Fitting root diameter counts with power law probabilities

G. J. Meijer*†

August 9, 2023

1 Segments

The probability can be described using a multi-segmented power-law curves that is C^0 continuous. n is the number of segments, separated by breakpoint values x_b (length $n-1$). The minimum and maximum observed values of x are given by x_{min} and x_{max}, which automatically form the lower bound of the first segment and upper bound of the final segment when maximising the probability.

2 Individual probabilities

The probability in each segment i is given by:

$$p = \alpha_i x^{\beta_i} \tag{1}$$

Because of C^0 continuity, at breakpoint location j:

$$\alpha_j x_{b,j}^{\beta_j} = \alpha_{j+1} x_{b,j}^{\beta_{j+1}} \tag{2}$$

so the multiplier α for the next segment satisfies:

$$\alpha_{j+1} = \alpha_j x_{b,j}^{\beta_j - \beta_{j+1}} \tag{3}$$

The probability in segment j ($2 \leq j \leq n$) can now be described as:

$$p_j = \alpha_1 x^{\beta_j} \prod_{i=2}^{j} x_{b,i-1}^{\beta_{i-1}-\beta_i} \tag{4}$$

or alternatively as:

$$p_j = \alpha_1 x^{\beta_1} \prod_{i=2}^{j} \left(\frac{x}{x_{b,i-1}} \right)^{\beta_i - \beta_{i-1}} \tag{5}$$

The log-transformed value of Equation 5 is:

$$\log(p_j) = \log(\alpha_1) + \beta_1 \log(x) + \sum_{i=2}^{j} (\beta_i - \beta_{i-1}) (\log(x) - \log(x_{b,i-1})) \tag{6}$$

*Department of Architecture and Civil Engineering, University of Bath, Bath BA2 7AY, UK.
†Corresponding author, gjm36@bath.ac.uk

Using the Heaviside function $H()$, the probability of any x can be written as:

$$\log(p) = \log(\alpha_1) + \beta_1 \log(x) + \sum_{i=1}^{n-1} (\beta_{i+1} - \beta_i)(\log(x) - \log(x_{b,i})) H[(\log(x) - \log(x_{b,i}))] \tag{7}$$

which is a function that can easily be vectorised in a numerical algorithm.

3 Multiplication constant α_1

The unknown value of α_1 should be set so that the total probability under the multi-segment curve (sum of total probabilities for each segment) is equal to one:

$$1 = \sum_i \int \alpha_i x^{\beta_i} dx \tag{8}$$

Therefore α_1 can be expressed in terms of the β and x_b vectors. Define vector γ as the vector containing all the boundaries of all segments:

$$\gamma = [x_{min}, \mathbf{x}_b, x_{max}] \tag{9}$$

The value of integral functions for each segment i:

$$I_i = \int_{x=\gamma_i}^{\gamma_{i+1}} x^{\beta_i} dx = \begin{cases} \frac{\gamma_{i+1}^{\beta_i+1} - \gamma_i^{\beta_i+1}}{\beta_i+1} & \text{when } \beta_i \neq -1 \\ \log(\gamma_{i+1}) - \log(\gamma_i) & \text{when } \beta_i = -1 \end{cases} \tag{10}$$

We can write:

$$1 = \sum_i \alpha_i I_i \tag{11}$$

Given that:

$$\frac{\alpha_i}{\alpha_1} = \prod_{j=1}^{i-1} x_{b,j}^{\beta_j - \beta_{j+1}} \tag{12}$$

We can rewrite α_1 as:

$$\alpha_1^{-1} = \sum_{i=1}^{n} I_i \prod_{j=1}^{i-1} x_{b,j}^{\beta_j - \beta_{j+1}} \tag{13}$$