第１回の補足資料

ベイズ則の読み方

正田 備也

masada@rikkyo.ac.jp

ベイズ則

ベイズ則とは

$$p(z|x) \propto p(x|z)p(z) \tag{1}$$

省略なしで書けば、例えば

$$p(z=s|x=a) \propto p(x=a|z=s)p(z=s) \tag{2}$$

- $\propto$ は、「左の値が右の値に比例する」という意味。
- 細かい注意：
  - たとえ $p(z|x)$ と書いていても、常に、x や z が特定の値（例えば $x=a$ と $z=s$ など）を取っている状況を考えている。
  - ベイズ則に限らず、いつでもこのことに注意する。

問題

ベイズ則を使って計算した結果、以下のようになったとする。

$$
\begin{aligned}
p(z=r|x=a) &\propto 0.01 \\
p(z=s|x=a) &\propto 0.02 \\
p(z=t|x=a) &\propto 0.03
\end{aligned}
\tag{3}
$$

左辺の 3 つの確率 $p(z=r|x=a)$, $p(z=s|x=a)$, $p(z=t|x=a)$ を、確率としてちゃんと求めるには、どうすればいいか？

- 上のままだと、$0.01+0.02+0.03 \neq 1$ と、足して 1 にならない。何に比例するかが分かっているだけの状態。

確率が計算できる場合

- もし、確率変数 z のとる値が、r, s, t の 3 種類の値で尽くされているのであれば、前のスライドの式 (3) のように比例関係さえ分かっていれば、確率を求めることができる。
- なぜなら・・・
  - 確率は、すべての場合にわたって和をとったものがぴったり 1 になっていなければならない。
  - だが、いまの例では、r, s, t ですべての場合がつくされている。
  - ということは、あとは、足して 1 になるように、3 つの値を一斉に何かで割ればいいだけ。

問題の答え(1/2)

- 確率変数 z のとる値が、r, s, t の 3 種類の値で尽くされているとする。

$$p(z = r|x = a) \propto 0.01$$

$$p(z = s|x = a) \propto 0.02$$

$$p(z = t|x = a) \propto 0.03$$

- 右辺の値をすべて足すと $0.01 + 0.02 + 0.03 = 0.06$
  - これを規格化定数と呼ぶ。（それで割れば確率になる、という値。）
- この値で、ベイズ則によって求められたそれぞれの値を割れば、確率、つまり、足して 1 になる値が得られる。

問題の答え(2/2)

$$p(z = r|x = a) = \frac{0.01}{0.01 + 0.02 + 0.03} = \frac{1}{6}$$
$$p(z = s|x = a) = \frac{0.02}{0.01 + 0.02 + 0.03} = \frac{1}{3}$$
$$p(z = t|x = a) = \frac{0.03}{0.01 + 0.02 + 0.03} = \frac{1}{2}$$

- 規格化定数を求めることがポイント。
  - 規格化定数で割ることによって、比例記号 $\propto$ を外せている。
  - つまり、等号 $=$ で答えを求めることができている。

確率が計算できない場合

- 規格化定数が求められない場合、つまり・・・
- 「ベイズ則の右辺をその値で割れば、確率が得られるよ、という値」が求められない場合、上のように確率を計算することは、できない。

問 規格化定数が計算できない場合とは、どのような場合か？