$$\prod_j p(x^{(j)} | y) \, p(y)$$

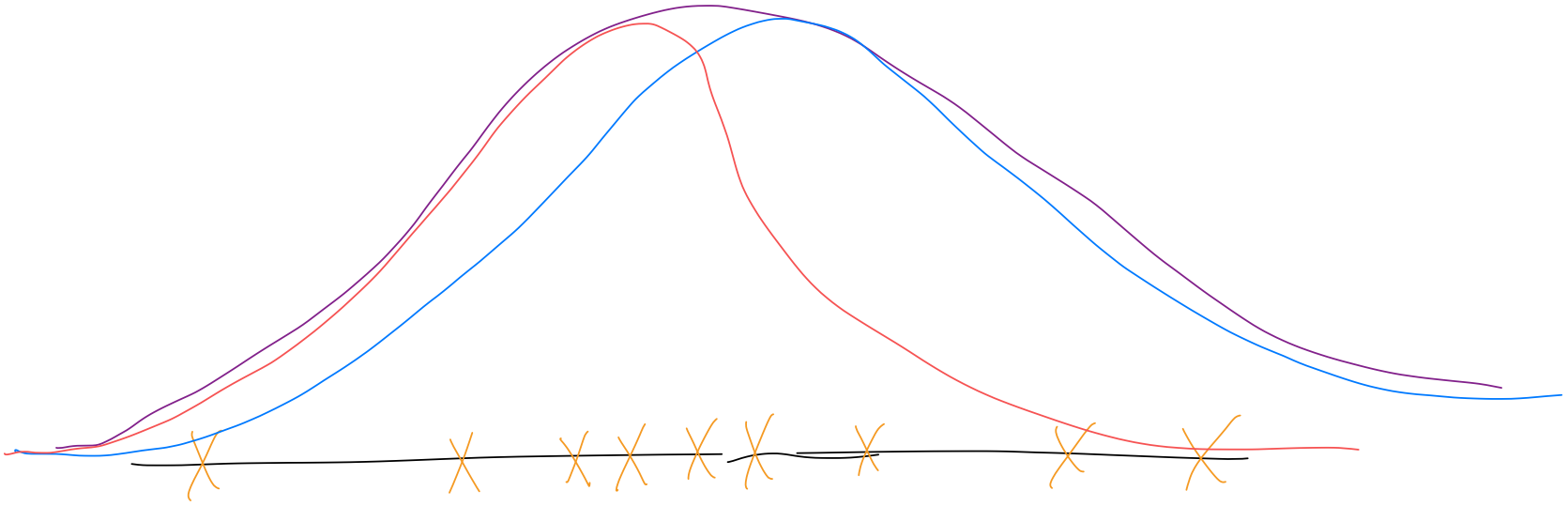
$$\log(p(y)) + \sum_j \log(p(x^{(j)} | y))$$

$$= \frac{\#\text{ of times } j^{th} \text{ attribute was } b \text{ and } y \text{ was } c + 1}{\#\text{ of samples with class } c + \#\text{ of choices for feature } j}$$

$p(\text{sunny} \mid \text{golf})$

$p(\text{cloudy} \mid \text{golf})$

$p(\text{rainy} \mid \text{golf})$



$N(\mu, \sigma^2)$

likelihood

$p(D \mid \omega_1, \ldots \omega_m)$

(

MLE: maximum likelihood estimation

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

$$p(D \mid \mu, \sigma) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \prod_{i=1}^{N} e^{\frac{-(x_i-\mu)^2}{2\sigma^2}}$$

$$3x^2 + 4y + 6xy + 10$$

$$\frac{\partial}{\partial x} = 6x + 0 + 6y$$

$$\mathcal{L}(\mu) = p(D|\mu,\sigma) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N \prod_{i=1}^{N} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$\ln \mathcal{L}(\mu) = \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N + \sum_{i=1}^{N} \ln\left(e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}\right)$$

$$\ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^N + \sum_{i=1}^{N} -\frac{(x_i-\mu)^2}{2\sigma^2}$$

$$\partial/\partial\mu \ln \mathcal{L}(\mu) = 0 + \sum_{i=1}^{N} -\frac{2(x_i-\mu)(-1)}{2\sigma^2}$$

$$0 = \sum_{i=1}^{N} \frac{-2(x_i - \mu)(-1)}{2\sigma^2}$$

$$0 = \frac{2}{2\sigma^2} \sum (x_i - \mu)$$

$$0 = \frac{1}{\sigma^2} \sum_i (x_i - \mu)$$

$$0 = \sum_i^N (x_i - \mu)$$

$$= -N\mu + \sum_{i=1}^{N} x_i$$

$$N\mu = \sum_{i=1}^{N} x_i \quad \Rightarrow \mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

$$\log \mathcal{L}(\sigma) = \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N + \sum_{i=1}^{N} \frac{-(x_i-\mu)^2}{2\sigma^2}$$

$$= N\log\left(\frac{1}{\sigma}\cdot\frac{1}{\sqrt{2\pi}}\right) + \text{''}$$

$$= -N\log(\sigma) - N\left(\log\left(\sqrt{2\pi}\right)\right) + \text{''}$$

$$= -N\log(\sigma) - N\left(\log\left(\sqrt{2\pi}\right)\right) + \sum_{i=1}^{N} -(x_i-\mu)^2\left(\tfrac{1}{2}\right)\left(\sigma^{-2}\right)$$

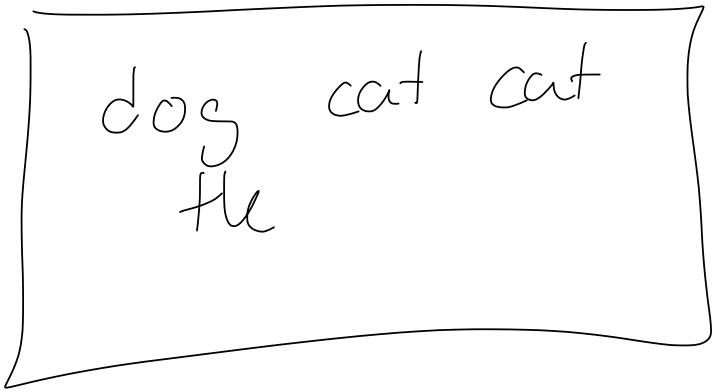
$$= -N\frac{1}{\sigma} + \frac{1}{2}\sum_{i=1}^{N} -(x_i-\mu)^2\left(-2\sigma^{-3}\right)$$

$$0 = -N\frac{1}{\sigma} - \frac{1}{\sigma^3}\sum_{i=1}^{N} -(x_i-\mu)^2$$

$$\frac{N}{\sigma} = \frac{1}{\sigma^3} \sum_{i=1}^{N} -(x_i - \mu)^2$$

$$N\sigma^2 = \sum_{i=1}^{N} -(x_i - \mu)^2$$

$$\sigma^2 = \frac{\sum_{i=1}^{N} -(x_i - \mu)^2}{N}$$



[cant word a , cant word b, cant word c, ...

...]

$$P(x \mid y = c) = \frac{\#\text{ of times this word occured in documents of class}}{\#\text{ of total word cnt for that class}}$$