ニューラルネットワークを用いたケインズ『確率論』の検討

新井 一成[*1]
Email: koffice@u-gakugei.ac.jp

*1：サイエンスフロンティアラボ

◎Key Words　ディープラーニング，意思決定論，Python

1. はじめに

本研究の目的は，確率的意思決定モデルの検討である。不確実な現代において，教育における意思決定論の価値は高まりつつある。サヴェジ，サイモン，カーネマン=トゥベルスキー，ギルボア=シュマイドラー等多分野の研究者が意思決定モデルの構築を試みたが，未だ合意に至る結論は出ていない。

本研究では，決定理論や確率概念が明確に確立する以前の，ケインズ『確率論』(1921)[1]に着目する。

ケインズは計算も比較もできない非数値的確率に着目し，多くの論理式からなる公理系の構築を試みた。そして，それらの論理式にはブールの挑戦問題などの19世紀の諸課題を含み，数理的に解くことが困難である。

本研究はPythonによるニューラルネットワークを用いて論理式の解の範囲を検討する。ケインズ『確率論』の意思決定メカニズムを解明することを通じて，高等教育における初学者向けに他の決定モデルとの違いがわかる教材を開発する。

2. ケインズ『確率論』とその現代的解釈

『雇用・利子および貨幣の一般理論』(1936)で知られる経済学者 J.M.ケインズの研究範囲は，たんに経済学にとどまるものではなかった。ケインズは19世紀までの分析哲学・数学・論理学について幅広く研究し，その集大成ともいえるのが『確率論』(1921)である。

2.1 20世紀初頭の確率理論

19世紀以前, 確率論の体系は確立されていなかったが, ラプラス『確率の哲学的試論』(1814)によって体系づけられた。いわゆる古典確率である。ラプラスは確率を次のように定義した。確率とは，m 個の「ある結果が起こる場合」と，n 個の「可能な場合」の商 m/n である。ラプラスの確率は，現代でいうところの客観的確率と主観的確率の両方の性質を持つ。これは，m や n をどのように計測するか決める必要があるからだ。

「サイコロを１回ふったとき１の目の出る確率」を例に考える。サイコロの目が全部で６種類あり，それらが出る確率はそれぞれ同様に確からしいので，サイコロの１の目の出る確率は 6 通り中１通りなので確率を 1/6 とする。これは主観的確率の考え方である。一方，サイコロを 6 万回ふったところ，約１万回，１の目が出た。そこで回数の相対頻度をとって 10000/60000 = 1/6 とする。これは客観的確率の考え方である。主観的確率はベイズによる先駆的研究ののち，20 世紀に入りラムジーやデ・フィネッティにより掘り下げられた。またサヴェジにより意思決定論の分野に取り入れられ，プロスペクト理論の成立に大きな影響を与えた。また客観的確率はエリスとヴェンによる先駆的研究ののち，20 世紀に入りミーゼスが極限を用いた頻度理論を完成させた。また20世紀にはこれらの議論とは別にコルモゴロフによる公理論的確率が提唱され，確率の定義そのものが修正され現代数学との親和性が高まった。このように，ケインズの『確率論』は，確率論の百花繚乱であった20世紀初頭の過渡期に執筆されており，現在主流となっている確率概念とは大きく離れたものであった。

2.2 ケインズ『確率論』と論理確率

『確率論』で提唱される確率はいわゆる論理確率である。確率とは，ある命題から別の命題に至るまでの，ひとの推論に付随するという解釈を基本とする。つまり，前提命題と結論命題の間の，必ずしも確実とは言い切れない推論の信頼性に関する合理的信念の度合いこそが確率であるとし，ケインズはこれを「確率-関係(Probability-relation)」と呼んだ。現代的に解釈するにはこれは写像にあたる。

