

ベイズ統計学入門

FooQoo

2019/1/14

1 確率の復習

1.1 確率の定義

事象 A の起こる確率 p を次のように一般化する。ただし起こりうる全ての事象は等確率で発生する。

$$p = \frac{\text{事象 } A \text{ の起こる場合の数}}{\text{起こりうる全ての場合の数}} \quad (1)$$

起こりうる全ての事象の集合 U を標本空間と呼ぶ。 U の個数に対する、事象 A が発生する要素の個数で割った値を「数学的確率」とする。一方で、何回も試行を繰り返した時、得られる事象 A の割合を「統計的確率」と呼ぶ。

1.2 同時確率

事象 A と B を考える。 A と B が同時に起こる事象を $A \cap B$ と表す。この確率を、 $P(A \cap B)$ と表し、同時確率と呼ぶ。

1.3 条件付き確率

ある事象 A が起こった条件のもとで事象 B の起こる確率を、条件付き確率と呼ぶ。条件付き確率は、 $P(B|A)$ で表し、以下の式で計算される。

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad (2)$$

確率の定義のみ用いた証明は以下。

$$P(B|A) = \frac{n_{A \cap B}}{n_A}, P(A \cap B) = \frac{n_{A \cap B}}{n_U}, P(A) = \frac{n_A}{n_U} \text{ より} \quad (3)$$

$$P(B|A) = \frac{n_{A \cap B}}{n_A} \quad (4)$$

$$= \frac{n_U}{n_A} \frac{n_{A \cap B}}{n_U} \quad (5)$$

$$= \frac{1}{\frac{n_A}{n_U}} \frac{n_{A \cap B}}{n_U} \quad (6)$$

$$= \frac{P(A \cap B)}{P(A)}$$

1.4 ベイズの定理

同時確率と条件付き確率を整理すると以下の式が導出される。

$$P(A \cap B) = P(A)P(B|A) \quad (7)$$

この式は A と B の役割を入れ替えても成立する。

$$P(A \cap B) = P(B)P(A|B) \quad (8)$$

以上より、 $P(B) \neq 0$ について解くと、次の式が得られる。

$$P(A)P(B|A) = P(B)P(A|B) \quad (9)$$

$$P(A|B) = \frac{P(B)P(A|B)}{P(A)} \quad (10)$$

以上がベイズの定理である。 $P(A|B)$ を $P(B|A)$ の逆確率と呼ぶ。また、何も条件がない時に A の起こる確率 $P(B)$ を B の条件がどれぐらい A に影響するかを定量的に与えている公式がベイズの定理とみなすことができる。 B をデータと解釈するなら、データ取得前と後との確率の関係を示しているとみなせる。その点で $P(A)$ を事前確率、 $P(A|B)$ を事後確率と呼ぶ。

ここでベイズの定理における A を原因や過程 (Hypothesis) とし、 B を A のもとで得られる結果やデータ (Data) と解釈する。この解釈のためにベイズの定理を次のように書き換える。

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)} \quad (11)$$

この定式化によって、原因が与えられた時にその結果を議論するところを、結果から原因をたどるように変換することができる。この意味で $P(H|D)$ を原因の確率と呼ぶ。

2 ベイズの展開公式

データ D が得られた時の原因が H_1, H_2, H_3 とする。これら 3 つの原因に重複がない場合、すなわち排反であるとする。この場合、データが与えられる確率 $P(D)$ は以下の式で与えられる。この時、確率の乗法定理を適用することでベイズの基本公式の変形が得られる。

$$P(D) = P(D \cup H_1) + P(D \cup H_2) + P(D \cup H_3) \quad (12)$$

$$P(D) = P(D|H_1)P(H_1) + P(D|H_2)P(H_2) + P(D|H_3)P(H_3) \quad (13)$$

上の結果をベイズの定理に代入する。

$$P(H_1|D) = \frac{P(D|H_1)P(H_1)}{P(D|H_1)P(H_1) + P(D|H_2)P(H_2) + P(D|H_3)P(H_3)} \quad (14)$$

以上ベイズの展開公式を一般化する。

$$P(H_i|D) = \frac{P(D|H_i)P(H_i)}{P(D|H_1)P(H_1) + P(D|H_2)P(H_2) + \dots + P(D|H_n)P(H_n)} \quad (15)$$

ベイズの展開公式で、 $P(D|H_1)$ を原因 H の尤度と呼ぶ。これは、 H_i のもとでデータの得られるもっともらしい確率を表す。尤度かけられる $P(H_i)$ を事前確率と呼ぶ。これはデータ D の影響をまだ考慮していな

い、分析前の原因 H_i の起こる確率である。原因の確率 $P(H_i|D)$ を事後確率と呼ぶ。ベイズの基本公式を用いてデータ D を考慮して得られた分析後の原因 H_i の確率だからである。

ベイズ理論の計算は以下の 3 ステップに分かれている。

1. モデル化し、それから尤度を算出する。
2. 事前確率を設定。
3. ベイズの展開公式を用いて事後確率を算出する。

ここで事前確率がわかっていないケースがある。数学的にはここで計算が止まってしまうが、経験や常識を用いて対処する。例えば、「何も条件がないのなら事前確率として、選択確率は等確率」とする。このような「何も情報がなければ確率は同等」とすることを理由不十分の原則と呼ぶ。一度事後確率が計算されれば、次回はその確率を事前確率とすることで、学習することができる。これを、ベイズ更新という。この時、データの入力順によって結果が変動するようことが考えられる。しかし、データが同じであれば解析順序によらないことが保証される。これをベイズ理論の逐次合理性と呼ぶ。

MAP 推定は事後確率が一番大きいことが一番よく起こると推定する。これに対して、選択肢に対して損失がある場合、期待損失を求めることで意思決定に役立てることができる。