Classification with generative models

- Fit a distribution to each class separately
- Use Bayes’ rule to classify new data

What distribution to use? Are Gaussians enough?

Exponential families of distributions

We’ve described some distributions for one-dimensional data. What about higher dimensions?

1. Naive Bayes: Treat coordinates as independent.
   For $x = (x_1, \ldots, x_d)$, fit separate models $P(y_i)$ to each $x_i$, and assume
   \[
   \Pr(x_1, \ldots, x_d) = \Pr_1(x_1)\Pr_2(x_2) \cdots \Pr_d(x_d).
   \]
   This assumption is typically inaccurate.

2. Multivariate Gaussian.
   Model correlations between features: we’ve seen this in detail.

3. Graphical models.
   Arbitrary dependencies between coordinates.