ビリーフとベイズ

植野真臣
電気通信大学
情報理工学研究科
情報数理工学プログラム

今後のスケジュール（予定）

4月８日　授業の概要とガイダンス

4月15日　ベイズの定理

4月22日　ベイズはどのように誕生したか？

5月13日　ベイズはコンピュータ、人工知能の父である！！

5月20日　ビリーフとベイズ

5月27日　6月3日　6月10日　尤度と最尤推定

6月17日　6月24日　7月1日　ベイズ推定と事前分布、階層ベイズ、因果推論

7月8日　自宅でオンデマンド授業

7月15日　自宅でオンデマンド授業

7月22日　7月29日　確率的グラフィカルモデル、ベイジアンネットワークと機械学習

8月 5日　テストと総括

1. 頻度論による確率

コイン を何百回も投げて表が出た回数（頻度）を数えて，その割合を求めることを考えよう．いま，投げる回数をn とし，表の出た回数n_1とすると，$n \longrightarrow \infty$のとき，$\frac{n_1}{n} \to \frac{1}{2}$

となることが予想される．このように，何回も実験を繰り返してn 回中，事象A が n_1 回出たとき，$\frac{n_1}{n}$ をA の確率と解釈するのが頻度主義である．

しかし，この定義では真の確率は無限回実験をしなければならない．

ベイズの定理(一般化された記述)

データXが得られたときのC_iの確率

$$P(C_i|X) = \frac{P(C_i)P(X|C_i)}{\sum_{i=1}^{n} P(C_i)P(X|C_i)}$$

が成り立つ.

ベイズの定理(一般化された記述)

データXが得られたときのC_iの確率

事後確率

$$\boxed{P(C_i|X)} = \frac{P(C_i)P(X|C_i)}{\sum_{i=1}^{n} P(C_i)P(X|C_i)}$$

が成り立つ.

ベイズの定理(一般化された記述)

データXが得られたときのC_iの確率

事前確率

事後確率

$$\boxed{P(C_i|X)} = \frac{\boxed{P(C_i)}P(X|C_i)}{\sum_{i=1}^{n} P(C_i)P(X|C_i)}$$

が成り立つ.

ベイズの定理(一般化された記述)

データXが得られたときのC_iの確率

事前確率

データの出る確率（尤度）

事後確率

$$\boxed{P(C_i|X)} = \frac{\boxed{P(C_i)}\boxed{P(X|C_i)}}{\sum_{i=1}^{n} P(C_i)P(X|C_i)}$$

が成り立つ.

ベイズの定理(一般化された記述)

データXが得られたときのC_iの確率

事前確率　データの出る確率（尤度）

事後確率

$$\boxed{P(C_i|X)} = \frac{\boxed{P(C_i)}\boxed{P(X|C_i)}}{\boxed{\sum_{i=1}^{n} P(C_i)P(X|C_i)}}$$

C_iについて定数

が成り立つ.

再掲

キリストの弟子たちはキリストの復活を望んでいました。あまりに臨みが強すぎて少し似ているだけの人でもキリストに見えてしまうことがあります。弟子がキリストの復活を見たと証言する事象をA, 実際にキリストが復活したという事象をBとする. $\mathrm{P}(A|\mathrm{B}) = 1.0$, $\mathrm{P}(A|\neg\mathrm{B}) = 0.5$, $\mathrm{P}(B) = 0.000001$ とする. ある弟子がキリストの復活を見たと証言したとき、本当にキリストが復活した確率を求めてみよう.

回答

$$\mathrm{P}(A|\mathrm{B}) = 1.0,\ \mathrm{P}(A|\neg\mathrm{B}) = 0.5,$$

$$\mathrm{P}(B) = 0.000001$$

$$\mathrm{P}(B|A) = \frac{P(B)\mathrm{P}(A|\mathrm{B})}{\sum_{B\neg B} P(B)\mathrm{P}(A|\mathrm{B})}$$

$$= \frac{0.000001 \times 1.0}{0.000001 \times 1.0 + (1 - 0.000001) \times 0.5}$$

$$\fallingdotseq 0.000002$$

約２倍になった！！

再掲　例題1-2

この後、30人の弟子が独立にキリストの復活を見たと証言した。本当にキリストが復活した確率を求めてみよう.

$\mathrm{P}(A|\mathrm{B}) = 1.0$, $\mathrm{P}(A|\neg\mathrm{B}) = 0.5$, $\mathrm{P}(B) = 0.000002$とする.

回答

$$\mathrm{P}(A|\mathrm{B}) = 1.0,\ \mathrm{P}(A|\neg\mathrm{B}) = 0.5,$$

$$\mathrm{P}(B) = 0.000002 \text{より}$$

$$\mathrm{P}(B|\mathrm{A}) = \frac{\mathrm{P}(B)\mathrm{P}(A|B)^{30}}{\mathrm{P}(B)\mathrm{P}(A|B)^{30} + \mathrm{P}(\neg B)\mathrm{P}(A|\neg B)^{30}}$$

$$= \frac{0.000002\times 1.0^{30}}{0.000002\times 1.0^{30} + (1 - 0.000002)\times 0.5^{30}}$$

$$= 0.99953$$

主観確率としてのビリーフ(信念)

例えば，以下のような主観確率の例がある.

1. 第三次世界大戦が20XX 年までに起こる確率が0.01
2. 明日，会社の株式の価格が上がる確率が0.35
3. 来年の今日，東京で雨が降る確率が0.5

ベイズ統計では，これらの主観確率は個人の意思決定のための信念として定義され，ビリーフ(belief)と呼ばれる．当然，頻度論的確率を主観確率の一種とみなすことができるが，その逆は成り立たない.

