# Distance metric

- Acceptation region is defined by the distance metric
- Reasonable distance metrics depend on the data format, e.g.
  - Euclidean distance for low dimensional numerical outputs
  - L1 distance for data containing outliers
  - Quantiles or Wasserstein distance for distributions

# Data dimensionality

- Dimensionality of the data often causes problem when running LFI methods
  - In high dimensions it becomes increasing improbable to generate data close to the observed data
- The (current) standard approach is still to use summary statistics $S(\cdot)$
  - E.g $d(x, x^o) \approx d(S(x), S(x^o)) \quad (\approx \rho(x, x^o))$
- Sufficient statistics usually do not exist
- How to choose them?