Mean Average Path Length

- "Small World Effect": i\(\leftrightarrow\)j \(\rightarrow\) i\(\leftrightarrow\)e\(\rightarrow\)j

undirected graph:

\[ \ell = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i<j} d_{ij} \]

formula also counts 0 distances from 1 to i: \( \frac{1}{2} n(n+1) = \frac{1}{2} n(n-1) + n \)

- Expression allowing for disconnected components (where \(d_{ij}\) can occur): harmonic mean:

\[ \ell^{-1} = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i<j} \frac{1}{d_{ij}} \]
Transitivity / Clustering Coefficient

- **Clustering coefficient (whole graph):**
  \[ C = \frac{3 \times \text{number of triangles in the network}}{\text{number of connected triples of vertices}} = \frac{6 \times \text{number of triangles in the network}}{\text{number of paths of length two}} \]

- **Clustering coefficient (Watts-Strogatz-version, for node i):**
  \[ C_i = \frac{\text{number of triangles connected to vertex } i}{\text{number of triples centered on vertex } i} = \frac{|\{e_{ikl}\} | v_k,v_j \in N_i|}{k_i(k_i-1)} \]
  (see Introduction, \( k_i = \text{degree of node } i \))

- **Clustering coefficient (Watts-Strogatz-version, for whole graph):**
  \[ C = \frac{1}{n} \sum_i C_i \]
  mean of ratio instead of ratio of means
**Transitivity / Clustering Coefficient**

- **Clustering coefficient (whole graph):**
  \[
  C = C^{(I)} = \frac{3 \times \text{number of triangles in the network}}{\text{number of connected triples of vertices}}
  = \frac{6 \times \text{number of triangles in the network}}{\text{number of paths of length two}}
  \]

- **Clustering coefficient (Watts-Strogatz-version, for node i):**
  \[
  C_i = \frac{\text{number of triangles connected to vertex } i}{\text{number of triples centered on vertex } i}
  = \frac{|\{e_{ijkl} \mid v_k, v_j \in N_i\}|}{k_i(k_i-1)}
  \]
  (see Introduction, \(k_i = \text{degree of node } i\))

  Clustering coefficient (Watts-Strogatz-version, for whole graph):
  \[
  C = C^{(2)} = \frac{1}{n} \sum_i C_i
  \]
  mean of ratio instead of ratio of means

**Example:**

\[
C^{(I)} = \frac{3 \times 1}{8} = 0.375
\]

\[
C^{(2)} = \frac{1}{n} \sum_i C_i
\]

\[
C^{(2)} = \frac{1}{5} (1 + 1 + 1/6 + 0 + 0) = 13/30 = 0.433333
\]
Transitivity / Clustering Coefficient

Example:

\[
C_l = \frac{3 \times \text{number of triangles in the network}}{8} = 0.375
\]

\[
C^2 = \frac{1}{n} \sum_i C_i \quad \text{with} \quad C_i = \frac{\text{number of triangles connected to vertex } i}{\text{number of triples centered on vertex } i}
\]

\[
C^2(1/5) (1 + 1 + 1/6 + 0 + 0) = 13/30 = 0.433333
\]

Transitivity / Clustering Coefficient

Example:

\[
C_l = \frac{3 \times \text{number of triangles in the network}}{8} = 0.375
\]

\[
C^2 = \frac{1}{n} \sum_i C_i \quad \text{with} \quad C_i = \frac{\text{number of triangles connected to vertex } i}{\text{number of triples centered on vertex } i}
\]

\[
C^2(1/5) (1 + 1 + 1/6 + 0 + 0) = 13/30 = 0.433333
\]

