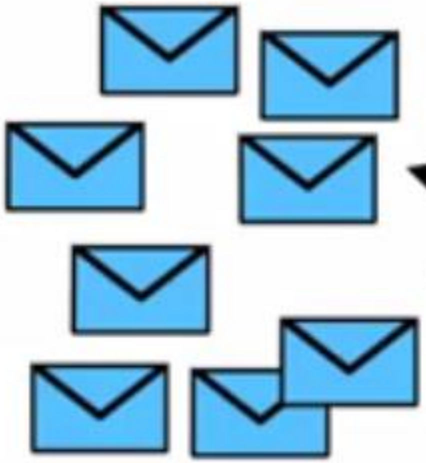
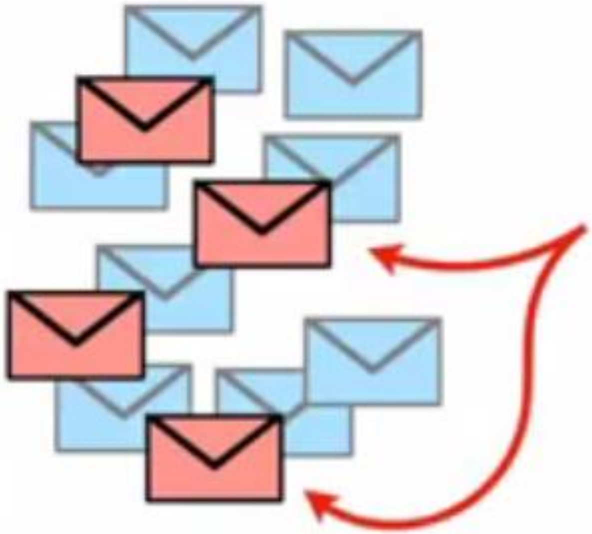
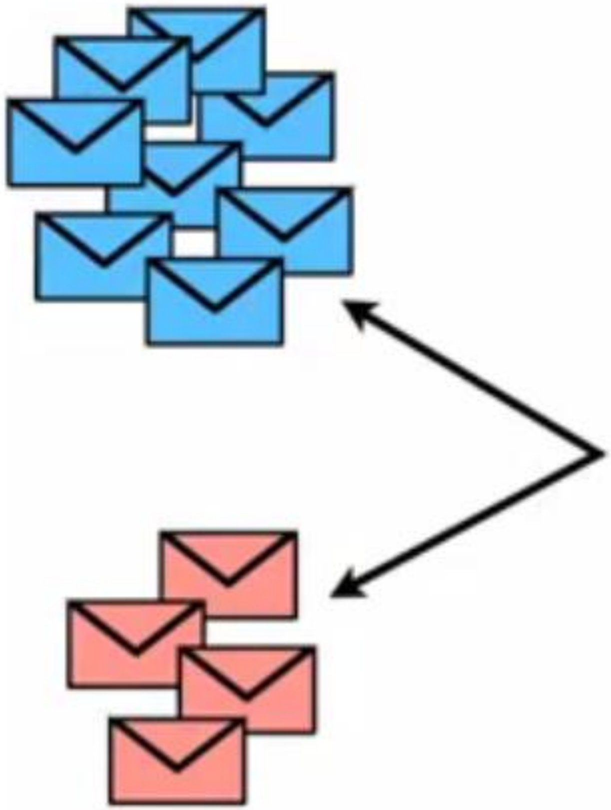
Naive Bayes



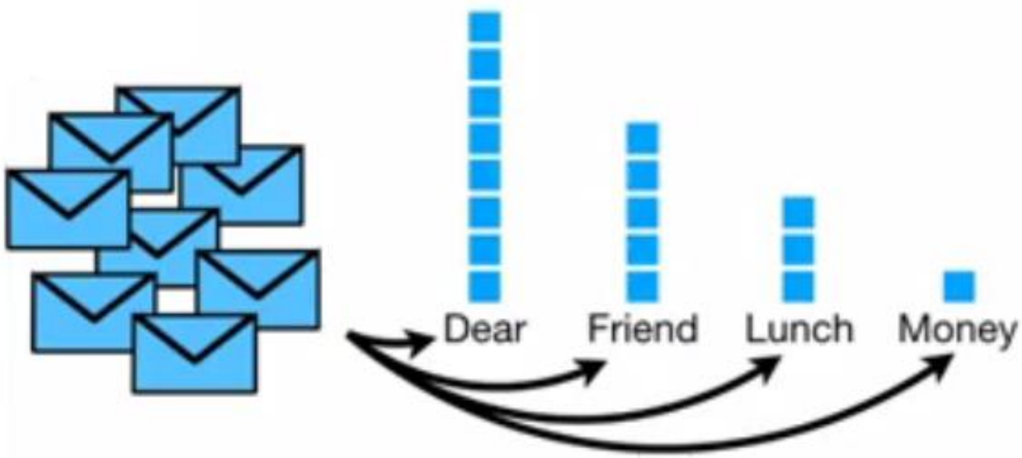
Imagine we received **normal messages** from friends and family…



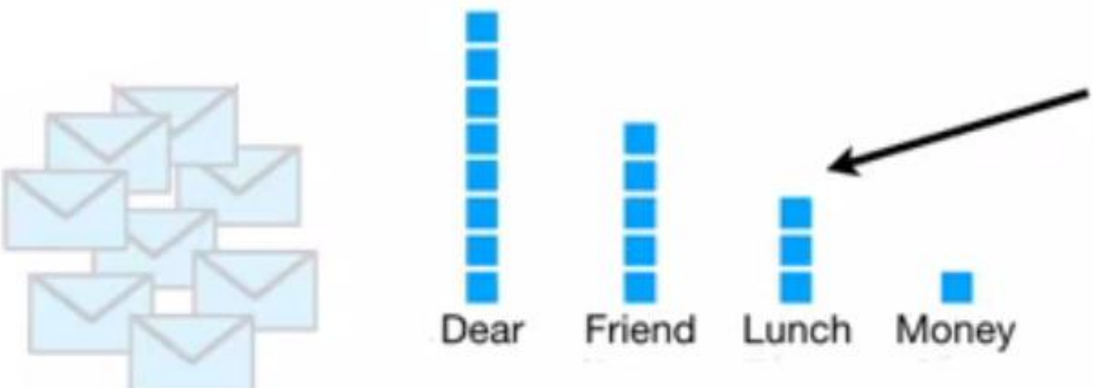
…and we also received **spam** (unwanted messages that are usually scams or unsolicited advertisements)…



