Infinite & undefined Expectation

Pareto distributions
in Computer Science
Expected value over countably infinite sets

\[ S = \text{a countably infinite subset of } R \]
\[ S = \{s_1, s_2, \ldots\} \]
\[ X = \text{a random variable which gets values in } S \]
\[ E(X) = \sum_{i=1}^{\infty} s_i \Pr(X(\omega) = s_i) \]
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

\[ P(X = i) = \frac{1}{zi^3} ; \quad z = \sum_{i=1}^{\infty} \frac{1}{i^3} < \infty \]  

Distribution is well defined

\[ E(X) = \sum_{i=1}^{\infty} \frac{i}{zi^3} = \sum_{i=1}^{\infty} \frac{1}{zi^2} < \infty \]  

Expectation is finite

\[ E[X^2] = \sum_{i=0}^{\infty} \frac{i^2}{zi^3} = \sum_{i=0}^{\infty} \frac{1}{zi} = \infty \]  

Variance is infinite

Consider next the distribution

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

Distribution is well defined but

\[ E(X) = \sum_{i=1}^{\infty} \frac{i}{zi^2} = \sum_{i=1}^{\infty} \frac{1}{zi} = \infty \]  

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

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

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

\[ \sum_{i=1}^{\infty} iP(X = i) = \infty \]
Consider a game with both wins and losses 
\( i \in \{0,-1,+1,-2,+2,\ldots\} \)

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

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

\[
\sum_{i=-\infty}^{\infty} iP(X = i) \text{ is undefined}
\]
Expectation over pos and neg integers: the good case

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

\[ E(X) = \sum_{i=1}^{\infty} iP(X = i) + \sum_{i=-1}^{\infty} iP(X = i) = \frac{1}{Z} \left( \sum_{i=1}^{\infty} \frac{1}{i^2} - \sum_{i=1}^{\infty} \frac{1}{i^2} \right) = \frac{c-c}{Z} = 0 \]
A symmetric distribution on pos and neg integers, the bad case

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

\[ E(X) = \sum_{i=-\infty}^{\infty} iP(X = i) = \frac{1}{Z} \left( \sum_{i=1}^{\infty} \frac{1}{i} - \sum_{i=1}^{\infty} \frac{1}{i} \right) = \frac{\infty - \infty}{Z} = \text{undefined} \]
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:
\[1/1, 1/2, 1/3, 1/4, \ldots\quad -1/1, -1/2, -1/3, -1/4, \ldots\]
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)

\[
\frac{1}{1} \frac{1}{1} - \frac{1}{2} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} - \frac{1}{3} + \frac{1}{5} - \frac{1}{4} + \frac{1}{6} + \frac{1}{7} - \frac{1}{5} + \frac{1}{8} - \frac{1}{6} + \frac{1}{9} + \frac{1}{10} - \frac{1}{7} \\
+ \frac{1}{11} - \frac{1}{8} + \frac{1}{12} + \frac{1}{13} - \frac{1}{9} + \frac{1}{14} - \frac{1}{10} + \frac{1}{15} + \frac{1}{16} - \frac{1}{11} + \frac{1}{17} - \frac{1}{12} + \frac{1}{18} = 0.3919
\]
Let $X$ be a random variable whose probability distribution is defined as:

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

$\alpha = 2.5$

Finite Expectation

Infinite 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:**

- $\alpha = 2.0$
- Number of trajectories: 50
- Number of data points: 50000

Undefined expectation
Undefined Variance
Let $X$ be a random variable whose probability distribution is defined as:

$$P(X = i) = \begin{cases} \frac{1}{|i|^2} & \text{if } i \neq 0 \\ 0 & \text{if } i = 0. \end{cases}$$

Simulation parameters:
- $\alpha = 1.5$
- Number of trajectories: 50
- Number of data points: 50000

$\lambda = 1.5$

Undefined Expectation, Undefined Variance
Pareto Distribution
Pareto: the contin. version of \( \frac{1}{k} \)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>( x_m &gt; 0 ) scale (real)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha &gt; 0 ) shape (real)</td>
<td></td>
</tr>
</tbody>
</table>

| 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

![Probability density function](image1)

![Cumulative distribution function](image2)
Moments

Raw

$E(x)$
$E(x^2)$

$E(x^d)$

Centered

$E(x - \mu) = 0$

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

$E((x - \mu)^2)$
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
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:

  \[ C^2 = \frac{Var(S)}{E(S)^2} = \frac{E(S^2) - E(S)^2}{E(S)^2} = \frac{E(S^2)}{E(S)^2} - 1 \]
$R = \text{load}$

$\alpha = \text{Pareto parameter}$