file:

dataSeries_RDEAAA.dat

From


There are 6 columns and 1700 rows, each column represent a different stochastic process.
• Which are kind of statistics underlying such processes?
• Which are the probability distributions describing the data?
• Are there any (auto) correlations? (memory effects)
• How can one build similar behaviors?

Suggestions
- see the series as results of random walks
- study the increments \( x(t+h)-x(t) \)

PS you need a computer