

Infinite & undefined Expectation

Pareto distributions

in Computer Science

Expected value over countably infinite sets

S = a countably infinite subset of \mathbb{R}

$S = \{s_1, s_2, \dots\}$

X = a random variable which gets values in S

$$E(X) \doteq \sum_{i=1}^{\infty} s_i \Pr(X(\omega) = s_i)$$

Recall some facts about series:

1.

$$\sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6}$$

2.

$$\sum_{i=1}^{\infty} \frac{1}{i} = \infty$$

3.

$$\sum_{i=1}^{\infty} \left(\frac{1}{i}\right)^d = \begin{cases} \infty & \text{if } 0 < d \leq 1 \\ \text{Finite} & \text{if } d > 1 \end{cases}$$

Consider the distribution

*distribution is over
the natural numbers = positive integers*

$$P(X = i) = \frac{1}{zi^3}; \quad z = \sum_{i=1}^{\infty} \frac{1}{i^3} < \infty \quad \textbf{Distribution is well defined}$$

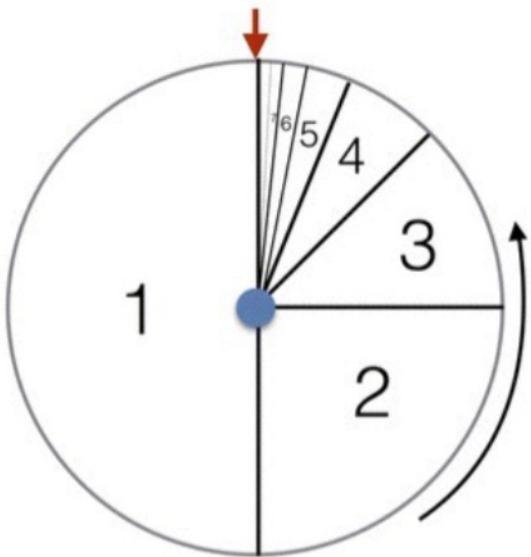
$$E(X) = \sum_{i=1}^{\infty} \frac{i}{zi^3} = \sum_{i=1}^{\infty} \frac{1}{zi^2} < \infty \quad \textbf{Expectation is finite}$$

$$E[X^2] = \sum_{i=0}^{\infty} \frac{i^2}{zi^3} = \sum_{i=0}^{\infty} \frac{1}{zi} = \infty \quad \textbf{Variance is infinite}$$

Consider next the distribution

$$P(X = i) = \frac{1}{zi^2}; \quad z = \sum_{i=1}^{\infty} \frac{1}{i^2} < \infty$$

$$E(X) = \sum_{i=1}^{\infty} \frac{i}{zi^2} = \sum_{i=1}^{\infty} \frac{1}{zi} = \infty \quad \begin{aligned} &\textbf{Distribution is well} \\ &\textbf{defined but} \\ &\textbf{Expectation is infinite} \end{aligned}$$



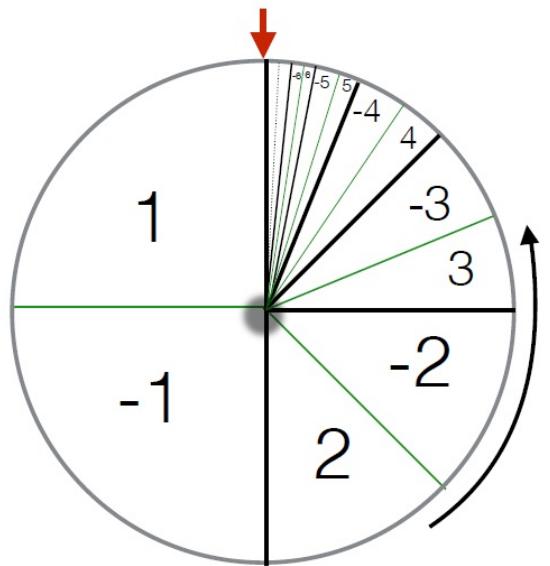
(b) A wheel with
Infinitely many outcomes

$$P(X = i) = \frac{6}{\pi^2 i^2};$$

$$\sum_{i=1}^{\infty} P(X = i) = 1$$

$$\sum_{i=1}^{\infty} i P(X = i) = \infty$$

Participation in this game is worth any price
(on the long term)



