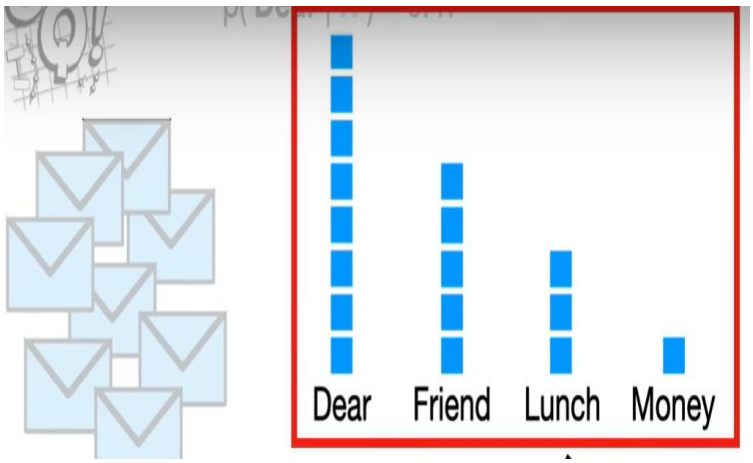
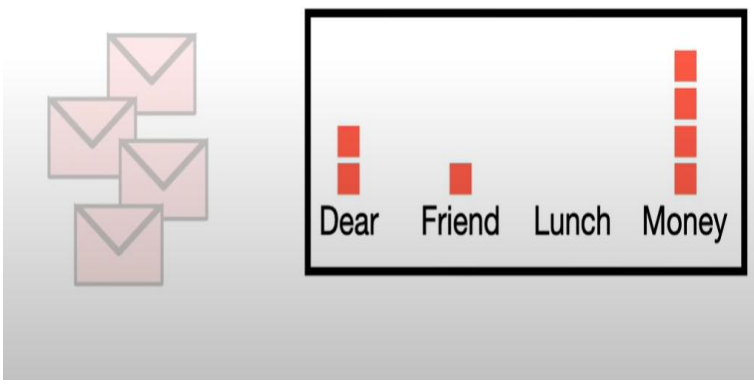



$$p(\textbf{Dear} \mid \textbf{N}) = 0.47$$
$$p(\textbf{Friend} \mid \textbf{N}) = 0.29$$
$$p(\textbf{Lunch} \mid \textbf{N}) = 0.18$$
$$p(\textbf{Money} \mid \textbf{N}) = 0.06$$

$$p(\textbf{Dear} \mid \textbf{S}) = 0.29$$
$$p(\textbf{Friend} \mid \textbf{S}) = 0.14$$
$$p(\textbf{Lunch} \mid \textbf{S}) = 0.00$$
$$p(\textbf{Money} \mid \textbf{S}) = 0.57$$

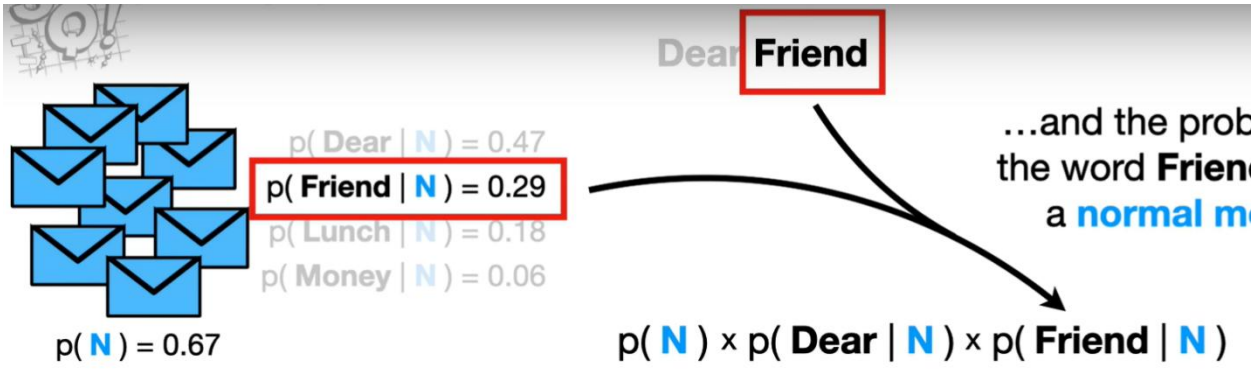
For example, since **8** of the **12** messages are **normal messages**, our initial guess will be **0.67**.