<table>
<thead>
<tr>
<th>network</th>
<th>type</th>
<th>n</th>
<th>m</th>
<th>(t)</th>
<th>(k)</th>
<th>(C^1)</th>
<th>(C^2)</th>
<th>(r)</th>
<th>Ref(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>film actors</td>
<td>undirected</td>
<td>449,933</td>
<td>25,548,452</td>
<td>115,434</td>
<td>3,487</td>
<td>2.5</td>
<td>0.20</td>
<td>0.78</td>
<td>0.238</td>
</tr>
<tr>
<td>company directors</td>
<td>undirected</td>
<td>7,473</td>
<td>55,392</td>
<td>14,414</td>
<td>4,601</td>
<td>0.59</td>
<td>0.88</td>
<td>0.74</td>
<td>106, 235</td>
</tr>
<tr>
<td>math course</td>
<td>undirected</td>
<td>2,533</td>
<td>49,486</td>
<td>7,937</td>
<td>3,375</td>
<td>0.15</td>
<td>0.34</td>
<td>0.12</td>
<td>105, 152</td>
</tr>
<tr>
<td>physics course</td>
<td>undirected</td>
<td>52,899</td>
<td>240,330</td>
<td>9,270</td>
<td>0.16</td>
<td>0.45</td>
<td>0.56</td>
<td>0.36</td>
<td>311, 315</td>
</tr>
<tr>
<td>biology course</td>
<td>undirected</td>
<td>13,833</td>
<td>4,800,000,000</td>
<td>15,353</td>
<td>4,922</td>
<td>0.088</td>
<td>0.60</td>
<td>0.12</td>
<td>311, 315</td>
</tr>
<tr>
<td>telephone call</td>
<td>undirected</td>
<td>47,000,000,000</td>
<td>8,900,000,000</td>
<td>3,160</td>
<td>0.25</td>
<td>0.13</td>
<td>0.30</td>
<td>0.16</td>
<td>13</td>
</tr>
<tr>
<td>email messages</td>
<td>directed</td>
<td>59,410</td>
<td>8,300</td>
<td>14,141</td>
<td>0.45</td>
<td>0.31</td>
<td>0.16</td>
<td>0.26</td>
<td>311, 315</td>
</tr>
<tr>
<td>email addresses</td>
<td>directed</td>
<td>14,861</td>
<td>57,529</td>
<td>3,832</td>
<td>0.25</td>
<td>0.17</td>
<td>0.11</td>
<td>0.06</td>
<td>321</td>
</tr>
<tr>
<td>student relationships</td>
<td>directed</td>
<td>573</td>
<td>477</td>
<td>14,616</td>
<td>10,016</td>
<td>0.005</td>
<td>0.006</td>
<td>0.00</td>
<td>0.008</td>
</tr>
<tr>
<td>social contacts</td>
<td>undirected</td>
<td>2,810</td>
<td>32</td>
<td>0.44</td>
<td>0.04</td>
<td>0.55</td>
<td>0.60</td>
<td>0.16</td>
<td>311, 315</td>
</tr>
<tr>
<td>WWW ref web</td>
<td>directed</td>
<td>2,693,804</td>
<td>1,497,135</td>
<td>5,550</td>
<td>11,274</td>
<td>0.11</td>
<td>0.29</td>
<td>0.00</td>
<td>0.067</td>
</tr>
<tr>
<td>WWW authors</td>
<td>directed</td>
<td>2,035,306,000,000</td>
<td>2,133,000,000</td>
<td>10,460</td>
<td>1,073</td>
<td>0.39</td>
<td>0.27</td>
<td>0.07</td>
<td>13</td>
</tr>
<tr>
<td>citation network</td>
<td>directed</td>
<td>76,329</td>
<td>1,619,946</td>
<td>8,57</td>
<td>0.47</td>
<td>0.13</td>
<td>0.35</td>
<td>0.15</td>
<td>314</td>
</tr>
<tr>
<td>Roger’s Theorums</td>
<td>directed</td>
<td>1,002</td>
<td>9,510</td>
<td>4,949</td>
<td>0.48</td>
<td>0.27</td>
<td>0.04</td>
<td>0.00</td>
<td>2,014</td>
</tr>
<tr>
<td>.word co-occurrences</td>
<td>undirected</td>
<td>469,902</td>
<td>17,000,000,000</td>
<td>70,012</td>
<td>27</td>
<td>0.04</td>
<td>0.13</td>
<td>0.20</td>
<td>119, 157</td>
</tr>
<tr>
<td>information</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>internet</td>
<td>undirected</td>
<td>10,697</td>
<td>31,992</td>
<td>5,984</td>
<td>3,313</td>
<td>2.5</td>
<td>0.09</td>
<td>0.39</td>
<td>0.189</td>
</tr>
<tr>
<td>power grid</td>
<td>undirected</td>
<td>1,461</td>
<td>6,594</td>
<td>2,677</td>
<td>10,930</td>
<td>0.19</td>
<td>0.08</td>
<td>0.00</td>
<td>0.003</td>
</tr>
<tr>
<td>travel routes</td>
<td>undirected</td>
<td>527</td>
<td>10,403</td>
<td>66,787</td>
<td>2,86</td>
<td>0.69</td>
<td>0.002</td>
<td>0.003</td>
<td>368</td>
</tr>
<tr>
<td>software packages</td>
<td>directed</td>
<td>1,439</td>
<td>1,723</td>
<td>1,293</td>
<td>2,43</td>
<td>0.079</td>
<td>0.034</td>
<td>0.002</td>
<td>318</td>
</tr>
<tr>
<td>software classes</td>
<td>directed</td>
<td>1,377</td>
<td>2,263</td>
<td>1,813</td>
<td>1,53</td>
<td>0.032</td>
<td>0.012</td>
<td>0.019</td>
<td>144</td>
</tr>
<tr>
<td>electronic circuits</td>
<td>undirected</td>
<td>240,067</td>
<td>53,248</td>
<td>43,145</td>
<td>11.15</td>
<td>0.018</td>
<td>0.000</td>
<td>0.015</td>
<td>155</td>
</tr>
<tr>
<td>peer-to-peer network</td>
<td>directed</td>
<td>889</td>
<td>1,266</td>
<td>1,472</td>
<td>2,28</td>
<td>0.012</td>
<td>0.011</td>
<td>0.008</td>
<td>6, 354</td>
</tr>
<tr>
<td>technical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>metropolitan network</td>
<td>directed</td>
<td>765</td>
<td>4,386</td>
<td>8,964</td>
<td>2,569</td>
<td>2.2</td>
<td>0.099</td>
<td>0.067</td>
<td>0.234</td>
</tr>
<tr>
<td>protein interactions</td>
<td>directed</td>
<td>2,115</td>
<td>2,209</td>
<td>2,12</td>
<td>1,487</td>
<td>2.4</td>
<td>0.072</td>
<td>0.071</td>
<td>0.154</td>
</tr>
<tr>
<td>marine food web</td>
<td>directed</td>
<td>135</td>
<td>589</td>
<td>4,43</td>
<td>2,05</td>
<td>0.16</td>
<td>0.23</td>
<td>0.26</td>
<td>204</td>
</tr>
<tr>
<td>food service web</td>
<td>directed</td>
<td>92</td>
<td>497</td>
<td>1,584</td>
<td>1,50</td>
<td>0.20</td>
<td>0.087</td>
<td>0.028</td>
<td>273</td>
</tr>
<tr>
<td>social network</td>
<td>directed</td>
<td>307</td>
<td>2,159</td>
<td>7,98</td>
<td>3,09</td>
<td>0.18</td>
<td>0.23</td>
<td>0.28</td>
<td>418</td>
</tr>
</tbody>
</table>

Table II: Basic statistics for a number of published networks. The properties measured are type of network, directed or undirected; total number of vertices, total number of edges, mean degree; \(z\) mean vertex shortest distance; \(\ell\) mean of degree distribution if the distribution follows a power law (or 0 if not; in most cases are given for directed graphs); clustering coefficient \(C^1\) from Eq. (3); clustering coefficient \(C^2\) from Eq. (6); and degree correlation coefficient \(r\). Last column gives the citation(s) for the network in the bibliography. Blank entries indicate unavailable data.
### ILE II Basic statistics for a number of published networks. The properties measured are type of graph, directed or undirected; total number of vertices and edges; percent vertex and degree distribution; mean degree distribution; standard deviation of degree distribution; number of components; and number of largest components. The results are given for directed graphs only (i.e., columns 3 and 4 are not used). The vertex and degree distributions follow a power law (p = 1.0). The number of vertices is shown in parentheses. The columns are: column 1: graph type; column 2: number of vertices; column 3: number of edges; column 4: percent vertex distribution; column 5: percent degree distribution; column 6: vertex distribution; column 7: degree distribution; column 8: number of components; column 9: number of largest components.

<table>
<thead>
<tr>
<th>Network Type</th>
<th>N</th>
<th>M</th>
<th>p</th>
<th>q</th>
<th>r</th>
<th>c1</th>
<th>c2</th>
<th>c3</th>
<th>R(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directorship</td>
<td>100</td>
<td>500</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>2.0</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Company</td>
<td>1000</td>
<td>5000</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>5.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>University</td>
<td>10000</td>
<td>50000</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>10.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

### ILE II Basic statistics for a number of published networks. The properties measured are type of graph, directed or undirected; total number of vertices and edges; percent vertex and degree distribution; mean degree distribution; standard deviation of degree distribution; number of components; and number of largest components. The results are given for directed graphs only (i.e., columns 3 and 4 are not used). The vertex and degree distributions follow a power law (p = 1.0). The number of vertices is shown in parentheses. The columns are: column 1: graph type; column 2: number of vertices; column 3: number of edges; column 4: percent vertex distribution; column 5: percent degree distribution; column 6: vertex distribution; column 7: degree distribution; column 8: number of components; column 9: number of largest components.