レオナード･ジミー･サヴェッジ

1917 –1971

シカゴ大学、コロンビア大学

主観確率の数学理論の構築

「アメリカが５年以内に戦争する確率」

「核兵器が使用される確率」

「ギャンブルの意思決定」

2. ビリーフ(信念)

つぎの二つの賭けを考えよう.

1. もしキリストが復活していれば1 万円もらえる.

2. 赤玉n 個, 白玉$100-n$ 個が入っている合計100 個の玉が入っている壺の中から一つ玉を抜き出し, それが赤玉なら1 万円もらえる.

どちらの賭けを選ぶかといわれれば, 2 番目の賭けで赤玉が100 個ならば, 誰もが迷わず2 番目の賭けを選ぶだろうし, 逆に$n=0$ ならば, 1 番目の賭けを選ぶだろう. この二つの賭けがちょうど同等になるようにn を設定することができれば, $\frac{n}{100}$があなたの「キリストが復活した」ビリーフになる. このように, ベイズ統計における確率の解釈「ビリーフ」は頻度主義の確率で扱える対象を拡張でき, 個人的な信念やそれに基づく意思決定をも合理的に扱えるツールとなる.

１と2の期待効用が等価になるようにBeliefが求められる

$u(a,\theta)$: 行動a, 変数θのときの効用関数

期待効用

$$\mathrm{E}[u(a,\theta)] = \int u(a,\theta)p(\theta)d\theta$$

データxを所与としたとき，予測分布$p(y|x) = \int p(y|\theta)p(\theta|x)d\theta$を用いて

$$\mathrm{E}[u(a,y)] = \int u(a,y)p(y|x)dy$$

例1 では，もともとのキリストが復活する確率 $P(B)$ が，弟子の報告により$P(B \mid A)$ にビリーフが更新されていることがわかる．すなわち，弟子の証言によって事前のビリーフが事後のビリーフに更新されたのである．このとき，ベイズ統計では，

弟子の証言を「エビデンス」（evidence）と呼び，

事前のビリーフを「**事前確率**」（prior probability），事後のビリーフを「**事後確率**」（posterior probability）と呼ぶ

デシジョンツリー

ライファとシュレイファー

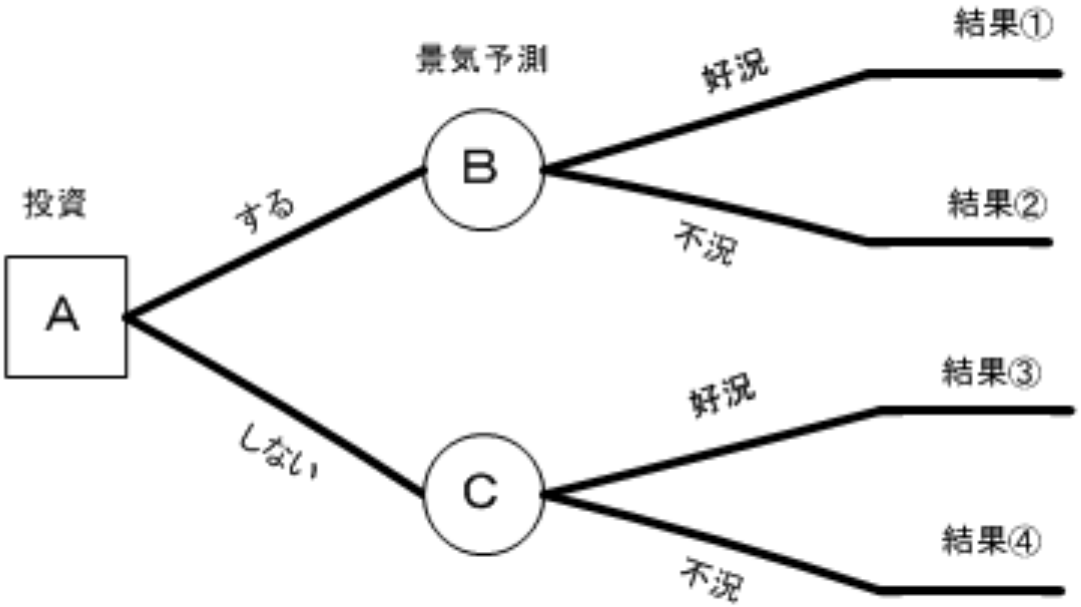
ベイズ意思決定理論

期待効用最大化　$\Sigma P(Xi)P(Ui)$

デシジョンツリーの開発

経済学や経営学の分野で大ブレーク

デシジョンツリー



- 意思決定ノード：意思決定者がコントロールできる行動で、「□」で表わす。
- イベントノード：意思決定者がコントロールできない事象で「○」で表わす。
- 結果ノード＝リンク先で表す
- 結果ノード：結果価値を得る最終点開いたリンクで表わす。

