

2018 年度大学院 都市防災学特論

講義題目 建築情報学：確率論の基礎

【補講 1】情報概論

コンテンツとキーワード

1. 確率論を学ぶ必要性を考える
 - (1) 工学の目的と方法論
 - (2) 不確定性の多面的意味
蓋然性、乱雑性、偶然性、信頼性、認識論的不確定性、現象論的不確定性、
統計学的不確定性
 - (3) 統計学における確率の 4 つの考え方
頻度論的確率論、数学的確率論、主観的確率論、ゲーム論的確率論
2. 情報と不確定性の関係
 - (1) 情報に関する 2 つのモデル
DIKW ヒエラルキー、シャノン
 - (2) シャノンによる情報量の定義
自己情報量
 - (3) 2 進数の持つ情報量
bit、byte、word、long word
 - (4) シャノンによる情報エントロピーの定義
情報エントロピー、平均情報量
3. 情報の意味を考えよう
 - (1) 演歌歌詞と JPOP 歌詞のエントロピーを計算してみよう
 - (2) 熱力学的エントロピーと情報エントロピーの関係について
4. 本講のまとめ

1. 確率論を学ぶ必要性を考える

(1.1) 工学の目的と方法論

建築を学ぶ者が、なぜ確率論を学修する必要があるのか考えたことがあるだろうか。建築は芸術性が重要視されるとは言え、その目的と方法論は工学の主幹をなぞっており、その一分野を構成していることは疑いがない。図 1 は工学が扱う「対象」「境界領域」「目的」「方法論」「最終目標」の相互関係を示したものであり、建築学にも通用する。図では対象が【複

雑な社会】となっているが、建築構造学を学ぶ者は【構造物】、建築住居学を学ぶ者は【住宅】、都市計画学を学ぶ者は【都市・地域社会】と置き換えることで自分の問題に意識転換できるであろう。要は、暮らしやすさの実現という目的を達成するために、状態(現象)をモデル化し、シミュレーションすることで最適解を発見するという方法論は同じである。そのプロセスにおいて状態のモデル化が必須である。物理学で言うモデル化は観測事象のノイズを捨て去った理想状態を数学的に記載すること

であるが、工学は現実の状態を再現する必要性があることから、理想状態のみをモデルと称して扱うわけにはいかない。たとえば、構造物の地震応答をバネーマス系の振動モデルで単純化し、2階常微分方程式の解として求めることは良くあることではあるが、解が実応答と乖離することは経験済みである。単純モデルには記載しにくいノイズ(不確定性)が現象には必ず付随する。工学は実像(実社会)を対象とするのだから、理想状態に基づく決定論的予測は厳密には成立しない。ノイズ(不確定性)を含めた実像が対象となり、それを扱う。だから、モデル化に際し不確定性を扱う理論(確率論)が必要となってくる。

(1.2) 不確定性の多面的意味

不確定性について、もう少し理解を深めよう。不確定性(uncertainty)とは、不確実さの程度という意味において、以下の2つを含蓄する。

- 1) 蓋然性(probability)
- 2) 乱雑性(randomness)

蓋然性とは確からしさの程度を言い、哲学的意味合いを含んでいる。それに対して、乱雑性とは規則性のなさを意味し、科学的意味合いを含んでいる。似たような言葉に、偶然性(contingency)があり、これは偶発性を意味し、蓋然性と乱雑性の両方の意味にまたがった言葉である。

われわれは、不確定性という言葉をもっと一つの意味で使うこともある。それは信頼性(reliability)であり、対象とする事象の信憑性あるいは証明可能性を言う。

このように、不確定性は本質的に異なる多面的意味を持っている。われわれはこれらを包含した意味で不確定性を扱うが、いま扱っている事象の不確定性の本質はそのうちのどれに該当するかを、常に意識しておく必要があるし、そのことが事象を扱う際に混乱を防ぐ唯一の方法である。

