# Introduction to Heavy-Tailed Distributions, Self-Similar Processes, and Long-Range Dependence

Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu

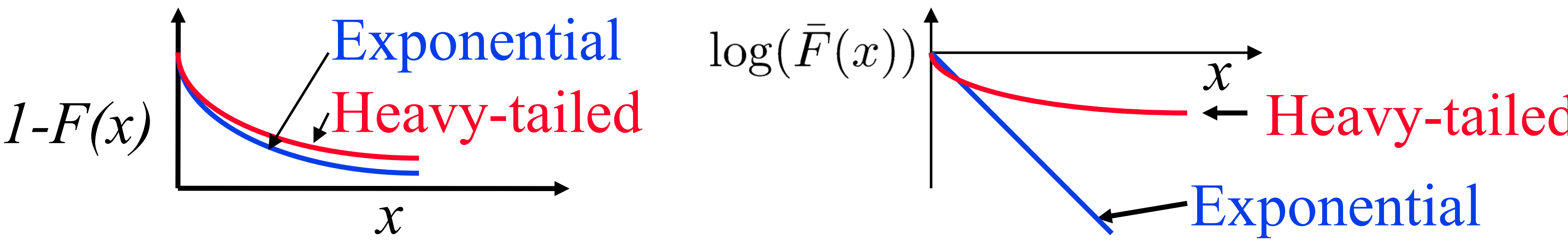
These slides are available on-line at:

http://www.cse.wustl.edu/~jain/cse567-15/


1. Heavy-Tailed Distributions (HTDs)
2. How to Check for Heavy Tail?
3. Self-Similar Processes
4. Long Range Dependence (LRD)
5. Generating LRD Sequences
6. Self-Similarity vs. LRD
7. Hurst Exponent Estimation

- Note: These slides are based on R. Jain, "The Art of Computer Systems Performance Analysis," 2nd Edition (in preparation).

# Heavy-Tailed Distributions (HTDs)




- CCDF decays slower than the exponential distribution

$$P(X > x) = 1 - F(x) = \bar{F}(x) = e^{-\lambda x}$$

- CCDF = Complementary cumulative distribution function

$$\bar{F}(x) = 1 - F(x)$$

- For heavy tailed distributions, CCDF is slower by some power of $x$ $\quad \bar{F}(x) \rightarrow c x^n e^{-\lambda x}$
- Very large values possible

# Examples of HTD Variables

- Many real-world phenomenon have been found to follow heavy tailed distributions.
  - Distribution of wealth.
    One percent of the population owns 40% of wealth.
  - File sizes in computer systems
  - Connection durations
  - CPU times of jobs
  - Web pages sizes
- Significant impact on buffer sizing in switches and routers.

# Example 38.1

- Weibull distribution

$$F(x) = 1 - e^{-(x/a)^b} \qquad a > 0, b > 0$$

$$\lim_{x\to\infty} \frac{x^n e^{-\lambda x}}{1-F(X)} = \frac{x^n e^{-\lambda x}}{e^{-(x/a)^b}} = e^{((x/a)^b - \lambda x) + n \ln x} = 0 \text{ for } 0 < b < 1$$

- Other examples of heavy tailed distributions are Cauchy, log-normal, and t-distributions.

# Power Tailed Distributions

- A subset of heavy tailed distributions

$$\bar{F}(x) \Rightarrow \frac{c}{x^{\alpha}}$$

- CCDF approaches a power function for large x
  - For such distributions:
    all moments E[$x^l$] for all values of $l \geq \alpha$ are infinite.
  - If $\alpha \leq 2$, $x$ has infinite variance
  - If $\alpha \leq 1$, the variable has infinite mean
  - $\alpha$ is called the **tail index**.

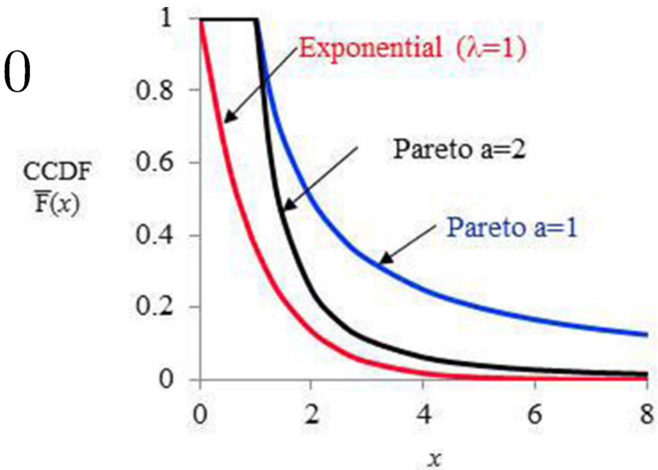
- All power tailed distributions are heavy tailed but all heavy tailed distributions are not power tailed.

