Hazen Formula

\[ F_i = \left( \frac{i - \alpha}{n - 2\alpha + 1} \right) \quad \text{(CDF)} \]

\[ f_i = \frac{dF_i}{dt} = \frac{F_{i+1} - F_i}{t_{i+1} - t_i} \quad \text{(PDF)} \]

\[ = \frac{1}{(t_{i+1} - t_i)} \cdot \left\{ \frac{(i+1) - \alpha - (i - \alpha)}{n - 2\alpha + 1} \right\} \]

\[ = \frac{1}{(t_{i+1} - t_i)(n - 2\alpha + 1)} \]

Hazard rate

\[ \lambda_i = \frac{f_i}{(1 - F_i)} \]

\[ = \frac{1}{(t_{i+1} - t_i)} \left( \frac{1}{n - 2\alpha + 1} \right) \]

\[ = \frac{1}{1 - \left( \frac{i - \alpha}{n - 2\alpha + 1} \right)} \]

\[ = \frac{1}{(t_{i+1} - t_i)(n - \alpha - i + 1)} \]
Extreme value PDF is given by

\[ f(t) = \frac{1}{b} e^{-(t-\tau)/b} \cdot e^{-(t-\tau)/\lambda} \]

CDF

\[ F(t) = \int_{-\infty}^{t} \frac{1}{b} e^{y - (y-\tau)/b} \cdot e^{-(y-\tau)/\lambda} \, dy \]

Let \( y = e^{(t-\tau)/b} \),

\[ t' = -\infty, \quad y = 0 \]
\[ t' = t, \quad y = e^{(t-\tau)/b} \]

\[ F(t) = \int_{0}^{e^{(t-\tau)/b}} \frac{1}{b} y \cdot e^{-y} \cdot \frac{1}{b} \, dy \]

\[ = e^{-\frac{1}{b} \cdot e^{(t-\tau)/b}} \]

\[ = 1 - e^{-(t-\tau)/\lambda} \quad (-\infty < t < \infty) \]

\[ R(t) = 1 - F(t) = e^{-(t-\tau)/\lambda} \quad (-\infty < t < \infty) \]

\[ \lambda(t) = \frac{f(t)}{1 - F(t)} = \frac{\frac{1}{b} e^{(t-\tau)/b} \cdot e^{-(t-\tau)/\lambda}}{e^{-e^{(t-\tau)/\lambda}}} \]

\[ = \frac{1}{b} e^{(t-\tau)/\lambda} \]
\[ R(t) = 1 - F(t) = e^{-(t-\tau)/b} \]

\[ H(t) = -\ln R(t) = e^{(t-\tau)/b} \]

\[ h(t) = \frac{dH(t)}{dt} = \frac{1}{b} e^{(t-\tau)/b} \]

Notes:

1. In this problem, you must have noticed that we needed to extend the limit to \( t < 0 \), in order for the distribution to add up to 1. Many distributions commonly used by the reliability engineers have this trouble, e.g., Normal distribution, \textit{Two-Sided}. The reason being that the approximations necessary to take a physical distribution to these limiting distributions is not always consistent with \( t > 0 \). requirement.

2. Two solutions are possible.

   a. Integrate to 1: \[ 1 = m \int_{-\infty}^{\infty} f(t) \, dt \]

   Find the constant \( m \) and then do rest of the calculations.

   b. Use classical formula with \( t < 0 \) part included, but remember that all rates calculated at \( t \to 0 \) is inaccurate. As \( t \) gets larger (the interesting region), the results are better.
Data Analysis

'Fish in a river' distribution

\[ f(t) = \frac{x_0}{\sqrt{4\pi D t^3}} e^{-\frac{x_0^2}{4D t}} \]

\[ F(t) = \int_0^t f(\tau) d\tau \]

\[ F(t) = \frac{x_0}{\sqrt{4\pi D}} \int_0^t \frac{e^{-\frac{x_0^2}{4D \tau}}}{\tau^{3/2}} d\tau \]

\[ \frac{ut}{4DP} = \frac{u^2}{2} \Rightarrow u = \frac{x_0}{\sqrt{4DP}} \]

\[ \frac{u}{u^2} du = 2u du \]

\[ d\tau = -\frac{4D}{x_0^2} p^{3/2} \cdot 2 \cdot \frac{x_0}{\sqrt{4DP}} dw \]