A wheel with
Infinitely many outcomes
both positive and negative

Consider a game with both wins and losses
 $i \in \{0, -1, +1, -2, +2, \dots\}$

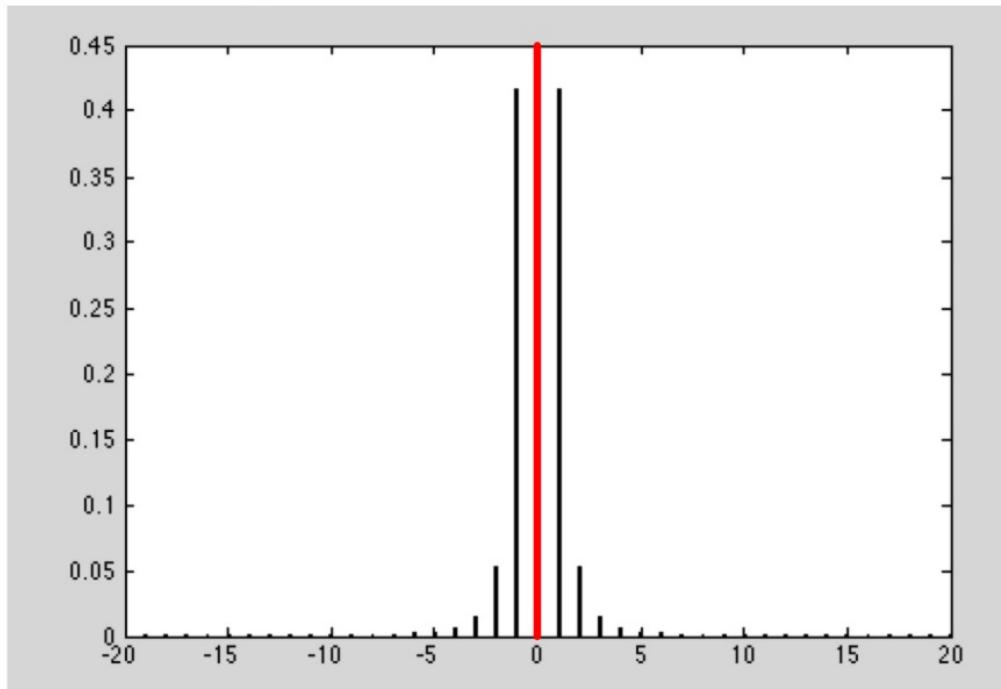
$$P(X = i) = \begin{cases} \frac{1}{Z} \frac{1}{i^{1.5}} & \text{if } i \neq 0 \\ 0 & \text{if } i = 0 \end{cases}, \quad Z = 2 \sum_{i=1}^{\infty} \frac{1}{i^{1.5}}$$

$$\sum_{i=-\infty}^{\infty} P(X = i) = 1$$

$$\sum_{i=\infty}^{\infty} iP(X = i) \text{ is undefined}$$

Expectation over pos and neg integers: the good case

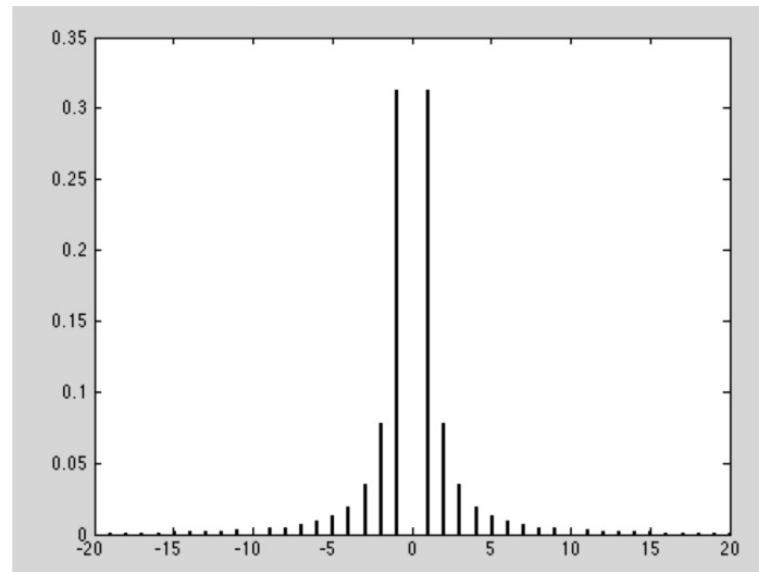
$$P(X = i) = \begin{cases} 0 & \text{if } i = 0 \\ \frac{1}{Z|i|^3} & \text{if } i \neq 0 \end{cases} ; \quad Z = 2 \sum_{i=1}^{\infty} \frac{1}{|i|^3} < \infty$$



$$E(X) = \sum_{i=1}^{\infty} iP(X = i) + \sum_{i=-1}^{-\infty} iP(X = i) = \frac{1}{Z} \left(\sum_{i=1}^{\infty} \frac{1}{i^2} - \sum_{i=1}^{\infty} \frac{1}{i^2} \right) = \frac{c - c}{Z} = 0$$