# Example 38.2

- Pareto distribution:

$$F(x) = 1 - x^{-\alpha} \qquad 1 \le x \le \infty, \alpha > 0$$

$$\bar{F}(x) = 1 - F(X) = \frac{1}{x^{\alpha}}$$




- pdf: $f(x) = \frac{d}{dx}F(x) = \alpha x^{-\alpha-1}$
- It's $l^{\text{th}}$ moment is:

$$\text{For } l \ne \alpha: \quad E[x^l] = \int_1^\infty x^l f(x)dx = \int_1^\infty \alpha x^{l-\alpha-1}dx = \frac{\alpha}{l-\alpha}x^{l-\alpha}\Big|_1^\infty = \begin{cases} \infty & l > \alpha \\ \frac{\alpha}{\alpha-l} & l < \alpha \end{cases}$$

$$\text{For } l = \alpha: \quad E[x^l] = \int_1^\infty \alpha x^{-1}dx = \alpha \ln x|_1^\infty = \infty$$

- All moments for $l \ge \alpha$ are infinite.
- For $2 \ge \alpha > 1$ variance and higher moments are infinite. For $1 \ge \alpha$ variance does not exist.
- For $1 \ge \alpha > 0$, even mean is infinite.

# Effect of Heavy Tail

- A random variable with HTD can have very large values with finite probabilities resulting in many outliers.
- Sampling from such distributions results in mostly small values with a few very large valued samples.
- Sample statistics (e.g., sample mean) may have a large variance $\Rightarrow$ sample sizes required for a meaningful confidence are large.
- Sample mean generally under-estimates the population mean.
- Simulations with heavy-tailed input require very long time to reach steady state and even then the variance can be large.

  $$|\bar{x}_n - \mu| \approx cn^{\frac{1}{\alpha}-1}$$

  c is some constant.

# Effect of Heavy Tail (Cont)

- The number of observations required to reach k-digit accuracy:

$$\frac{|\bar{x}_n - \mu|}{\mu} \leq 10^{-k} \qquad \frac{c n^{\frac{1}{\alpha}-1}}{\mu} \leq 10^{-k} \qquad n \geq 10^{\frac{\log(c/\mu)+k}{1-\frac{1}{\alpha}}}$$

- Assuming c=1, $\mu$=1, $10^{11}$ observations are required for a single decimal digit accuracy (*k=1*) if $\alpha$=1.1.
- Central limit theorem applies only to observations from distributions with finite variances.
  - For heavy-tailed distributions with infinite variance, the central limit theorem does not apply.
  - The sample mean does not have a normal distribution even after a large number of samples.
  - Confidence interval formulas mentioned earlier can not be used.

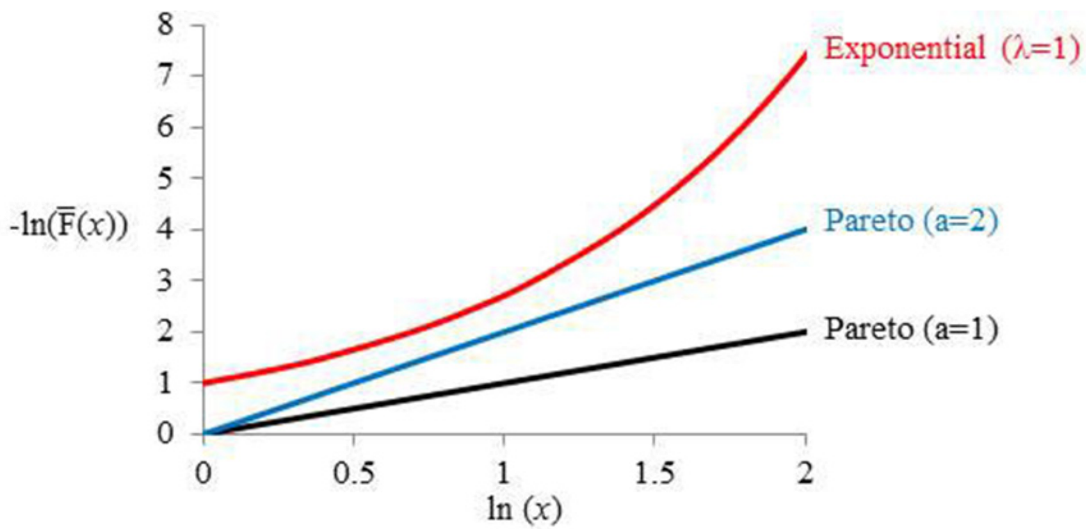
# Effect of Heavy Tail (Cont)

