Density Estimation

Many approaches make assumptions about the probability distribution of our data.

- naive Bayes $\quad p(x_1, x_2, \dots, x_D \mid \omega) = \prod_{i=1}^{D} p(x_i \mid \omega)$

Parametric Model

These are models $\qquad$ Parameters: μ, σ

$$p(X=x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

(1) How do we find this term?

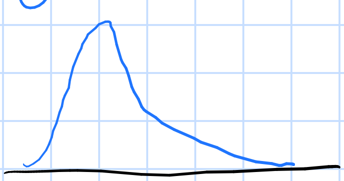
$$p(Y \mid X) = \frac{p(X \mid y)\, p(y)}{\underbrace{P(X)}_{\text{Total probability Thm}}}$$

$p(y)$ → Easy to estimate

(2) How good of a fit is our model to the data?

- KS - test

(1)



(2)

Maximum Likelihood Estimator

$$\hat{\theta} = \arg\max_{\theta \in \Phi} \{ p(D \mid \theta) \} = \arg\max_{\theta \in \Phi} \{ \log p(D \mid \theta) \}$$

- Choose the $\hat{\theta}$ that best describes the data.
- Why do we take the log probability?

$$p(D \mid \theta) = \prod_{i=1}^{n} p(x_i \mid \theta) \Rightarrow \log p(D \mid \theta) = \sum_{i=1}^{n} \log p(x_i \mid \theta)$$

Maximize the LL

$$\nabla_\theta L(\theta) = \sum_{i=1}^{n} \nabla_\theta \log p(x_i|\theta) = 0$$

$n = 75$

ex/ Gaussian distribution $\theta = \mu$, Σ is known

$$\log p(x_i|\theta) = -\frac{1}{2}\log\{(2\pi)^d |\Sigma|\} - \frac{1}{2}(x_i - \theta)^T \Sigma^{-1}(x_i - \theta)$$

μ

$$\nabla_\theta \log p(x_i|\theta) = 0 = \Sigma^{-1}(x_i - \theta)$$

$$\sum_{i=1}^{n} \Sigma^{-1}(x_i - \theta) = 0$$

$$\Sigma^{-1}\left(\sum_{i=1}^{n}(x_i - \theta)\right) = 0$$

then $\theta = \frac{1}{n}\sum_{i=1}^{n} x_i$

Bayesian Estimation

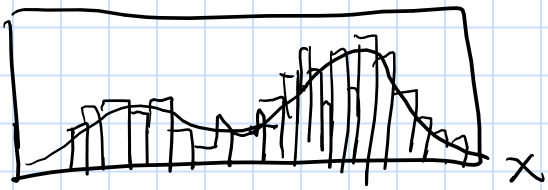
$$\theta_{BE} = \arg\max\{p(\theta|D)\} = \arg\max\{p(\theta)p(D|\theta)\}$$

Nonparametric Density Estimation

Problems w/ parametric models

- Assume the form of the probability distribution
- Most distributions are unimodal

Density Estimation 101



$$P_i = \frac{n_i}{N \Delta i}$$

$$P = \int_R p(x')\, dx'$$

↳ smoothed or averaged version of $p(x)$

$$\int_R p(x')\, dx' \simeq p(x) \cdot V$$

$x \to$ point in R

$V \to$ volume enclosed by the region R

$k \to$ # of samples that fall in a region R

$$\Rightarrow \boxed{p_n(x) = \frac{k/n}{V}}$$

$\mathbb{E}[k] = n\,p(x)$

$V \to 0$?

$n \to \infty$?