A symmetric distribution on pos and neg integers, the bad case

$$P(X = i) = \begin{cases} 0 & \text{if } i = 0 \\ \frac{1}{Zt^2} & \text{if } i \neq 0 \end{cases} ; \quad Z = 2 \sum_{i=1}^{\infty} \frac{1}{i^2} < \infty$$



$$E(X) = \sum_{i=1}^{\infty} iP(X = i) + \sum_{i=-1}^{-\infty} iP(X = i) = \frac{1}{Z} \left(\sum_{i=1}^{\infty} \frac{1}{i} - \sum_{i=1}^{-\infty} \frac{1}{i} \right) = \frac{\infty - \infty}{Z} = \text{undefined}$$

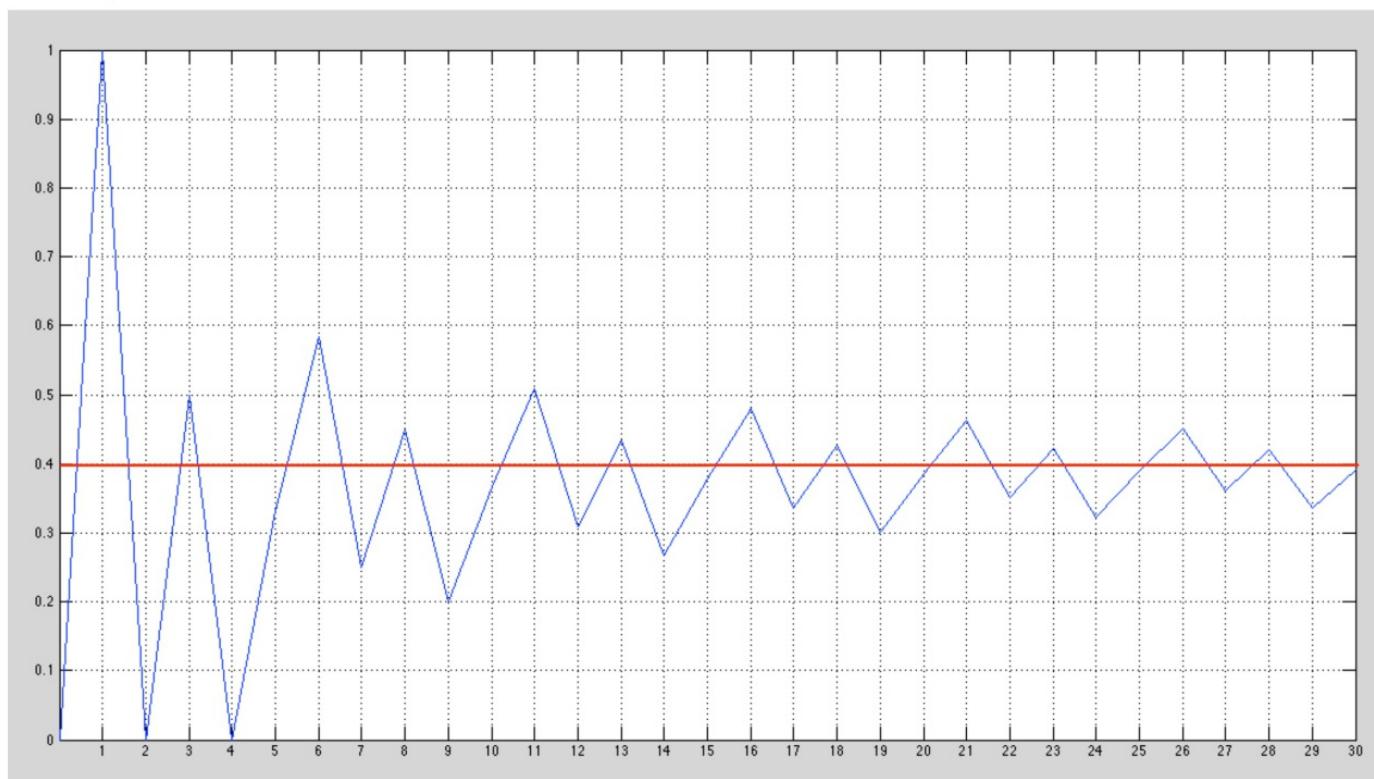
Undefined limit means you can get the limit of your choice by changing the order of summation.