\[ F(t) = \frac{x_0}{\sqrt{4\pi D}} \int_0^{\infty} \frac{2}{\sqrt{\pi}} \cdot 2 \cdot p^{3/2} \cdot e^{-u^2} du \]

\[ = \frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-u^2} du \]

\[ = \frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-u^2} du \]

\[ = + \left[ 1 - \text{erf} \left( \frac{x_0}{\sqrt{4D t}} \right) \right] \]
\[ \Lambda(t) = \frac{f(t)}{1 - F(t)} = \frac{\frac{z_0}{(4\pi Dt)^{3/2}} \cdot e^{-\frac{z_0^2}{4Dt}}}{\text{erf} \left( \frac{z_0}{\sqrt{4Dt}} \right)} \]

\[ R(t) = 1 - F(t) = \text{erf} \left( \frac{z_0}{\sqrt{4Dt}} \right) \]

\[ \Phi(t) = -\ln R(t) = -\ln \left[ \text{erf} \left( \frac{z_0}{\sqrt{4Dt}} \right) \right] \]
### 2.1 Empirical CDF for the 4 datasets

![Empirical CDF](image)
Dataset 1
Probability plot for Normal distribution

Dataset 1
Probability plot for Lognormal distribution

Dataset 1
Probability plot for Weibull distribution

Lognormal fits the dataset
Dataset 2

Probability plot for Normal distribution

Probability plot for Lognormal distribution

Probability plot for Weibull distribution

Lognormal fits the data best
Dataset 3
Probability plot for Normal distribution

Probability plot for Lognormal distribution

Probability plot for Weibull distribution

Lognormal fits the data best.
Dataset 4
Probability plot for Normal distribution

Probability plot for Lognormal distribution

Probability plot for Weibull distribution

$\xrightarrow{\text{Lognormal fit}}$
the data besit
2.2 Lognormal distribution fitting -

Dataset

<table>
<thead>
<tr>
<th>Dataset</th>
<th>M</th>
<th>( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-11.157</td>
<td>2.8931</td>
</tr>
<tr>
<td>2</td>
<td>-7.5088</td>
<td>2.0912</td>
</tr>
<tr>
<td>3</td>
<td>-7.5737</td>
<td>1.5184</td>
</tr>
<tr>
<td>4</td>
<td>-7.9421</td>
<td>3.0669</td>
</tr>
</tbody>
</table>

MATLAB command used -

```matlab
param = mle (Dataset1, 'distribution', 'logn');
mu1 = param(1);
sigma1 = param(2);  % Similarly for other datasets
```

2.3 Kolmogorov-Smirnov goodness of fit test -

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Ksstat</th>
<th>CV</th>
<th>P</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.050652</td>
<td>0.141954</td>
<td>0.967622</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0.121377</td>
<td>0.157558</td>
<td>0.220666</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0.1876</td>
<td>0.173741</td>
<td>0.02724</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0.179428</td>
<td>0.274941</td>
<td>0.0401965</td>
<td>0</td>
</tr>
</tbody>
</table>

MATLAB command used -

```matlab
[ch1, p1, ksstat1, cv1] = kstest (Dataset1, [Dataset1 logncdf (Dataset1, mu1, sigma1)], 0.05);
% Similarly for other datasets.
```
For dataset 3, even though \texttt{ksstat} > \texttt{cv}, the lognormal distribution is acceptable at 95\% confidence interval.
2.4 Likelihood Ratio Test -

Dataset 1 -

Lognormal vs. Normal

\[ L_r \rightarrow 310.0231 \]

\[ P \rightarrow 3.1643 \times 10^{-20} \]

Lognormal is more likely than Weibull & Normal.

Dataset 2 -

Lognormal vs. Normal

\[ L_r \rightarrow 216.9521 \]

\[ P \rightarrow 7.9266 \times 10^{-16} \]

Lognormal is more likely than Weibull & Normal.

Dataset 3 -

Lognormal vs. Normal

\[ L_r \rightarrow 206.1109 \]

\[ P \rightarrow 3.1137 \times 10^{-25} \]

Lognormal is more likely than Weibull & Normal.

Dataset 4 -

Lognormal vs. Normal

\[ L_r \rightarrow 95.7436 \]

\[ P \rightarrow 1.4395 \times 10^{-13} \]

Lognormal is more likely than Weibull & Normal.