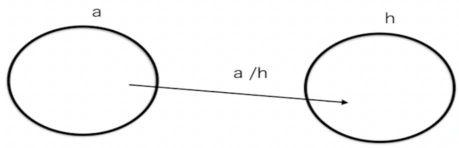

図 1　著者によるケインズ確率の図示
（前提命題を a，結論命題を h,確率関係を α とすると $\alpha=a/h$）

古典数学において，前提の真偽から結論の真偽は論理的に導くことができるが，哲学的命題を含む一般の命題については論理的に導ける保証がない。ケインズはそうした一般的な命題さえも確率論の範疇に取り込もうとした。すなわちケインズは, 確率をたんに数学の一分野としてではなく，ひとの認識に関する分析哲学の一分野として扱おうと試みた。ケインズ以前にも, すでに確率の哲学的議論は多く行われてきたが，数理モデルの中に認識論的要素を取り入れた点が,『確率論』の画期的な点である。

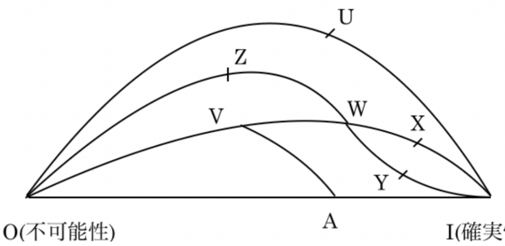

図 2　ケインズ(1921)による確率-関係の図示

ケインズ『確率論』が従来の理論の拡張系にあたることは，ケインズ自身が図示した図２からもうかがえる。図２において，O は確率 0 を，I は確率 1 を示している。数直線 O-A-I 上にある確率は数値的確率であり，必ず何らかの数値をもつ。その一方，曲線 O-V-W-X-I や曲線 O-Z-W-Y-I のように確率には複数のシリーズがあり，シリーズ上の確率は数値で表現することが不可能である。ただし数値表現が不可能であるものでも，他の確率との蓋然性の比較をすることが可能なものもある。例えば W の確率は数値表現できないが，V，Z の確率より大きく X，Y の確率より小さい。また U のように他のどの確率とも比較不可能な確率もある。これは現代的には O を下限, I を上限とする半順序束であると解釈できる。

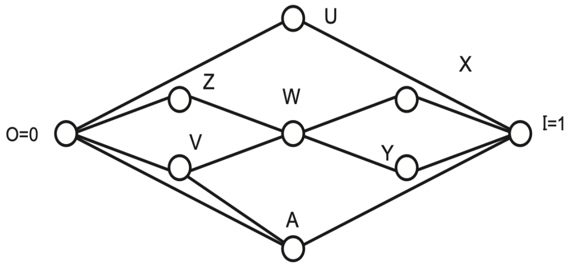

図 3　村田(2009)[2]による半順序束のハッセ図
（ただし内部に線形順序束 O-A-I を含む）

もう一つ，ケインズの確率には特徴的な点がある。それは, 確率の数値が決定するまでの間の推論自体に, 重みづけが行われる点である。推論の重みは有利な証拠と不利な証拠の比率によって決まり, 推論の重みが高いほど, 導き出された確率が「真の確率」である可能性が高まることを意味する。