You have at your disposal two infinitely large sums with shrinkingly small pieces:

$$1/1, 1/2, 1/3, 1/4, \dots \quad -1/1, -1/2, -1/3, -1/4, \dots$$

Suppose you want the limit to be 0.4, by alternating between positives and negatives you can get arbitrarily close to 0.4 (or to any other number)

$$\begin{aligned} & 1/1 - 1/1 + 1/2 - 1/2 + 1/3 + 1/4 - 1/3 + 1/5 - 1/4 + 1/6 + 1/7 - 1/5 + 1/8 - 1/6 + 1/9 + 1/10 - 1/7 \\ & + 1/11 - 1/8 + 1/12 + 1/13 - 1/9 + 1/14 - 1/10 + 1/15 + 1/16 - 1/11 + 1/17 - 1/12 + 1/18 = 0.3919 \end{aligned}$$



← → C webwork.cse.ucsd.edu/misc/expected.html
Recent changes - Se My Page CSE103 AWS Googles TWiki UCSD

Let X be a random variable whose probability distribution is defined as:

$$P(X = i) = \begin{cases} \frac{1}{|i|^\alpha} & \text{if } i \neq 0 \\ 0 & \text{if } i = 0. \end{cases}$$

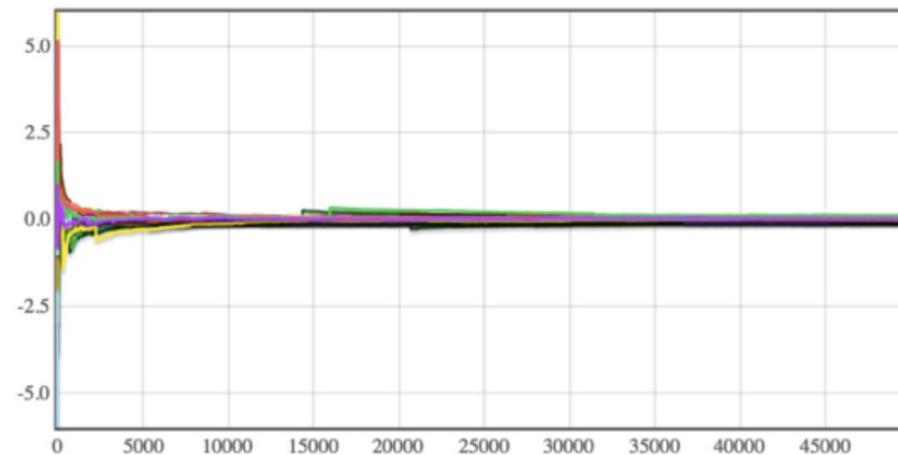
Simulation parameters:

α : 2.5

Number of trajectories: 50

Number of data points: 50000

Run



finite expectation
and Variance

Let X be a random variable whose probability distribution is defined as:

$$P(X = i) = \begin{cases} \frac{1}{|i|^\alpha} & \text{if } i \neq 0 \\ 0 & \text{if } i = 0. \end{cases}$$

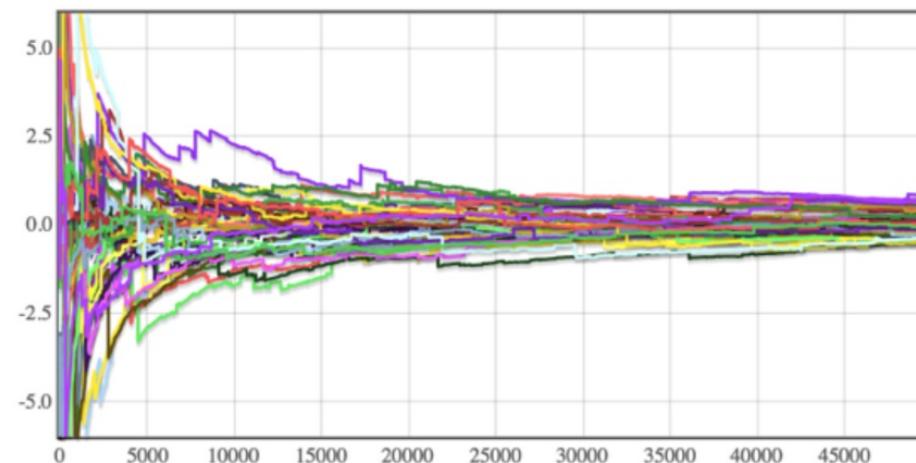
Simulation parameters:

α : 2.0

Number of trajectories: 50

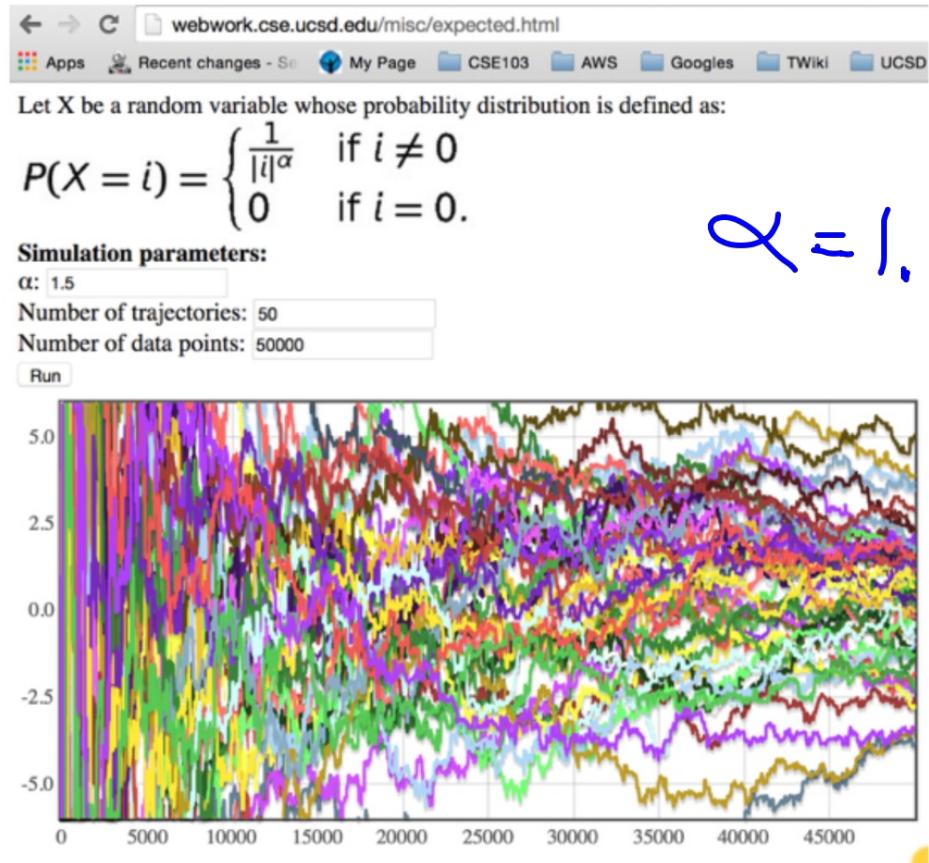
Number of data points: 50000

Run



$$\alpha = 2$$

Finite Expectation
Infinite Variance



$$\alpha = 1.5$$

Undefined
Expectation,

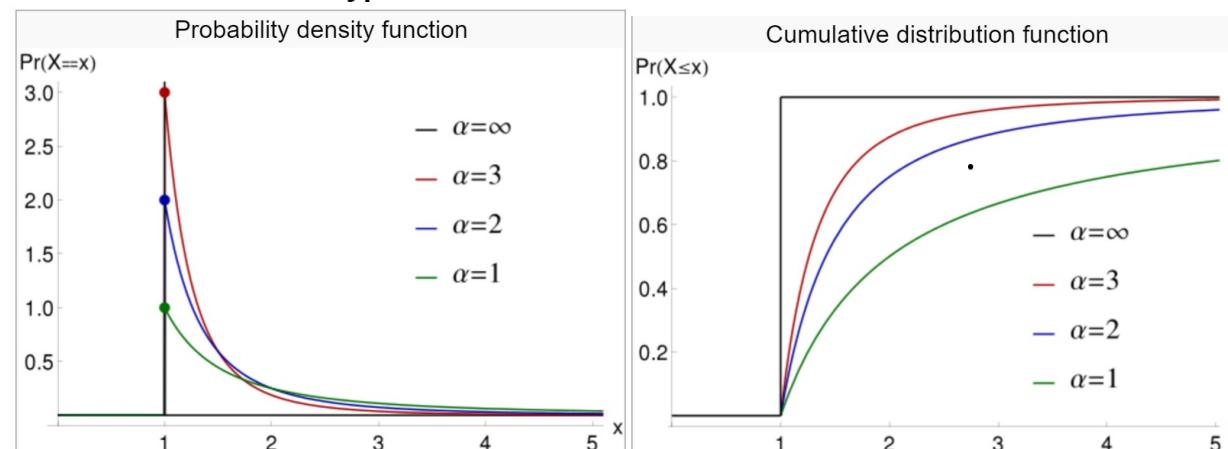
Pareto Distribution

Pareto: the contin.

Version of $\frac{1}{i^k}$

Parameters	$x_m > 0$ scale (real) $\alpha > 0$ shape (real)
Support	$x \in [x_m, +\infty)$
PDF	$\frac{\alpha x_m^\alpha}{x^{\alpha+1}}$ for $x \geq x_m$
CDF	$1 - \left(\frac{x_m}{x}\right)^\alpha$ for $x \geq x_m$
Mean	$\begin{cases} \infty & \text{for } \alpha \leq 1 \\ \frac{\alpha x_m}{\alpha-1} & \text{for } \alpha > 1 \end{cases}$
Variance	$\begin{cases} \infty & \text{for } \alpha \in (0, 2] \\ \frac{x_m^2 \alpha}{(\alpha-1)^2 (\alpha-2)} & \text{for } \alpha > 2 \end{cases}$