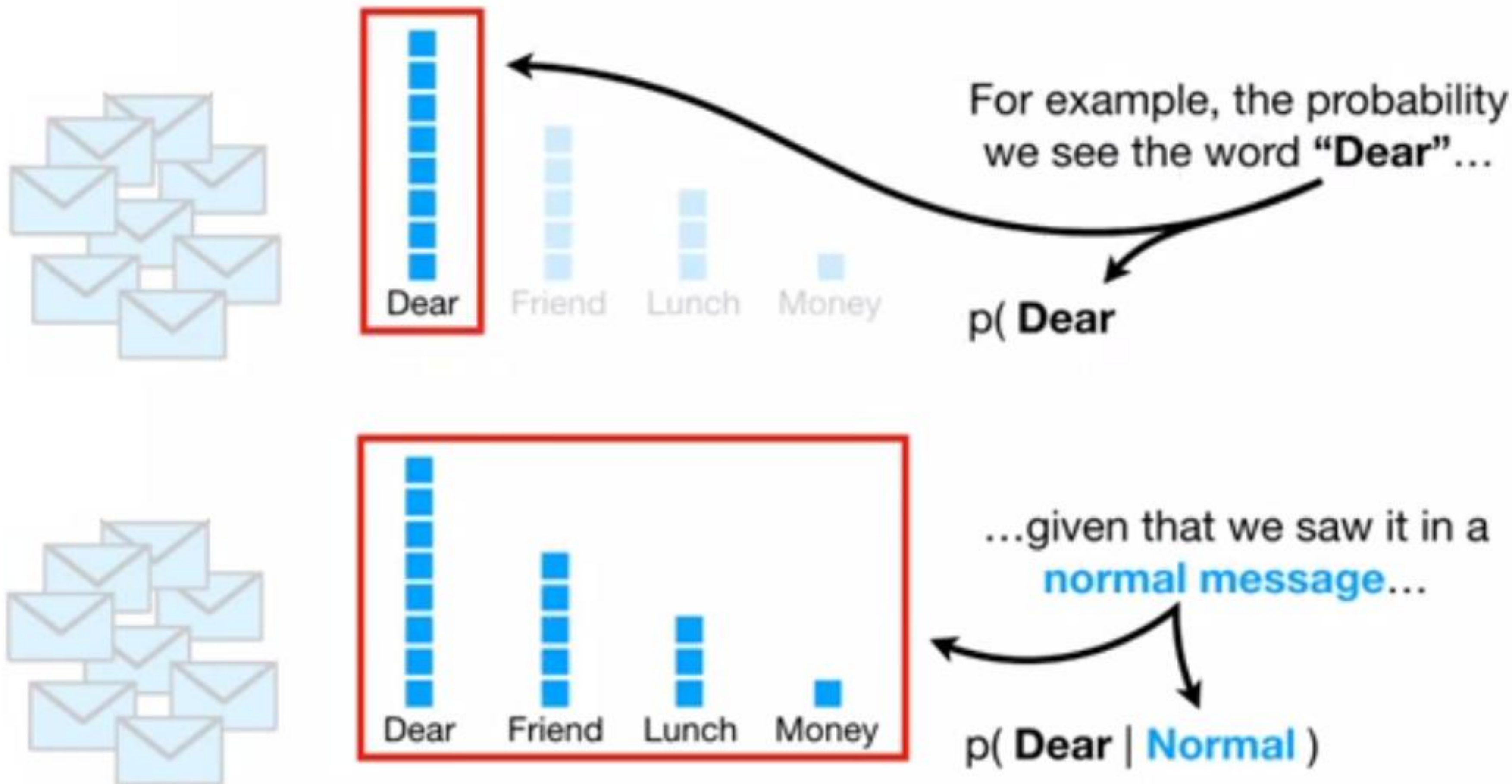
...and we wanted to filter out the **spam** messages.

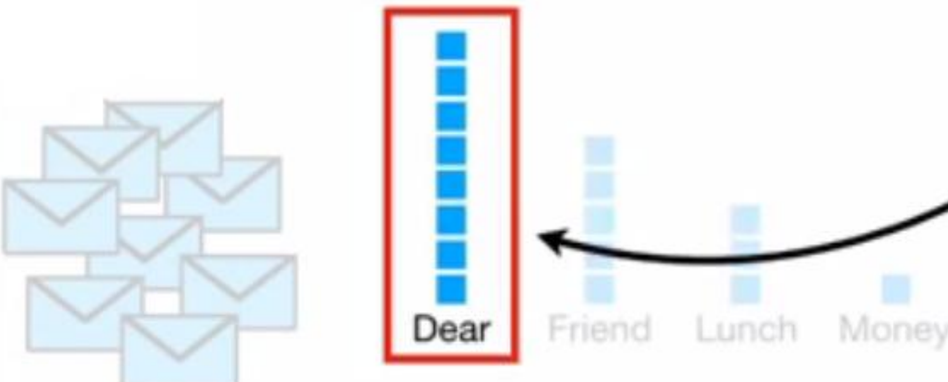


histogram of all the words that occur in the **normal messages** from friends and family.



We can use the histogram to calculate the probabilities of seeing each word, given that it was in a **normal message**.

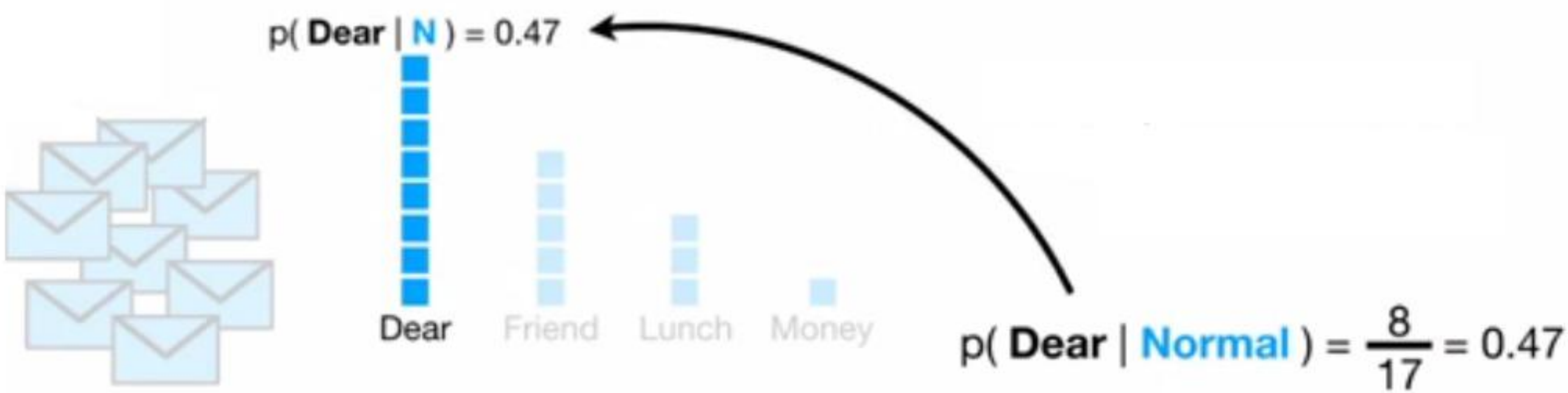

Dear
Friend
Lunch
Money
For example, the probability we see the word "Dear"...
p(Dear
Dear
Friend
Lunch
Money
...given that we saw it in a normal message...
p(Dear | Normal)

