Standard and Mean Deviation

See Technical Incerto Chapter 4 by Nassim Taleb for context.

Jensen's Inequality

Jensen's inequality explains why **STD ≥ MAD**

Standard Deviation

Defined as:

$$\sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N}} \quad (1)$$

$$\sqrt{\frac{\sum d_i^2}{N}} \quad (2)$$

That is, the square root of the average squared deviation.

$$\sqrt{\mathrm{average}(\mathrm{d}_\mathrm{i}^2)} \quad (3)$$

Mean Deviation

Defined as:

$$\frac{\sum_{i=1}^{N}|x_i - \bar{x}|}{N} \quad (4)$$

$$\frac{\sum\sqrt{d_i^2}}{N} \quad (5)$$

That is, the average of the square root of the squared deviations.

$$\mathrm{average}\left(\sqrt{d_i^2}\right) \quad (6)$$

So, the difference between these functions is around where we apply the square root function.

The square root is a concave function:

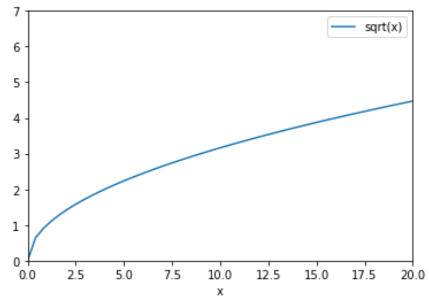

Figure 1: Square Root of x

STD ≥ MAD because of Jensen's inequality:

- The square root function is concave
- So, *Sqrt of Average* ≥ *Average of Sqrt*
- Standard Deviation ≥ Mean Deviation

Furthermore:

- **STD = MAD if and only if all deviations have the same magnitude**.
- This helps with intuition for why STD >> MAD when the fourth moment (kurtosis), is high.

Author: https://github.com/FergM/

Datapoint Weights

For simplicity of notation, in this section d_i refers to the absolute deviation of the i_{th} datapoint (from the mean).

Standard Deviation

Defined as:

$$\sqrt{\frac{\sum d_i^2}{N}} \tag{7}$$

Expands to:

$$\sqrt{\frac{d_1 d_1 + d_2 d_2 + d_3 d_3 + \cdots + d_N d_N}{N}} \tag{8}$$

Each datapoint is weighted by its own size:

$$\sqrt{\frac{w_1 d_1 + w_2 d_2 + w_3 d_3 + \cdots + w_N d_N}{N}} \tag{9}$$

Mean Deviation

Defined as:

$$\frac{\sum d_i}{N} \tag{10}$$

Expands to:

$$\frac{d_1 + d_2 + d_3 + \cdots + d_N}{N} \tag{11}$$

Each datapoint takes a relative weight of 1:

$$\frac{1 \cdot d_1 + 1 \cdot d_2 + 1 \cdot d_3 + \cdots + 1 \cdot d_N}{N} \tag{12}$$

Extreme Example

Many small deviations and one massive deviation.

- 10^6 Datapoints
- One outlier = 10^6
- All others = -1

Standard Deviation

$$\sqrt{\frac{1 + 1 + 1 + \cdots + (10^6 \cdot 10^6)}{N}} \tag{13}$$

$$\approx \sqrt{\frac{10^6 + (10^6 \cdot 10^6)}{10^6}} \tag{14}$$

$$\approx \sqrt{10^6} = 1000 \tag{15}$$

Mean Deviation

$$\frac{1 + 1 + 1 + \cdots + 10^6}{N} \tag{16}$$

$$\approx \frac{10^6 + 10^6}{10^6} \approx 2 \tag{17}$$

Author: https://github.com/FergM/

Mean Absolute Deviation
Chapter 4

Github: github.com/FergM/fattails

Twitter: @MFergal

LINK TO CODE

- https://github.com/FergM/fattails/blob/main/notebooks/NB29%20-%20Std%20vs%20MAD%20Efficiency.ipynb

Changing Deviation Granularity

- Daily to Annual

Changing Granularity

- Suppose we generate the following daily sequence
  - $annual\ return = d_1 + d_2 + d_3 + \cdots + d_{365}$
  - No covariance between days
  - Identical daily distribution

Changing Granularity

- Suppose we generate the following daily sequence
  - $annual\ return = \ d_1 + d_2 + d_3 + \cdots + d_{365}$
  - No covariance between days
  - Identical daily distribution
- Then
  - $Var(annual\ return)$
  - $= Var(d_1 + d_2 + d_3 + \cdots + d_{365})$
  - $= Var(d_1) + Var(d_2) + Var(d_3) + \cdots + Var(d_{365})$
  - $= N \cdot Var(d)$

Changing Granularity

Assumptions:

- No covariance between days
- Identical daily distribution

- Then
  - $Var(annual\ return)$
  - $= N \cdot Var(d)$
- So
  - $Std(annual\ return)$
  - $= \sqrt{365} \cdot Std(daily\ return)$

Efficiency Calculation

Normalise the standard deviation:

- $\frac{\sigma}{\mathbf{E}(\boldsymbol{\sigma})}$

Efficiency Calculation

Normalise the standard deviation:

- $\frac{\sigma}{\mathbf{E}(\boldsymbol{\sigma})}$

Calculate the Variance of this normalised dispersion:

- $Var\left(\frac{\sigma}{\mathbf{E}(\boldsymbol{\sigma})}\right) = \frac{Var(\sigma)}{\mathbf{E}(\sigma)^2}$