- **M/PT/1 queue**: Poisson arrivals and power-tailed service time
  - pdf of queue length $f(n) \rightarrow c(\rho)/n^{\alpha}$
    where $c(\rho)$ is a function of the traffic intensity $\rho$.
  - If $\alpha \leq 1$, the mean service time is infinite and so are the traffic intensity and the mean queue length.
  - If $\alpha \leq 2$, the service time has infinite variance, and so does the queue length.
- **PT/M/1 queue**:
  - Tail index $\alpha \leq 1$, the mean inter-arrival time is infinite.
  - For $1 < \alpha \leq 2$, the variance of the inter-arrival time is infinite.
- Heavy tailed-ness also implies **predictability**:
  - If a heavy tailed task has run a long time, it is expected to run for an additional long time.

$$\lim_{x \to \infty} E[X - x | X > x] = \infty$$

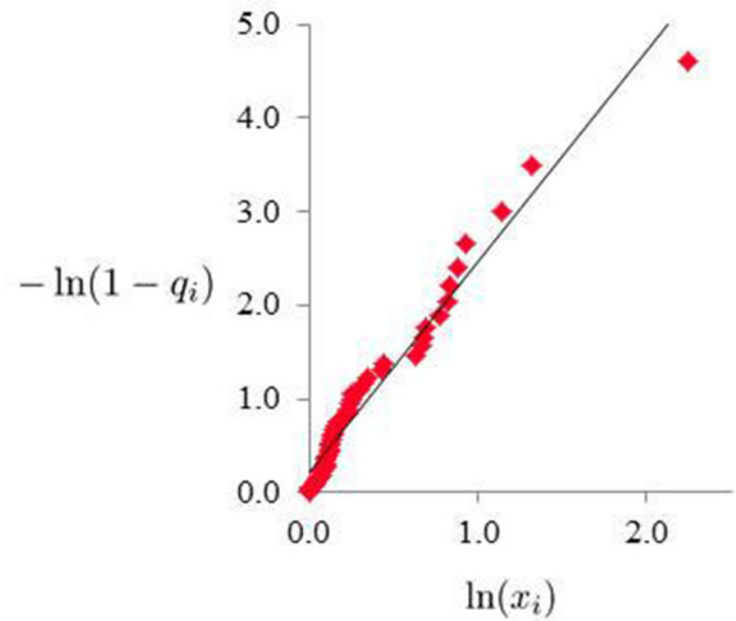
# How to Check for Heavy Tail?

- Make a Q-Q plot on a log-log graph assuming a Pareto distribution
- $F(x) = 1-x^{-\alpha} \quad x=(1-F)^{-1/\alpha}$
- On a log-log graph: $\ln x = (-1/\alpha) \ln (1-F)$
- Find $\alpha$ from the slope of the best-fit line. $\alpha \geq 1 \Rightarrow$ Heavy Tailed



# Example 38.3

- Check if this set of 50 observations has a heavy tail: 2.426, 1.953, 1.418, 1.080, 3.735, 2.307, 1.876, 1.110, 3.131, 1.134, 1.171, 1.141, 2.181, 1.007, 1.076, 1.131, 1.156, 2.264, 2.535, 1.001, 1.099, 1.149, 1.225, 1.099, 1.279, 1.052, 1.051, 9.421, 1.346, 1.532, 1.000, 1.106, 1.126, 1.293, 1.130, 1.043, 1.254, 1.118, 1.027, 1.383, 1.288, 1.988, 1.561, 1.106, 1.256, 1.187, 1.084, 1.968, 1.045, 1.155



# Example 38.3 (Cont)

| $x_i$ | Rank $R_i$ | Quantile $q_i = \frac{R_i - 0.5}{n}$ | $-\ln(1-q_i)$ | $\ln(x_i)$ | $-\ln(1-q_i) \times \ln(x_i)$ |
|---|---|---|---|---|---|
| 2.426 | 46 | 0.910 | 0.886 | 2.408 | 2.134 |
| 1.953 | 40 | 0.790 | 0.669 | 1.561 | 1.044 |
| 1.418 | 36 | 0.710 | 0.349 | 1.238 | 0.432 |
| 1.080 | 10 | 0.190 | 0.077 | 0.211 | 0.016 |
| 1.735 | 49 | 0.970 | 1.318 | 3.507 | 4.620 |
| 1.307 | 45 | 0.890 | 0.836 | 2.207 | 1.845 |
| 1.876 | 39 | 0.770 | 0.629 | 1.470 | 0.925 |
| 1.110 | 16 | 0.310 | 0.105 | 0.371 | 0.039 |
| … | … | … | … | … | |
| 1.084 | 11 | 0.210 | 0.236 | 0.081 | 0.019 |
| 1.968 | 41 | 0.810 | 1.661 | 0.677 | 1.124 |
| 1.045 | 6 | 0.110 | 0.117 | 0.044 | 0.005 |
| 1.155 | 24 | 0.470 | 0.635 | 0.144 | 0.092 |
| | | Sum | 17.308 | 49.654 | 37.006 |
| | | Sum of Squares | 14.781 | 95.774 | 164.977 |
| | | Average | 0.346 | 0.993 | 0.740 |