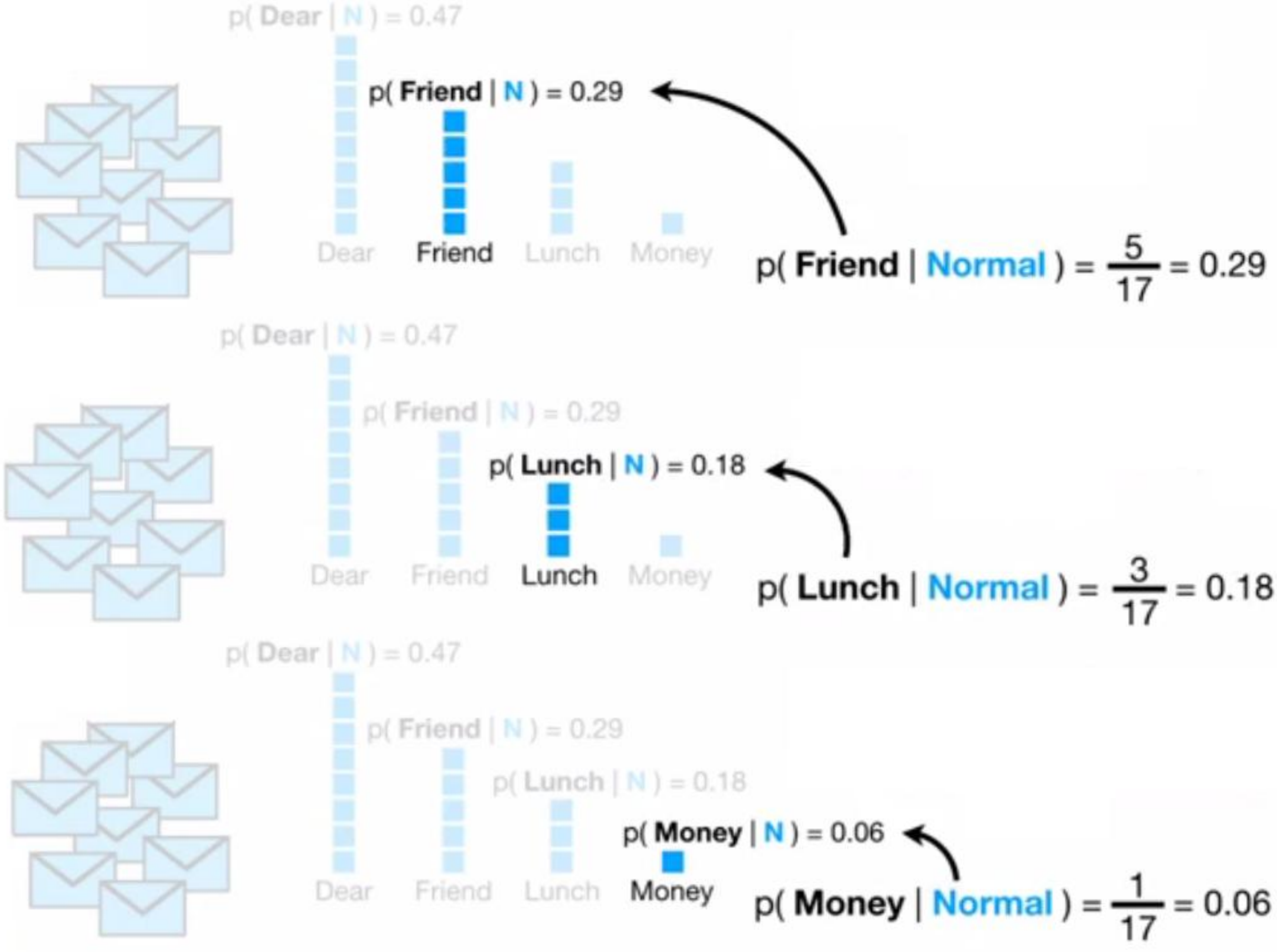


8, the total number of times **Dear** occurred in **normal messages**...

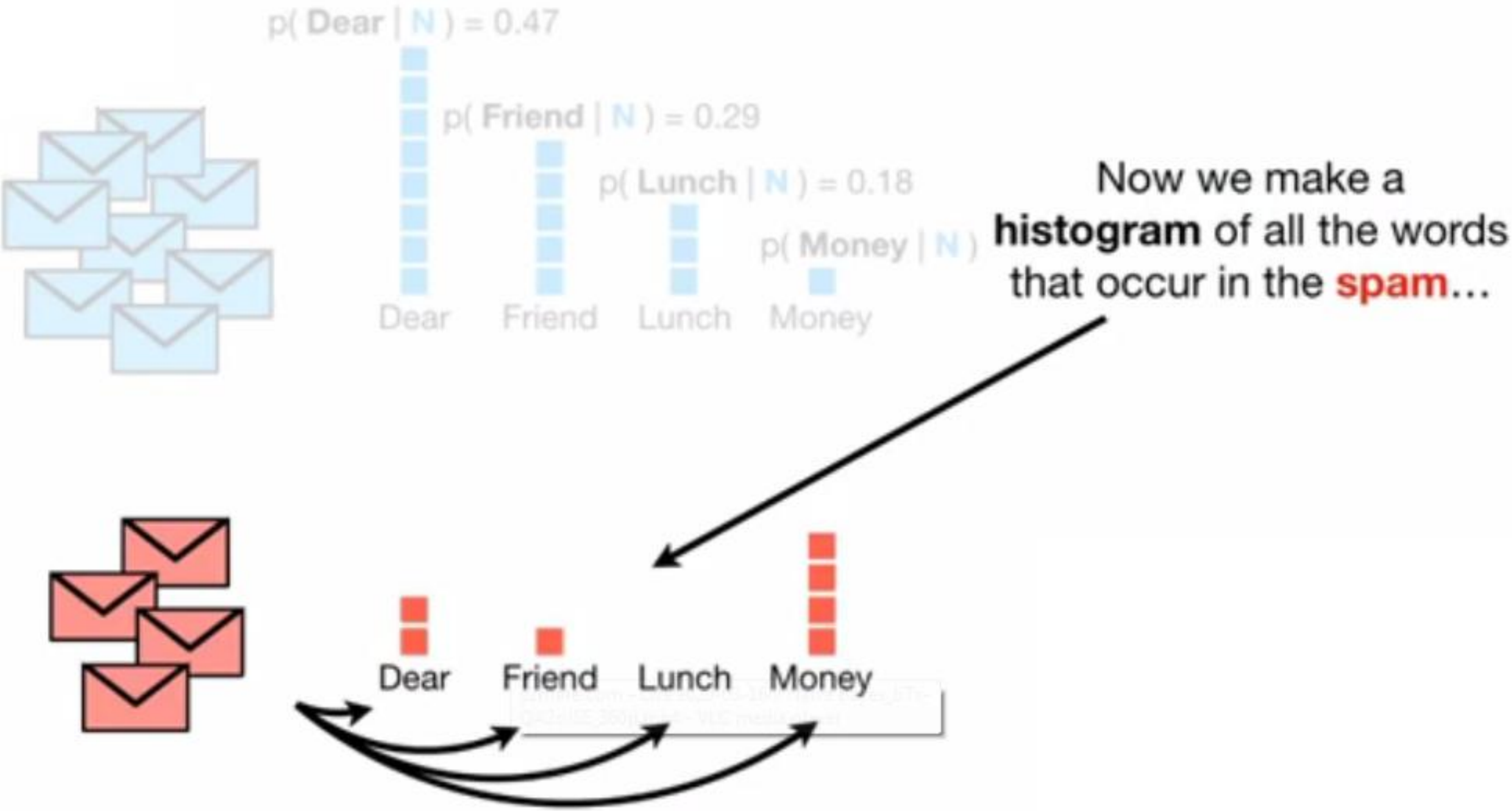
$$p(\textbf{Dear} \mid \textbf{Normal}) = \frac{8}{17}$$