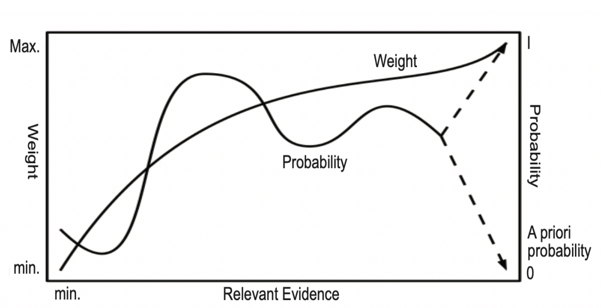

図 4 オドンネル(1989)[3]による推論の重みモデル

これは後世のシャノン(1948)[4]の情報エントロピー理論や，デンプスター＝シェーファー(1967,1976)[5][6]の証拠理論とも関連が深い。

2.3　『確率論』におけるブールの挑戦問題

『確率論』は５部 33 章からなる，500 ページ以上の大作である。第１部で確率関係と推論の重みに関する概要を示した。第２部では多くの論理式を用いて公理系を構築し, 確率関係によって表現可能な範囲を模索した。第３部は推論の重みを掘り下げ, 帰納および類比(アナロジー)との関連を探った。第４部ではパース(1883)[7]やムーア(1903)[8]といった分析哲学の大家を引き合いに出しながら, 確率による意思決定モデルの構築を目指した。第５部は第２部で構築した自身の確率体系を，当時主流であった統計的頻度説や主観的確率論のなかに位置付けようとした。このように『確率論』の議論する範囲は多岐にわたるが，本研究では第２部 17 章「逆確率ならびに平均に関する若干の問題」におけるブールの挑戦問題(Boole’s Challenge Probrem)に着目する。これはブールが 1851 年に方程式を雑誌に掲載し，読者に方程式の解を求めるよう出題したのが初出である。ケイリー，ウィルブラハム，マッコールらが解を提出し, それをもとにブールは『思考の法則』(1854)[9]の中で正解を提示した。

$$\frac{[1c_1(1-p_1)-u][1-c_2(1-p_2)-u]}{1-u}=\frac{(u-c_1\,p_1)(u-c_2\,p_2)}{(c_1\,p_1+c_1\,p_2-u)}$$

式 1　ブールの挑戦問題

ブールは式 1 をもって, 解はc_1p_1とc_2p_2より大きく, かつ$1-c_1(1-p_1)$と$1-c_2(1-p_2)$と$c_1p_1+c_2p_2$の間にあるとした。しかし，ケインズはその正解こそ間違いであり，自身の公理系を用いて正しい解は$u=(c_1p_1+c_2p_2)-y$であると主張した。とくに，公理系において論理式(24)で表される加法定理の誤った適用により，ブールは正しい解に辿り着けなかったと考察した。なおブレディ(2004)[10]による別解釈もあり，雑誌に提出された解が正しいのか, ブールが正しいのかケインズが正しいのか今なお結論は出ていない。挑戦問題に関する学説史は拙稿(2016)[11]に譲る。『確率論』におけるケインズの解が正しいかどうかによって，ケインズの提案する確率の公理が, 挑戦問題などのほかの立場から現れた確率論上の課題を解くために有用かどうかの判断につながるため，誰の解答が正しかったのか明らかにする数学的意義は大きい。

しかし，議論の整理に必要な前提条件が倫理学や認識論といった多くの哲学的分野にまたがっており, かつ, ケインズ自身が示した論理式の展開が極めて難解であるため, 解析的に解を導出することは非常に困難といえる。

3.　ニューラルネットワークを用いたモデリング

前章ではケインズ『確率論』が置かれた状況を整理した。確率論史において重要な意味をもつブールの挑戦問題の正しい解を，学説史的に解決したり解析的に解決したりすることは現状難しそうであることがわかった。そこで，シミュレーションによる解決を提案する。

3.1　『確率論』の課題

ケインズ『確率論』の最大の問題点は, 第 1 部で「確率」とともに提示された「推論の重み」に関する議論が，第２部以降でほとんど記述されていない点にある。これは, 数値化できない確率を導く際に，数値化できない推論の重みを割り振ることで複雑化した確率関係を用いて，現実世界の，理論上は数値化できるとされる確率論の諸問題の解決に適用しようとしたことに起因している。ケインズは自らの公理系の有用性を示そうとするあまり，「確率」とともに提示した「推論の重み」の提示を控えざるを得ず, そのため『確率論』全体で議論の簡明さが失われてしまったと推察される。そこで, 前提命題から結論命題を導くメカニズムの可視化を試みる。