# Example 38.3 (Cont)

$$b_1 = \frac{\Sigma xy - n\overline{xy}}{\Sigma x^2 - n\overline{x}^2}$$

$$\frac{1}{\alpha} = \frac{37.006 - 50(0.346)(0.993)}{95.774 - 50(0.993)^2} = \frac{37.006 - 17.188}{95.774 - 49.311} = \frac{19.818}{46.463} = 0.427$$

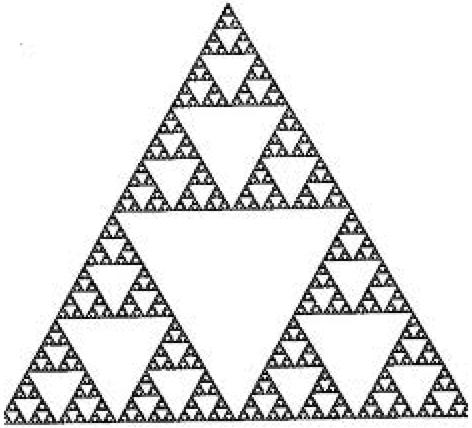
$$\alpha = 2.34$$

# Exercise 38.6

- Check if the following set of 50 observations has a power tail: 2.24, 1.67, 1.86, 1.12, 1.31, 1.63, 1.83, 7.87, 34.75, 4.24, 1.60, 2.51, 1.67, 8.04, 1.81, 5.47, 2.85, 2.05, 4.51, 1.85, 6.15, 1.86, 3.47, 2.84, 11.71, 3.02, 12.88, 1.36, 2.10, 18.85, 1.17, 31.43, 5.70, 1.76, 1.04, 1.29, 1.24, 2.50, 1725.34, 28.14, 1.43, 4.06, 1.56, 3.77, 1.00, 3.03, 118.59, 5.40, 1.01, 1.38

# Self-Similarity





- When zoomed, the sub objects have the same shape as the original object
- Also called Fractals
- Latin "fractus" = "fractional" or "broken"
  ⇒Traditional Euclidean geometry can not be used to analyze these objects because their perimeter is infinite.

# Self-Similar Processes

- Scaling in time = scaling in magnitude

$$x_{at} \sim a^H x_t \qquad \forall a > 0$$

- Statistical similarity ⇒ Similar distributions with similar mean and variance

$$E[x_{at}] = a^H E[x_t]$$

$$\text{Var}[x_{at}] = a^{2H} \text{Var}[x_t]$$

  - Similar variance ⇒ Self-similar in the second order
  - Similar higher order moments ⇒ Self-similarity of higher orders
  - All moments similar ⇒ strictly self-similar.

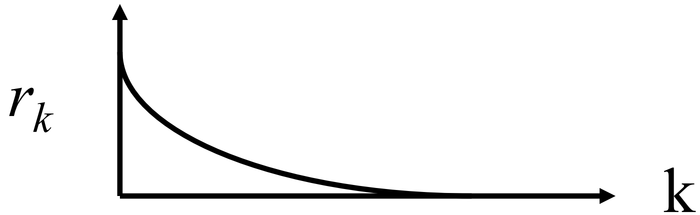
# Example 38.4

- Consider the white noise process $e_t$ with zero mean and unit variance: $e_t = z \sim N(0,1)$
  Here $z$ is the unit normal variate.
- Consider the process $x_t$: $x_t = t^H e_t = t^H z$
- For this process:

$$x_{at} = (at)^H z = a^H t^H z = a^H x_t$$

  - Therefore, $x_t$ is a self-similar process.
  - H = Hurst exponent
- Harold Edwin Hurst, a hydrologist, who was studying optimum dam sizing for reservoirs along Nile River in Egypt

