

The likelihood is

$$\mathcal{L} = \prod_{i=1, N} \{(1 - P_b)P(y_i, x_i|\text{'good' distribution}) + P_b P(y_i, x_i|\text{'bad' distribution})\} . \quad (1)$$

Here, $P(y_i, x_i|\text{'good' distribution})$ is the probability of obtaining the data point (y_i, x_i) , assuming that the point is drawn from the ‘good’ distribution (i.e. the straight line). Similarly for the ‘bad’ distribution. P_b is the probability that any given point belongs to the bad distribution.