

ベイズの定理は、「あることから  $X$  についての確信の度合いが、関連するデータ  $D_1, D_2 \dots$  の出現によって更新されていく」という状況の説明によく用いられる。

$$P(X | D_1) = \frac{P(X) \times P(D_1 | X)}{P(X) \times P(D_1 | X) + P(X^c) \times P(D_1 | X^c)}$$

ここで、 $P(X)$  を  $X$  についての**事前確率**と考えるなら、関連するデータ  $D_1$  が得られることによって、**事後確率**  $P(X | D_1)$  に更新されるということである。

(例) 迷惑メールフィルタ

最近のウイルス対策ソフトは、受信したメールが迷惑メールであるかどうかの判定をおこなってくれるものが多い。迷惑メールかどうかの判定に、ベイズの定理が用いられていることがある。

たとえば、あるメールのタイトルに「完全無料」という迷惑メールのタイトルに多く含まれる単語があったとする。このメールが迷惑メールかどうかを判定するには次のようにおこなう。

$P(X)$  あるメールが迷惑メールであるという事前確率

$D_1$  「完全無料」という迷惑メールのタイトルに多く含まれる単語

$P(D_1 | X)$  迷惑メールのタイトルに「完全無料」という単語が含まれる確率

$P(D_1 | X^c)$  迷惑メールでない普通のメールのタイトルに「完全無料」という単語が含まれる確率

ベイズの定理を用いて、これらの  $P(X | D_1)$  を計算し、このメールが迷惑メールであるかどうかを判定する。

実際には、他の単語(「出会える」など)や、疑わしい URL へのリンクがあれば、それらが  $D_2, D_3 \dots$  となって、より正確な判定をおこなうことを可能とする。この場合、 $P(X | D_1)$  を事前確率として、事後確率  $P(X | D_1, D_2)$  をベイズの定理によって求めていく。

$$P(X | D_1, D_2) = \frac{P(X | D_1) \times P(D_2 | X, D_1)}{P(X | D_2) \times P(D_2 | X, D_1) + P(X^c | D_2) \times P(D_2 | X^c, D_1)}$$

なお、迷惑メールのタイトルに「完全無料」という単語が含まれる確率  $P(D_1 | X)$  は過去に受信した迷惑メールから求められる。メールソフトに学習させた結果である。