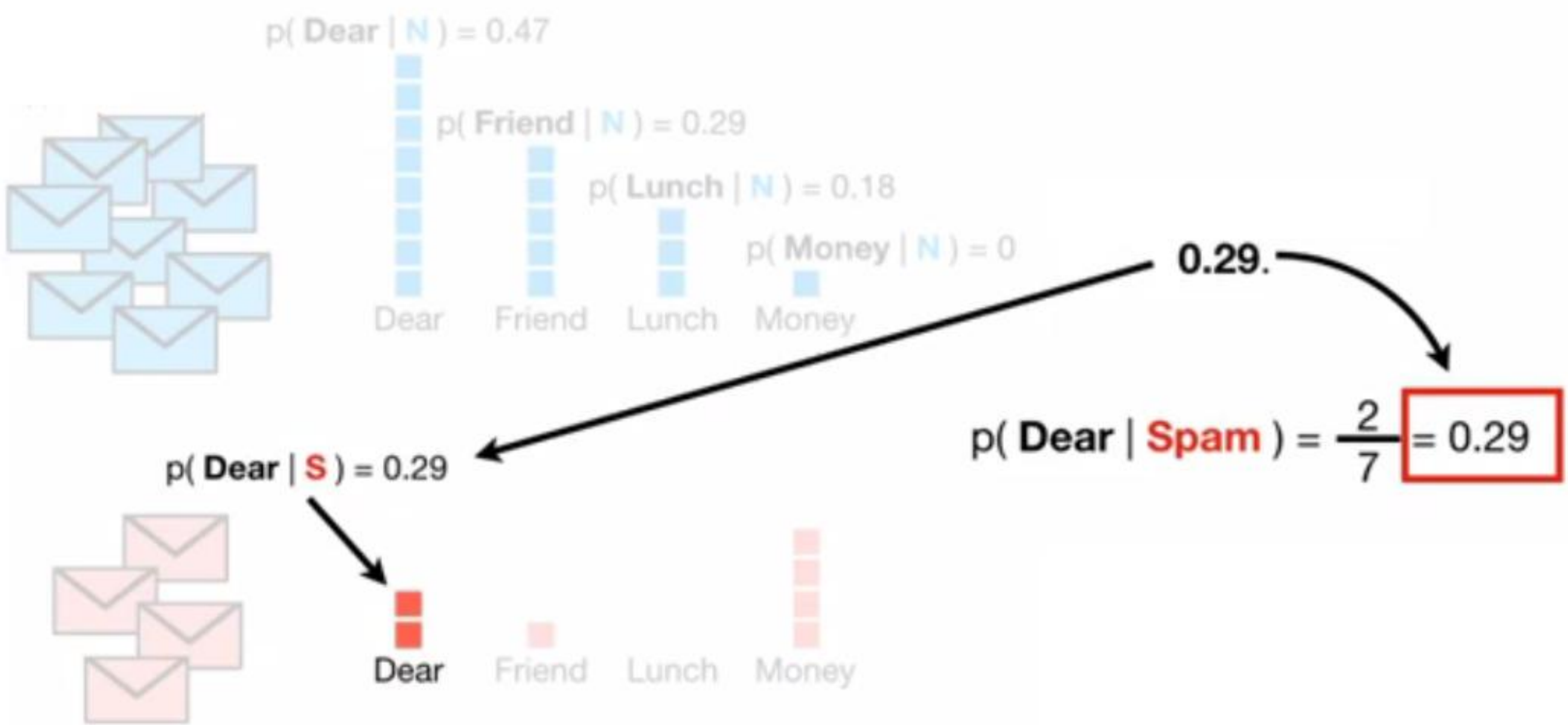
divided by **17**, the total number of words in all of the **normal messages**.


p(Dear | N) = 0.47
Dear
Friend
Lunch
Money
p(Dear | Normal) = $\frac{8}{17}$ = 0.47

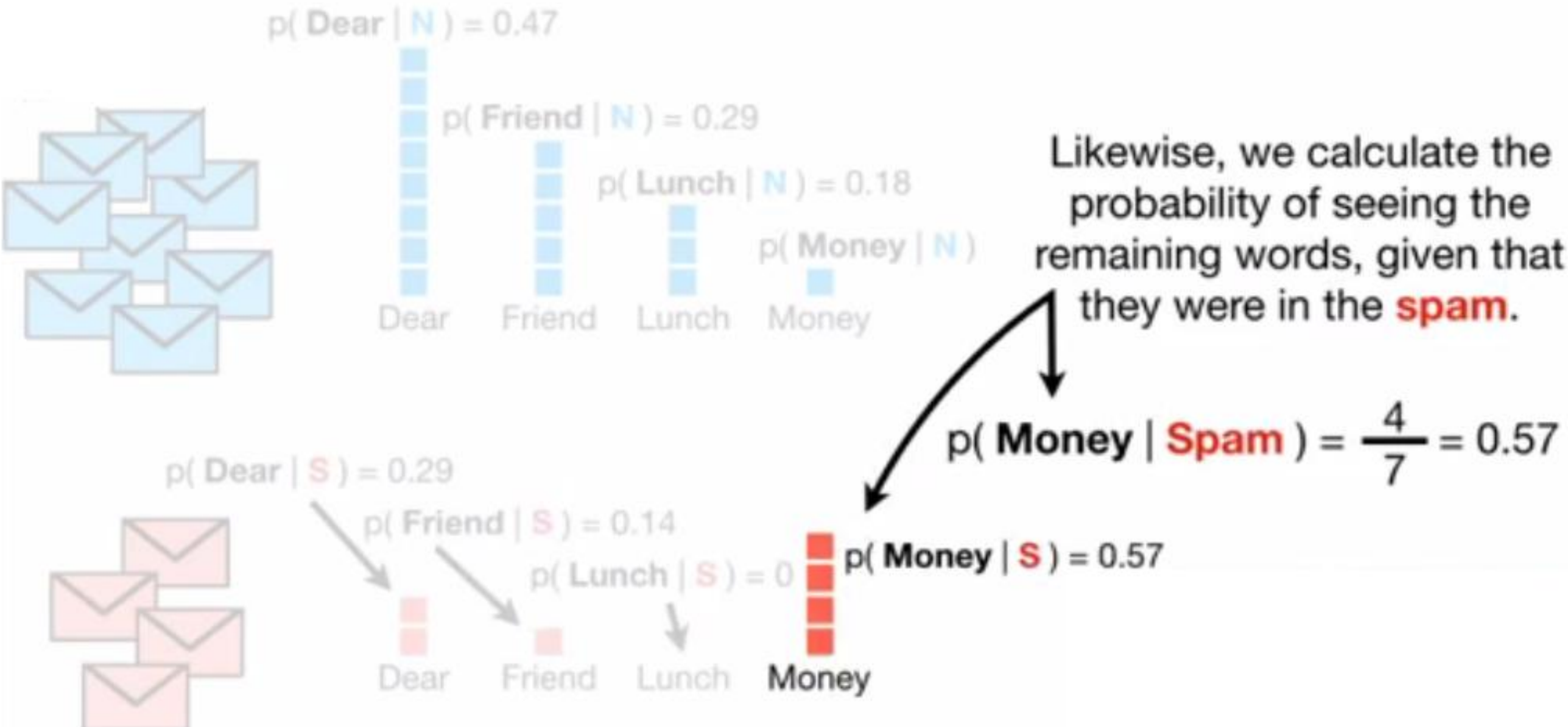


p(Dear | N) = 0.47
p(Friend | N) = 0.29
Dear
Friend
Lunch
Money
p(Friend | Normal) = 5/17 = 0.29
p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
Dear
Friend
Lunch
Money
p(Lunch | Normal) = 3/17 = 0.18
p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N) = 0.06
Dear
Friend
Lunch
Money
p(Money | Normal) = 1/17 = 0.06


p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N)
Dear
Friend
Lunch
Money
Now we make a **histogram** of all the words that occur in the **spam**…
Dear
Friend
Lunch
Money

