Laplace Additive Smoothing

-> There are two reasons why you use laplace additive smoothing

(1) Because you do not want values to be zero

(2) you can adjust bias-variance tradeoff because of $n\alpha$ at denomingator.

$$P(\text{feature}|\text{class}) = \frac{\text{count} + \alpha}{n + \alpha \cdot K}$$

when α value is zero or close to zero, there is lot of variance in output

$P(y/x) = P(y)\, P(f_1|y)\, P(f_2|y)\, P(f_3|y) \ldots$

let say its, $\frac{0}{500}$ for first case i.e 0

but for another sample, it is $\frac{1}{500}$.

So even with small difference in input, there is high change in output (high variance)

when value of α is high,

let say $\alpha = 1000$ $\quad \frac{1+1000}{500+2000} = \frac{1001}{2500} \approx \frac{2}{5}$

$\alpha = 100000$, $\quad \frac{1+10000}{500+20000} \approx \frac{1}{2}$

as you increase value of α, it approches to $1/2$, high bias. As you reduce the value of α, it approaches to high variance