3.2 Prologによる確率関係の整理

Prolog(Programing Logic)とは論理プログラミング言語のひとつであり，述語論理学の発想に基づいて設計されている，コンピュータ上で記号処理を行う言語である。すなわち，R 言語や Python といった処理の手順を記述する一般的な手続き型プログラミングとは異なり，手続きをこちらで決めることなく，値を宣言的に決めた後，望む結果について質問することでプログラミングができる。Prolog のプログラムは事実(ホーン節)・規則(ルール節)・質問(ゴール節)の 3 つの役割に分けられる。

```
ksmall(0,prob(u/h)).
ksmall(0,prob(z/h)).
ksmall(0,prob(v/h)).
ksmall(0,prob(a/h)).
ksmall(prob(z/h),prob(w/h)).
ksmall(prob(v/h),prob(w/h)).
ksmall(prob(v/h),prob(a/h)).
ksmall(prob(w/h),prob(x/h)).
ksmall(prob(w/h),prob(y/h)).
ksmall(prob(u/h),1).
ksmall(prob(x/h),1).
ksmall(prob(y/h),1).
ksmall(prob(a/h),1).
```

```
keynesorder(X,Z) :-
ksmall(X,Z).
keynesorder(X,Z) :-
ksmall(X,Y) , keynesorder(Y,Z).
```

```
?- keynesorder(prob(x/h),prob(y/h)).
false.

?- keynesorder(prob(w/h),X).
X = prob(x/h) ;
X = prob(y/h) ;
X = 1 ;
X = 1 ;
false.
```

図 5 拙稿(2010)[12]によるケインズ確率関係の表現
(上から順にホーン節，ルール節，ゴール節)

図 5 は，図 3 の半順序束による確率関係の解釈に基づき著者がプログラムしたものである。ゴール節で確率の順序関係を質問すると，自身と比較可能なものがある場合全てを網羅的に解として返す。また任意の 2 つの確率の順序関係を確認することが可能で，順序比較が不可能な場合には Prolog は false を返す。Prolog により，ケインズの確率関係については可視化することが可能といえる。

3.3 区間推定による意思決定モデルの検討

確率関係については，半順序束の性質を用いることで，Prolog を用いた可視化が可能となった。次に，確率関係と推論の重みの関係性について，区間推定の観点から整理する。

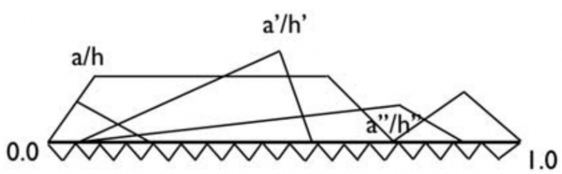

図 6 キーバーグ(2000)[13]による確率関係の解釈

図 6 はベイズ確率の信念関数のモデリングを行う SIPTA プロジェクトにおいて，キーバーグが示した『確率論』の解釈である。原著にある図 2 との相違点は，確率関係をつねに上限・下限をもつ概念として扱っている点にある。上限と下限の数値が一致する場合にのみ，ケインズのいう確率は数値解をもつという。これは統計学上の区間推定・点推定の関係にあたる。推論の重みが最大のとき，上限と下限は一致する。キーバーグは図 2 の数値解の数直線集合 O-A-I の代わりにノコギリ状の集合を想定した。これは数直線上を，任意の微小区間において区間推定可能な集合として扱うことで，つねに点推定可能な集合とみなしている。本研究もキーバーグの解釈に従う。この解釈によれば，推論の重みが最も軽いとき，確率関係は上限 1，下限 0 の区間推定で表現できる。反対に推論の重みが最大のとき，確率関係はある数値のまわりの点推定で表現できる。この解釈に従えば，推論の重みとは，確率関係を決定する際の誤差分布のようなものと考えられる。

