Today, we will:

- Discuss **Particle Statistics** in **Section 8.2**
- Discuss an example particle size distribution in detail – *see class handout*

The data shown below are from the class handout, so you can follow along.

<table>
<thead>
<tr>
<th>$j$ = class (bin number)</th>
<th>$D_{p, \text{min}, j}$ (lower limit)</th>
<th>$D_{p, \text{max}, j}$ (upper limit)</th>
<th>$D_{pj}$ (middle value)</th>
<th>$\Delta D_{p,j}$ = class width</th>
<th>$n_j$ = frequency (count per class)</th>
<th>$n_j/\Delta D_{p,j}$ = count per class width</th>
<th>$f(D_{p,j}) = n_j/(\Delta D_{p,j} n_t)$ = fraction per class width</th>
<th>probability in this class $= f(D_{p,j}) \times \Delta D_{p,j} = n_j/n_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td>2.5</td>
<td>3</td>
<td>104</td>
<td>34.667</td>
<td>0.0347</td>
<td>0.104</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>160</td>
<td>80.000</td>
<td>0.0800</td>
<td>0.0800</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>2</td>
<td>$^\ast$ 161</td>
<td>80.500</td>
<td>0.0805</td>
<td>0.161</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>9</td>
<td>8.5</td>
<td>1</td>
<td>75</td>
<td>75.000</td>
<td>0.0750</td>
<td>0.0750</td>
</tr>
<tr>
<td>5</td>
<td>9</td>
<td>10</td>
<td>9.5</td>
<td>1</td>
<td>67</td>
<td>67.000</td>
<td>0.0670</td>
<td>0.0670</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>14</td>
<td>12</td>
<td>4</td>
<td>186</td>
<td>46.500</td>
<td>0.0465</td>
<td>0.186</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>16</td>
<td>15</td>
<td>2</td>
<td>61</td>
<td>30.500</td>
<td>0.0305</td>
<td>0.061</td>
</tr>
<tr>
<td>8</td>
<td>16</td>
<td>20</td>
<td>18</td>
<td>4</td>
<td>79</td>
<td>19.750</td>
<td>0.0198</td>
<td>0.079</td>
</tr>
<tr>
<td>9</td>
<td>20</td>
<td>35</td>
<td>27.5</td>
<td>15</td>
<td>90</td>
<td>6.000</td>
<td>0.0060</td>
<td>0.09</td>
</tr>
<tr>
<td>10</td>
<td>35</td>
<td>50</td>
<td>42.5</td>
<td>15</td>
<td>17</td>
<td>1.133</td>
<td>0.0011</td>
<td>0.017</td>
</tr>
<tr>
<td>11</td>
<td>50</td>
<td>100</td>
<td>75</td>
<td>50</td>
<td>0</td>
<td>0.000</td>
<td>0.0000</td>
<td>0</td>
</tr>
</tbody>
</table>

Minimum diameter in class $j$

Number of particles in class $j$

Fraction of particles in class $j$ divided by class width: $n_j/\Delta D_{p,j}$

Probability in class $j$

Number of particles in class $j$ divided by class width: $n_j/\Delta D_{p,j}$

Middle diameter in class $j$: $D_{pj} = (D_{p,\text{min},j} + D_{p,\text{max},j}) / 2$

Number of particles in class $j$ divided by class width: $n_j/\Delta D_{p,j}$

Maximum diameter in class $j$

Width of class $j$: $\Delta D_{p,j} = D_{p,\text{max},j} - D_{p,\text{min},j}$

**See the handout — also available as an Excel file on the website.**
Plot 1 - Simple histogram

Plot 2 - Modified histogram

PDF = $f(D_p, j)$

Plot 3 - Histogram of fraction per class width

Plot 4 - Probability density function (PDF), linear scale

Plot 5 - PDF, log scale

Plot 6 - Cumulative distribution function

Plot 4 = PDF

$F(D_p) = \int_0^{D_p \max} f(D_p, j) \, dD_p$

Area = 1

This is true for any PDF.
Our PDF is not symmetric — Not Gaussian

However, it turns out that many particle distributions in air pollution work are log-normal — see plot 5

Cumulative Distribution Function $F(D_p)$

For $D_p = a$, $F(a) = \int_0^a f(D_p) \, dD_p$ by definition

The probability that $D_p = a$

$F(b) - F(a) = \int_0^b f(D_p) \, dD_p - \int_0^a f(D_p) \, dD_p$

$= \text{Prob. that a particle diameter lies between } a \text{ and } b$

$F(\infty) = 1$ (Prob. that $0 < D_p < \infty$ is 100% !)