<table>
<thead>
<tr>
<th>Network Type</th>
<th>N</th>
<th>M</th>
<th>p</th>
<th>q</th>
<th>r</th>
<th>c1</th>
<th>c2</th>
<th>c3</th>
<th>R(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directorship</td>
<td>100</td>
<td>500</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>2.0</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Company</td>
<td>1000</td>
<td>5000</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>5.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>University</td>
<td>10000</td>
<td>50000</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>10.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

### ILE II Basic statistics for a number of published networks. The properties measured are type of graph, directed or undirected; total number of vertices and edges; percent vertex and degree distribution; mean degree distribution; standard deviation of degree distribution; number of components; and number of largest components. The results are given for directed graphs only (i.e., columns 3 and 4 are not used). The vertex and degree distributions follow a power law (p = 1.0). The number of vertices is shown in parentheses. The columns are: column 1: graph type; column 2: number of vertices; column 3: number of edges; column 4: percent vertex distribution; column 5: percent degree distribution; column 6: vertex distribution; column 7: degree distribution; column 8: number of components; column 9: number of largest components.

<table>
<thead>
<tr>
<th>Network Type</th>
<th>N</th>
<th>M</th>
<th>p</th>
<th>q</th>
<th>r</th>
<th>c1</th>
<th>c2</th>
<th>c3</th>
<th>R(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directorship</td>
<td>100</td>
<td>500</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>2.0</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Company</td>
<td>1000</td>
<td>5000</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>5.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>University</td>
<td>10000</td>
<td>50000</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>10.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

### ILE II Basic statistics for a number of published networks. The properties measured are type of graph, directed or undirected; total number of vertices and edges; percent vertex and degree distribution; mean degree distribution; standard deviation of degree distribution; number of components; and number of largest components. The results are given for directed graphs only (i.e., columns 3 and 4 are not used). The vertex and degree distributions follow a power law (p = 1.0). The number of vertices is shown in parentheses. The columns are: column 1: graph type; column 2: number of vertices; column 3: number of edges; column 4: percent vertex distribution; column 5: percent degree distribution; column 6: vertex distribution; column 7: degree distribution; column 8: number of components; column 9: number of largest components.

<table>
<thead>
<tr>
<th>Network Type</th>
<th>N</th>
<th>M</th>
<th>p</th>
<th>q</th>
<th>r</th>
<th>c1</th>
<th>c2</th>
<th>c3</th>
<th>R(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directorship</td>
<td>100</td>
<td>500</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>2.0</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Company</td>
<td>1000</td>
<td>5000</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>5.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>University</td>
<td>10000</td>
<td>50000</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>10.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

### ILE II Basic statistics for a number of published networks. The properties measured are type of graph, directed or undirected; total number of vertices and edges; percent vertex and degree distribution; mean degree distribution; standard deviation of degree distribution; number of components; and number of largest components. The results are given for directed graphs only (i.e., columns 3 and 4 are not used). The vertex and degree distributions follow a power law (p = 1.0). The number of vertices is shown in parentheses. The columns are: column 1: graph type; column 2: number of vertices; column 3: number of edges; column 4: percent vertex distribution; column 5: percent degree distribution; column 6: vertex distribution; column 7: degree distribution; column 8: number of components; column 9: number of largest components.

<table>
<thead>
<tr>
<th>Network Type</th>
<th>N</th>
<th>M</th>
<th>p</th>
<th>q</th>
<th>r</th>
<th>c1</th>
<th>c2</th>
<th>c3</th>
<th>R(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directorship</td>
<td>100</td>
<td>500</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>2.0</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Company</td>
<td>1000</td>
<td>5000</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>5.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>University</td>
<td>10000</td>
<td>50000</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>10.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

### ILE II Basic statistics for a number of published networks. The properties measured are type of graph, directed or undirected; total number of vertices and edges; percent vertex and degree distribution; mean degree distribution; standard deviation of degree distribution; number of components; and number of largest components. The results are given for directed graphs only (i.e., columns 3 and 4 are not used). The vertex and degree distributions follow a power law (p = 1.0). The number of vertices is shown in parentheses. The columns are: column 1: graph type; column 2: number of vertices; column 3: number of edges; column 4: percent vertex distribution; column 5: percent degree distribution; column 6: vertex distribution; column 7: degree distribution; column 8: number of components; column 9: number of largest components.

<table>
<thead>
<tr>
<th>Network Type</th>
<th>N</th>
<th>M</th>
<th>p</th>
<th>q</th>
<th>r</th>
<th>c1</th>
<th>c2</th>
<th>c3</th>
<th>R(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directorship</td>
<td>100</td>
<td>500</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
<td>2.0</td>
<td>0.2</td>
<td>0.1</td>
<td>0.0</td>
</tr>
<tr>
<td>Company</td>
<td>1000</td>
<td>5000</td>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
<td>5.0</td>
<td>0.1</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>University</td>
<td>10000</td>
<td>50000</td>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
<td>10.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>
### Degree Distribution

**Notation:**

\[ p(k) = p_k = \text{fraction of nodes having degree } k \]

**Cumulative distribution:**

\[ P_k = \sum_{k'=k}^{\infty} p_{k'} \]

**Power law:**

\[ p_k \sim k^{-\alpha} \]

\[ P_k \sim \sum_{k'=k}^{\infty} k'^{-\alpha} \sim k^{-(\alpha - 1)} \]

**Exponential:**

\[ p_k \sim e^{-k/\kappa} \]

\[ P_k = \sum_{k'=k}^{\infty} p_{k'} \sim \sum_{k'=k}^{\infty} e^{-k'/\kappa} \sim e^{-k/\kappa} \]
Degree Distribution

- Notation: $p(k) = p_k = \text{fraction of nodes having degree } k$

- Cumulative distribution:
  \[ P_k = \sum_{k' = k}^{\infty} p_{k'} \]

- Power law:
  \[ p_k \sim k^{-\alpha} \]
  \[ \Rightarrow P_k \sim \sum_{k' = k}^{\infty} k'^{-\alpha} \sim k^{-(\alpha - 1)} \]

- Exponential:
  \[ p_k \sim e^{-k/\kappa} \]
  \[ \Rightarrow P_k = \sum_{k' = k}^{\infty} p_{k'} \sim \sum_{k' = k}^{\infty} e^{-k'/\kappa} \sim e^{-k/\kappa} \]
"Power law" == "Scale free":

- $f(x) = x^\alpha$ is only solution to functional equation formalizing scale freedom $f(ax) = b f(x)$

- In other words: change of scale $\rightarrow f$ still "looks the same"

Other point of view:

Although we can compute the expectation $E(k) = \sum_k k \cdot k^{-\alpha}$ if $\alpha > 1$, the variance (error bars) $\text{Var}(k) = \sum_k (k - E(k))^2 \cdot k^{-\alpha}$ diverges $\rightarrow$ we "cannot be sure about $k$"

- "no characteristic scale" $\rightarrow$ "scale free"

Examples:

- Power law: citation NW, WWW, Internet, metabolic NW, telephone call NW, human sexual contact NW etc.