例題

研究投資には100億円が必要

研究が成功すると利益が200億円見込める

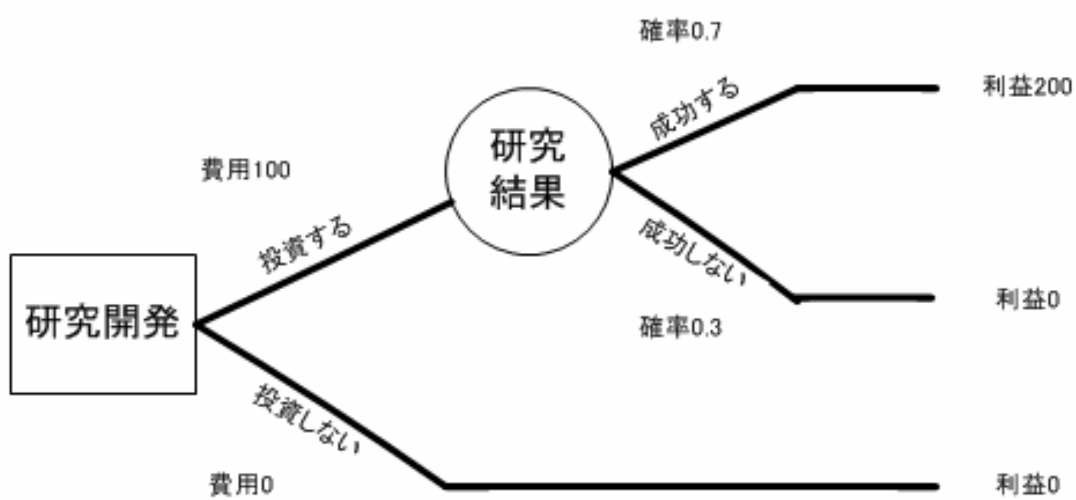
研究が成功する割合は70%である

例題

研究投資には100億円が必要

研究が成功すると利益が200億円見込める

研究が成功する割合は70％である

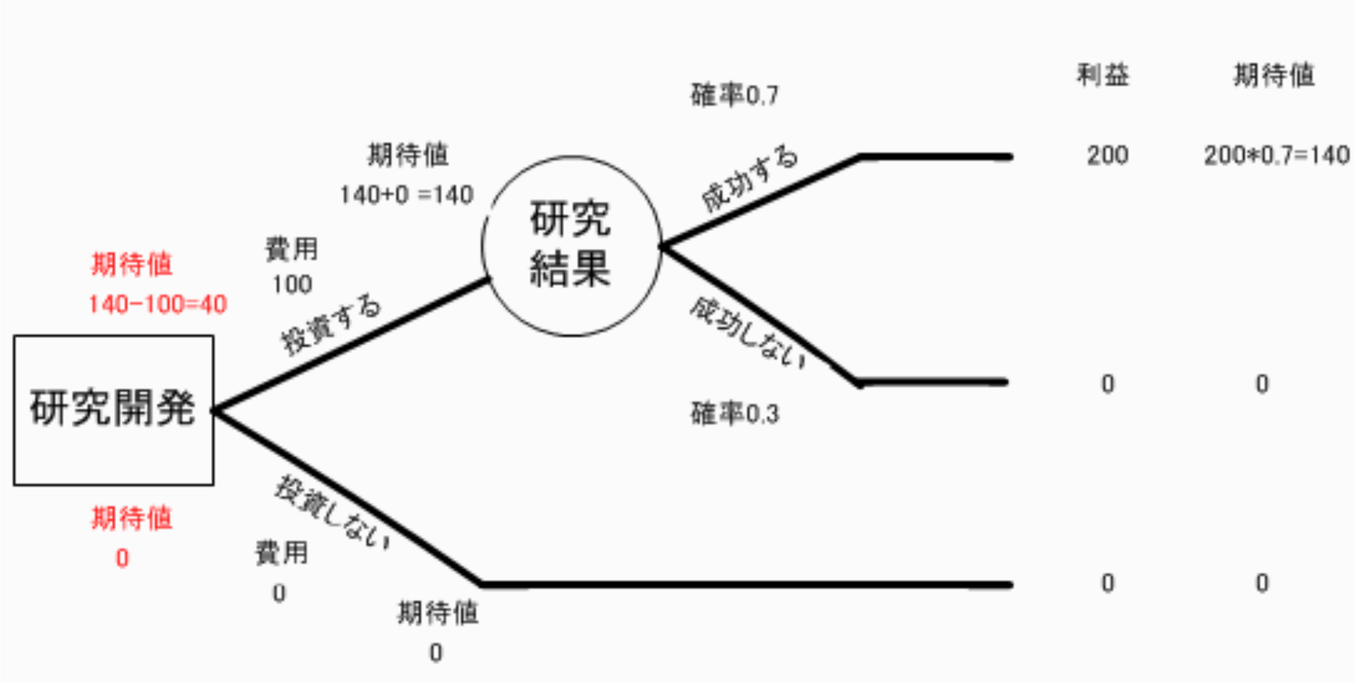


例題

研究投資には100億円が必要

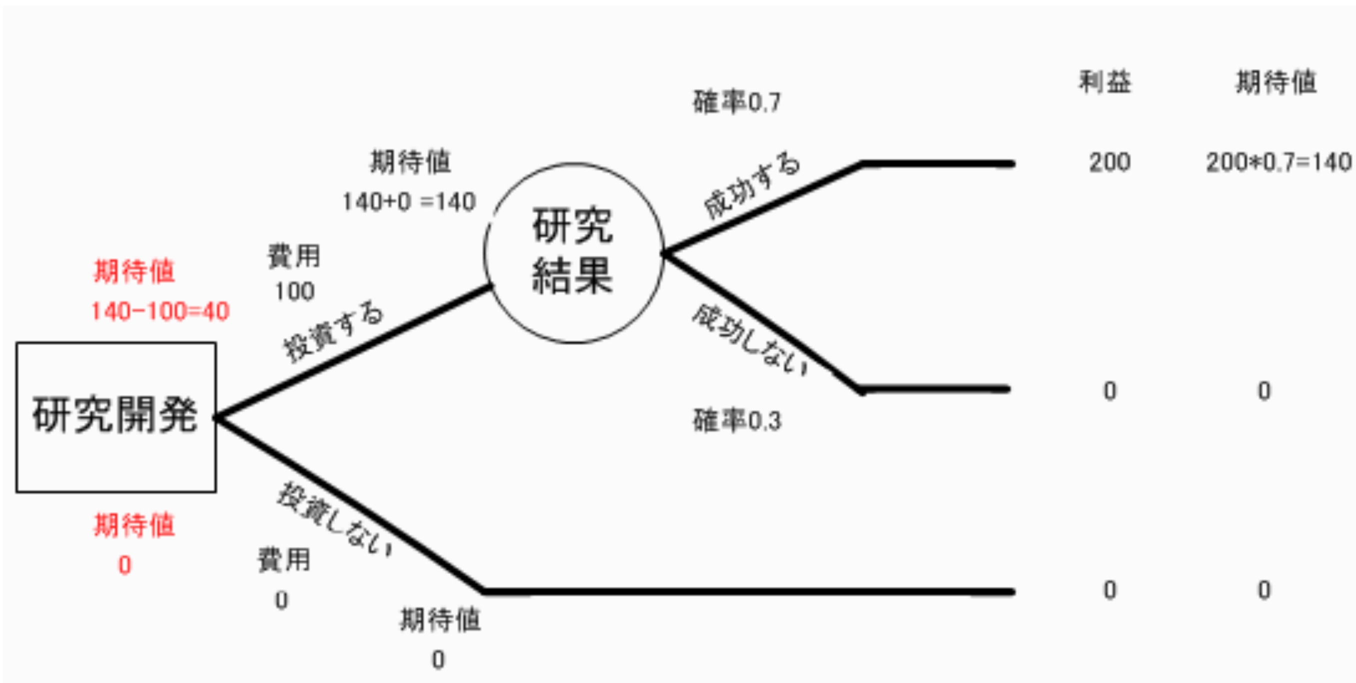
研究が成功すると利益が200億円見込める

研究が成功する割合は70%である



結論

投資したほうが40億円期待値が高いので投資せよ。



演習問題1

A社を100億で買収するかどうかを検討する.

A社は確率 0.6で10年で150億が見込まれるが, 確率0.1で100億, 確率0.3で倒産し利益0となる.

買収するのがよいか？　それともしないことがよいか？

演習問題2

100億の投資資金がある.

A社かB社に投資するかどうかを考えている.

A社に投資するには50億必要であり、その年の景気がよくなれば60億利益が出る. 景気が悪くなれば30億利益が出る. 一方、B社に投資するには100億必要であり、その年の景気がよくなれば160億利益が出る. 景気が悪くなれば80億利益が出る. 投資資金の余った分は、銀行に預け、確実に10%の利子がつく. 景気が良くなる確率を0.5とし、A社、B社に投資するか、もしくはどちらにも投資しない場合のどれが良いかを考えよ。

情報を考える