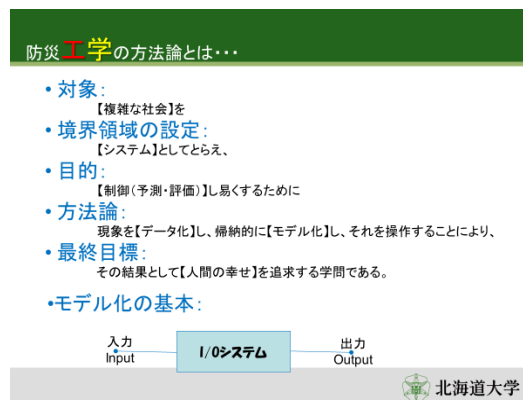


図1 工学の目的と方法論

このことを、技術者である A. H-SANG と W.H.TANG¹⁾はその著書の中で、以下のように定義している。

1) 偶然的不確定性（現象論的不確定性）：aleatory uncertainty：現象本来が持つ不安定性に起因する不確定性。上記の乱雑性に近い概念である。

2) 認識論的不確定性：epistemic uncertainty：モデル化に当たっての知識不十分さに起因する不確定性。工学的に重要な不確定性の考え方であり、上記の蓋然性に近い概念である。

さらに、柴田明徳²⁾はその著書において、以下の不確定性についても触れている。

3) 統計的不確定性：statistical uncertainty：サンプルの不十分さに起因する不確定性を言い、工学的には実験観測の適正さあるいは難しさにつきまとう非常に重要な不確定性である。上記の信頼性に近い概念である。

(1.3) 統計学における確率の4つの考え方

確率論を純粋の数学として扱う場合は、以下の立場が明確にされている。

1) 頻度論的確率：試行結果に基づき発生確率値を与える（コルモゴロフの測度論に集約）

2) 数学的確率：基本は等可能性とし、発生確率に数値を与える

3) 主観的確率：自身の信念に立脚し確率値を与える（ベイズの方法に発展）

4) ゲーム論的確率：確率計算を表に出さず、ゲーム戦略論から確率法則を証明する方法
どれも工学的に利用できる方法論・立場であり、結果の信頼性を確保するためにどの立場を取るかは、その研究者の研究哲学に係ってくる重大なテーマである。

【設問1】上記の1)3)4)は工学的手法とも言える。それぞれの代表的事例を紹介し、それぞれのメリット・デメリットを理解せよ。

【解答例】

- 1) 頻度論的確率論の立場：通常のばらつきを観測データより考慮するモデル。平均値と分散が主要パラメータ。
- 3) 主観的確率：ベイズの方法によるもの。サンプル数が確保しにくい稀な現象を扱う場合、主観確率を事前分布として、観測値を尤度とする事後確率分布を求めるアプローチ。地震被害を扱う研究室は、観測事例が少ないためこの立場をとることが多い。
- 4) ゲーム論的確率：イベントツリーによるシナリオ追跡型モデル（被害評価モデル）には分かりやすいモデルとなる。

【設問2】現象論的不確定性（バラツキ）と認識論的不確定性（モデル不十分性）との本質的違いを母集団と標本分布の関係から説明しなさい。

設問2の解答は、次回の講義で解説予定。

¹⁾ A. H-S. Ang, W. H. Tang：改訂 土木・建築のための確率・統計の基礎, pp.506, 丸善株式会社, 2007.

²⁾ 柴田明徳：確率的手法による構造安全性の解析 ―確率の基礎から地震災害予測まで―, pp.275, 森北出版株式会社, 2005.