# Short Range Dependence (SRD)




- Sum of Autocorrelation function is finite.
- Example 38.5: AR(1) with zero mean:

$$x_t = a_1 x_{t-1} + e_t$$

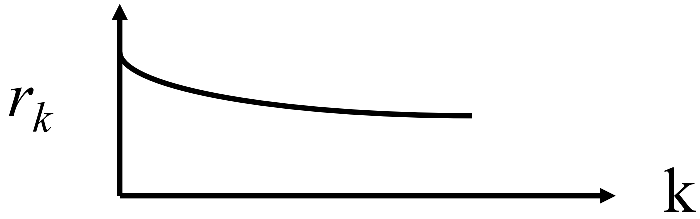
- For this process, the autocorrelation function decreases exponentially:

$$Cor[x_t, x_{t-k}] = r_k = a_1^{|k|}$$

- Sum of autocorrelations is finite (provided $|a_1| < 1$ ):

$$\sum_{k=0}^{\infty} r_k = \frac{1}{1 - a_1}$$

# Long Range Dependence (LRD)




- Sum of Autocorrelation function is infinite $\sum r_k = \infty$

Alternative Definition:

- Limiting tail behavior of the autocorrelation:

$$r(k) \to k^{2H-2} L(k) \qquad k \to \infty$$

Here, L(x) is a ***slowly varying function*** of *x*.

*L(ax)/L(x)* tends to 1 as *x* approaches infinity.

Constants and logarithms are examples of slowly varying functions.

# Examples of Processes with LRD

- Aggregation of a large number on-off processes with heavy-tailed on-times or heavy-tailed off times results in long-range dependence.
- File sizes have a long-tailed distribution
  $\Rightarrow$ Internet traffic has a long range dependence.
- Connection durations have also been found to have a heavy-tailed distribution $\Rightarrow$ traffic has a long range dependence
- UNIX processes have been found to have a heavy-tailed distribution $\Rightarrow$ resource demands have LRD
- Congestion and feedback control mechanisms such as those used in Transmission Control Protocol (TCP) increase the range of dependence in the traffic.

# Effect of Long Range Dependence

- Long-range dependence invalidates all results for queueing theory obtained using Poisson processes, e.g., Buffer sizes required to avoid overflow may be off by thousands times.

# Self-Similarity vs. LRD

- Self-similarity $\neq$ Long-range dependence
- Self-similar process can be short-range dependent or long-range dependent
- Self-similar processes with $\frac{1}{2} < H < 1$ have long range dependence.
- Self-similar processes with $0 < H \leq \frac{1}{2}$ have short range dependence.
- ARIMA(p, d, q) with integer valued d are SRD.
- FARIMA(p, d+$\delta$, q) with $0<\delta<\frac{1}{2}$ have long-range dependence.

# FARIMA Models and LRD

- Fractional Auto-regressive Integrated Moving Average (FARIMA) processes exhibit LRD for certain values of *d*.
- Consider FARIMA(0, 0.25, 0): $(1-B)^{0.25}x_t = e_t$

$$\begin{aligned} x_t &= (1-B)^{-0.25}e_t \\ &= e_t - (-0.25)e_{t-1} + \frac{-0.25(-0.25-1)}{(1)(2)}e_{t-2} + \frac{-0.25(-0.25-1)(-0.25-2)}{(1)(2)(3)}e_{t-3} + \cdots \\ &= e_t + 0.25e_{t-1} + \frac{0.25(0.25+1)}{(1)(2)}e_{t-2} + \frac{0.25(0.25+1)(0.25+2)}{(1)(2)(3)}e_{t-3} + \cdots \\ &= e_t + 0.25e_{t-1} + 0.16e_{t-2} + 0.12e_{t-3} + 0.10e_{t-4} + \cdots + 0.04e_{t-10} \end{aligned}$$

The coefficient of $e_{t-k}$ is $\frac{0.25(0.25+1)\cdots(0.25+k-1)}{(1)(2)\cdots(k)} = \frac{\Gamma(0.25+k)}{\Gamma(0.25)\Gamma(k+1)}$

Here, $\Gamma()$ is the Gamma function: $\Gamma(p+1) = p\Gamma(p)$

It is a generalization of factorial. For integer *p*, $\Gamma(p+1) = p!$

For example, $\Gamma(3) = 2, \quad \Gamma(2) = 1, \quad \Gamma(1) = 1, \quad \Gamma(0) = \infty$

# FARIMA Models and LRD

- Consider FARIMA(0, $\delta$, 0) with -1/2 < $\delta$ <1/2 and $\delta \neq 0$.

$$x_t = (1-B)^{-\delta} e_t = \sum_{k=0}^{\infty} b_k e_{t-k}$$

Where:

$$b_k = \frac{\Gamma(k+\delta)}{\Gamma(\delta)\Gamma(k+1)}$$

Since $e_t$ is Gaussian, $x_t$ is also Gaussian.