１．昨日、親友に会いました。

２．昨日、大谷翔平に会いました。

どちらがニュースバリューが高い？

効用関数

$$u(a,\theta) = -\log P(\theta)$$

θの生起確率が低いほど，θの起こったという情報（ニュース価値）が大きくなる.

期待効用関数

$$-\sum_{\theta} P(\theta)\log P(\theta)$$

エントロピーと等価.

どちらが情報量が高い？

１．昨日、親友に会いました。

２．昨日、大谷翔平に会いました。

明日親友にあう確率=1/2

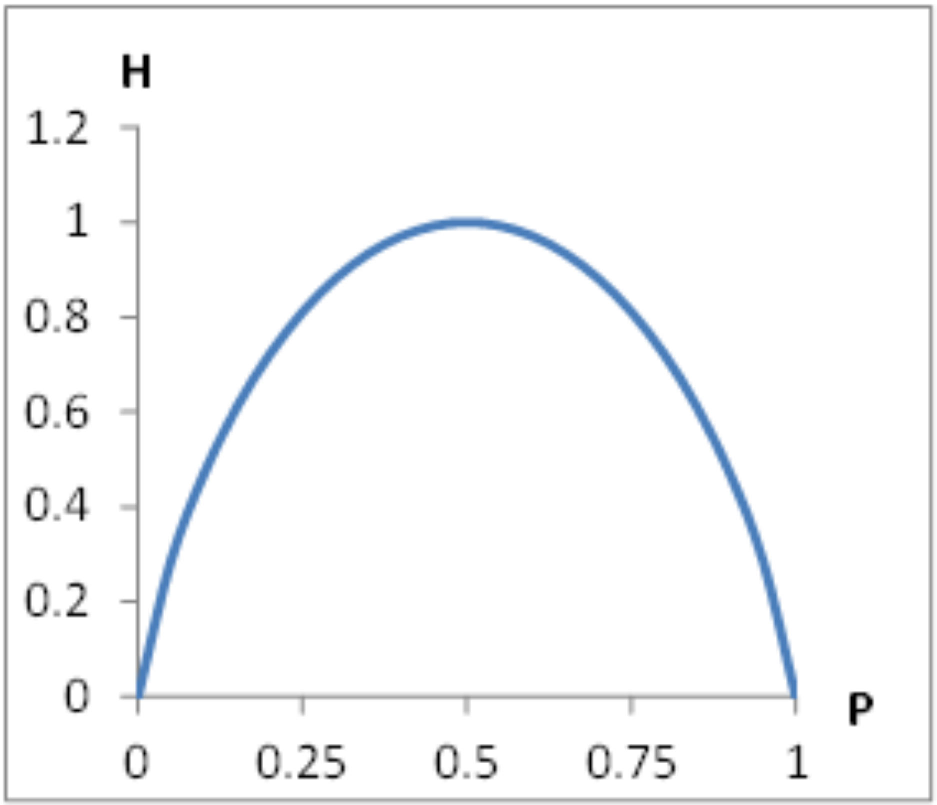
明日大谷翔平にあう確率1/10000

$$
\begin{gathered}
-P(\text{親友に会う})\log P(\text{親友に会う}) \\
-P(\text{親友に会わない})\log P(\text{親友に会わない}) \\
= -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1.0
\end{gathered}
$$

$$
\begin{gathered}
-P(\text{大谷翔平に会う})\log P(\text{大谷翔平に会う}) \\
-P(\text{大谷翔平に会わない})\log P(\text{大谷翔平に会わない}) \\
= -\frac{1}{10000}\log_2\frac{1}{10000} - \frac{9999}{10000}\log_2\frac{9999}{10000} = 0.00173
\end{gathered}
$$