2. 情報と不確定性の関係

不確定性は情報と深い関係を持っており、数学的に扱うことができる。確率論に入る前に、不確定性と情報との深い関係について、理解しておこう。

(2.1) 情報に関する 2 つのモデル

われわれは情報を扱う。しかし、情報にもその持つ意味・内容や信頼性等々において色々なレベルがあり、扱い方が変わってくる。大きく、情報には以下の二つの見方がありそれぞれを情報モデルとして扱っている。

1) 情報の意味・価値に関する構造モデル：R. Ackoff(1989)の DIKW ヒエラルキー³

2) 情報の形式に関する確率モデル：C. Shannon(1948)の情報理論⁴

テキストは、岡田のホームページからダウンロードできる（終了）。

DIKW ヒエラルキーについて、簡単に説明する（図 2 参照）。情報の価値にはレベルがあり、情報工学ではこれを DIKW ヒエラルキーと呼んでいる。最下位レベルは Data であり、事象から抽出されたままの評価されていないメッセージを言う。このレベルではセンサーやフィルター等を通しての取得（モニタリング）方法がテーマとなることが多い。Data を加工処理することで意思決定に役立つ Information

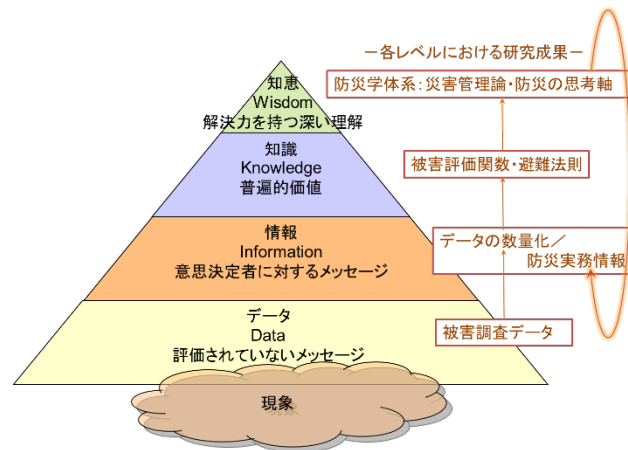


図 2 DIKW ヒエラルキー

に変換される。意思決定者へのメッセージであり、意思決定内容の明確化がそのプロセスにおいて重要な作業となる。その上位にある Knowledge は Information を分析し傾向などを見いだした普遍的な価値を言う。所謂一般法則の導出に代表される。最上位の Wisdom は次の意思決定に進むための様々な解決力を持つ深い理解を意味する。

【設問 3】

- 1) 参考文献 3) を要約しなさい。
- 2) 文献 3) は地震防災分野における DIKW の事例紹介である。各自の研究分野における DIKW の例を挙げなさい。
- 3) また、自分の研究領域の将来像（特に W）について考察しなさい。

³ 岡田成幸、中嶋唯貴：モニタリング技術が防災の形を進化させる、強震観測とモニタリング技術が災害時に果たすべき役割，日本建築学会東海大会パネルディスカッション，2012.

⁴ アブラムソン：情報理論入門、好学社、1969.

(2.2) シャノンによる情報量の定義

参考文献 4) を参照しつつ、情報の数学的扱いについて解説する。シャノンは、ある情報を得た際のそのことによる情報量を、以下の 2 つの要請を充たすことを条件に、自己情報量 I を定義した。

要請 1 : 発生する確率が低いこと（珍しいこと）が分かったときの方が、情報量が多い。

要請 2 : 情報量は足し算で増えていく。

$$\text{情報量 } I = -\log_2 p \quad \cdots (1)$$

ここに、 p は情報の生起確率を意味する。

シャノンは、なぜこのような「情報量」を定義したのであろうか。まず要請 1 はどのような意味を持つのであろうか。シャノンは、「分からない状態」から何らかの情報を得て「分かる状態」になる、これが情報を得ることの意味であると考えた。「分からない状態」と「分かる状態」との差が大きいほど情報量が大きいと考える。情報とはその状態の変化の大きさと捉えたのである。要請 2 は「情報量は加法性を持つ」という意味である。これを理解するために、情報を小出しにすることを考えてみよう。以下、甘利俊一⁵⁾の著書を参考に解説する。

