M-SCORE: AN INDICATOR
QUANTIFYING INDIVIDUAL'S
SCIENTIFIC RESEARCH OUTPUT

FEI SHU
MCGILL UNIVERSITY
WHY

- Triple Helix Model
- RISIS infrastructure
- Relational Quality Assessment & Management Model
- Decision Tree Analysis

Do policy makers like such complex concepts?
- Have these complex models or indicators been used?

Simple indicators is beautiful
- Impact factor
- H-index
PARETO DISTRIBUTION

- Bibliometric Distribution
  - Publication distribution
  - Journal coverage distribution
  - Word occurrence distribution
  - Citation distribution

- Double logarithm return:
  - Returns a linear line

Diagram:
- Graph showing the relationship between X and Y on a double logarithmic scale.
PARETO DISTRIBUTION

- \( \log(Y) = a \cdot \log(X) + N \)
  - \( A \) is negative
  - \( N \) is a fixed value
- \( \log(Y) = - \log(X) + N \)
- \( \log(Y) + \log(X) = N \)
- \( \log(X \cdot Y) = N \)
- \( X \cdot Y = N \)
PARETO DISTRIBUTION

- $X \cdot Y = N$
- $X_1 \cdot Y_1 = X_2 \cdot Y_2 = X_3 \cdot Y_3 = X_4 \cdot Y_4$
  
  $= X_5 \cdot Y_5 = \ldots$

- Citation distribution
  - 1 paper was cited 200 times
  - 10 papers were cited 20 times
  - 20 papers were cited 10 times
  - 200 papers were cited 1 time
CITATION DISTRIBUTION

- $X \times Y = 200$
  - 1 paper was cited 200 times
  - 10 papers were cited 20 times
  - 20 papers were cited 10 times
  - 200 papers were cited 1 time

- $X \times Y = 100$
  - 1 paper was cited 100 times
  - 10 papers were cited 10 times
  - 100 papers were cited 1 time
M-SCORE

- The maximum value of M where $M = P \times C$
- Area CBPO
- Triangle 2C-2P-O
  - Double the area CBPO
  - Reflect the major research impact
  - Exclude the outliers
    - Highly cited papers
    - Lowly cited or uncited papers
  - $P$ and $C$ should be in the range of $(0, \frac{1}{2} P_{\text{max}})$ and $(0, \frac{1}{2} C_{\text{max}})$
<table>
<thead>
<tr>
<th>Name</th>
<th>Number of Publications</th>
<th>Total Citations Received</th>
<th>h-index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spink, Amanda</td>
<td>199</td>
<td>6,358</td>
<td>40</td>
</tr>
<tr>
<td>Belkin, Nicholas</td>
<td>104</td>
<td>2,921</td>
<td>27</td>
</tr>
<tr>
<td>Saracevic, Tefko</td>
<td>78</td>
<td>3,886</td>
<td>23</td>
</tr>
<tr>
<td>Cronin, Blaise</td>
<td>168</td>
<td>2,047</td>
<td>23</td>
</tr>
<tr>
<td>Marchionini, Gary</td>
<td>122</td>
<td>2,288</td>
<td>23</td>
</tr>
<tr>
<td>Bates, Marcia.</td>
<td>45</td>
<td>2,218</td>
<td>19</td>
</tr>
<tr>
<td>Borgman, Christine</td>
<td>88</td>
<td>1,923</td>
<td>19</td>
</tr>
<tr>
<td>McCain, Katherine</td>
<td>64</td>
<td>1,311</td>
<td>17</td>
</tr>
<tr>
<td>White, Howard</td>
<td>32</td>
<td>2,463</td>
<td>16</td>
</tr>
<tr>
<td>Fidel, Raya</td>
<td>37</td>
<td>1,064</td>
<td>14</td>
</tr>
</tbody>
</table>

- Top 10 LIS scholars from (Cronin & Meho, 2006)
  - Scopus data
  - Three tie in h-index of 23
  - Two tie in h-index of 19
<table>
<thead>
<tr>
<th>Name</th>
<th>Number of Publications</th>
<th>Total Citations</th>
<th>h-index</th>
<th>M-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spink, Amanda</td>
<td>199</td>
<td>6,358</td>
<td>40</td>
<td>1,682</td>
</tr>
<tr>
<td>Belkin, Nicholas</td>
<td>104</td>
<td>2,921</td>
<td>27</td>
<td>1,235</td>
</tr>
<tr>
<td>Saracevic, Tefko</td>
<td>78</td>
<td>3,886</td>
<td>23</td>
<td>1,510</td>
</tr>
<tr>
<td>Cronin, Blaise</td>
<td>168</td>
<td>2,047</td>
<td>23</td>
<td>672</td>
</tr>
<tr>
<td>Marchionini, Gary</td>
<td>122</td>
<td>2,288</td>
<td>23</td>
<td>600</td>
</tr>
<tr>
<td>Bates, Marcia.</td>
<td>45</td>
<td>2,218</td>
<td>19</td>
<td>880</td>
</tr>
<tr>
<td>Borgman, Christine</td>
<td>88</td>
<td>1,923</td>
<td>19</td>
<td>678</td>
</tr>
<tr>
<td>McCain, Katherine</td>
<td>64</td>
<td>1,311</td>
<td>17</td>
<td>330</td>
</tr>
<tr>
<td>White, Howard</td>
<td>32</td>
<td>2,463</td>
<td>16</td>
<td>497</td>
</tr>
<tr>
<td>Fidel, Raya</td>
<td>37</td>
<td>1,064</td>
<td>14</td>
<td>506</td>
</tr>
</tbody>
</table>

- Differentiate research impact
- Favours scholars with a large set of frequently cited papers
  - McCain: 10 papers were cited at least 33 times
  - Fidel: 11 papers were cited at least 46 times
CONCLUSION

- M-score is Based on the property of the pareto distribution
  - Could be used to measure the individual’s

- Fix the inconsistency of h-index addressed by Waltman, L., & van Eck, N. J. (2012)

- Need to be tested more

- Time value could be added

- Could be normalized