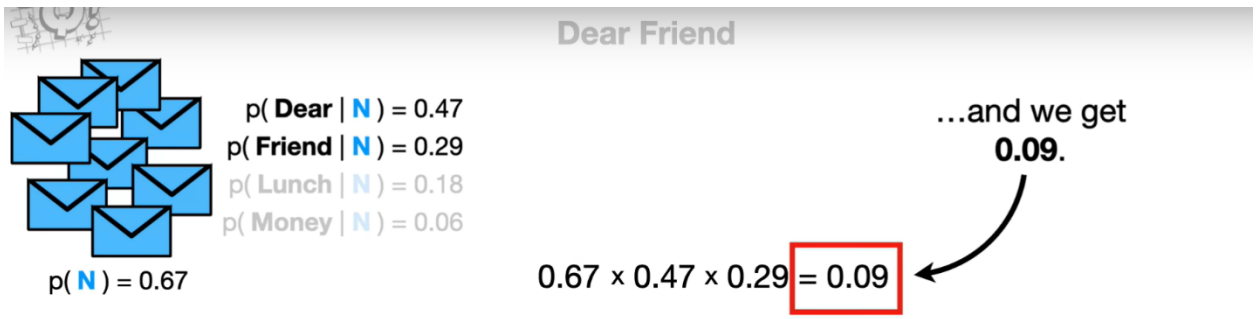
$$p(\text{N}) = \frac{8}{8+4} = 0.67$$

And since **4** of the **12** messages are **spam**, our initial guess will be **0.33**.

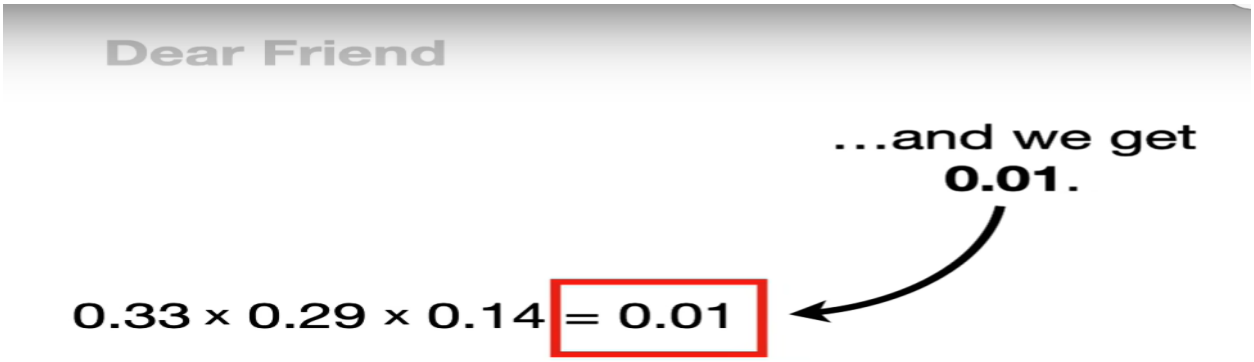
$$p(\text{S}) = \frac{4}{4+8} = 0.33$$

# Dear Friend



$$p(S) \times p(\text{Dear} \mid S) \times p(\text{Friend} \mid S)$$




We **eliminate the denominator** in Naive Bayes **because it's the same (constant) for all classes** — we only care which class gives a higher probability.