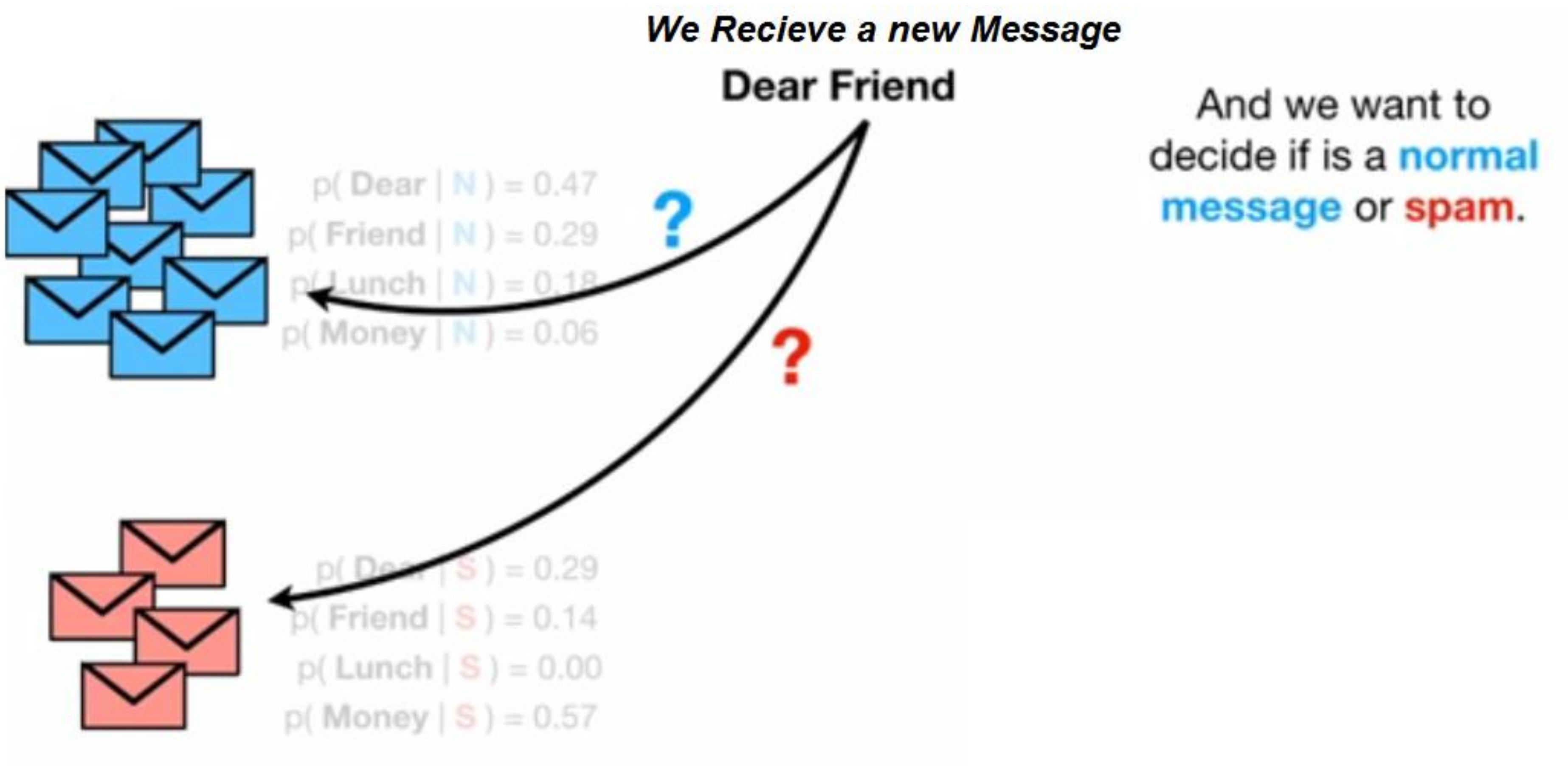


p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N) = 0
Dear Friend Lunch Money
0.29.
p(Dear | Spam) = 2/7 = 0.29
p(Dear | S) = 0.29
Dear Friend Lunch Money

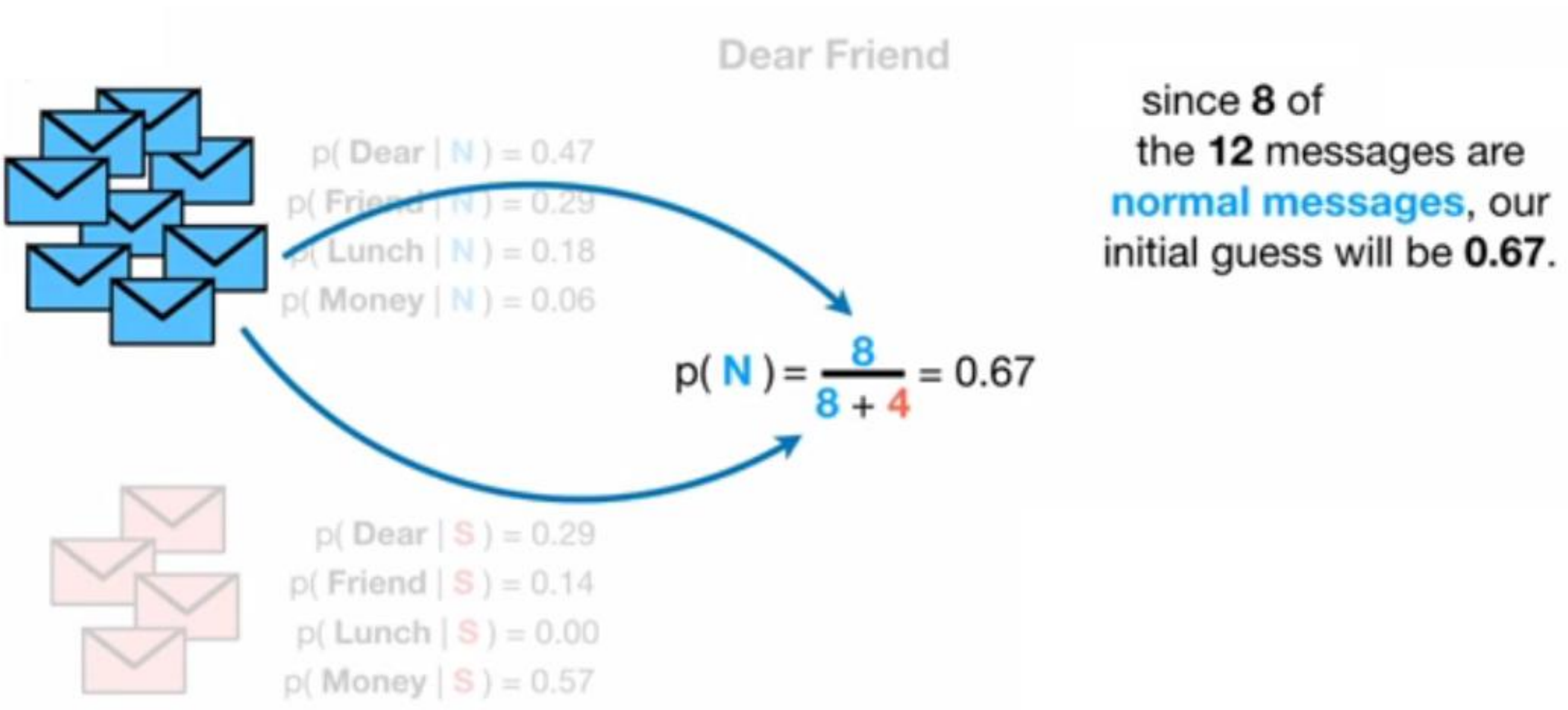

p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N)
Dear
Friend
Lunch
Money
Likewise, we calculate the probability of seeing the remaining words, given that they were in the spam.
p(Money | Spam) = 4/7 = 0.57
p(Dear | S) = 0.29
p(Friend | S) = 0.14
p(Lunch | S) = 0
p(Money | S) = 0.57
Dear
Friend
Lunch
Money

