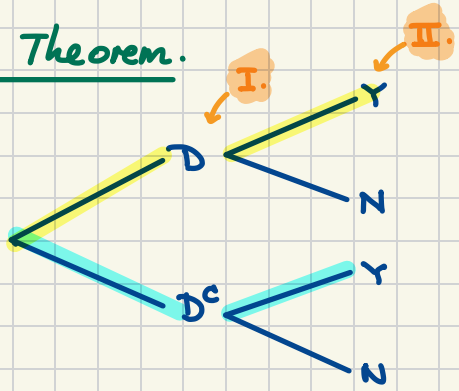
Bayes Theorem.

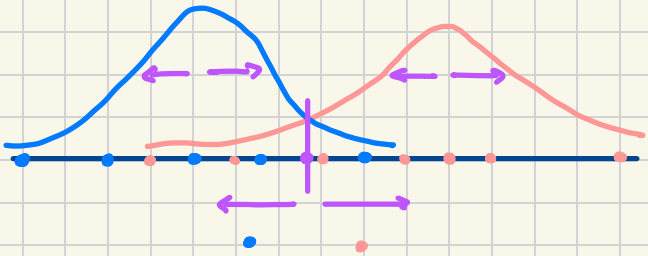


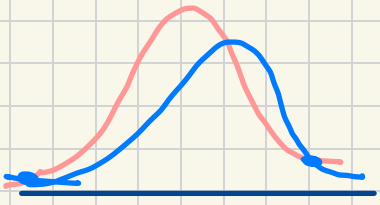
In m362k:

$$\mathbb{P}[D \mid Y] = ?$$

$$\mathbb{P}[D \mid Y] = \frac{\mathbb{P}[D \cap Y]}{\mathbb{P}[Y]} = \frac{\mathbb{P}[D] \cdot \mathbb{P}[Y \mid D]}{\mathbb{P}[D] \cdot \mathbb{P}[Y \mid D] + \mathbb{P}[D^c] \cdot \mathbb{P}[Y \mid D]}$$

Linear Discriminant Analysis.





Linear Discriminant Analysis w/ $p=1$.

Fisher 1936.

Goal: Classify observations into 1 of K classes ($K \geq 2$), i.e., figure out

$$p_k(x) = \mathbb{P}[Y=k \mid X=x]$$ posterior probability

Environment:

- π_k ... prior probability that a randomly chosen observation falls into category $k = 1..K$
- $f_k(x)$ ... density fction of X for observations from class $k = 1..K$

choice, model

$f_k(x)dx$ ... the probability that X fallso into $(x, x+dx)$ for points in class $k = 1..K$

Then,

Bayes Theorem

$$\mathbb{P}[Y=k \mid X=x] = \frac{\mathbb{P}[Y=k] \cdot \mathbb{P}[X=x \mid Y=k]}{\mathbb{P}[X=x]}$$

$\mathbb{P}[X=x \mid Y=k]$: $x \in (x, x+dx)$, $f_k(x)dx$

$\mathbb{P}[X=x]$: The Law of Total Probability

$$\Rightarrow \quad p_k(x) = \frac{\pi_k \cdot f_k(x)}{\sum_{l=1}^{K} \pi_l \cdot f_l(x)}$$

$$\Rightarrow \text{Classify into } k = \underset{l=1..K}{\operatorname{argmax}} \{p_l(x)\}$$

The Specifics of LDA ($p=1$).

The Choice: f_k are normal densities for each $k = 1..K$, i.e.,

$$f_k(x) = \frac{1}{\sigma_k \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_k}{\sigma_k}\right)^2} \quad \text{for all } x \in \mathbb{R}$$

w/ μ_k being the mean and
σ_k the std deviation for class k

Additional Assumption: Homogeneity: $\sigma_k = \sigma$, $k = 1..K$

$\Rightarrow$

$$p_k(x) = \frac{\pi_k \cdot f_k(x)}{\sum_{l=1}^{K} \pi_l f_l(x)}$$

Remember: We're looking for the k for which its $p_k(x)$ is MAXIMAL over all k.