

【No. 14】 あるコンピュータネットワーク上のサーバには、大量の電子メールが蓄えられており、これらのメールのうち 10% が迷惑メールで、残りの 90% が無害なメールである。また迷惑メール全体のうちで件名に「至急確認」の文字列を含んでいるものの割合は 80% であるが、無害なメールのうちで件名に「至急確認」の文字列をふくんでいるものの割合は 40% であることが分かっている。

いま、このサーバに接続してメールを立て続けに 2 通受信した。いずれのメールの件名にも「至急確認」の文字が含まれていたとき、これらのメールが両方とも迷惑メールである確率に最も近いのは次のうちではどれか。ただし、メールはサーバ上から無作為に選ばれるものとする。

1. 0.3% 2. 0.6% 3. 1.9% 4. 3.3% 5. 19.4%

題材は目新しいですが、典型的な条件付きの確率で、どの区分でも頻出の問題ですね。この問題では「2 通とも」というところがちょっとひねっていますが、1 通も 2 通も各メールは独立ですから変わりません。

解答

このサーバのメールは次の 4 つに分類される。

迷惑メール (0.1)	「至急確認」あり (0.8)	0.08
	「至急確認」なし (0.2)	0.02
無害なメール (0.9)	「至急確認」あり (0.4)	0.36
	「至急確認」なし (0.6)	0.54

これより、「至急確認」の文字が含まれているメールを 1 通受信したとき、それが迷惑メールである確率は、

$$\frac{0.08}{0.08 + 0.36} = \frac{2}{11}$$

とわかる。したがって、2 通受け取ったとき、両方とも迷惑メールである確率は、

$$\left(\frac{2}{11}\right)^2 = \frac{4}{121} = 3.3\%$$