3.4 Pythonによる推論の重みの表現

本研究では，キーバーグの解釈に基づくモデリングのため，Python を用いてニューラルネットワークの構築を行う。Pyhton は TensorFlow や Keras をはじめとするディープラーニング関連のライブラリが豊富であり，ブラウザ上の実行環境である Jupyter Notebook や Jupyter Lab をもつので検証環境に恵まれている。また数理モデルとしてニューラルネットワークを選択した。ニューラルネットワークは最初からある値の数値解を決めて収束させるのではなく，数値解を特に決めずに，逆誤差伝播によって徐々に値の修正を行うモデルである。これはデンプスター＝シェーファーの証拠理論の表現ともいえ，従って，ポストケインジアンであるオドンネルが提案した推論の重み解釈(図 4)とも矛盾しない。

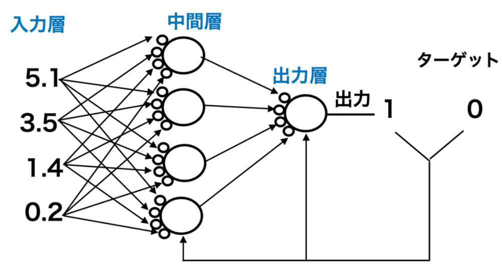

図 7 ニューラルネットワークのモデリング
(金丸(2018)[14]を参照して著者作成)

なお，高籔(1993)[15] のように『確率論』のニューラルネットワークによるモデリングはすでに存在する。近年，ニューラルネットワークは AI の意思決定に幅広く用いられている。AI は，ひとが最適解を算出するのが容易ではない事象や，一見すると主観的確率も客観的確率もよくわかっていない事象に対しても一定の解答を与える。このことから，『確率論』の意思決定モデルとしてニューラルネットワークを用いることは妥当であるといえる。『確率論』が 100 年後の現在に執筆さ

れていたら，ケインズは推論の重みの説明にニューラルネットワークを採用したのであろうか。

```python
def make_generator_model():
    kernel_size = (5, 5)
    noise_vector_dim = 100  # ノイズのベクトルサイズ

    model = tf.keras.Sequential()

    model.add(layers.Dense(units=32*32*128, use_bias=False,
                           input_shape=(noise_vector_dim,)))
    model.add(layers.BatchNormalization())
    model.add(layers.LeakyReLU())

    model.add(layers.Reshape(target_shape=(32, 32, 128)))
    assert model.output_shape == (None, 32, 32, 128)

    model.add(layers.Conv2DTranspose(64, kernel_size, strides=(2,2),
                                     padding='same', use_bias=False))
    model.add(layers.BatchNormalization())
    model.add(layers.LeakyReLU())
    assert model.output_shape == (None, 64, 64, 64)
```

図 8 ディープラーニングの例 (DCGAN)

3.5　今後の課題

具体的に，ニューラルネットワークを用いて『確率論』をどのようにモデリングすれば推論の重みを区間推定における信頼区間の誤差修正にあて，確率関係を点推定の数値解とすることができるか，現在検討中である。拙稿(2022)[16]で扱ったように，複数のディープラーニングを組み合わせることができれば, Prolog のプログラムについてもモチーフを生かすことができるだろう。ただし，塩沢(1983)[17]の指摘にあるように，束論によるモデリングと数値解をもつモデリングの両立は通常であれば困難を伴う。うまく工夫したい。

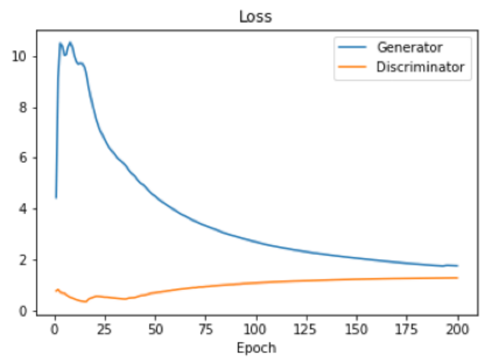

図 9 損失関数の収束

4.　ケインズ『確率論』の高等教育への応用

『確率論』の内容理解はどのような教材の開発につながるのか。検討をもって含意にかえたい。

4.1　教育対象

対象は大学生とする。事前に統計学や意思決定論について前提知識がなく，講義を通じて初めて触れると仮定する。通年ではなく，半期の講義を想定する。

4.2　教育目的

サヴェジやカーネマン=トゥベルスキーといった既存の確率的意思決定モデルの提示ののち，ケインズ『確率論』の意思決定メカニズムを提示する。両者の比較を通じて，以下の内容理解が進むことが望ましい。