- Exponential: power grid, railway NW

- Power law with exp. cut-offs: Movie co-actor NW
Degree Distribution

Examples:
- Power law: citation NW, WWW, Internet, metabolic NW, telephone call NW, human sexual contact NW etc.
- Exponential: power grid, railway NW
- Power law with exp. cut-offs: Movie co-actor NW

Maximum Degree

- "less or equal than one vertex with $k_{\text{max}}$"
  $\Rightarrow n p_{k_{\text{max}}} = 1 \Rightarrow$ for power law $p_k = k^\alpha$: $k_{\text{max}} \sim n^{1/\alpha}$ but: not very accurate estimation

- Other estimation:
  - prob $p$ of "exactly $m$ nodes with $k$ and rest of nodes smaller than $k$":
    \[ \binom{n}{m} p_k^m (1 - P_k)^{n-m} \]
  - $\Rightarrow$ prob of $k$ being the highest degree in graph:
    \[ h_k = \sum_{m=1}^{n} \binom{n}{m} p_k^m (1 - P_k)^{n-m} \]
    \[ = (p_k + 1 - P_k)^n - (1 - P_k)^n \]
  - $\Rightarrow$ expected highest degree:
    \[ k_{\text{max}} = \sum_k k h_k \]
Maximum Degree

"less or equal than one vertex with $k_{\text{max}}$\n$np_{k_{\text{max}}} = 1 \Rightarrow$ for power law $p_k = k^{-\alpha}$; \( k_{\text{max}} \sim n^{1/\alpha} \)
but: not very accurate estimation

Other estimation:
$\cdot$ prob $p$ of "exactly $m$ nodes with $k$ and rest of nodes smaller than $k$":
$$\binom{n}{m} p_k^m (1 - P_k)^{n - m}$$
$\Rightarrow$ prob of $k$ being the highest degree in graph:
$$h_k = \sum_{m=1}^{n} \binom{n}{m} p_k^m (1 - P_k)^{n - m}$$
$$= (p_k + 1 - P_k)^n - (1 - P_k)^n$$
$\Rightarrow$ expected highest degree:
$$k_{\text{max}} = \sum_k k h_k$$
Maximum Degree

**less or equal than one vertex with $k_{\text{max}}$**
\[ np_{k_{\text{max}}} = 1 \rightarrow \text{for power law } p_k = k^{-\alpha}; \ k_{\text{max}} \sim n^{1/\alpha} \]
but: not very accurate estimation

**Other estimation:**
- prob $p$ of „exactly $m$ nodes with $k$ and rest of nodes smaller than $k$“:
  \[ \binom{n}{m} p_k^m (1 - P_k)^{n-m} \]
- prob of $k$ being the highest degree in graph:
  \[ h_k = \sum_{m=1}^{n} \binom{n}{m} p_k^m (1 - P_k)^{n-m} \]
  \[ = (p_k + 1 - P_k)^n - (1 - P_k)^n \]
- expected highest degree:
  \[ k_{\text{max}} = \sum_k k h_k \]
Maximum Degree

- since $h_k$ is small for small $k$ and also for large $k$ take as $k_{\text{max}}$ the modal value of $h_k$ 

  modal value: \[
  \frac{d}{dk} h_k = 0
  \]

  Using $dP_k/dk = p_k$ we get

  \[
  \frac{d}{dk} h_k = n \left[ \left( \frac{dP_k}{dk} - p_k \right) (p_k + 1 - P_k)^{n-1} + p_k (1 - P_k)^{n-1} \right] = 0
  \]

  or $k_{\text{max}}$ is a solution of

  \[
  \frac{dp_k}{dk} \approx -np_k^2
  \]

  (assuming: $p_k$ is small for $k > k_{\text{max}}$ and that $np_k \ll 1$ and that $P_k \ll 1$)

  \[\text{we get for power law } p_k \sim k^{-\alpha} \text{ that } k_{\text{max}} \sim n^{1/(\alpha-1)}\]

Maximum Degree

- "less or equal than one vertex with $k_{\text{max}}$"

  \[np_k = 1 \rightarrow \text{for power law } p_k = k^{\alpha} \sim n^{\alpha} \text{ but not very accurate estimation}\]

  \[\text{Other estimation:}\]

  - prob $p$ of "exactly $m$ nodes with $k$ and rest of nodes smaller than $k$": 

    \[
    \binom{n}{m} p_k^m (1 - P_k)^{n-m}
    \]

  - prob of $k$ being the highest degree in graph:

    \[
    h_k = \sum_{m=1}^{n} \binom{n}{m} p_k^m (1 - P_k)^{n-m} = (p_k + 1 - P_k)^n - (1 - P_k)^n
    \]

  - \text{expected highest degree:}

    \[k_{\text{max}} = \sum_k kh_k\]

Maximum Degree

- since $h_k$ is small for small $k$ and also for large $k$ take as $k_{\text{max}}$ the modal value of $h_k$ 

  modal value: \[
  \frac{d}{dk} h_k = 0
  \]

  Using $dP_k/dk = p_k$ we get

  \[
  \frac{d}{dk} h_k = n \left[ \left( \frac{dP_k}{dk} - p_k \right) (p_k + 1 - P_k)^{n-1} + p_k (1 - P_k)^{n-1} \right] = 0
  \]

  or $k_{\text{max}}$ is a solution of

  \[
  \frac{dp_k}{dk} \approx -np_k^2
  \]

  (assuming: $p_k$ is small for $k > k_{\text{max}}$ and that $np_k \ll 1$ and that $P_k \ll 1$)

  \[\text{we get for power law } p_k \sim k^{-\alpha} \text{ that } k_{\text{max}} \sim n^{1/(\alpha-1)}\]
since $h_k$ is small for small $k$ and also for large $k$ \( \Rightarrow \) take as $k_{\text{max}}$ the modal value of $h_k$ \( \Rightarrow \)