$F(a) - F(a) = \int_a^a f(D_p) \, dD_p = 0$

Zero prob. that $D_p = a$ exactly
Cumulative Distribution Function,

\[ F(D_p) \]

\[ 0 \]

\[ 0.5 \]

\[ 1.0 \]

\[ D_{p,50} = \text{median diameter} \]

\[ (\text{half of } D_p < D_{p,50}) \]

\[ (\text{half } D_p > D_{p,50}) \]

\[ D_{p,50} \]

\[ D_{p,50} = D_{\text{max at which } F(D_p) = 0.500} \]

\[ j=4 \]

\[ D_{p,50} = 9.0 \mu m \]

\[ \text{not } 8.5 \mu m \text{ (D}_{p,j}\text{)} \]

\[ \text{middle value} \]

\[ \text{max value in the class } j=4 \]

50% of the particles in this sample are < 9.0 \mu m in diameter.

50% --- --- --- --- ---

**Statistics:**

Arithmetic mean diameter = \( D_{p,\text{am (number)}} \)

based on number of particles

[Later on we will define \( D_{p,\text{am (mass)}} \)]
For list data:

\[
D_{p,\text{am}} = \frac{\sum_{i=1}^{n_t} D_{p, i}}{n_t}
\]

For grouped data:

Assume each particle in a particular bin has \( D_p = D_{p, j} \) (m.i. value)

\[
D_{p,\text{am}} = \frac{\sum_{j=1}^{J} n_j D_{p, j}}{n_t}
\]

Here, we get \( D_{p,\text{am}} = 11.2 \, \mu m \)

\[\text{NOTICE:} \quad D_{p,\text{am}} \neq D_{p,\text{mohin}} \quad \text{(since the POP is not Gaussian)}\]

**Standard Deviation**

**For list data**, \( \sigma \) is:

\[
\sigma = \left[ \frac{\sum_{i=1}^{n_t} (D_{p, i} - D_{p,\text{am}})^2}{n_t - 1} \right]^{\frac{1}{2}}
\]

**For grouped data**, \( \sigma \) is:

\[
\sigma = \left[ \frac{\sum_{j=1}^{J} n_j (D_{p, j} - D_{p,\text{am}})^2}{n_t - 1} \right]^{\frac{1}{2}}
\]

Here, we calculate \( \sigma = 7.93 \, \mu m \)

\( D_{p,\text{mode}} = \) high point on the POP (the most frequent class)

\( D_{p,\text{mode}} = 7.0 \, \mu m \)

Here \( D_{p,\text{mode}} < D_{p,\text{moe}} < D_{p,\text{am}} \)
Geometric mean \( D_{p, gm} \) [we multiply instead of add in this]

\[ D_{p, gm} = \left( \frac{D_{p_1}}{n_1}, \frac{D_{p_2}}{n_2}, \frac{D_{p_3}}{n_3}, \ldots, \frac{D_{p_{nt-1}}}{n_{nt-1}}, \frac{D_{p_{nt}}}{n_{nt}} \right)^{\frac{1}{n_t}} \]

For grouped data,

\[ D_{p, gm} = \left( \frac{D_{p_1}^{n_1}, D_{p_2}^{n_2}, D_{p_3}^{n_3}, \ldots, D_{p_j}^{n_j}}{n_t} \right)^{\frac{1}{n_t}} \]

For \( j = 1, \ldots, t \)

OR, after some algebra

\[ D_{p, gm} = \exp \left( \frac{1}{n_t} \sum_{j=1}^{n_t} n_j \ln D_{p,j} \right) \]

Here, \( 0_{p, gm} = 8.98 \ \text{mm} \)

Geometric standard deviation \( \sigma_g \)

Define \( \sigma_g \) using its natural log.

\[ \ln \sigma_g = \text{standard deviation in terms of } \ln (D_{p,j}) \text{, not } D_{p,j} \text{ itself.} \]

\[ \sigma_g = e^{\ln \sigma_g} \]
\[ \ln \sigma_g = \left[ \frac{\sum_{j=1}^{J} n_j \left( \ln \left( \frac{p_{0,j,i}}{\sigma_{am}} \right) \right)^2}{n_t - 1} \right]^{\frac{1}{2}} \]

For our sample data, \( \sigma_g = 1.97 \text{ mm} \)

Where \( \left[ \ln \left( \frac{p_{0,j,i}}{\sigma_{am}} \right) \right]_{am} = \frac{\sum_{j=1}^{J} n_j \ln \left( \frac{p_{0,j}}{\sigma_{am}} \right)}{n_t} \)