Pareto Type I



Moments

Raw

$$E(x)$$

$$E(x^2)$$

⋮

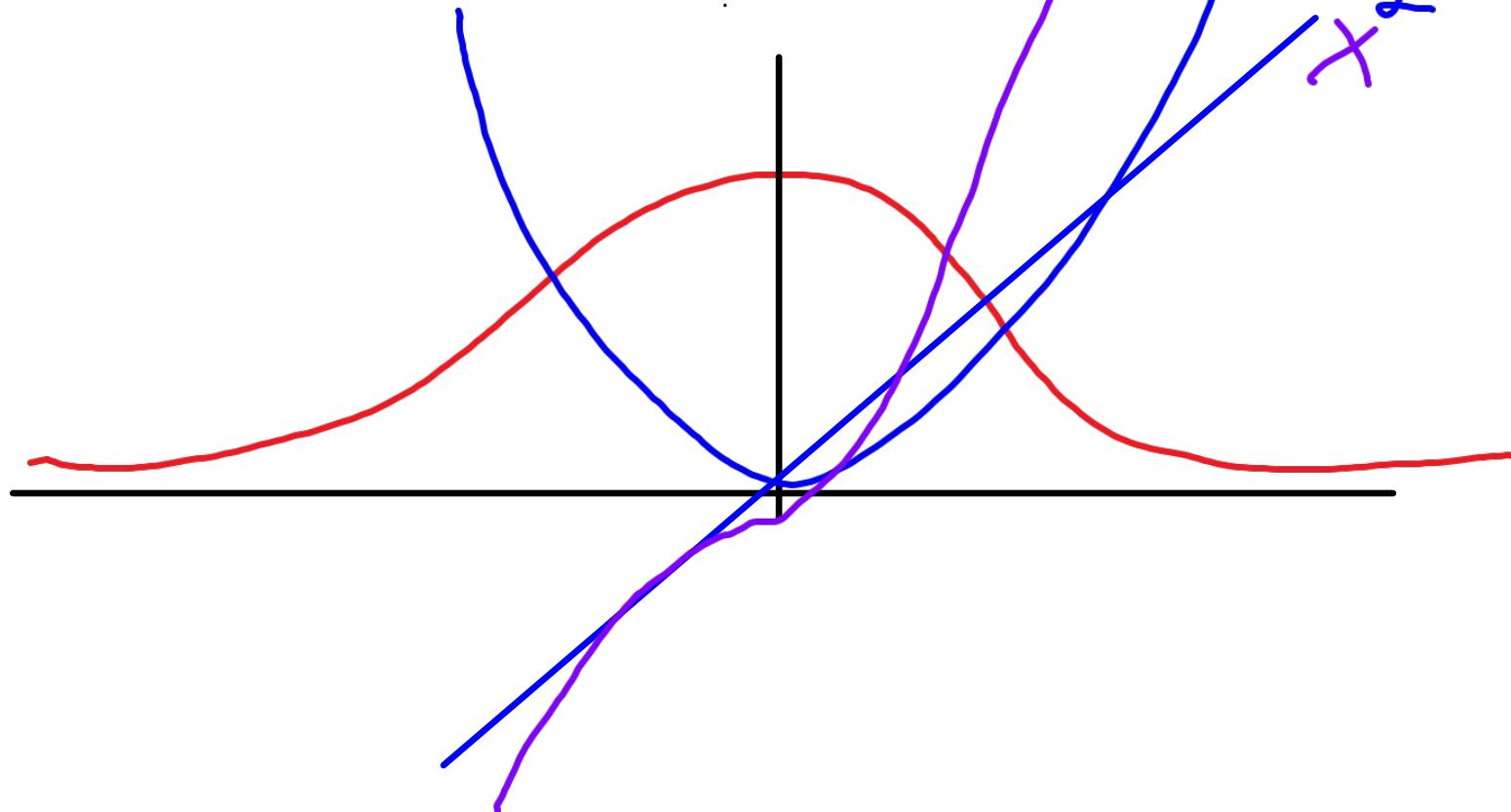
$$E(x^d)$$

centered

$$E(x-\mu) = 0$$

$$\text{Var}(x) = E((x-\mu)^2)$$

$$E((x-\mu)^d)$$



Light and Heavy tail distributions

Light tails

- Exponential, normal
- Exponential tails
- All moments are finite.
- $P(X > 2a|X > a) = P(X > a)$
- Exponential: The time until a program completes does not depend on how long it ran.

Heavy tails

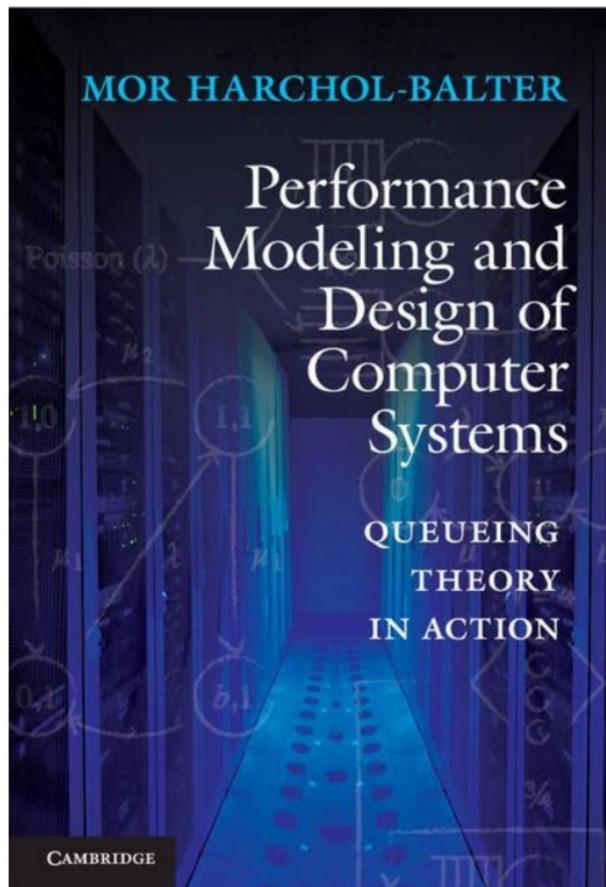
- Pareto
- Power law.
- Some moments are infinite.
- $P(X > 2a|X > a) = \text{constant}$
- Pareto: Decreasing failure rate: the longer a program has run, the longer before it finishes / crashes.
- “Elephant and mice”

Examples of heavy tail distributions

- In the world:
 - Wealth Distribution:
 - Richest 1% of the US population owns 35% of the wealth.
 - Poorest 60% of the US population owns 5% of the wealth.
 - Size of cities
 - Size Of earthquakes.
 - Frequency of words.
- In computer Science
 - Job run time.
 - Sizes of files in web sites.
 - Internet nodes out-degree
 - Number of packets in an IP flow.

