Summary

- Expectation-Maximization (EM) algorithm
  - General algorithm for computing ML estimate of mixture models
  - Hill-climbing, so can only converge to a local maximum (depending on initial points)
- E-step: “augment” data by predicting values of useful hidden variables
- M-step: exploit the “augmented data” to improve estimate of parameters (“improve” is guaranteed in terms of likelihood)
- “Data augmentation” is probabilistic ➔ Split counts of events probabilistically