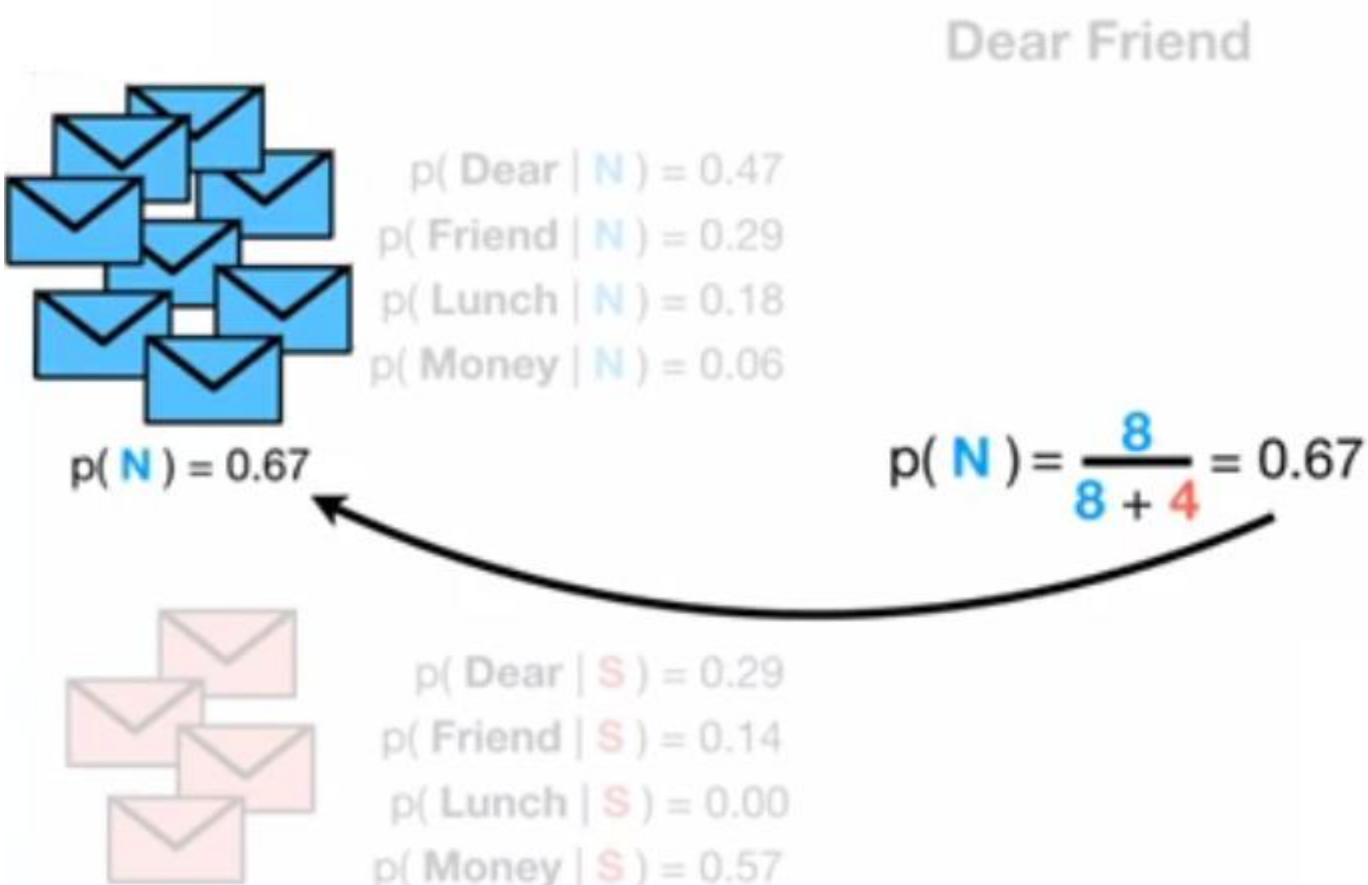
p(**Dear** | N) = 0.47
p(**Friend** | N) = 0.29
p(**Lunch** | N) = 0.18
p(**Money** | N) = 0.06

p(**Dear** | S) = 0.29
p(**Friend** | S) = 0.14
p(**Lunch** | S) = 0.00
p(**Money** | S) = 0.57



We Recieve a new Message
Dear Friend
p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N) = 0.06
?
?
p(Dear | S) = 0.29
p(Friend | S) = 0.14
p(Lunch | S) = 0.00
p(Money | S) = 0.57
And we want to decide if is a normal message or spam.


Dear Friend
p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N) = 0.06
p(Dear | S) = 0.29
p(Friend | S) = 0.14
p(Lunch | S) = 0.00
p(Money | S) = 0.57
p(N) = 8 / (8 + 4) = 0.67
since 8 of the 12 messages are normal messages, our initial guess will be 0.67.



Dear Friend
p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N) = 0.06
p(N) = 0.67
p(N) = 8 / (8 + 4) = 0.67
p(Dear | S) = 0.29
p(Friend | S) = 0.14
p(Lunch | S) = 0.00
p(Money | S) = 0.57

Now we multiply that initial guess by probability that the word **Dear** occurs in a **normal message**...

...and the probability that the word **Friend** occurs in a **normal message**.

p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N) = 0.06

p(N) = 0.67

$$p(\,N\,) \times p(\,\textbf{Dear} \mid N\,) \times p(\,\textbf{Friend} \mid N\,)$$

p(Dear | S) = 0.29
p(Friend | S) = 0.14
p(Lunch | S) = 0.00
p(Money | S) = 0.57

p(**Dear** | N) = 0.47
p(**Friend** | N) = 0.29
p(**Lunch** | N) = 0.18
p(**Money** | N) = 0.06

p(N) = 0.67

p(**Dear** | S) = 0.29
p(**Friend** | S) = 0.14
p(**Lunch** | S) = 0.00
p(**Money** | S) = 0.57

Now we multiply that initial guess by probability that the word **Dear** occurs in a **normal message**...

...and the probability that the word **Friend** occurs in a **normal message**.

p(N) × p(**Dear** | N) × p(**Friend** | N)

0.67 × 0.47 × 0.29

Now we multiply that initial guess by probability that the word **Dear** occurs in a **normal message**...

...and the probability that the word **Friend** occurs in a **normal message**.

