

学習とロジステック曲線について

リスク認知から意思決定に至る数学モデル

人類はどのようにして意思決定をするのか

最も身近な日常生活の中での意思決定モデルとして、「天気予報の降水確率」がある。降水確率 0%では、雨は絶対に降らないのか？降水確率が何%になると、人は傘を持って外出するのか？

社会生活において、飛行機事故、原子力発電所事故、自然災害、医療事故等の様々なリスクがあるが、リスクを認知して意思決定に至る数学的モデルを調査した。

松本睦郎（札幌啓成高等学校 講師）

Episode 1 降水確率何%で傘を持参するか？

降水確率とは、1mm 以上の雨が降る確率である。ある特定地域の降水確率が 30%であるとは、「30%という予報が 100 回発表されたとき、30 回は 1mm 以上の降水がある。」という過去のデータから求めた値である。降水確率 0%とは、5%未満をすべて意味するので、全く雨が降らない訳ではない。100%は、土砂降りの雨が降るという意味でもない。

人は、降水確率は何%になると傘を持参する意思決定をするのだろうか？

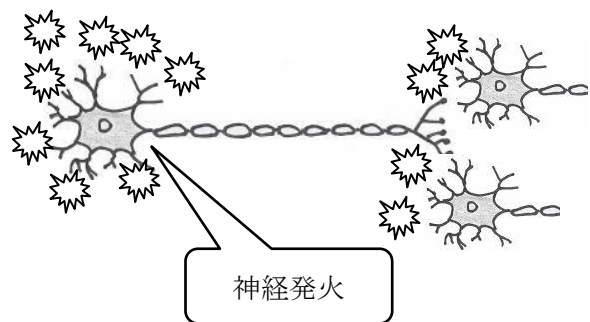
確率と意思決定には、何かの数学的関係があるようである。

Episode 2 神経細胞の学習モデル

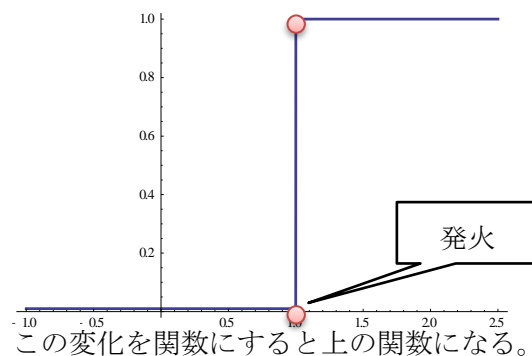


脳細胞には、多くの神経細胞（ニューロン）が存在しネットワークを構成している。（上図）

1つのニューロンが刺激を受け取ったとき、刺激の総和が少なくてそのニューロン固有の閾値を越えなければ、受け取った刺激を無視する。細胞にとって、小さい刺激を無視することは、生命を維持する上で重要である。



入力刺激の総和がニューロン固有の閾値を超えると神経細胞は反応して、軸索を通じて他のニューロンに刺激を伝達する。この状態を発火と呼ぶ。



Episode 3 イソップ寓話

「羊飼いの少年とオオカミ」の数理

寓話の内容

羊飼いの少年が、退屈しのぎに「オオカミが来た！」と嘘をついて騒ぎを起こす。だまされた大人たちは武器を持



って出てくるが、徒労に終わる。少年が繰り返し同じ嘘をついたので、本当に狼が現れた時には大人たちは信用せず、誰も助けに来なかった。そして村の羊は全て狼に食べられてしまった。『ウィキペディア (Wikipedia) より』

この寓話を例にリスクの認知学習過程から意思決定に到達する数学的モデルを作成する。

H_1 : 「羊飼いの少年がウソツキである。」 事象

H_2 : 「羊飼いの少年が正直である。」 事象

Z_1 : 「オオカミが来る。」 事象

Z_2 : 「オオカミが来ない。」 事象

原因 \ 結果	Z_1 : オオカミ来る	Z_2 : オオカミ来ない
H_1 : ウソツキ少年	$\frac{1}{3}$	$\frac{2}{3}$
H_2 : 正直少年	$\frac{3}{4}$	$\frac{1}{4}$

$$P(Z_1|H_1) = \frac{1}{3}, P(Z_2|H_1) = \frac{2}{3}$$

$$P(Z_1|H_2) = \frac{3}{4}, P(Z_2|H_2) = \frac{1}{4}$$

条件付き確率をこのように過去のビックデータから初期設定する。(↑)



