Rutherford scattering data

The following data is from Geiger and Marsden's 1913 paper where y is the number of scattering events.

θ	y
150	22.2
135	27.4
120	33.0
105	47.3
75	136
60	320
45	989
37.5	1760
30	5260
22.5	20300
15	105400

Let $d\sigma$ be the differential cross section for Rutherford scattering.

$$d\sigma = \frac{Z^2\alpha^2(\hbar c)^2}{4E^2(1-\cos\theta)^2}d\Omega$$

$d\sigma$ is an *unnormalized* probability mass function hence

$$\Pr(\theta=\theta_k) = f(k) = C\,d\sigma\Big|_{\theta=\theta_k}$$

where C is a normalization constant. Let C absorb the constants in $d\sigma$ and write

$$f(k) = \frac{C}{(1-\cos\theta_k)^2}$$

To find C let

$$x_i = \frac{1}{(1-\cos\theta_i)^2}$$

By total probability

$$C\sum_i x_i = 1$$

It follows that

$$C = \frac{1}{\sum_i x_i}$$

Hence the scattering probability for angle θ_k is

$$f(k) = \frac{x_k}{\sum_i x_i}$$

Predicted values $\hat{y}_k$ are computed as

$$\hat{y}_k = f(k)\sum_i y_i = \frac{x_k\sum_i y_i}{\sum_i x_i}$$

The following table shows the predicted values $\hat{y}$.

θ	y	$\hat{y}$
150	22.2	34.1
135	27.4	40.7
120	33.0	52.7
105	47.3	74.9
75	136	216
60	320	474
45	989	1383
37.5	1760	2778
30	5260	6608
22.5	20300	20471
15	105400	102162

The coefficient of determination R^2 measures how well predicted values fit the data.

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} = 0.999$$

The result indicates that $d\sigma$ explains 99.9% of the variance in the data.