**modal value:** \( \frac{d}{dk} h_k = 0 \)

Using \( \frac{dP_k}{dk} = p_k \) we get

\[
\frac{d}{dk} h_k = n \left[ \left( \frac{dp_k}{dk} - p_k \right) (p_k + 1 - P_k)^{n-1} + p_k (1 - P_k)^{n-1} \right] = 0
\]

or $k_{\text{max}}$ is a solution of

\[
\frac{dp_k}{dk} \approx -np_k^2
\]

(assuming: $p_k$ is small for $k > k_{\text{max}}$ and that $np_k \ll 1$ and that $P_k \ll 1$)

→ we get for power law $p_k \sim k^{-\alpha}$ that $k_{\text{max}} \sim n^{1/(\alpha-1)}$
Network Resilience

- What happens if nodes are removed? (interesting e.g. for vaccination effects in disease spreading in human contact networks)

- For power law networks:
  - remove random nodes: no effect on mean distances
  - remove high degree nodes: drastic effect

- Interpretations:
  - Internet is easy to attack
  - Internet is not easy to attack

Network Resilience

- What happens if nodes are removed? (interesting e.g. for vaccination effects in disease spreading in human contact networks)

- For power law networks:
  - remove random nodes: no effect on mean distances
  - remove high degree nodes: drastic effect

- Interpretations:
  - Internet is easy to attack
  - Internet is not easy to attack

Mixing Patterns

- What happens if nodes are removed? (interesting e.g. for vaccination effects in disease spreading in human contact networks)

- For power law networks:
  - remove random nodes: no effect on mean distances
  - remove high degree nodes: drastic effect

- Interpretations:
  - Internet is easy to attack
  - Internet is not easy to attack

- Ecological NW, Internet, some social NW:
  - Assortative Mixing (Homophily): Nodes attach to similar nodes / nodes of same class OR
  - Disassortative Mixing (Heterophily): Nodes attach to nodes of different classes (almost n-partite behavior)

- Diassortativity:
  - Food Web: Plants ↔ Herbivores ↔ Carnivores but few Plants ↔ Plants etc.
  - Internet: Backbones provider ↔ ISP ↔ end user but few ISP ↔ ISP etc.

- Assortativity:
  - Social NW
**Mixing Patterns**

- Ecological NW, Internet, some social NW:
  - **Assortative Mixing (Homophily):** Nodes attach to similar nodes / nodes of same class OR
  - **Disassortative Mixing (Heterophily):** Nodes attach to nodes of different classes (almost n-partite behavior)

- **Diassortativity:**
  - Food Web: Plants ↔ Herbivores ↔ Carnivores but few Plants ↔ Plants etc.
  - Internet: Backbones provider ↔ ISP ↔ end user but few ISP ↔ ISP etc.

- **Assortativity:**
  - Social NW

---

**Mixing Patterns**

\[
\mathbf{E} = \begin{array}{c|cccc}
\text{women} & \text{black} & \text{hispanic} & \text{white} & \text{other} \\
\hline
\text{men} & 506 & 32 & 69 & 26 \\
\text{hispanic} & 23 & 308 & 114 & 38 \\
\text{white} & 26 & 46 & 509 & 68 \\
\text{other} & 10 & 14 & 47 & 32 \\
\end{array}
\]

TABLE III Couples in the study of Catania et al. [85] tabulated by race of either partner. After Morris [302].

- measure mixing: analogous to modularity: mixing matrix \( \mathbf{e} = \frac{\mathbf{E}}{\|\mathbf{E}\|} \)

\[ P(j|i) = \frac{e_{ij}}{\sum_j e_{ij}}, \quad \sum_j P(j|i) = 1 \]

\[ \sum_j e_{ij} = 1, \quad \sum_j P(j|i) = 1 \]

---

**Mixing Patterns**

\[
\mathbf{E} = \begin{array}{c|cccc}
\text{women} & \text{black} & \text{hispanic} & \text{white} & \text{other} \\
\hline
\text{men} & 506 & 32 & 69 & 26 \\
\text{hispanic} & 23 & 308 & 114 & 38 \\
\text{white} & 26 & 46 & 509 & 68 \\
\text{other} & 10 & 14 & 47 & 32 \\
\end{array}
\]

TABLE III Couples in the study of Catania et al. [85] tabulated by race of either partner. After Morris [302].

- measure mixing: analogous to modularity: mixing matrix \( \mathbf{e} = \frac{\mathbf{E}}{\|\mathbf{E}\|} \)

\[ P(j|i) = \frac{e_{ij}}{\sum_j e_{ij}}, \quad \sum_j P(j|i) = 1 \]

\[ \sum_j e_{ij} = 1, \quad \sum_j P(j|i) = 1 \]
TABLE III Couples in the study of Catania et al. [85] tabulated by race of either partner. After Morris [302].

<table>
<thead>
<tr>
<th></th>
<th>women</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>black</td>
<td>hispanic</td>
<td>white</td>
<td>other</td>
</tr>
<tr>
<td>men</td>
<td>506</td>
<td>32</td>
<td>69</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>308</td>
<td>114</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>46</td>
<td>509</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>14</td>
<td>47</td>
<td>32</td>
</tr>
</tbody>
</table>

\[
E = \begin{pmatrix}
\text{black} & 506 & 32 & 69 & 26 \\
\text{hispanic} & 23 & 308 & 114 & 38 \\
\text{white} & 26 & 46 & 509 & 68 \\
\text{other} & 10 & 14 & 47 & 32
\end{pmatrix}
\]

**measure mixing: analogous to modularity: mixing matrix**

\[
e = \frac{E}{\|E\|}
\]

\[
P(j|i) = e_{ij} / \sum_j e_{ij}, \quad \sum_i e_{ij} = 1, \quad \sum_j P(j|i) = 1
\]

**measure mixing: analogous to modularity: mixing matrix**

\[
e = \frac{E}{\|E\|}
\]

\[
P(j|i) = e_{ij} / \sum_j e_{ij}, \quad \sum_i e_{ij} = 1, \quad \sum_j P(j|i) = 1
\]
Mixing Patterns

$E = \begin{array}{c|cccc} \text{men} & \text{black} & \text{hispanic} & \text{white} & \text{other} \\ \hline \text{women} & 506 & 22 & 69 & 26 \\ \text{black} & 23 & 308 & 114 & 38 \\ \text{hispanic} & 26 & 46 & 509 & 68 \\ \text{white} & 10 & 14 & 47 & 32 \\ \end{array}$

TABLE III Couples in the study of Catania et al. [85] tabulated by race of either partner. After Morris [90].