『オオカミが来たゾー』

と嘘をついたとき (条件付き)

オオカミが来る確率

$\frac{1}{3}$, 来ない確率 $\frac{2}{3}$

$$P(Z_1|H_1) = \frac{1}{3}, P(Z_2|H_1) = \frac{2}{3}$$

$$\frac{P(Z_1 \cap H_1)}{P(H_1)} = \frac{1}{3}, \frac{P(Z_2 \cap H_1)}{P(H_1)} = \frac{2}{3}$$

この少年が「ウソツキ」か「正直」である確率が「等しいもの」とすると。

$$P(H_1) = P(H_2) = \frac{1}{2}$$

$$\therefore P(Z_1 \cap H_1) = \frac{1}{6}, P(Z_2 \cap H_1) = \frac{1}{3}$$

同様にして、

$$P(Z_1|H_2) = \frac{3}{4}, P(Z_2|H_2) = \frac{1}{4}$$

$$\therefore P(Z_1 \cap H_2) = \frac{3}{8}, P(Z_2 \cap H_2) = \frac{1}{8}$$

ベイズの定理より

$$P(H_1|Z_1) = \frac{P(H_1 \cap Z_1)}{P(Z_1)} = \frac{P(H_1 \cap Z_1)}{P(H_1 \cap Z_1) + P(H_2 \cap Z_1)} = \frac{\frac{1}{6}}{\frac{1}{6} + \frac{3}{8}}$$

$$P(H_1|Z_2) = \frac{P(H_1 \cap Z_2)}{P(Z_2)} = \frac{P(H_1 \cap Z_2)}{P(H_1 \cap Z_2) + P(H_2 \cap Z_2)} = \frac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{8}}$$

$$\therefore P(H_1|Z_1) = \frac{4}{13} = 0.30769230769$$

これは、「オオカミが来た。」(結果)が発生したとき、羊飼いの少年がウソツキである確率を意味する。

実際に、オオカミが来たのでウソツキ確率の変化は、

$$0.5 \rightarrow 0.30769230769$$

$$\therefore P(H_1|Z_2) = \frac{8}{11} = 0.727272727$$

これは、「オオカミが来なかった。」(結果)が発生したとき、羊飼いの少年がウソツキである確率を意味する。実際に、オオカミが来ないので、ウソツキ確率の変化は

$$0.5 \rightarrow 0.727272727$$

初期には、少年がウソツキか正直か不明で、等確率と設定したが、「オオカミが来るゾー。」と村人に宣言した結果、オオカミが来なかったときのウソツキの確率が、0.5 から 0.72 へアップする。オオカミが来たときは、ウソツキの確率は、0.5 から 0.3 へとダウンしている。

ウソをつけば、信用は減少する。



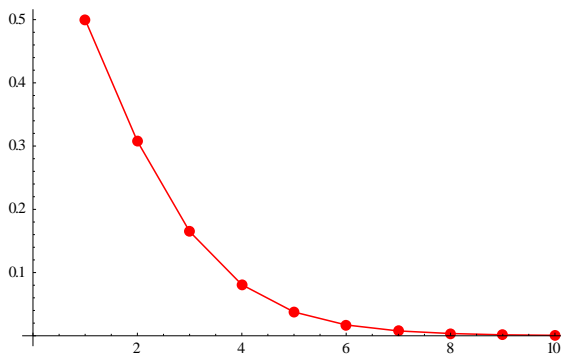
Episode4 確率漸化式を作成する

羊飼いの少年が「オオカミが来るゾー。」と村人の n 回宣言したときの、ウソツキ確率の漸化式を作成する。
 n 回目のウソツキ確率を w_n と設定する。
つまり、村人から見た少年がウソツキである確率。

原因 \ 結果	Z_1 :オオカミ来る	Z_2 :オオカミ来ない
$P(H_1) = w_1$	p_{11}	p_{12}
$P(H_2) = 1 - w_1$	p_{21}	p_{22}

