Many spam filter algorithms are based on Bayes’ rule. As a very simplified example, suppose that:

- 15% of spam messages that you receive contain the word “greeting”
- 1% of non-spam messages that you receive contain the word “greeting”
- 25% of the messages that you receive are spam.

Consider selecting an email message that you receive at random, and define the following events:

\[ G = \{ \text{message contains the word “greeting”} \} \]
\[ M = \{ \text{message is spam} \}. \]

1. Express the given values .15, .01, and .25 as probabilities in terms of these events.

2. Set up a table of hypothetical counts (with a total of 1000 or 10,000 messages) that could be used to answer the following questions.

3. Determine the probability that a randomly selected e-mail message contains the word “greeting.”

4. Given that a randomly selected e-mail message contains the word “greeting,” what is the probability that the message is spam?

5. How many times more likely is a message to be spam if it contains the word “greeting” than if it does not contain the word “greeting”?