「情報量の加法性」の解説

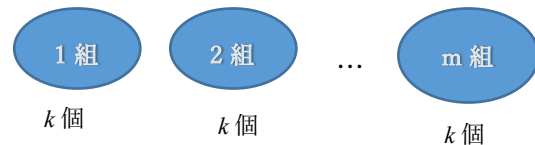
情報群 n 個を k 個からなる m 組に分ける。

1 組 : $A_1, A_2, \dots, A_k$

2 組 : $A_{k+1}, A_{k+2}, \dots, A_{2k}$

...

m 組 : $A_{(m-1)k+1}, A_{(m-1)k+2}, \dots, A_n$



ここで、情報 A_i が持つ情報量 $I_i = f(i)$ と表記すると、

情報 0 : n 番目にある = 情報量 $f(n) = f(mk)$

情報 1 : m 組目にある = 情報量 $f(m)$

情報 2 : m 組目の k 番目にある = 情報量 $f(k|m)$

いきなり正解の情報を与えた場合の情報量は $f(n)$ に相当し、小出しの場合の情報量は、 $f(m)$ を出し、次に $f(k)$ を出すことに相当する。これを式で表現すると、

$$\text{情報 0} = \text{情報 1} + \text{情報 2} = f(mk) = f(m) + f(k|m) \cdot f(m) = f(m) + f(k)$$

$$\text{i.e. } f(mk) = f(m) + f(k)$$

情報量 I は加法性が成立し、かつ上記等式（変数 m, k の積が和の形式に変換）が成立する関数型は対数であることが分かる。

⁵⁾ 甘利俊一：情報理論，ちくま学芸文庫，2011.

【設問 4】 p はその情報の生起確率である。なぜ、生起確率の負の対数が情報量になるのかを説明しなさい。

【解答例】

生起する確率が小さい事象が発生した際、それを知ったときに得られる情報は大きいと考えた（要請 1：たまたま知った事象は情報量が大きい。いつでも得られる情報は情報として価値が小さいと考えた）からである。情報を生起確率の関数と考えることで情報量を数学的に扱える。そして、生起確率が小さいほど情報量が大きくなる関係は生起確率の逆数 $1/p$ と捉えることで、その関係性を表現できる。対数には加法性（分数を足し算で表現することができるので）があるので、理論展開が簡単になる。

【設問 5】 なぜ対数の底を 2 にしたのかを説明しなさい。

【解答例】

対数の底を 2 にしたのは、生起確率が 50% ($p=1/2$) の時（生起性に関して一番曖昧な状態の時）に、情報量を 1 に正規化するためである。また、このように定義することで、単位を bit（2 進数表記）に定めることができる。すなわち、情報量を 2 進数の bit 符号がいくつ必要かという、情報量伝送に関する物理的意味も明確とすることができるからである。

(2.3) 2 進数の持つ情報量

1～10 の数字は 2 進数で以下のように表現される。

0=0000 1=0001 2=0010 3=0011 4=0100 5=0101 6=0110 7=0111 8=1000
9=1001 10=1010

2 進数の基本単位を bit と言い、0 又は 1 で数値表現される。10 進数 0～9 の数字が持つ情報量を 2 進数で表現するには何ビット必要かを計算することで、以下のように数字の情報量を計算することができる。

10 の数字の生起確率を等確率と考えると

$$p = 1/10$$

$$\therefore I = -\log_2(10) = 3.322[\text{bit}]$$

すなわち、3.4bit 相当の情報量を持っているという計算結果となり、3bit では足りず、4 ビットでは余るという結果を与えている。

なお、コンピュータが扱う情報量の基本単位は以下で与えている。

bit（ビット）：1or0

byte（バイト）：8bit

word（ワード）：32bit

long word：double precision 64bit

(2.4) シャノンによる情報エントロピーの定義

一つの情報の持つ「価値」を自己情報量 I で定義した。その情報が伝えところの事象の生起確率が情報の価値を決め、珍しい情報（生起確率が小さな情報）ほど、情報量が大きくなるように、シャノンは定義した。次に、その情報が入っている世界（情報の発信源・情報源）の「価値」を情報エントロピー（平均情報量）として以下のように定義した。

$$H(p_1, p_2, \dots, p_n) = -\sum_{i=1}^n p_i \cdot \log_2 p_i \quad \cdots (2)$$