二値の変数Aの情報量

$$H(A) = -\sum_{A=0,1} P(A)\log P(A)$$



不確かさ
わからなさ
の指標

期待効用関数としての利得情報量

θの確率分布が$Q(\theta)$から$P(\theta)$に変化したときの変化量を効用とする効用関数は

$$u(a,\theta) = (-\log Q(\theta)) - (-\log P(\theta))$$

期待効用は

$$\sum_{\theta} P(\theta)\left[(-\log Q(\theta)) - (-\log P(\theta))\right]$$

$=\sum_{\theta} P(\theta) \log \frac{P(\theta)}{Q(\theta)}$　　$Q(\theta)$と$P(\theta)$の差を示す。θが連続量の場合は

$=\int_{\theta} P(\theta) \log \frac{P(\theta)}{Q(\theta)}$

利得情報量

θの確率分布が$Q(\theta)$から$P(\theta)$に変化したときの変化量を効用とする効用関数は

$$u(a,\theta) = (-\log Q(\theta)) - (-\log P(\theta))$$

期待効用は

$$\sum_{\theta} P(\theta)\left[(-\log Q(\theta)) - (-\log P(\theta))\right]$$

$=\sum_{\theta} P(\theta) \log \frac{P(\theta)}{Q(\theta)}$　　$Q(\theta)$と$P(\theta)$の差を示す。

θが連続量の場合は

$=\int_{\theta} P(\theta) \log \frac{P(\theta)}{Q(\theta)}$

情報理論では、カルバックライブラー情報量と呼ぶ。

EVSI

EVSI= $E[u(a,\theta)P(\theta|X)] - E[u(a,\theta)P(\theta)]$

データがあるときとないときの期待効用の差　つまり　データの予想価値

Expected value of sample information(EVSI)と呼ばれる.

情報理論では、相互情報量に一致する

$$EVSI = H(A) - H(A|X)$$

情報理論では

$$H(A) - H(A|X)$$

相互情報量(mutual information)

と呼ぶ

通信路容量(Channel capacity)とは通信路の相互情報量の上限

ベイズでは　データXを得ることにより期待される情報

フィッシャー情報量

事後分布$P(\theta|X)$のθを少しだけ変化させて$\theta+h$にする。

このときの情報量利得の効用関数は

$u(a,\theta)=(-\log P(\theta|X))-(-\log P(\theta+h|X))$

期待効用関数は

$$\lim_{h\to 0}\int_{\theta} P(\theta+h|X)\log\frac{P(\theta+h|X)}{P(\theta|X)}$$

は　フィッシャー情報量に一致する。

証明は　上の期待効用関数を二次の項までテイラー展開すればよい

例題2

被害者Xはある日狙撃された。この事象をEとしよう。

命中率8割のスナイパーAと2割のスナイパーBのどちらかが犯人であることが分かっている。今、どちらが犯人かは全くわからない。

それぞれが犯人である確率を求めよ。

回答

AかBかわからないので

$$P(A) = P(B) = \frac{1}{2}$$

$$P(A|E) = \frac{\frac{1}{2} \cdot 0.8}{\frac{1}{2} \cdot 0.8 + \frac{1}{2} \cdot 0.2} = \frac{0.4}{0.5} = 0.8$$

$$P(B|E) = \frac{\frac{1}{2} \cdot 0.2}{\frac{1}{2} \cdot 0.8 + \frac{1}{2} \cdot 0.2} = \frac{0.1}{0.5} = 0.2$$

例題つづき

そのあとさらに２発Xに銃弾が打たれたが２発とも外れた。この事象をEとしてそれぞれが犯人である確率を求めよ。

回答

$$P(A) = 0.8, P(B) = 0.2$$

$$P(A|E) = \frac{0.8 \cdot (1 - 0.8) \cdot (1 - 0.8)}{0.8 \cdot (1 - 0.8) \cdot (1 - 0.8) + 0.2 \cdot (1 - 0.2) \cdot (1 - 0.2)} = \frac{0.032}{0.032 + 0.128} = 0.2$$

$$P(B|E) = \frac{0.2 \cdot (1 - 0.2) \cdot (1 - 0.2)}{0.8 \cdot (1 - 0.8) \cdot (1 - 0.8) + 0.2 \cdot (1 - 0.2) \cdot (1 - 0.2)} = \frac{0.128}{0.032 + 0.128} = 0.8$$

例題つづき

新たな容疑者としてスナイパーCが浮上してきた。Cの命中率は4割である。A,B,Cの誰が犯人かわからない。最初に命中、そのあと2回外れたデータより、それぞれが犯人である確率を求めよ。