- **first measure for Assortativity:**
  \[ Q = \frac{\sum_i P(i|i) - 1}{N - 1} \]

  **issues:** Asymmetry of $E \rightarrow$ two values; Not respecting size of classes

- **second measure for Assortativity:** (cmp. Modularity)
  \[ r = \frac{\text{Tr} e - \| e^2 \|}{1 - \| e^2 \|} \]
Mixing Patterns

• Special example: „class“ of nodes determined by degree
  → nodes attached to nodes with same or different degree?
    Both variants occur in real world NW

• Degree correlation measures:
  1) mean degree of neighbors of node with degree k:
     → if assortative mixing: curve should be increasing
     → Internet: curve decreases → diassortativity
  2) Pearson correlation for node degrees $k_i$ and $k_j$ of
     adjacent nodes $i$ and $j$
### Community and Group Structure

- **Is NW well clustered?** → see Parts on Clustering

**example:** friendship NW in US school:

---

### Navigability of NW

- Milgram showed: short paths exist
  - BUT: How do people find them?

  → see Part „Social Networks in Time and Space“

---

### Component Structure

- Does a giant component exist?

  → see section on random graphs

---

**Table: Network Data**

<table>
<thead>
<tr>
<th>Network</th>
<th>Type</th>
<th>$n$</th>
<th>$m$</th>
<th>$\langle k \rangle$</th>
<th>$\langle k \rangle$</th>
<th>$\langle e \rangle$</th>
<th>$\langle e \rangle$</th>
<th>$\langle d \rangle$</th>
<th>$\langle d \rangle$</th>
<th>$\langle z \rangle$</th>
<th>$\langle z \rangle$</th>
<th>$\langle f \rangle$</th>
<th>$\langle f \rangle$</th>
</tr>
</thead>
<tbody>
<tr>
<td>film actors</td>
<td>undirected</td>
<td>449</td>
<td>25,514</td>
<td>1334</td>
<td>3.4</td>
<td>2.3</td>
<td>0.20</td>
<td>0.38</td>
<td>0.78</td>
<td>0.26</td>
<td>20,416</td>
<td>105,323</td>
<td></td>
</tr>
<tr>
<td>company directors</td>
<td>undirected</td>
<td>7403</td>
<td>55,392</td>
<td>14.44</td>
<td>4.6</td>
<td>-</td>
<td>0.59</td>
<td>0.88</td>
<td>20.4</td>
<td>0.84</td>
<td>95,323</td>
<td>105,323</td>
<td></td>
</tr>
<tr>
<td>math membership</td>
<td>undirected</td>
<td>2532</td>
<td>846,490</td>
<td>3.92</td>
<td>7.7</td>
<td>-</td>
<td>0.15</td>
<td>0.24</td>
<td>0.12</td>
<td>0.19</td>
<td>197,182</td>
<td>105,323</td>
<td></td>
</tr>
<tr>
<td>physics membership</td>
<td>undirected</td>
<td>2532</td>
<td>846,490</td>
<td>3.92</td>
<td>7.7</td>
<td>-</td>
<td>0.15</td>
<td>0.24</td>
<td>0.12</td>
<td>0.19</td>
<td>197,182</td>
<td>105,323</td>
<td></td>
</tr>
<tr>
<td>biology membership</td>
<td>undirected</td>
<td>152</td>
<td>11,803</td>
<td>15.53</td>
<td>4.9</td>
<td>-</td>
<td>0.98</td>
<td>0.60</td>
<td>0.12</td>
<td>0.96</td>
<td>311,313</td>
<td>311,313</td>
<td></td>
</tr>
<tr>
<td>telephone cell graph</td>
<td>undirected</td>
<td>47,000,000</td>
<td>0.16</td>
<td>2.1</td>
<td>0.9</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>0.8</td>
<td>8,9</td>
<td>8,9</td>
<td></td>
</tr>
<tr>
<td>small messages</td>
<td>directed</td>
<td>55,912</td>
<td>57,392</td>
<td>5.38</td>
<td>15.5</td>
<td>-</td>
<td>0.17</td>
<td>0.89</td>
<td>0.12</td>
<td>0.86</td>
<td>301,313</td>
<td>301,313</td>
<td></td>
</tr>
<tr>
<td>email addresses books</td>
<td>directed</td>
<td>57,392</td>
<td>57,392</td>
<td>5.38</td>
<td>15.5</td>
<td>-</td>
<td>0.17</td>
<td>0.89</td>
<td>0.12</td>
<td>0.86</td>
<td>301,313</td>
<td>301,313</td>
<td></td>
</tr>
<tr>
<td>student relationships</td>
<td>undirected</td>
<td>57,392</td>
<td>57,392</td>
<td>5.38</td>
<td>15.5</td>
<td>-</td>
<td>0.17</td>
<td>0.89</td>
<td>0.12</td>
<td>0.86</td>
<td>301,313</td>
<td>301,313</td>
<td></td>
</tr>
<tr>
<td>sexual contacts</td>
<td>undirected</td>
<td>2,410</td>
<td>2,410</td>
<td>2.2</td>
<td>2.2</td>
<td>-</td>
<td>0.32</td>
<td>2.2</td>
<td>2.2</td>
<td>2.2</td>
<td>265,266</td>
<td>265,266</td>
<td></td>
</tr>
<tr>
<td>WWW inf. etc.</td>
<td>directed</td>
<td>369,304</td>
<td>1,097,135</td>
<td>5.55</td>
<td>11.27</td>
<td>-</td>
<td>0.11</td>
<td>0.29</td>
<td>0.007</td>
<td>0.14</td>
<td>14,34</td>
<td>14,34</td>
<td></td>
</tr>
<tr>
<td>WWW AlClearly</td>
<td>directed</td>
<td>252,564</td>
<td>2,130,000</td>
<td>10.46</td>
<td>16.14</td>
<td>-</td>
<td>0.11</td>
<td>0.29</td>
<td>0.007</td>
<td>0.14</td>
<td>14,34</td>
<td>14,34</td>
<td></td>
</tr>
<tr>
<td>citation network</td>
<td>directed</td>
<td>783,359</td>
<td>676,158</td>
<td>8.57</td>
<td>5.8</td>
<td>-</td>
<td>0.13</td>
<td>0.35</td>
<td>0.15</td>
<td>0.35</td>
<td>241</td>
<td>241</td>
<td></td>
</tr>
<tr>
<td>Roger’s Thesaurus</td>
<td>directed</td>
<td>1,023</td>
<td>5,133</td>
<td>4.99</td>
<td>4.87</td>
<td>-</td>
<td>0.13</td>
<td>0.35</td>
<td>0.15</td>
<td>0.35</td>
<td>241</td>
<td>241</td>
<td></td>
</tr>
<tr>
<td>word co-occurrences</td>
<td>undirected</td>
<td>461,962</td>
<td>17,000,000</td>
<td>70.13</td>
<td>2.7</td>
<td>-</td>
<td>0.13</td>
<td>0.35</td>
<td>0.15</td>
<td>0.35</td>
<td>119,157</td>
<td>119,157</td>
<td></td>
</tr>
<tr>
<td>Internet</td>
<td>undirected</td>
<td>10,407</td>
<td>31,992</td>
<td>5.98</td>
<td>3.31</td>
<td>-</td>
<td>0.39</td>
<td>0.19</td>
<td>0.18</td>
<td>0.19</td>
<td>146,146</td>
<td>146,146</td>
<td></td>
</tr>
<tr>
<td>power grid</td>
<td>undirected</td>
<td>4,941</td>
<td>6,094</td>
<td>2.87</td>
<td>10.95</td>
<td>-</td>
<td>0.10</td>
<td>0.18</td>
<td>0.08</td>
<td>0.08</td>
<td>416</td>
<td>416</td>
<td></td>
</tr>
<tr>
<td>train routes</td>
<td>undirected</td>
<td>587</td>
<td>104,203</td>
<td>66.79</td>
<td>2.86</td>
<td>-</td>
<td>0.69</td>
<td>0.19</td>
<td>0.18</td>
<td>0.18</td>
<td>866</td>
<td>866</td>
<td></td>
</tr>
<tr>
<td>software packages</td>
<td>undirected</td>
<td>1,249</td>
<td>1,723</td>
<td>2.42</td>
<td>2.2</td>
<td>-</td>
<td>0.11</td>
<td>0.18</td>
<td>0.18</td>
<td>0.18</td>
<td>213</td>
<td>213</td>
<td></td>
</tr>
<tr>
<td>software classe</td>
<td>undirected</td>
<td>1,237</td>
<td>2,133</td>
<td>1.61</td>
<td>2.5</td>
<td>-</td>
<td>0.03</td>
<td>0.12</td>
<td>0.19</td>
<td>0.19</td>
<td>366</td>
<td>366</td>
<td></td>
</tr>
<tr>
<td>electronic circuits</td>
<td>undirected</td>
<td>1,297</td>
<td>31,149</td>
<td>4.34</td>
<td>11.05</td>
<td>-</td>
<td>0.09</td>
<td>0.35</td>
<td>0.15</td>
<td>0.35</td>
<td>155</td>
<td>155</td>
<td></td>
</tr>
<tr>
<td>peer-to-peer network</td>
<td>undirected</td>
<td>880</td>
<td>1,056</td>
<td>1.42</td>
<td>5.28</td>
<td>-</td>
<td>0.11</td>
<td>0.36</td>
<td>0.15</td>
<td>0.36</td>
<td>6,354</td>
<td>6,354</td>
<td></td>
</tr>
<tr>
<td>metabolic network</td>
<td>undirected</td>
<td>785</td>
<td>3,466</td>
<td>9.64</td>
<td>1.8</td>
<td>-</td>
<td>0.25</td>
<td>0.28</td>
<td>0.38</td>
<td>0.38</td>
<td>114</td>
<td>114</td>
<td></td>
</tr>
<tr>
<td>protein interactions</td>
<td>undirected</td>
<td>3,115</td>
<td>2,539</td>
<td>2.12</td>
<td>0.9</td>
<td>-</td>
<td>0.22</td>
<td>0.71</td>
<td>0.12</td>
<td>0.71</td>
<td>222</td>
<td>222</td>
<td></td>
</tr>
<tr>
<td>marine food web</td>
<td>directed</td>
<td>135</td>
<td>568</td>
<td>4.43</td>
<td>2.05</td>
<td>-</td>
<td>0.16</td>
<td>0.23</td>
<td>0.36</td>
<td>0.36</td>
<td>294</td>
<td>294</td>
<td></td>
</tr>
<tr>
<td>freshwater food web</td>
<td>directed</td>
<td>92</td>
<td>997</td>
<td>3.84</td>
<td>1.9</td>
<td>-</td>
<td>0.20</td>
<td>0.87</td>
<td>0.32</td>
<td>0.32</td>
<td>272</td>
<td>272</td>
<td></td>
</tr>
<tr>
<td>natural network</td>
<td>directed</td>
<td>257</td>
<td>2,359</td>
<td>7.68</td>
<td>3.07</td>
<td>-</td>
<td>0.14</td>
<td>0.28</td>
<td>0.20</td>
<td>0.28</td>
<td>146,421</td>
<td>146,421</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** Basic statistic for a network of the networks. The properties measured are: type of graph, directed or undirected; total number of vertices $n$; total number of edges $m$; average degree $\langle k \rangle$; maximum vertex distance $d$; exponent $\gamma$ of degree distribution; if the distribution follows a power law ($\gamma = \gamma$); if not, in/out degree are given for directed graphs; clustering coefficient $C^{(3)}$ from Eq (3); clustering coefficient $C^{(3)}$ from Eq (8); and degree correlation coefficient $r$. So last column gives the citation(s) for the network in the bibliography. Black entries indicate unavailable data.
**Random Graph Models: Poisson Graph**