1 回前のウソツキ (H_1) の確率を w_{n-1} とすると、
(I) オオカミが来るとき。

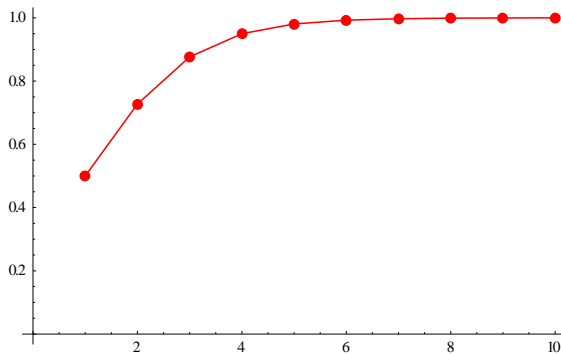
$$w_n = \frac{P(Z_1 \cap H_1)}{P(Z_1 \cap H_1) + P(Z_1 \cap H_2)} = \frac{p_{11} \times w_{n-1}}{p_{11} \times w_{n-1} + p_{21}(1 - w_{n-1})}$$



村人から見た少年がウソツキである確率は、減少していく。

(II) オオカミが来ないとき。

$$w_n = \frac{P(Z_2 \cap H_1)}{P(Z_2 \cap H_1) + P(Z_2 \cap H_2)} = \frac{p_{11} \times w_{n-1}}{p_{11} \times w_{n-1} + p_{21}(1 - w_{n-1})}$$



村人から見た少年がウソツキである確率は増加していく。

Episode5 情報量とロジスティック曲線について

「情報量」とは、結果 (オオカミが来る。来ない。) を得た場合の確率の比の自然対数とする。

$$\log\left(\frac{p_{11}}{p_{21}}\right), \log\left(\frac{p_{12}}{p_{22}}\right)$$

オオカミが来るときの情報量 (ウソツキの確信を減らす方向) マイナス方向へ延長

$$\log\left(\frac{1}{\frac{3}{3}}\right) = \log\left(\frac{4}{9}\right) = -0.81093$$

オオカミが来ないときの情報量 (ウソツキの確信を増やす方向) プラス方向へ延長

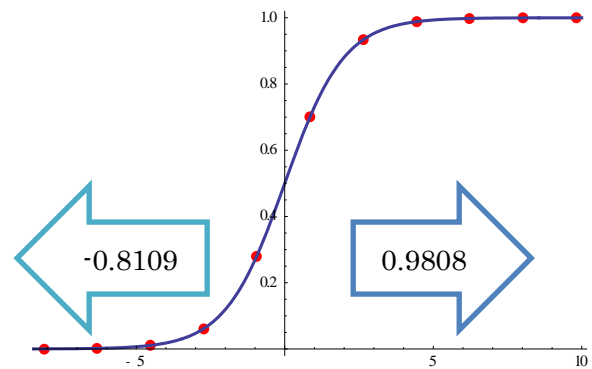
$$\log\left(\frac{2}{\frac{3}{1}}\right) = \log\left(\frac{8}{3}\right) = 0.980829$$

「情報投入量: L_n 」を定義する。

$$L_n = (Z_1 \text{ の回数}) \times \log \frac{p_{11}}{p_{21}} + (Z_2 \text{ の回数}) \times \log \frac{p_{12}}{p_{22}}$$

