Continuous Valued data & Bayes Rule

So far, we've been seeing categorical data (features). what if we have continuous data? Sometimes situation warrant using alternate techniques: → we assume normal/gaussian distribution, → then we can use pdf [p. density fn] to obtain a value proportional to probability $p(x|y)$

→ Norm pdf with x as single var

$$p(x|y) \propto \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}}$$

→ x is vctr of features, we need multivari form

$$p(x|\mu,\Sigma) \propto \frac{1}{\sqrt{(2\pi)^D|\Sigma|}} e^{\frac{(x-\mu)^T\Sigma^{-1}(x-\mu)}{2}}$$

D = dimentionality of data
Σ = cov mat of data
$|\Sigma|$ = determinant
μ = vctr of means of features

Bayes law:

1) we calculate (μ,σ) for a particular class/feature.
2) we thus get $p(x|y)$ we want $p(y|x)$
3) we use Bayes law!:

$$p(y|x) = \frac{p(y)\,p(x|y)}{p(x)}$$

$p(y|x)$: posterior
$p(y)$: the prior [p of y in general] ∴ of samples that hav y → get directly from data
$p(x|y)$: likelihood of generating x given y
$p(x)$: evidence (class independent)

Naive Bayes Prob:

→ make an assumption: all features are conditionally independent (no dependency in features)

$$\prod_{k=1}^{D} p(f_k = x_k \mid y = i)$$

(one more weird formula. check slide 26)

class independent ∴ ignore it

- no denominator as we're approximating the numerator based on the conditionally independent assumption
- overall Bayes rule doesn't hold