回答

$$P(A)=\frac{1}{3}, P(B)=\frac{1}{3}, P(C)=\frac{1}{3}$$

$$P(A|E)$$

$$=\frac{\frac{1}{3}\cdot 0.8\cdot(1-0.8)\cdot(1-0.8)}{\frac{1}{3}\cdot 0.8\cdot(1-0.8)\cdot(1-0.8)+\frac{1}{3}\cdot 0.2\cdot(1-0.2)\cdot(1-0.2)+\frac{1}{3}\cdot 0.4\cdot(1-0.4)\cdot(1-0.4)}$$

$$=\frac{0.032}{0.032+0.128+0.144}=0.10526$$

$$P(B|E)$$

$$=\frac{\frac{1}{3}\cdot 0.2\cdot(1-0.2)\cdot(1-0.2)}{\frac{1}{3}\cdot 0.8\cdot(1-0.8)\cdot(1-0.8)+\frac{1}{3}\cdot 0.2\cdot(1-0.2)\cdot(1-0.2)+\frac{1}{3}\cdot 0.4\cdot(1-0.4)\cdot(1-0.4)}$$

$$=\frac{0.128}{0.032+0.128+0.144}=0.421053$$

$$P(C|E)$$

$$=\frac{\frac{1}{3}\cdot 0.4\cdot(1-0.4)\cdot(1-0.4)}{\frac{1}{3}\cdot 0.8\cdot(1-0.8)\cdot(1-0.8)+\frac{1}{3}\cdot 0.2\cdot(1-0.2)\cdot(1-0.2)+\frac{1}{3}\cdot 0.4\cdot(1-0.4)\cdot(1-0.4)}$$

$$=\frac{0.144}{0.032+0.128+0.144}=0.473684$$

スナイパーCがあやしい

尤度

スナイパーA,B,Cのデータパターン E=(命中、外れ、外れ)が出る確率 $P(E|A)$, $P(E|B)$, $P(E|C)$ を求めた。

これらを「尤度」と呼ぶ。事前確率を考えず、尤度だけを考えるフィッシャーたちの学派を尤度派と呼ぶ。

例題 3「*Parade*」雑誌の読者相談

1990年、ニュース雑誌「*Parade*」にてマリリン・ボス・サヴァント（IQ304:世界最高でギネス記録）が連載するコラム「マリリンにおまかせ」で、読者投稿があった.

＜投稿された相談＞

プレーヤーの前に閉じた3つのドアがあって、1つのドアの後ろには景品の新車が、2つのドアの後ろには、はずれを意味するヤギがいる。プレーヤーは新車のドアを当てると新車がもらえる。プレーヤーが1つのドアを選択した後、司会のモンティが残りのドアのうちヤギがいるドアを

開けてヤギを見せる。ここでプレーヤーは、最初に選んだドアを、残っている開けられていないドアに変更してもよいと言われる。ここでプレーヤーはドアを変更すべきだろうか？

マリリンの回答

「正解は『ドアを変更する』である。なぜなら、ドアを変更した場合には景品を当てる確率が2倍になるからだ」

→

著名な数学者を含む読者からの「彼女の解答は間違っている」との約1万通の投書が殺到し、本問題は大議論に発展した.

文献

Wikipedia: https://ja.wikipedia.org/wiki/

%E3%83%A2%E3%83%B3%E3%83%86%E3%82%A3%E3%83%BB%E3%83%9B%E3%83%BC%E3%83%AB%E5%95%8F%E9%A1%8C#

ヒント

$P(E|A$が当たり$)=?$
$P(E|\mathrm{B}$が当たり$)=?$
$P(E|C$が当たり$)=?$

ベイズの定理を用いて

$P(A$が当たり$|E)=?$

$P(\mathrm{C}$が当たり$|E)=?$

ベイズでの解

A, B, Cのドアがあるとする．簡単のためAのドアを選んだとする．司会者がドアBをはずれとして開けたとし，この事象をEと書く．今，どのドアが当たりかはわからないので、それぞれのドアが当たりの確率は

$P(A\text{が当たり})=\frac{1}{3}$, $P(B\text{が当たり})=\frac{1}{3}$, $P(C\text{が当たり})=\frac{1}{3}$.

$P(E|A\text{が当たり})=\frac{1}{2}$, $P(E|\mathrm{B}\text{が当たり})=0$, $P(E|C\text{が当たり})=1$

$$P(A\text{が当たり}|E)=\frac{P(A\text{が当たり})P(E|A\text{が当たり})}{P(A\text{が当たり})P(E|A\text{が当たり})+P(B\text{が当たり})P(E|B\text{が当たり})+P(C\text{が当たり})P(E|C\text{が当たり})}=\frac{1}{3}$$

$$P(\mathrm{C}\text{が当たり}|E)=\frac{P(C\text{が当たり})P(E|C\text{が当たり})}{P(A\text{が当たり})P(E|A\text{が当たり})+P(B\text{が当たり})P(E|B\text{が当たり})+P(C\text{が当たり})P(E|C\text{が当たり})}=\frac{2}{3}$$

ベイズは合理的な意思決定に利用できる.

例題4 (3 囚人問題)

ある監獄にアラン，バーナード，チャールズという3 人の囚人がいて，それぞれ独房に入れられている．3 人は近く処刑される予定になっていたが，恩赦が出て3 人のうち1人だけ釈放されることになったという．誰が恩赦になるかは明かされておらず，それぞれの囚人が「私は釈放されるのか?」と聞いても看守は答えない．囚人アランは一計を案じ，看守に向かって「私以外の2 人のうち少なくとも1 人は死刑になるはずだ．その者の名前が知りたい．私のことじゃないんだから教えてくれてもよいだろう?」と頼んだ.すると看守は「バーナードは死刑になる」と教えてくれた．それを聞いたアランは「これで釈放される確率が1/3 から1/2に上がった」とひそかに喜んだ．果たしてアランが喜んだのは正しいのか？

ヒント

アランが釈放されることをA,バーナードが釈放されることをB，チャールズが釈放されることをCと書く。今，誰が釈放されるかはわからないので、

$P(A)=\frac{1}{3}$, $P(B)=\frac{1}{3}$, $P(C)=\frac{1}{3}$. 看守の証言をEとする.

$P(E|A)=?$, $P(E|\mathrm{B})=?$, $P(E|C)=?$

$$P(A|E)=\frac{P(A)P(E|A)}{P(A)P(E|A)+P(B)P(E|B)+P(C)P(E|C)}$$

を求めよ。

回答

アランが釈放されることをA,バーナードが釈放されることをB，チャールズが釈放されることをCと書く。今，誰が釈放されるかはわからないので、

$P(A)=\frac{1}{3}, P(B)=\frac{1}{3}, P(C)=\frac{1}{3}$．看守の証言を$E$とする.