Queuing theory in CS

II	Necessary Probability Background	39
3	Probability Review	43
3.1	Sample space and events	43
3.2	Probability defined on events	44
3.3	Conditional probabilities and events	46
3.4	Law of total probability	48
3.5	Discrete versus continuous random variables	50
3.6	Probabilities and densities	52
3.6.1	Discrete: probability mass function	52
3.6.2	Continuous: probability density function	55
3.7	Expectation & variance	59
3.8	Normal distribution	64
3.8.1	Linear transformation property	66
3.8.2	Central limit theorem	69
3.9	Joint probabilities and independence	72
3.10	Conditional probabilities and expectations	74
3.11	Probabilities and expectations via conditioning	78
3.12	Linearity of expectation	81
3.13	Sum of a random number of random variables	84
3.14	Exercises	86



Queues and light/heavy tails

- Suppose we have K servers.
- If the arrival time of jobs is Poisson and the job size is exponentially distributed (light tailed), then, using a separate queue for each server and random assignment gives good performance.
- If arrival times are Poisson, but distribution of job sizes is heavy tailed, then using a queue for each server and random assignment gives very poor performance (infinite expected waiting time)
 - Why? Small jobs are stuck behind large jobs for a long time, even if another server is free.

Alternatives to FCFS with random assignment

- Round-robin instead of random: a small improvement.
- Join-Shortest-Queue (JSQ): good improvement for light tails job size distribution – not good for heavy tails.
- If job-size known in advance:
 - Size-Interval-Task-Assignment (SITA): small jobs go to server1, larger to server2,... (Express Lane)
 - Least-Work-Left (LWL): Job goes to the server for which the remaining work before job starts to execute is the shortest. (Greedy selfish strategy)

Measuring the weight of the tail

- We consider “bounded Pareto” where there is a maximal job size (S)
- S is a positive random variable
- The ***variance*** of S is: $Var(S) = E(S^2) - E(S)^2$
- We are interested in the relation between the STD and the mean – A unit-less quantity.

The ***variation coefficient*** of S is:

$$C^2 \doteq \frac{Var(S)}{E(S)^2} = \frac{E(S^2) - E(S)^2}{E(S)^2} = \frac{E(S^2)}{E(S)^2} - 1$$

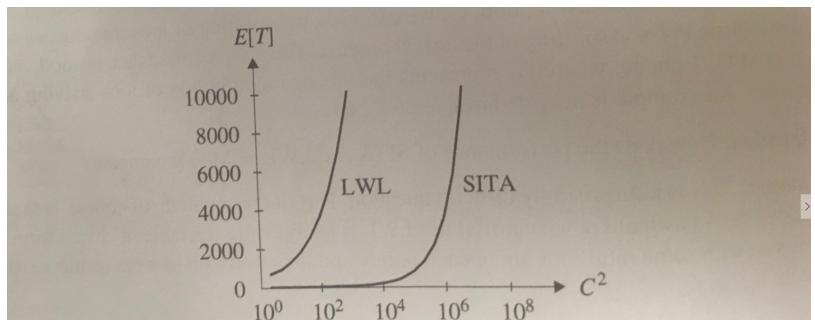


Figure 24.3. Expected response time, $E[T]$, for SITA and LWL versus C^2 in a 2-server system with Bounded Pareto job size distribution with $\alpha = 1.4$ and resource requirement $R = 0.95$.

$R = \text{load}$
 $\alpha = \text{Pareto parameter.}$

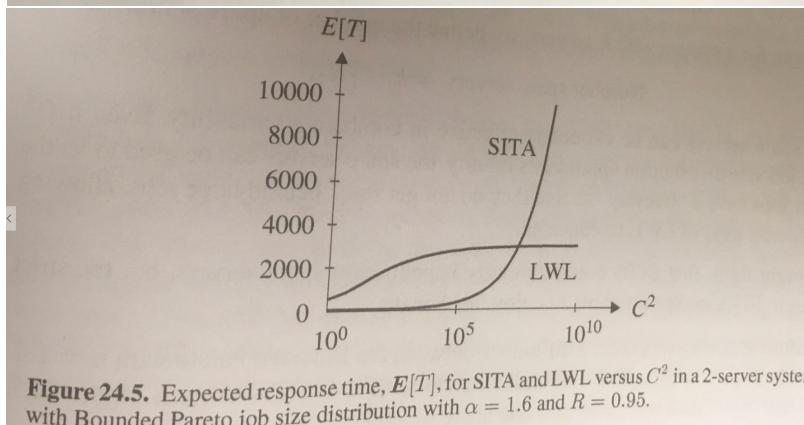


Figure 24.5. Expected response time, $E[T]$, for SITA and LWL versus C^2 in a 2-server system with Bounded Pareto job size distribution with $\alpha = 1.6$ and $R = 0.95$.