$e_t$ is Gaussian Noise, $x_t$ is fractional Gaussian Noise (fGn)

# FARIMA and LRD (Cont)

- The autocovariance of the FARIMA(0,$\delta$,0) sequence is:

$$E[x_t x_{t-k}] = \begin{cases} \sigma^2 \frac{\Gamma(1-2\delta)}{\Gamma^2(1-\delta)} & k = 0 \\ \sigma^2 \frac{\Gamma(k+\delta)\Gamma(1-2\delta)}{\Gamma(k-\delta+1)\Gamma(\delta)\Gamma(1-\delta)} & k \neq 0 \end{cases}$$

- Autocorrelation at lag k: $r_k = \dfrac{\Gamma(k+\delta)\Gamma(1-\delta)}{\Gamma(k-\delta+1)\Gamma(\delta)}$

- Stirling's approximation: $\Gamma(p+1) \approx \sqrt{2\pi p}\left(\dfrac{p}{e}\right)^p$

- For large $k$, $r_k$ tends to $c|k|^{2\delta-1}$ where $c = \dfrac{\Gamma(1-\delta)}{\Gamma(\delta)}$

- Recall that for LRD: $r(k) \to k^{2H-2}L(k) \qquad k \to \infty$
  $\Rightarrow$ 2H-2=2$\delta$-1, that is, H=$\delta$+1/2.
  $\Rightarrow$ A FARIMA(0,$\delta$,0) sequence has LRD if 0<$\delta$<1/2.

# Generating LRD Sequences

- Generate the FARIMA(p, d+$\delta$, q) LRD sequence
- FARIMA(p, d+$\delta$, q) = ARIMA(p, d, q) with $e_t$ replaced by fractional Gaussian noise generated by FARIMA(0, $\delta$, 0)
- ARIMA(p,d,q) is given by

$$\phi(B)(1-B)^d x_t = a_0 + \psi(B)e_t$$

- It can be generated by one of the following two methods:

1. Using previous values of $x_t$:

$$x_t = a_0 + (1-\phi(B)(1-B)^d)x_t + \psi(B)e_t$$

# Generating LRD Sequences (Cont)

2. Converting the model to a moving average model using a Taylor series expansion:

$$x_t = \frac{\psi(B)}{\phi(B)(1-B)^d} e_t = \left( \sum_{i=0}^{m} c_i B^i \right) e_t = \sum_{i=0}^{m} c_i e_{t-i}$$

Here $c_i$ are coefficients of the Taylor series expansion and $m$ is selected large enough so that $c_i$ for $i > m$ are negligible.

- Generate a white noise sequence $e_i \sim N(0, 1)$
- Generate a FARIMA(0, δ, 0) sequence $y_i$ using a moving average of a large number $m$ of $e_i$:

$$y_i = \sum_{k=0}^{m} \frac{\Gamma(k+\delta)}{\Gamma(k+1)\Gamma(\delta)} e_{i-k}$$

# Generating LRD Sequences (Cont)

- Generate a FARIMA(p, d+$\delta$, q) sequence $x_i$ by generating a usual ARIMA(p, d, q) as in Step 1 above with the white noise $e_i$ replaced by $y_i$

$$x_t = a_0 + \frac{\psi(B)}{\phi(B)(1-B)^d} y(i)$$

- m=100 or m=1000 has been found to provide good results.

# Example 38.6

- Generate a FARIMA(0,0.25,0) Sequence

$$\begin{aligned} x_t &= (1-B)^{-0.25} e_t \\ &= e_t + 0.25 e_{t-1} + \frac{0.25(0.25+1)}{2} e_{t-2} + \frac{0.25(0.25+1)(0.25+2)}{6} e_{t-3} + \cdots \\ &= e_t + 0.25 e_{t-1} + 0.16 e_{t-2} + 0.12 e_{t-3} + 0.10 e_{t-4} + \cdots + 0.04 e_{t-10} \end{aligned}$$

- Generate 60 N(0,1) random numbers for $e_{-9}$ thru $e_{50}$
- The numbers are: 0.376, 0.789, -0.629, 0.102, 0.240, -0.909, 0.706, 0.019, -0.646, -2.030, -0.140, -1.816, 1.373, -0.723, -1.486, -0.984, -0.392, -0.323, 0.214, 0.652, -0.148, 0.499, -0.226, -1.878, 0.975, 0.273, -0.080, 0.040, 1.607, -0.154, -0.601, -0.468, 0.199, 1.129, 0.299, 1.332, -0.760, 0.980, -0.134, 1.378, -1.059, 0.364, -0.715, 0.769, -1.671, -0.346, 0.195, 0.157, -0.038, -1.253, -0.773, -0.910, 0.304, 1.146, 1.630, 0.578, 1.349, 0.615, -0.396