- 確率的意思決定には複数のモデリングが可能であること。
- どのモデルにも長所と短所があり，一概に優位性を決められないこと。
- そもそも「確率」の捉え方に複数の立場があり，高校までに学習した，場合の数による数え上げは古典確率にすぎないこと。
- ひとの確率認識には曖昧な点があり，それは AI も同じであること。
- ひとの意思決定の仕方と AI の意思決定の仕方は大きく異なるが，得られる結果が似通うこともあること。
- 上記の理解を通じて，受講前よりも AI の仕組みに詳しくなり，AI を身近なものとして感じること。
- ひとや AI が計算可能な範囲はごくわずかであり，世の中には計算が困難な事象が非常に多いことを知ること。

参考文献

(1) Keynes.J.M.(1921),*A Treatise on Probability*, London;Macmilan. 佐藤隆三訳(2010)『確率論(ケインズ全集 第八巻)』,東洋経済新報社

(2) 村田晴紀(2010)「ケインズ『確率論』における「短期確率加重関数」の研究 ―不確実性下における意思決定の考察―」, 東京学芸大学 2009 年度修士論文 .

(3) O'Donnell, R. M.(1989)KEYNES: Philosophy, Economics & Politics, Palgrave Macmillan.

(4) Shannon, ClaudeE.(1948)"A Mathematical Theory of Communication,"The Bell System Technical Journal, Vol. 27, pp. 379-423,623-656.

(5) Dempster.A.P(1967), Upper and lower probabilities induced by a multi-valued mapping, The Annals of mathematical statistics,38,325-339.

(6) Shafer.G(1976),*A Mathematical Theory of Evidence,*Princeton Univ.Press,Princeton.

(7) Peirce.C.S(1883),"A Theory of Probable Inference",*Johns Hopkins Studies in Logic*.

(8) Moore.G.E(1903), *Principia Ethica,*深谷昭三訳(1977)『倫理学原理』,三和書房.

(9) Boole.G(1854),*An Investigation of the Laws of Thought,on which are founded the Mathematical Theories on Probabilities,*Dorer Publications.

(10) Brady.M.E(2004) *J.M.Keynes' Theory of Decision Making, Induction, and analogy The Role of Interval Valued Probability in His Approach*, Xlibris.

(11) 新井一成・高籔学(2016)「ケインズ『確率論』における「ブールの挑戦問題」の扱われ方の整理 -M.E.ブレディ(2004)を参考に-」『東京学芸大学紀要』67 号,人文社会科学系 II pp.143-149.

(12) 新井一成,高籔学,村田晴紀(2010),「Prolog を用いた教材の提案 -意思決定論の視点から-」『2010 PC Conference 論文集』

(13) Kyburg.Jr.H.E .and Imprecise Probabilities Project(1998-2000),"Interval-Valued Probabilities",SIPTA.

(14) 金丸隆志(2018),『Raspberry Pi ではじめる機械学習 基礎からディープラーニングまで』,講談社.

(15) 高籔学(1993),「証券需要価格の粘着性--ケインズ「確率論」再考」，『修道商学』34 (1), pp.183-202,広島修道大学商経学会.

(16) Arai.K,Hosokawa.M and Kunishima.M(2022),"Analysis of fuzzy images by close-up photography of large industrial drones", International Symposium on Affective Science and Engineering ISASE2022 (0), 1-4, 日本感性工学会.

(17) 塩沢由典(1983),「『確率論』からみたケインズ」,『別冊経済セミナー ケインズ生誕 100 年』, pp.76-82, 日本評論社.