これは発信者によるその情報群の持つ不確かさ（エントロピー）を数量表現したものである。情報量 $I_i (= -\log_2 p_i)$ はその情報が伝えるところの事象が生じたときの価値（情報量）を表すものであったが、未だ事象が生起していないときにその情報が持つ価値はその発生確率を乗算した量（ $= -p_i \cdot \log_2 p_i$ ）で算定される。よって、その発信源がもつ総量としての「価値」は情報群の期待値（平均情報量）として、すなわち式（2）で示される自己情報量 I_i に確率 p_i で平均したもの（ p_i を乗じて情報数 n で総和する）として算定される。

情報源の価値がなぜ不確かさ（エントロピー）で表すことができるのか、を考えてみよう。

不確かさ(エントロピー)が情報であることの解説

情報の不確かさ（エントロピー）は情報量 I に生起確率を乗算することと定義した。これは情報量に生起確率をかければ、情報量の平均値（その情報量の平均的発生確率で重み付けした量）を求めることに等しい。この平均的情報量（エントロピー）は情報源が持つ不確かさ（あいまいさ）と理解することができる。

例題（天気予報のエントロピー）で考えよう。地方ごとに天気予報の確率が以下のよう発表された。

	晴れの確率	曇りの確率	雨の確率
地方 A	0.5	0.3	0.2
地方 B	0.6	0.3	0.1
地方 C	1/3 (=0.33)	1/3 (=0.33)	1/3 (=0.33)

地方 A の情報エントロピーは式（2）より

$$\begin{aligned}
 H_A &= -0.5 \times \log_2(0.5) - 0.3 \times \log_2(0.3) - 0.2 \times \log_2(0.2) \\
 &= 0.5 \times 1 + 0.3 \times 1.7 + 0.2 \times 2.3 \\
 &= 1.5 \text{ bit}
 \end{aligned}$$

同様に、地方 B の情報エントロピー： $H_B = 1.3 \text{ bit}$

地方 C の情報エントロピー： $H_C = 1.6 \text{ bit}$

∴ エントロピーの大小は $C > A > B$ となり、地方 C の予測が一番難しい（情報の不確かさが一番大きい）ということになる。

ここで、情報を受け取ったという意味を考えてみよう。

一般的に情報量とは以下のように言うことができる。

情報量 = （情報を受け取る前の不確かさ）－（情報を受け取った後、尚残る不確かさ）

天気予報の場合、情報がない状態は地方 C に等しい（すなわち、発生確率は等確率）。

よって、地方 A については、

$$\text{情報量}_A = 1.6 \text{ bit} - 1.5 \text{ bit} = 0.1 \text{ bit}$$

すなわち地方 A において情報（天気予報）を受け取ったことにより、情報源が持つ 1.6 bit

のエントロピー（不確かさ）が情報を得た事により 0.1 減少した。すなわち、0.1bit の情報量を獲得したということになる。

一方、地方 B については、

$$\text{情報量}_B = 1.6\text{bit} - 1.3\text{bit} = 0.3\text{bit}$$

地方 B においては、情報を受け取った後で、0.3 ビット情報源の不確かさが小さくなった。すなわち、0.3bit の情報量を獲得した。

地方 C については、

$$\text{情報量}_C = 1.6\text{bit} - 1.6\text{bit} = 0\text{bit}$$

地方 C については、何ら情報を得ていないということになる。

情報源が持つ情報の不確かさの期待値をエントロピーで表すことにより、不確かさが bit 数で表現することが可能となり、さらに情報を受け取ったことによりその不確かさがどれだけ低減するかという算定を通して、得た情報量の価値が数量化できる。これが、エントロピーが情報に等しいという理由である。