# Example 38.6 (Cont)

- Use the above equation to get $x_1$ through $x_{50}$
- The numbers are: -2.16, -0.77, -2.23, 0.61, -0.93, -1.85, -1.68, -1.15, -0.96, -0.27, 0.33, -0.18, 0.32, -0.23, -1.90, 0.54, 0.27, -0.02, 0.04, 1.58, 0.23, -0.43, -0.48, 0.19, 1.16, 0.60, 1.61, -0.23, 1.08, 0.22, 1.63, -0.48, 0.51, -0.51, 0.74, -1.53, -0.63, -0.10, 0.01, -0.19, -1.35, -1.19, -1.36, -0.29, 0.87, 1.70, 0.97, 1.70, 1.18, 0.20
- Note:
  1. m = 10 is used for illustration only.
     Need to use m = 1000
  2. A sample of 50 observations is too small to study long-range dependence.

# Exercise 38.9

- Generate a sample of 50 observations for a long-range dependent process with a Hurst exponent of 0.65. Use the following sequence of 60 unit normal variates: 0.24, -0.91, 0.71, 0.02, -0.65, -2.03, -0.14, -1.82, 1.37, -0.72, -1.49, -0.98, -0.39, -0.32, 0.21, 0.65, -0.15, 0.50, -0.23, -1.88, 0.98, 0.27, -0.08, 0.04, 1.61, -0.15, -0.60, -0.47, 0.20, 1.13, 0.30, 1.33, -0.76, 0.98, -0.13, 1.38, -1.06, 0.36, -0.71, 0.77, -1.67, -0.35, 0.20, 0.16, -0.04, -1.25, -0.77, -0.91, 0.30, 1.15, 1.63, 0.58, 1.35, 0.61, -0.40, -1.60, 0.02, 0.55, -1.45

# Hurst Exponent Estimation

**Variance-time plot** (Similar to the method of batch means)

1. Start with $m$=1
2. Divide the sample of size $n$ in to non-overlapping subsequences of length m. There will be $j = \lfloor n/m \rfloor$ such subsequences.
3. Take the sample mean of each subset

$$\bar{x}_{km} = \frac{1}{m} \sum_{i=(k-1)m+1}^{i=km} x_i \qquad k = 1, 2, 3, \ldots, j$$

4. Compute the overall mean:

$$\bar{\bar{x}}_m = \frac{1}{j} \sum_{k=1}^{k=j} \bar{x}_{km}$$

# Hurst Parameter Estimation (Cont)

5. Compute the variance of the sample means

$$s_m^2 = \frac{1}{j-1} \sum_{k=1}^{k=j} (\bar{x}_{k,m} - \bar{\bar{x}}_m)^2$$

6. Repeat steps 2 through 5 for m=1, 2, 3, …
7. Plot variance $s_m^2$ as a function of the subsequence size *m* on a log-log graph
8. Fit a simple linear regression to log(var) vs. log (*m*).
9. The slope of the regression line is 2H-2.
10. That is, the Hurst exponent is $1+a_1/2$, where $a_1$ is the slope of the regression line.

Note: 1. H estimate using this variance time plot method is biased

2. If a process is non-stationary, it may not be self-similar or have LRD, but may result in Hurst exponent between 0.5 and 1

# Example 38.7

- Determine the Hurst exponent for the data of Example 38.6
- Batch size $m$=1: 50 batches, batch mean $\bar{x}_i = x_i$

  Overall mean $\bar{\bar{x}}_1 = \frac{1}{50}\sum_{i=1}^{50} x_i = -0.21$

  Variance of batch means $Var[\bar{x}_1] = \frac{1}{49}\sum_{i=1}^{50}(x_i - \bar{\bar{x}}_1)^2$
- Batch size m=2: 25 batches, batch mean $\bar{x}_{2i} = \frac{1}{2}(x_{2i-1} + x_{2i})$

  Overall mean $\bar{\bar{x}}_2 = \frac{1}{25}\sum_{i=1}^{25} \bar{x}_{2i} = -1.4627$

