DATA MINING 2
Maximum Likelihood Estimation

Riccardo Guidotti

a.a. 2020/2021

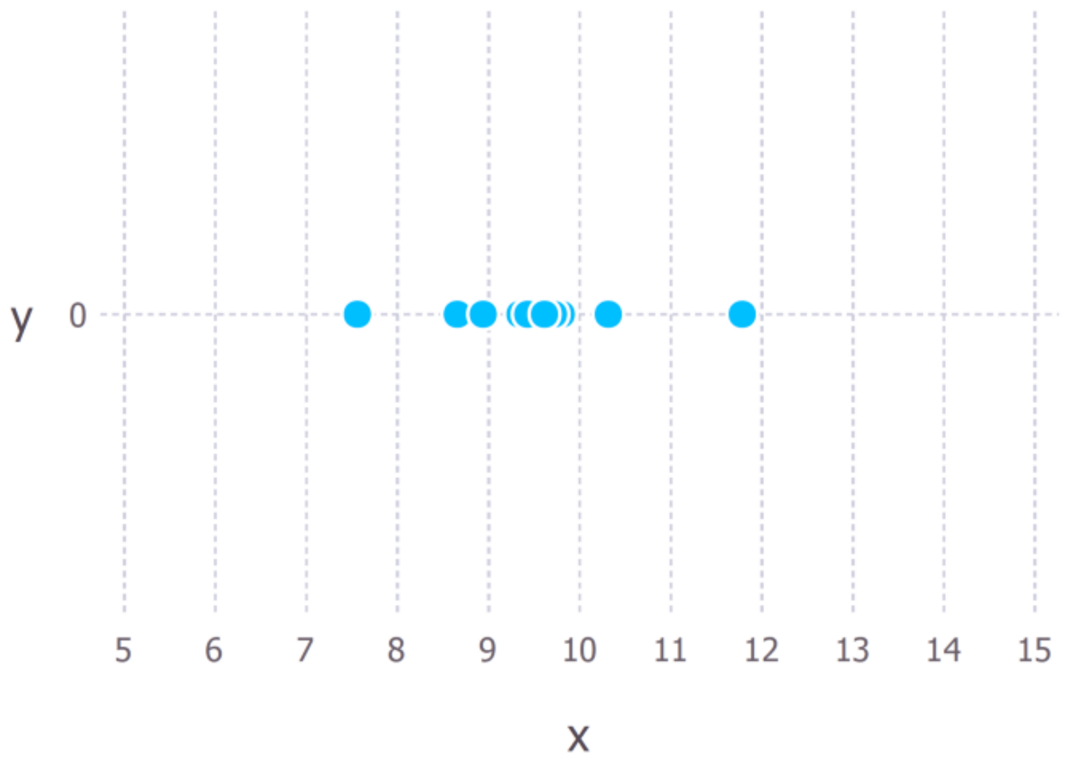
UNIVERSITÀ DI PISA

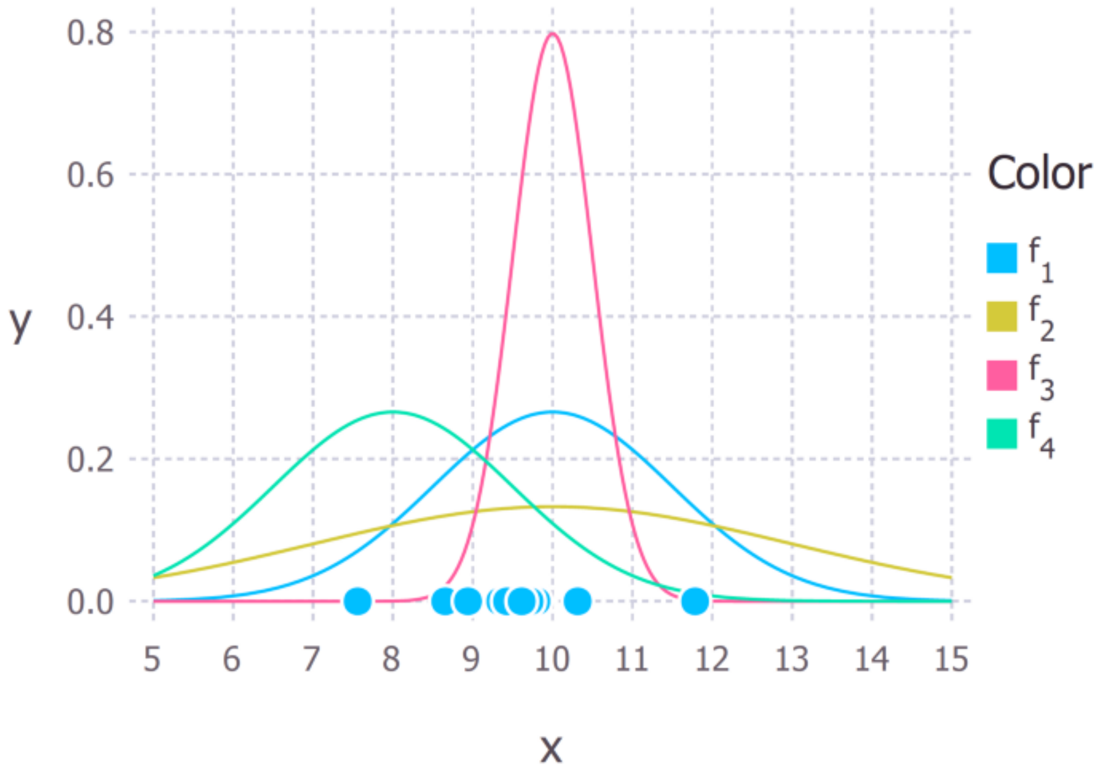
Intuition

- Maximum Likelihood Estimation (MLE) is a method that determines values for the parameters of a model.
- The parameter values are found such that they maximize the **likelihood** that the process described by the model produced the data that were actually observed.

Which model fit best?

- Normal Gaussian distribution
- Parameters: *mean* and *standard deviation*





Calculating the MLE

- Example: we have three data points 9, 9.5, 11
- We want to calculated the total probability of observing all the data, i.e. the joint probability distribution of all observed data points.
- Assumption: each data point is generated independently from the others.
- If the events are independent, then the total probability of observing all the data is the product of observing each data point individually (i.e. the product of the marginal probabilities).

Calculating the MLE

- Probability of observing a single data point *x*

$$P(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Parameters