【設問 6】 アルファベットの持つ情報エントロピーを計算しなさい。

【解答例】

アルファベットは (a,b,c,⋯,x,y,z,bl) の 26 文字と空白文字からなるので、27 文字のもつ情報量を計算してみよう。アルファベットの生起確率を等確率 ($p=1/27$) とみなすと

$$\text{アルファベット一文字の持つ情報量 } I = -\log_2 p = -\log_2(1/27) = 4.756$$

アルファベットの世界が持つ情報エントロピーは同様に生起確率を等確率($1/27$)とみなすと、

$$\begin{aligned} H(\text{アルファベット}) &= \sum_{i=1}^n p_i [-\log_2 p_i] \\ &= n \times 1/27 [-\log_2(1/27)] = 27/27 \times 4.756 = 4.756 \quad [\text{bit}] \end{aligned}$$

すなわち、アルファベットを表現するには 4 ビットでは足りない。したがって、アルファベットは 1byte で表現される。

【設問 7】 ひらがなは 46 文字である。ひらがなのもつ情報量を計算し、必要なビット数を示せ。

【解答例】

ひらがなの生起確率を等確率 ($p=1/46$) とすると

$$\text{ひらがなの情報量 } I = -\log_2(1/46) = \log 46 / \log 2 = 5.524 \quad [\text{bit}]$$

必要なビット数は 6bit。すなわち、コンピュータ上では、1byte (=8bit)

【設問 8】 漢字の情報量はどのように計算できるだろうか。漢字はコンピュータ上では 2byte (1word) で表現される (日本語の Shift JIS コードは 4 文字である)。2byte で割り当てることが可能な日本語の文字数を計算せよ。

【解答例】

2byte=16bit より $2^{16}=65,536$ 文字扱い可能
(註: 1byte=8bit では $2^8=256$ 文字しか扱えないので)

3. 情報の意味を考えよう

情報理論が示す「情報」の意味が理解できただろうか。我々が通常使う「情報」とは意味合いがかなり異なっていることに注意すべきである。以下例題を使って、その違いをはつき

りと意識してみよう。相反するものではなく、視点が異なることに気づくべきである。同じメッセージ（情報）を送信者の立場から「量」として捉えるか、受信者の立場から「内容」として捉えるかの違いである。

(3.1) 演歌歌詞と JPOP 歌詞のエントロピーを計算してみよう

【設問 9】 以下は、JPOP と演歌の歌詞（テキスト）である。それぞれの歌詞(情報源)の情報エントロピーを計算せよ。そして、両者のエントロピーを比較することにどのような意味があるかを考察せよ。

(ヒント) それぞれの歌詞のひらがな、カタカナ、漢字、アルファベットの出現確率からエントロピーを計算せよ。空白は無視し、ひらがなとカタカナの濁音と小文字は清音に置き換えそれぞれ 46 文字で扱う。またアルファベットの小文字は大文字扱いとする。

歌手：サザンオールスターズ

作詞：桑田佳祐

作曲：桑田佳祐

泣かした事もある 冷たくしてもなお

よりそう気持ちがあればいいのさ

俺にしてみりゃ これで最後の Lady

エリー My love so sweet

歌手：森進一

作詞：川内康範

作曲：猪俣公章

おふくろさんよ おふくろさん

空を見上げりゃ 空にある

雨の降る日は 傘になり

お前もいつかは 世の中の

傘になれよと 教えてくれた

あなたの あなたの真実

忘れはしない

[略解答]

まず、両歌詞の文字種に着目してみよう。ひらがな、カタカナ、漢字、アルファベットが混用されている。それぞれの文字種の出現確率を生起確率とする。ひらがなは「あ、い、う、え、お...」ごとに分母はひらがなの全文字数、分子はそれぞれの出現頻度とし、出現確率を計算する。

