

10.4 PROJECT

THE PROBABILITY OF SPAM FILTERING

According to the website statista.com, 28.5% of all email traffic in 2019 was made up of spam—those pesky, useless, and potentially dangerous messages that just clog our email inboxes. Most email servers these days can filter spam automatically. Spam messages often have certain suspicious phrases in the subject lines. For example, “You Have Been Selected” is one such phrase.

An incoming email is checked for key elements, such as this phrase, then the server decides whether to put the email in your mailbox or send it to the spam folder.

In this activity, you will estimate the probability that an email with a specific subject line is classified as spam.

Let $P(S)$ be the probability that an email you have received is spam and $P(S^c)$ be the probability that the email is not spam.

1. According to statista.com, what were the values of $P(S)$ and $P(S^c)$ in 2019?

Let's assume that 10% of all spam messages contain the word *selected* in the subject line. In order to simplify our notation, we will name the events as follows.

S = email is spam

S^c = email is not spam

W = subject line contains the word *selected*

W^c = subject line does not contain the word *selected*

2. Express the statement “10% of all spam messages contain the word *selected* in the subject line” as a conditional probability.
3. We also will assume that 0.5% of all nonspam messages also contain *selected* in the subject line. Express the previous statement as a conditional probability.

Since every message can be classified as either spam or not spam, the probability that any message has *selected* in the subject line is the following.

$$P(W) = P(W | S)P(S) + P(W | S^c)P(S^c)$$

4. Compute the value of $P(W)$.
5. Finally, determine the probability that an email is spam, knowing it has the word *selected* in the subject line. (**Hint:** Use Bayes' Theorem.)