ただし、(Z_1 の回数) + (Z_2 の回数) = n

点 (L_n, w_n) をグラフにすると、次のグラフになる。
このグラフは、ロジスティック曲線と一致する。



横軸 L_n : 情報投入量, 縦軸 w_n : ウソツキである確率
統計学的に、

$$w_n = \frac{1}{1 + b \times \text{Exp}(-L_n)}, b_n = \frac{1 - w_0}{w_0}$$

$$w_0 = 0.5$$

とすると、

$$f(x) = \frac{1}{1 + \text{Exp}(-x)}$$

のグラフと重なる。(↑) のグラフ

実際に 1 回目に「オオカミが来なかったときの情報量 = 0.980829」をロジスティック曲線へ代入すると、

$$f(0.980829) = \frac{1}{1 + \text{Exp}(-0.980829)} = 0.727273$$

となり、見事に P3 で求めた確率

$$\therefore P(H_1|Z_2) = \frac{8}{11} = 0.727272727$$

と一致する。

【例題】羊飼いの少年が、4 回「オオカミが来るゾ」と叫んだ場合、4 回の結果として、

1 回	2 回	3 回	4 回
来ない	来ない	来た	来ない

が得られた。羊飼いの少年がウソツキである確率を求めよ。

〔解答〕

「来ない」が 3 回、「来た」が 1 回なので、情報投入量は、

$$L_4 = 3 \times 0.980829 - 1 \times 0.81093 = 2.131557$$

をロジスティック関数へ代入すると、

$$f(2.131557) = \frac{1}{1 + \text{Exp}(-2.131557)} = 0.893933$$

一連の羊飼いの少年の言動から、ウソツキである確率は、0.893933 である。

Episode6 村人の意思決定の形成

村人が「少年がウソツキである。」「少年はウソツキでない。」という 2 つの意思決定をする過程を考えてみよう。

ウソツキである確率が 90% 以上のとき、

$$w_n \geq 0.9$$

村人は、「少年はウソツキである。」と意思決定する。

逆に、ウソツキ確率が、10% 以下のとき、

$$w_n \leq 0.1$$

村人は、「少年はウソツキではない。」と意思決定する。

0.9 を上側閾値。0.1 を下側閾値と呼ぶ。

各閾値に至る情報量を求める。

上側閾値の情報量は、

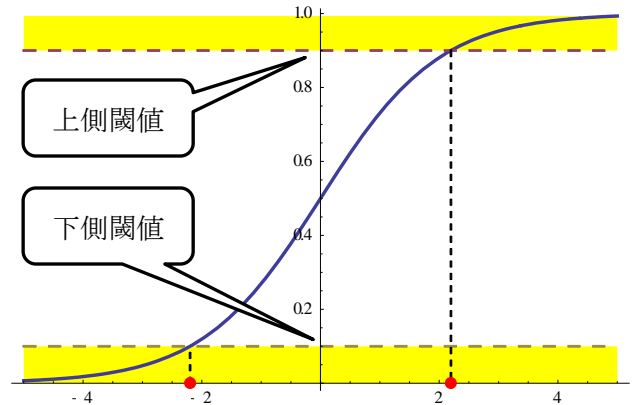
$$\frac{1}{1 + e^{-x}} = \frac{9}{10}$$

$$x = \log 9 = 2.19722$$

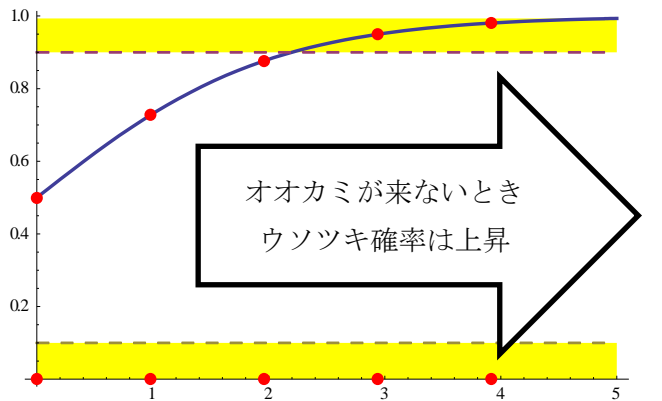
下側閾値は、

$$\frac{1}{1 + e^{-x}} = \frac{1}{10}$$

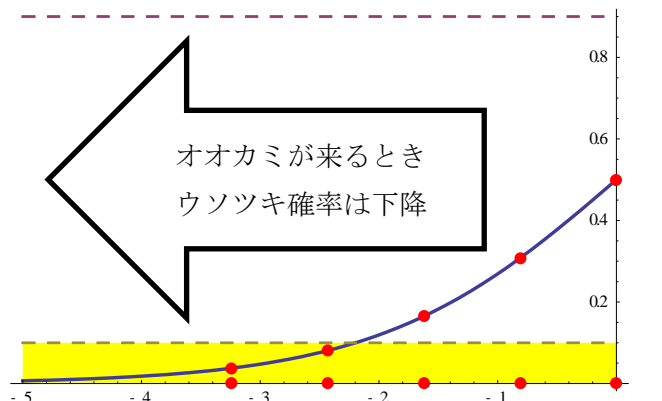