- \( G_{n,p} \): space of graphs with \( n \) nodes and each of the \( \frac{1}{2} \, n(n-1) \) edges appears with probability \( p \)

- \( p_k \): probability that a node has degree \( k \):
  \[
  p_k = \binom{n}{k} p^k (1 - p)^{n-k} \approx \frac{z^k e^{-z}}{k!}
  \]

for \( n \to \infty \) and holding the mean degree of a node \( z = p(n-1) \) fixed
(Poisson approximation of Binomial distribution)

\( \Rightarrow \) „Poisson random graphs“
Random Graph Models: Poisson Graph

- Given: property $Q_{i}$ (is connected, has diameter $xyz$ etc.) of $G_{n,p}$:
  "$G_{n,p}$ has property $Q$ with high probability": $P(Q|n,p) \rightarrow 1$ if $n \rightarrow \infty$
  (adapted from [2] which, in turn, is adapted from [3])

- In such models $G_{n,p}$ phase transitions exist for properties $Q$:
  "threshold function" $q(n)$ (with $q(n) \rightarrow \infty$ if $n \rightarrow \infty$) so that:
  $$\lim_{n \to \infty} P(Q|n,p) = \begin{cases} 0 & \text{if } \lim_{n \to \infty} p(n)/q(n) = 0 \\ 1 & \text{if } \lim_{n \to \infty} p(n)/q(n) = \infty \end{cases}$$
  (adapted from [3])

Random Graph Models: Poisson Graph

- Given: property $Q_{i}$ (is connected, has diameter $xyz$ etc.) of $G_{n,p}$:
  "$G_{n,p}$ has property $Q$ with high probability": $P(Q|n,p) \rightarrow 1$ if $n \rightarrow \infty$
  (adapted from [2] which, in turn, is adapted from [3])