p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N) = 0.06

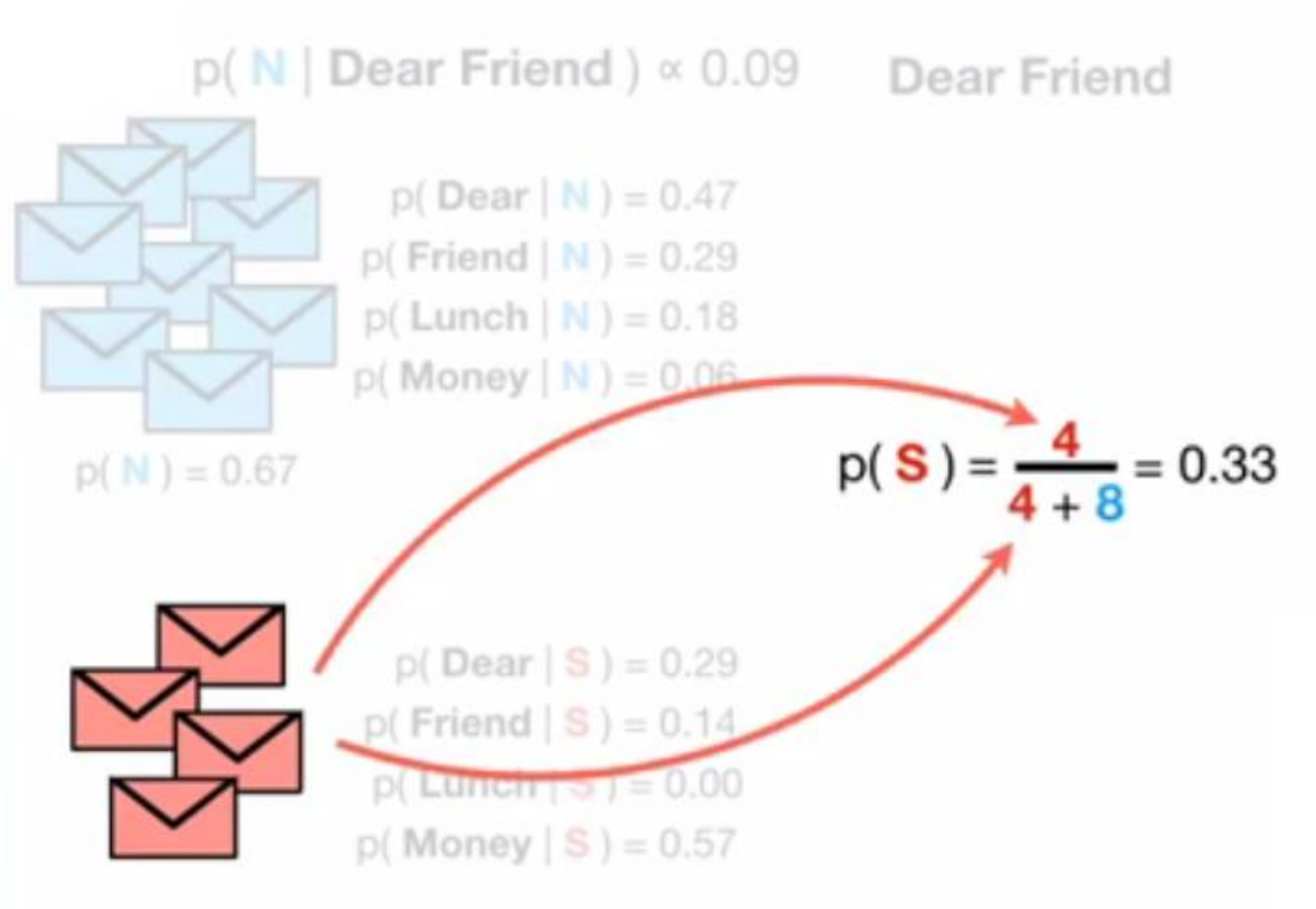
p(N) = 0.67

p(**N**) × p(**Dear** | **N**) × p(**Friend** | **N**)

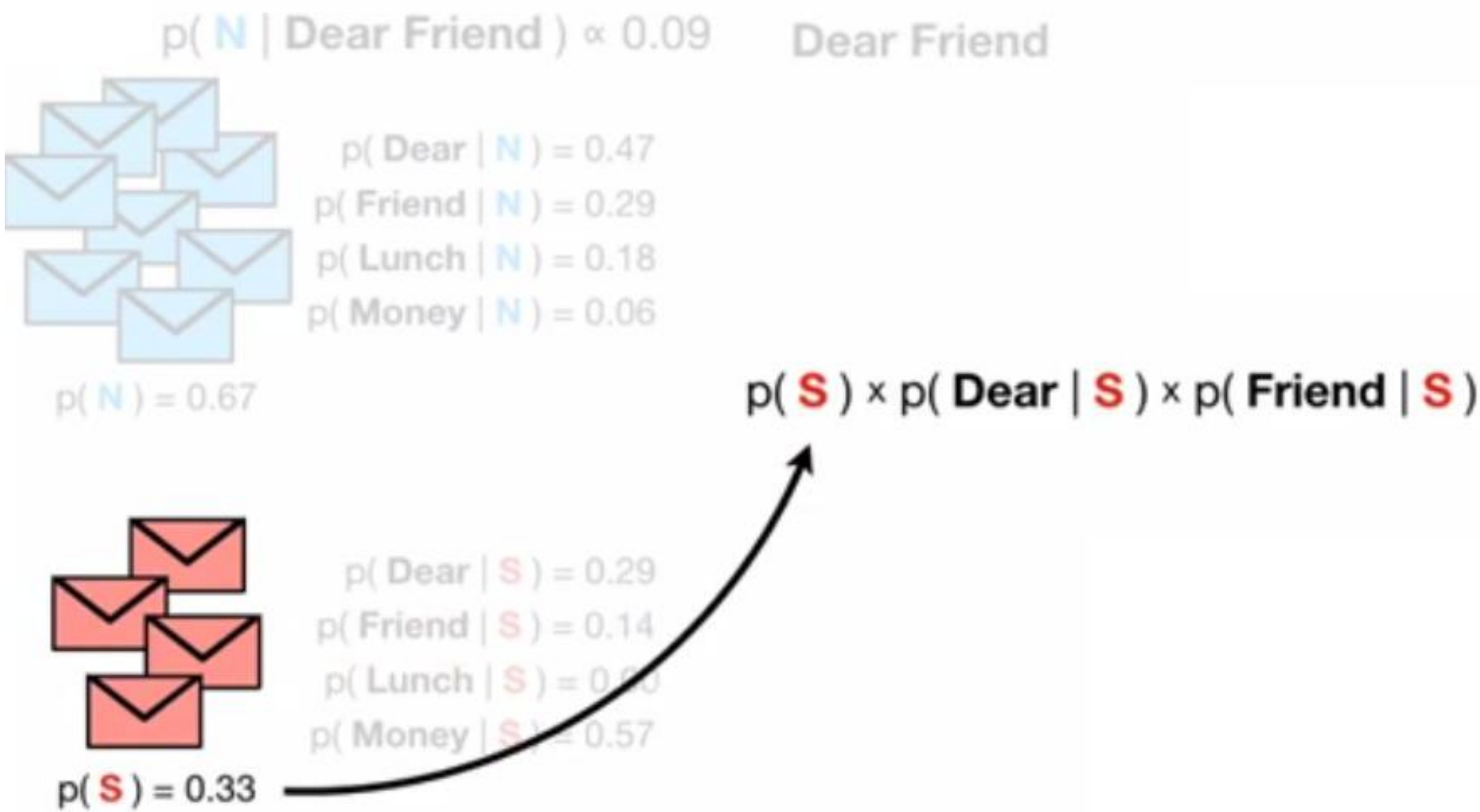
$0.67 \times 0.47 \times 0.29 \boxed{= 0.09}$ p(**N** | **Dear Friend**)

p(Dear | S) = 0.29
p(Friend | S) = 0.14
p(Lunch | S) = 0.00
p(Money | S) = 0.57

In a simple way, we can think of **0.09** as the score that **Dear Friend** gets if it is a **Normal Message**.



p(N | Dear Friend) ∝ 0.09
Dear Friend
p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N) = 0.06
p(N) = 0.67
p(S) = 4 / (4 + 8) = 0.33
p(Dear | S) = 0.29
p(Friend | S) = 0.14
p(Lunch | S) = 0.00
p(Money | S) = 0.57



p(N | Dear Friend) ∝ 0.09
Dear Friend
p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N) = 0.06
p(N) = 0.67
p(S) × p(Dear | S) × p(Friend | S)
p(Dear | S) = 0.29
p(Friend | S) = 0.14
p(Lunch | S)
p(Money | S) = 0.57
p(S) = 0.33

$p(\text{N} \mid \textbf{Dear Friend}) \propto 0.09$

Dear Friend

$p(\textbf{Dear} \mid \text{N}) = 0.47$

$p(\textbf{Friend} \mid \text{N}) = 0.29$

$p(\textbf{Lunch} \mid \text{N}) = 0.18$

$p(\textbf{Money} \mid \text{N}) = 0.06$

$p(\text{N}) = 0.67$

$$0.33 \times 0.29 \times 0.14 = 0.01 \propto p(\text{S} \mid \textbf{Dear Friend})$$

