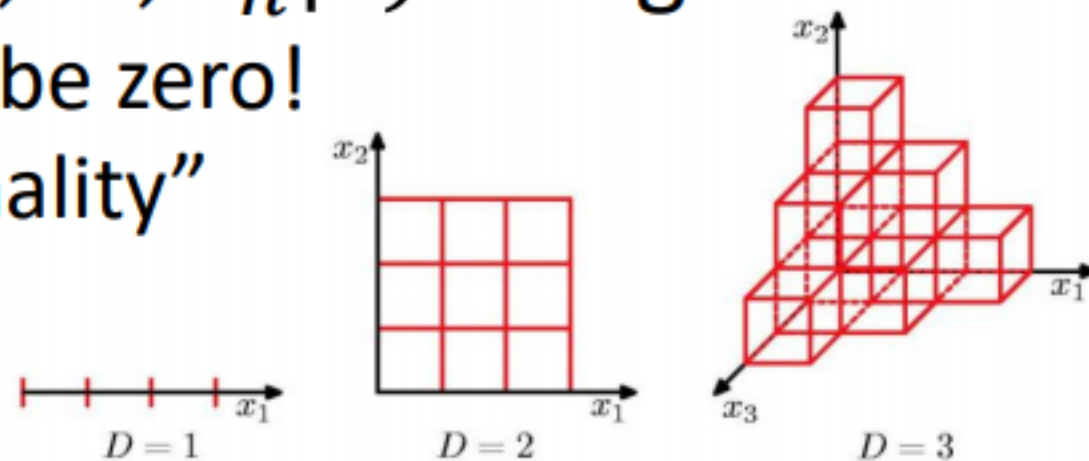# Naïve Bayes classifier: Why?

- Can't estimate $P(x_1, \ldots, x_n | c)$ using counts
  - Most counts would be zero!
  - "curse of dimensionality"

- What if $count(x_i = 1, c)$ is 0?
  - We would never assign class c to examples with $x_i = 1$
  - So use $P(x_i = 1 | c) \approx \dfrac{count(x_i = 1, c) + m\,\hat{p}}{count(c) + m}$

    Prior estimate for $P(x_i = 1 | c)$

  - m is a parameter
    - Interpretation: the number of "virtual" examples added to the training set