- Example: $$P(9, 9.5, 11; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(9-\mu)^2}{2\sigma^2}\right) \times \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(9.5-\mu)^2}{2\sigma^2}\right) \times \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(11-\mu)^2}{2\sigma^2}\right)$$

The Log Likelihood

- Maximum is find by differentiation, i.e., find the derivative of the function w.r.t. a variable, set it to zero and find the required value.
- Since the previous expression is not easy to differentiate, we simplify the calculus considering the natural logarithm of the expression.

$$\ln(P(x;\mu,\sigma)) = \ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{(9-\mu)^2}{2\sigma^2} + \ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{(9.5-\mu)^2}{2\sigma^2} + \ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{(11-\mu)^2}{2\sigma^2}$$

$$\ln(P(x;\mu,\sigma)) = -3\ln(\sigma) - \frac{3}{2}\ln(2\pi) - \frac{1}{2\sigma^2}\left[(9-\mu)^2 + (9.5-\mu)^2 + (11-\mu)^2\right]$$

The Log Likelihood

- This expression can be easily differentiated to find the maximum.

$$\frac{\partial \ln(P(x;\mu,\sigma))}{\partial \mu} = \frac{1}{\sigma^2}\left[9 + 9.5 + 11 - 3\mu\right].$$

$$\mu = \frac{9 + 9.5 + 11}{3} = 9.833$$

- The same can be done for the standard deviation.