(C) According to bayes rule $P(y|x) = \frac{P(x|y)\,P(y)}{P(x)}$

based on naive bayes rule $\Rightarrow \underbrace{P(y|x)}_{\text{Posterior}} = \underbrace{P(x|y)}_{\text{likelihood}} \overbrace{P(y)}^{\text{Prior}}$

if $P(y=i|x) > P(y=j|x) \Rightarrow \underline{y=i}$ else $\Rightarrow y=j$

Because we assume independance for all data incidence Then we can write the Probability for data X to be $y=1$ (spam) or $y=0$ (ham) can be written as:

$$P(y|X) = P(X,y) = P(x^1,y^1)\cdot P(X^2,y^2)\cdot\cdots\cdot P(X^m,y^m)$$

$\boxed{\text{m is total number of records of data}}$

$$= \prod_{i=1}^{m} P(x^i,y^i)$$

$\boxed{\text{Likelihood}}$

$$\overset{\text{naive}}{=} \prod_{i=1}^{m} P(X^i|y^i)\,P(y^i)$$

(why in next Page)*

$$= \prod_{i=1}^{m} P(y^i) \prod_{K=1}^{n} \theta_{c,K}^{x_K^i}$$

$\boxed{\text{we want to maximiz this equation. -this is constained optim}}$

for simplification we take the log (ln) from both sides. log is monotonic so the maximization result will hold

the constrained is: $\sum_{K=1}^{n} \theta_{c,K} = 1 \quad \forall c$