$$x = -\log 9 = -2.19722$$



村人の意思決定は、「ランダム・ウォーク問題」になる。



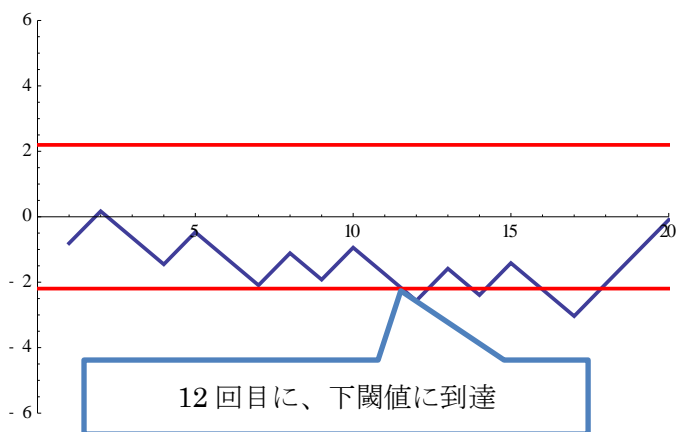
3 回連続して、オオカミが来ないとき上側閾値を超える。村人は、「少年がウソツキである。」と意思決定をする。(↑)



3 回連続して、オオカミが来るとき下側閾値を超える。
村人は、「少年がウソツキでない。」と意思決定をする。
(↑)

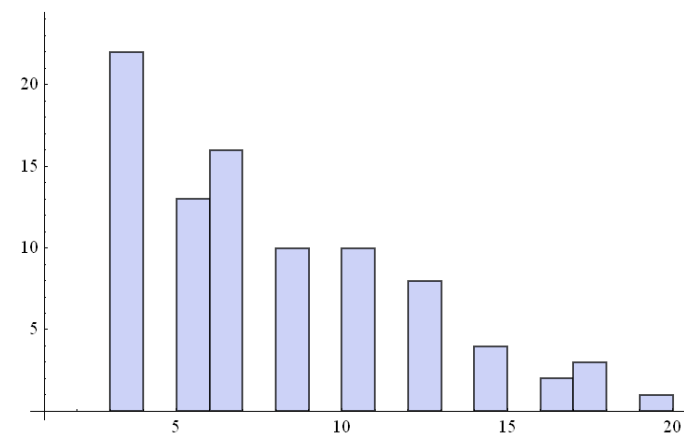
Episode7 シミュレーションしてみよう

現実的には、オオカミが「来た」「来ない」がランダムに発生するので、単純に確率 0.5 と仮定する。
村人が「少年はウソツキである。」「少年はウソツキでない。」と意思決定をするのに要する時間 (回数) はどの程度か? 計算により、上下閾値の到達回数の期待値を求めることは難しいので、乱数を利用してシミュレーションを 20 回実施した。(↓)



このケースは、12 回目に「少年はウソツキではない。」と意思決定をしたことになる。

100 回の試行を実施して、村人が意思決定するまでの回数をヒストグラムにした。(↓)



おおよそ、村人は 3 回～7 回程度で、意思決定をしていると予測される。

【補足例題】〔地震のシミュレーション〕

「30 年で首都直下型大地震の確率 80% である。」と報

道されている。

1 年間地震が発生しない確率を p とする。

30 年間地震が発生しない確率は、確率の積の法則により p^{30} となる。

$$p^{30} = \frac{2}{10}$$

この方程式を解くと、

$$0.947766$$

1 年間大地震が発生しない確率となる。

逆に、1 年間大地震が発生する確率は

$$1 - 0.947766 = 0.052234$$

となる。

「この低い確率から大地震は安心だ。」と言えるのでしょうか?

0~1 までの乱数を発生してシミュレーションをしてみます。30 個の乱数を発生させ、0.052234 よりも小さいときは、+1 カウントし 0.052234 よりも大きいときは 0 カウントする。1 行 30 個の数字を作成する。この操作を 10 回行うと、10 行 30 列の行列ができる。

```
matrix =
(
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
)
```

1 のあるところが、大地震が発生するものである。

この確率が高いか低いかは、村人である国民が判断しなければいけない。

確率は、社会生活にとってとても大切な数学的要因であることがわかる。

〔参考文献〕

- ベイズ統計学 (創元社) 松原望著
- 入門ベイズ統計 (東京図書) 松原望著
- ディープラーニングがわかる数学入門 (技術評論社) 涌井良幸・貞美著
- 基礎医学統計学 (南江堂) 加納克己、高橋秀人著
- 初等統計学 (培風館) PG ホエール著