- In such models $G_{n,p}$ phase transitions exist for properties $Q$:
  "threshold function" $q(n)$ (with $q(n) \rightarrow \infty$ if $n \rightarrow \infty$) so that:
  $$\lim_{n \to \infty} P(Q|n,p) = \begin{cases} 0 & \text{if } \lim_{n \to \infty} p(n)/q(n) = 0 \\ 1 & \text{if } \lim_{n \to \infty} p(n)/q(n) = \infty \end{cases}$$
  (adapted from [3])

Random Graph Models: Poisson Graph

- Given: property $Q$ (is connected, has diameter $xyz$ etc.) of $G_{n,p}$:
  "$G_{n,p}$ has property $Q$ with high probability": $P(Q|n,p) \rightarrow 1$ if $n \rightarrow \infty$
  (adapted from [2] which, in turn, is adapted from [3])

- In such models $G_{n,p}$ phase transitions exist for properties $Q$:
  "threshold function" $q(n)$ (with $q(n) \rightarrow \infty$ if $n \rightarrow \infty$) so that:
  $$\lim_{n \to \infty} P(Q|n,p) = \begin{cases} 0 & \text{if } \lim_{n \to \infty} p(n)/q(n) = 0 \\ 1 & \text{if } \lim_{n \to \infty} p(n)/q(n) = \infty \end{cases}$$
  (adapted from [3])

Random Graph Models: Poisson Graph

- Given: property $Q$ (is connected, has diameter $xyz$ etc.) of $G_{n,p}$:
  "$G_{n,p}$ has property $Q$ with high probability": $P(Q|n,p) \rightarrow 1$ if $n \rightarrow \infty$
  (adapted from [2] which, in turn, is adapted from [3])

- In such models $G_{n,p}$ phase transitions exist for properties $Q$:
  "threshold function" $q(n)$ (with $q(n) \rightarrow \infty$ if $n \rightarrow \infty$) so that:
  $$\lim_{n \to \infty} P(Q|n,p) = \begin{cases} 0 & \text{if } \lim_{n \to \infty} p(n)/q(n) = 0 \\ 1 & \text{if } \lim_{n \to \infty} p(n)/q(n) = \infty \end{cases}$$
  (adapted from [3])
Example: giant component / connectedness of $G_{n,p}$

- Let $u$ be the fraction of nodes that do not belong to giant component $X$
  $\Rightarrow$ probability for a given node $i$ to be not in $X$ $\approx u$

- probability for a given node $i$ (with assumed degree $k$) to be not in $X$
  $\Rightarrow$ probability that none of its neighbors is in $X$ $\approx u^k$

- $u$ (k fixed) $\Rightarrow u_k^k$ $\Rightarrow u = \sum_{k=0}^{\infty} p_k u^k = e^{-z} \sum_{k=0}^{\infty} \frac{(zu)^k}{k!} = e^{z(u-1)}$

- fraction $S$ of graph occupied by $X$ is $S = 1 - u$ $\Rightarrow$

\[ S = 1 - e^{-zS} \]
Example: giant component / connectedness of $G_{n,p}$

- Let $u$ be the fraction of nodes that do not belong to giant component $X$
  $\Rightarrow$ probability for a given node $i$ to be not in $X$

- probability for a given node $i$ (with assumed degree $k$) to be not in $X$
  $\Rightarrow$ probability that none of its neighbors is in $X$
  $\Rightarrow u^k$

- $u$ (k fixed) $\Rightarrow u^k$
  $\Rightarrow u = \sum_{k=0}^{\infty} \frac{p_k u^k}{k!} = e^{-z} \sum_{k=0}^{\infty} \frac{(zu)^k}{k!} = e^z(u-1)$

- fraction $S$ of graph occupied by $X$ is
  $S = 1 - u$ $\Rightarrow$
  $S = 1 - e^{-zS}$

Example: giant component / connectedness of $G_{n,p}$

- Let $u$ be the fraction of nodes that do not belong to giant component $X$
  $\Rightarrow$ probability for a given node $i$ to be not in $X$

- probability for a given node $i$ (with assumed degree $k$) to be not in $X$
  $\Rightarrow$ probability that none of its neighbors is in $X$
  $\Rightarrow u^k$

- $u$ (k fixed) $\Rightarrow u^k$
  $\Rightarrow u = \sum_{k=0}^{\infty} \frac{p_k u^k}{k!} = e^{-z} \sum_{k=0}^{\infty} \frac{(zu)^k}{k!} = e^z(u-1)$

- fraction $S$ of graph occupied by $X$ is
  $S = 1 - u$ $\Rightarrow$
  $S = 1 - e^{-zS}$

Result:

- $S = 1 - e^{-zS}$
- mean size $\langle s \rangle$ of smaller rest components (no proof): $\langle s \rangle = \frac{1}{1 - z + zS}$

- if the avg degree $z$ is larger than 1 ($= p \sim (1+\epsilon)/n$): $X$ exists
Example: giant component / connectedness of $G_{n,p}$

- Let $u$ be the fraction of nodes that do not belong to giant component $X$
  $\Rightarrow$ probability for a given node $i$ to be not in $X$

- probability for a given node $i$ (with assumed degree $k$) to be not in $X$
  $\Rightarrow$ probability that none of its neighbors is in $X$
  $\Rightarrow u^k$

- $u$ (k fixed) $\Rightarrow$ $u^k$
  $\Rightarrow u = \sum_{k=0}^{\infty} p_k u^k = e^{-z} \sum_{k=0}^{\infty} \frac{(zu)^k}{k!} = e^z(u-1)$

- fraction $S$ of graph occupied by $X$ is
  $S = 1 - u \Rightarrow S = 1 - e^{-zS}$

$\Rightarrow$ if the avg degree $z$ is larger than 1 (if $p \sim (1+\varepsilon)/n$): $X$ exists

$S = 1 - e^{-zS}$

mean size $<s>$ of smaller rest components (no proof): $\langle s \rangle = \frac{1}{1 - z + zS}$

$\Rightarrow$ if the avg degree $z$ is larger than 1 (if $p \sim (1+\varepsilon)/n$): $X$ exists
\[ S = 1 - e^{-zS} \]
\[ \text{mean size } <s> \text{ of smaller rest components (no proof): } <s> = \frac{1}{1 - z + zS} \]

\[ \text{if the avg degree } z \text{ is larger than 1 } \left( \equiv \frac{p}{n} \text{ is } 1+\epsilon \right) \text{: } X \text{ exists} \]