$P(E|A)=\frac{1}{2}, P(E|\mathrm{B})=0, P(E|C)=1$

$$P(A|E)=\frac{P(A)P(E|A)}{P(A)P(E|A)+P(B)P(E|B)+P(C)P(E|C)}=\frac{1}{3}$$

事前分布を変えてみよう

アランのそれぞれの事前確率は

$P(A) = \frac{3}{5}, P(B) = \frac{1}{5}, P(C) = \frac{1}{5}$

であった。この時、$P(A|E)$を求めよ。

回答

$$P(A)=\frac{3}{5},\ P(B)=\frac{1}{5},\ P(C)=\frac{1}{5}$$

$$P(E|A)=\frac{1}{2},\ P(E|\mathrm{B})=0,\ P(E|C)=1$$

$$P(A|E)=\frac{P(A)P(E|A)}{P(A)P(E|A)+P(B)P(E|B)+P(C)P(E|C)}$$

$$=\frac{\frac{3}{5}*\frac{1}{2}}{\frac{3}{5}*\frac{1}{2}+\frac{1}{5}*0+\frac{1}{5}*1}=\frac{3}{5}$$

$$P(C|E)=\frac{2}{5}$$

例題５

夜、一台のタクシーがひき逃げをしました。この市では、緑タクシーと青タクシーの二社が営業しています。事件とタクシー会社については次の情報が考えられています。

- 目撃者は、タクシーが青だったと証言している。事件当夜と同じ状況で目撃者の信頼性をテストした結果、この目撃者は青か緑かを80%の確率で正しく識別した。
- 市内を走るタクシーの８５％は緑タクシーで１５％が青タクシーである。

青タクシーがひき逃げした確率はいくらでしょうか？

解答

犯人が緑タクシーの場合をG, 青タクシーの場合をBとする.目撃者のタクシーが青だったという証言をEとする.

$$P(\mathrm{B}|\mathrm{E}) = \frac{P(\mathrm{E}|\mathrm{B})P(B)}{P(\mathrm{E}|\mathrm{G})P(G) + P(\mathrm{E}|\mathrm{B})P(B)}$$

$$= \frac{0.8 \times 0.15}{0.2 \times 0.85 + 0.8 \times 0.15}$$

$$\approx 0.41$$

人間の直感的予測

ハーバードなど有名文科系の大学生の大半が８０％と回答した。

人間の直感的予測では、事前分布が無視されやすい。

再掲：ベイズの定理(一般化された記述)

データXが得られたときのC_iの確率

事前確率　データの出る確率（尤度）

事後確率

$$\boxed{P(C_i|X)} = \frac{\boxed{P(C_i)}\boxed{P(X|C_i)}}{\boxed{\sum_{i=1}^{n} P(C_i)P(X|C_i)}}$$

C_iについて定数

一方、尤度のみを考える古典的統計学者は事前分布を完全に無視して（ないものとして）推論する。

例題6

以下の赤玉の出る確率はそれぞれいくらか？

1. 50個の赤玉と50個の白玉が入った壷から一つ玉を取り出す。
2. 赤玉と白玉が合わせて100個入った壷から一つ玉を取り出す。

例題６の回答

Beliefでは　１，２ともに赤玉と白玉の出る確率が等しいので1/2となる。頻度論では2の確率は求められない。

再掲：期待効用が等価になるように Beliefが求められる

$u(a,\theta)$: 行動a, 変数θのときの効用関数

期待効用

$$\mathrm{E}[u(a,\theta)] = \int u(a,\theta)p(\theta)d\theta$$

データxを所与としたとき，予測分布$p(y|x) = \int p(y|\theta)p(\theta|x)d\theta$を用いて

$$\mathrm{E}[u(a,y)] = \int u(a,y)p(y|x)dy$$

問題

人間は期待効用が正しく計算できない場合が多い・

例題7

リンダは31歳の独身女性。外交的で大変聡明である。専攻は哲学だった。学生時代には、差別や社会席偽の問題に強い関心を持っていた。また、反核運動に参加したことがある。

次のうちどちらの可能性が高いですか？

リンダは　銀行員である。

リンダは、銀行員でフェミニスト運動に熱心である。

文系学生のほとんどが「リンダは、銀行員でフェミニスト運動に熱心である。」と答えた。

理系学生は「リンダは、銀行員である。」と答える人が多かった。

人間の直感的予測

人間の直感的予測は確率計算ではなく

典型性との類似性に基づくようだ。

典型性が高いカテゴリはより高い可能性として推論される。

専門家の予測の怪しさ

Paul E Meehl (1986) Causes and effects of my disturbing little book, Journal of personality assessment 50 , 370-375

Paul E Meehl (2013) Clinical vs. Statistical Prediction:A theoretical analysis and a review of the Evidence. Echo Point Books & Media

臨床医の予測は単純な重回帰分析の予測に勝つことはない。

James Shanteu (1988) Psychological characteristisc and strategies of expert decision makers, Acta psychologica 68, 203-215

裁判官、監査人、病理学者、心理学者、ファイナンシャルプランナーの予測も同様の結果であった。

Json dna and Robyn M. Dawes (2004) The superiority of simple lternatives to regression for social science predictions, Journal of educational and behavioral science 29 317-331