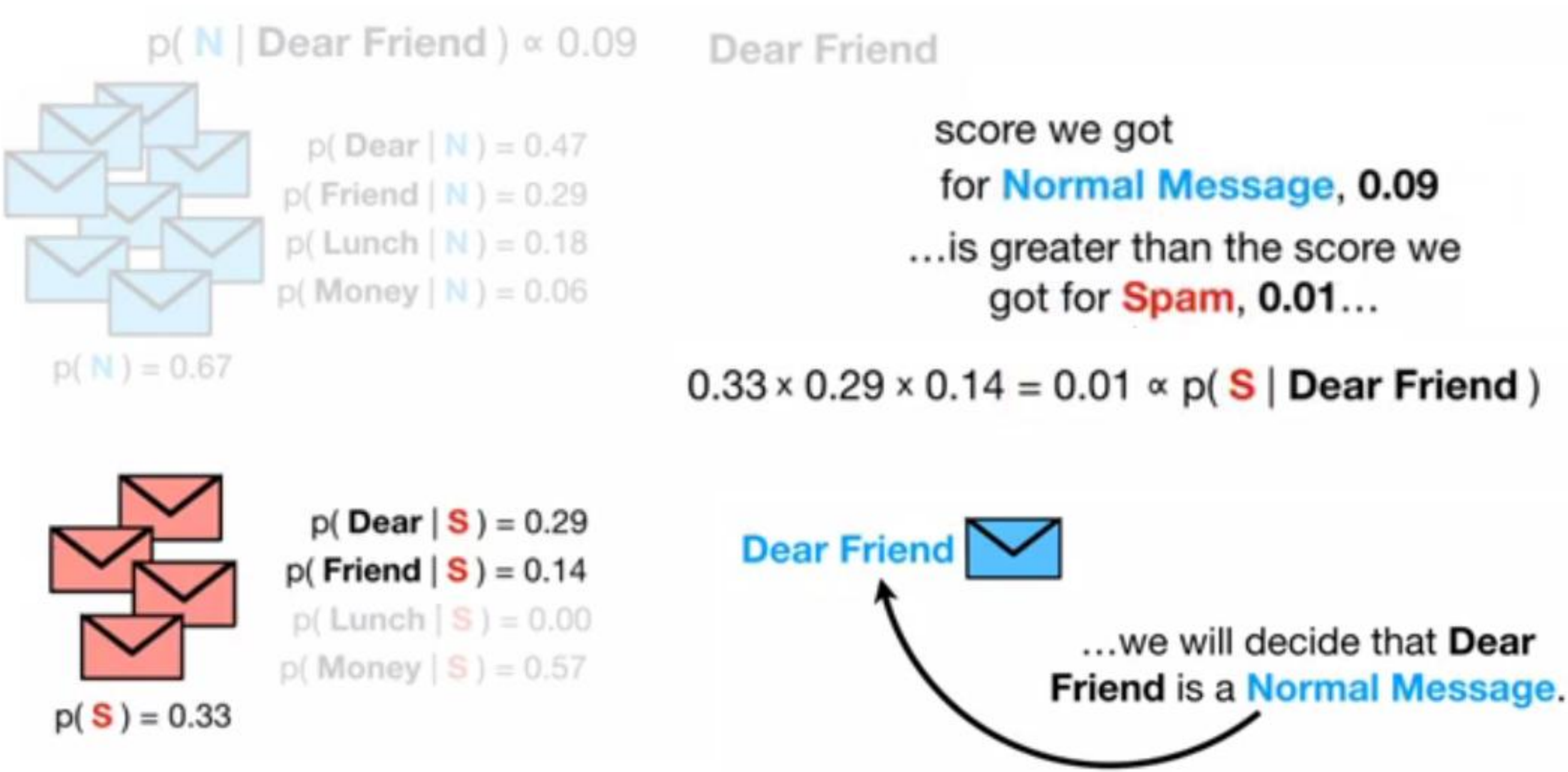
$p(\textbf{Dear} \mid \text{S}) = 0.29$

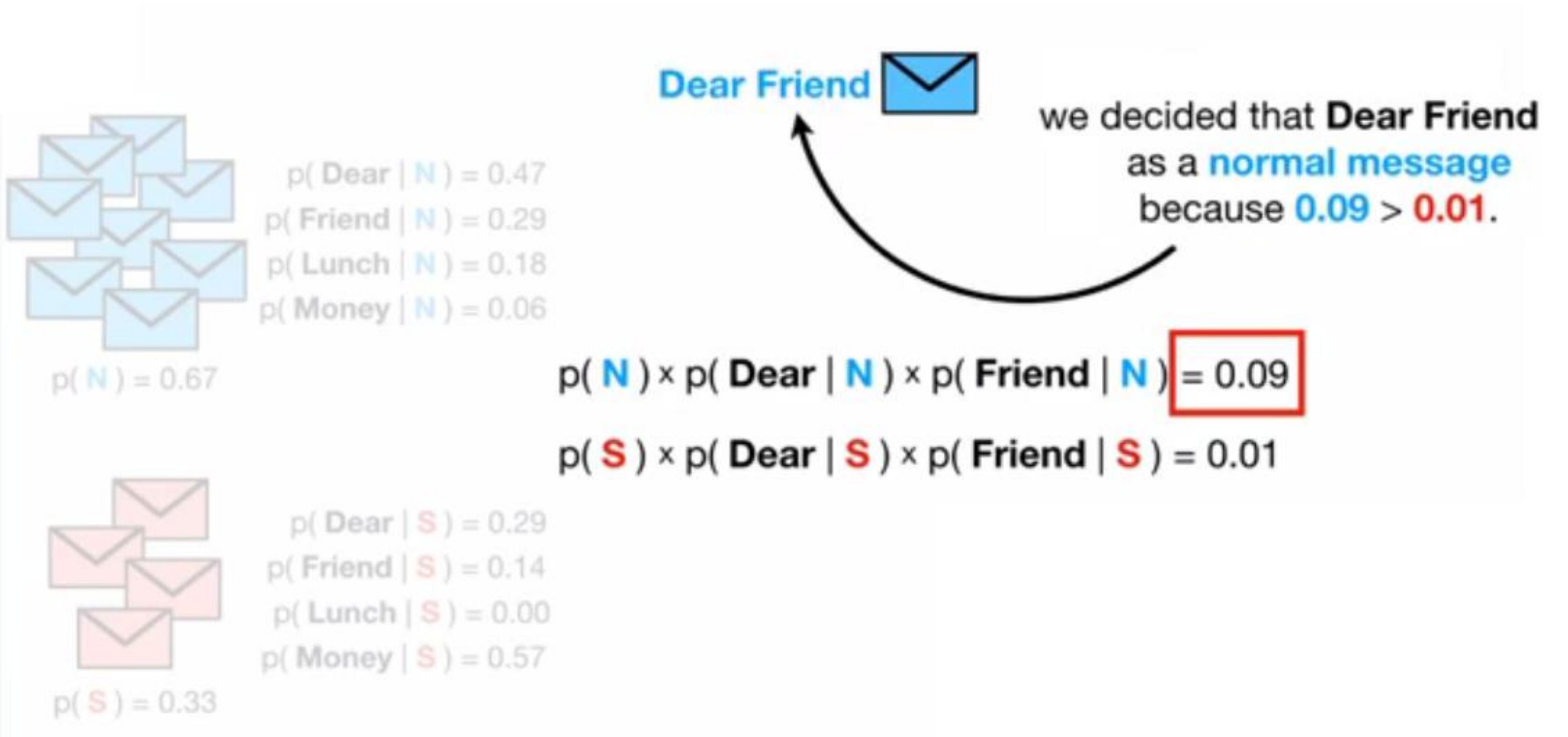
$p(\textbf{Friend} \mid \text{S}) = 0.14$

$p(\textbf{Lunch} \mid \text{S}) = 0.00$

$p(\textbf{Money} \mid \text{S}) = 0.57$

$p(\text{S}) = 0.33$


p(N | Dear Friend) ∝ 0.09
Dear Friend
p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N) = 0.06
p(N) = 0.67
score we got
for Normal Message, 0.09
...is greater than the score we got for Spam, 0.01...
0.33 × 0.29 × 0.14 = 0.01 ∝ p(S | Dear Friend)
p(Dear | S) = 0.29
p(Friend | S) = 0.14
p(Lunch | S) = 0.00
p(Money | S) = 0.57
p(S) = 0.33
Dear Friend
...we will decide that Dear Friend is a Normal Message.


Dear Friend
we decided that Dear Friend
as a normal message
because 0.09 > 0.01.
p(Dear | N) = 0.47
p(Friend | N) = 0.29
p(Lunch | N) = 0.18
p(Money | N) = 0.06
p(N) = 0.67
p(N) × p(Dear | N) × p(Friend | N) = 0.09
p(S) × p(Dear | S) × p(Friend | S) = 0.01
p(Dear | S) = 0.29
p(Friend | S) = 0.14
p(Lunch | S) = 0.00
p(Money | S) = 0.57
p(S) = 0.33