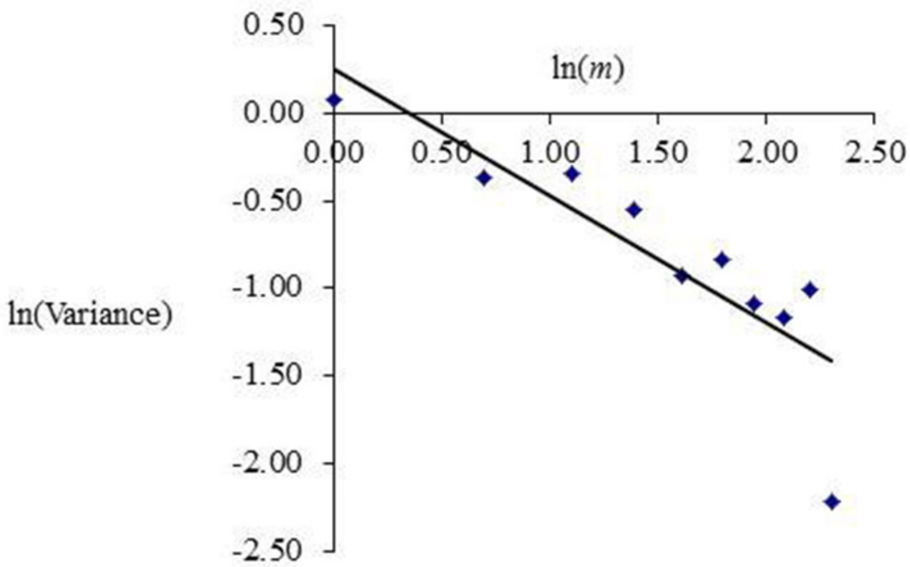
  Variance of batch means $var[\bar{x}_2] = \frac{1}{24}\sum_{i=1}^{25}(\bar{x}_{2i} - \bar{\bar{x}}_2)^2 = 0.52$

# Example 38.7 (cont)

| Batch Size $m$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| Overall Mean $\overline{\overline{x}}_m$ | -0.10 | -0.13 | -0.13 | -0.06 | -0.11 | 0.01 | -0.10 | 0.04 | 0.05 |
| Variance | 1.09 | 0.52 | 0.71 | 0.58 | 0.40 | 0.43 | 0.34 | 0.31 | 0.37 |
| ln( $m$ ) | 0.00 | 0.69 | 1.10 | 1.39 | 1.61 | 1.79 | 1.95 | 2.08 | 2.20 |
| ln(Variance) | 0.08 | -0.37 | -0.34 | -0.55 | -0.92 | -0.84 | -1.08 | -1.16 | -1.00 |

- Last few observations are not used for some batch sizes
  For example, m=9, only first 45 observations are used
  The overall mean is therefore different

# Example 38.7 (cont)



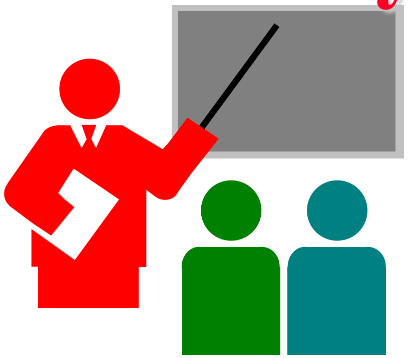

Slope = -0.73
H = 1 -0.73/2 = 0.635
ARIMA(0,0.25,0) ⇒ H=0.75

- Hurst exponent using variance time plot is biased
- If x_t is non-stationary, it may not be self-similar or may not have LRD, but it may still result in 0.5<H<1

# Exercise 38.10

- Estimate the Hurst exponent for the observations generated in Exercise 38.9

# Summary



1. Heavy tailed distributions: CCDF tail higher than exponential distribution
2. Self-Similar Process: $x_{at} \sim a^H x_t$
3. Long Range Dependence: $\sum r_k = \infty$
4. ARIMA(p,d+$\delta$,q) with 0<$\delta$<0.5 can be used to generate LRD sequences
5. Hurst parameter can be estimated with variance-time plots. For LRD 0.5<H<1.

# References


- Fractals, http://mathworld.wolfram.com/Fractal.html
- O. Sheluhin, S. M. Smolskiy, and A. V. Osin, "Self-Similar Processes in Telecommunications," Wiley, Chichester, England, 2007, 314 pp.
- P. Doukhan, G. Oppenheim, and M. S. Taqqu, (Eds), "Theory and Applications of Long Range Dependence," Birkhauser, Boston, MA.
- K. Park, W. Willinger, (Eds), "Self-Similar Network Traffic and Performance Evaluation," Wiley, New York, NY, 2000.
- R. J. Adler, R. E. Feldman, and M. S. Taqqu, (Eds), "A Practical Guide to Heavy Tails," Birkhauser, Boston, 1998