多くの場面で面接試験の将来予測はほとんど当たらない。

ダニエル　カーネマン　ファスト　アンド　スロー　早川書房

株の専門家の将来予測は　ほとんど当たっていない。

問題　Kahneman and Tversky 1979

あるくじがある。確率πで賞金xがあたり、確率1-πで賞金yがあたるくじを<x,y,π>とかく。

1. くじ　a1　<100万円, 0, 0.45>
2. くじ　a2　<50万円, 0, 0.9>

どちらのくじを選ぶか？

問題　Kahneman and Tversky 1979

あるくじがある。確率πで賞金xがあたり、確率1-πで賞金yがあたるくじを<x,y,π>とかく。

1. くじ　a1　<100万円, 0, 0.45>
2. くじ　a2　<50万円, 0, 0.9>

どちらのくじを選ぶか？

多くの人は　a1 < a2

問題　Kahneman and Tversky 1979

あるくじがある。確率πで賞金xがあたり、確率1-πで賞金yがあたるくじを<x,y,π>とかく。

1. くじ　a3 <100万円, 0, 0.001>
2. くじ　a4 <50万円, 0, 0.002>

どちらのくじを選ぶか？

問題　Kahneman and Tversky 1979

あるくじがある。確率πで賞金xがあたり、確率1-πで賞金yがあたるくじを<x,y,π>とかく。

1. くじ　a3　<100万円, 0, 0.001>
2. くじ　a4　<50万円, 0, 0.002>

どちらのくじを選ぶか？

多くの人は　a3 > a4

問題　Kahneman and Tversky 1979

まず1万円が与えられる。そのうえで以下のどちらのくじを選ぶか？

1. くじ　a5 <1万円, 0, 0.5>
2. くじ　a6 <5千円, 0, 1>

どちらのくじを選ぶか？

問題　Kahneman and Tversky 1979

まず1万円が与えられる。そのうえで以下のどちらのくじを選ぶか？

1. くじ　a5 <1万円, 0, 0.5>
2. くじ　a6 <5千円, 0, 1>

どちらのくじを選ぶか？

多くの人は　a5 < a6

問題　Kahneman and Tversky 1979

まず2万円が与えられる。そのうえで以下のどちらのくじを選ぶか？

1. くじ　a7　<-1万円, 0, 0.5>
2. くじ　a8　<-5千円, 0, 1>

どちらのくじを選ぶか？

問題　Kahneman and Tversky 1979

まず2万円が与えられる。そのうえで以下のどちらのくじを選ぶか？

1. くじ　a7　<-1万円, 0, 0.5>
2. くじ　a8　<-5千円, 0, 1>

どちらのくじを選ぶか？

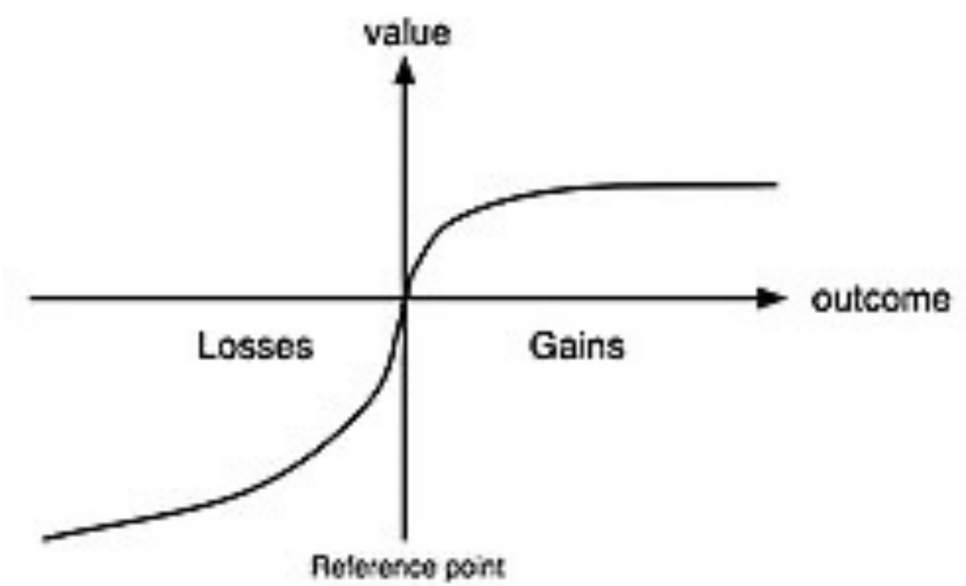
多くの人は　a7 > a8

説明

これら一連の結果が意味することは、人間は目の前に利益があると、利益が手に入らないというリスクの回避を優先し、損失を目の前にすると、損失そのものを回避しようとする傾向(損失回避性)があるということである。

「価値の大きさは金額に比例しない。金額が2倍になると　　、価値は2倍にはならず、2倍弱（1.6倍ぐらい）になる」こう考えると、「2倍の金額を半分の確率で得るよりも1倍の金額を確実に得る」ことの方が利益になるとわかる。また、「損害額を2倍にしても損害の価値（マイナス値）は2倍にはならない」のであれば、2倍の損害のリスクを半分の確率で負う方が利益になる、とわかる。
このように、「価値の大きさは金額に比例しない」というモデルを取ることで、説明が可能となる。

プロスペクト理論
Kahneman and Tversky 1979
効用関数は線形ではない



[引用 https://ja.wikipedia.org/wiki/%E3%83%97%E3%83%AD%E3%82%B9%E3%83%9A%E3%82%AF%E3%83%88%E7%90%86%E8%AB%96](https://ja.wikipedia.org/wiki/%E3%83%97%E3%83%AD%E3%82%B9%E3%83%9A%E3%82%AF%E3%83%88%E7%90%86%E8%AB%96)

まとめ

1. ベイズのbeliefは意思決定理論により決定される
2. 期待効用、期待損失により人間の意思決定を支援し、人工知能にも用いられる
3. 情報理論における様々な情報量もベイズでは意思決定理論である