それぞれの発生数 N を以下で表現する。

演歌歌詞のひらがな発生数： $N_E(H_{\text{あいうえお...}})$

演歌歌詞のカタカナ発生数： $N_E(K_{\text{アイウエオ...}})$

演歌歌詞の漢字発生数： $N_E(W_{\text{亜井字...}})$

演歌歌詞のアルファベット発生数： $N_E(A_{\text{abc...}}) = \emptyset$ (演歌歌詞には使われていないので空集合 $\emptyset$ となる)

JPOP 歌詞のひらがな発生数： $N_J(H_{\text{あいうえお...}})$

JPOP 歌詞のカタカナ発生数： $N_J(K_{\text{アイウエオ...}})$

JPOP 歌詞の漢字発生数： $N_J(W_{\text{亜井字...}})$

JPOP 歌詞のアルファベット発生数： $N_J(A_{\text{abc...}})$

次に発生確率 P を計算する。分母は

ひらがな=46

カタカナ=46

漢字=65536

アルファベット=26

以上から、たとえば演歌歌詞のひらがな (あ) の発生数が 2 とすると、

$$P_E(H_{\text{あ}}) = \frac{N_E(H_{\text{あ}})}{46} = \frac{2}{46}$$

以下、同様に計算し、発生確率を乗じて累積すれば良い。

カタカナ、漢字、アルファベットも同様に計算できる。それぞれの出現確率表は以下となる。

情報エントロピー	いとしのエリー	おふくろさん
ひらがな	2.639	2.726
カタカナ	0.240	
漢字	0.00195	0.00339
アルファベット	1.7288	
計	4.610 bit	2.729 bit

(3.2)熱力学エントロピーと情報エントロピーの関係について

熱力学エントロピー S とは；

ある気体分子がある条件 A (たとえば、容器 A の中に封じ込められた状態を想像しよう) の中で取り得る状態の数を W_A とする。その時の、熱力学的エントロピーは以下で表現される。

$$S_A = k \log_e W_A$$

同様に条件 B においては

$$S_B = k \log_e W_B$$

$$\therefore S = S_A + S_B = k \log_e (W_A \times W_B)$$

ここに、 k はボルツマン定数。

情報エントロピー H とは：

ある事象 A の発生確率を P_A とする。その時の、 A の情報量は以下で表現される。

$$I_A = \log_2 \frac{1}{P_A} = -\log_2 P_A$$

同様に、事象 B においては

$$I_B = \log_2 \frac{1}{P_B} = -\log_2 P_B$$

$\therefore$ 情報量 I は

$$I = I_A + I_B = -(\log_2 P_A + \log_2 P_B) = -\log_2 (P_A \times P_B)$$

熱力学エントロピーに合わせ、対数の底を e に変換すると

$$I = -\frac{\log_e (P_A \times P_B)}{\log_e 2} = -1.44 \log_e (P_A \times P_B)$$

すなわち、情報エントロピーではボルツマン定数に相当する定数は -1.44 ということになる。

なお、シャノンは、発生確率を考慮した平均情報エントロピーとして以下を定義した。

$$\therefore H(p_1, p_2, \dots, p_n) = -\sum_{i=1}^n p_i \cdot \log_2 P_i$$

(3.3) エントロピー増大の法則について

熱力学で言うところの「エントロピー増大の法則」とは、「自然界は、ほっておくと、 $S \rightarrow$ 大となる性質がある。すなわち、気体の状態数は大きくなり、無秩序状態となる」、ということである。

情報量（情報エントロピー）増大の法則とは、「自然界は、ほっておくと $I \rightarrow$ 大となる性質がある。すなわち、情報量は無秩序に大きくなる性質がある」。情報はいくらでも複製が可能であり、それを載せる媒体に制限がない限りいくらでもふくれあがる事を意味している。

この両者の増大法則を竹内啓⁶は以下のように述べている。

（以下、一部表現を変え転記する：－）宇宙の色々な局面において、「エントロピー増大の法則」と「情報量増大の法則」が共に働いていると思う。宇宙には偶然性という力が働いており、その偶然性には極限において完全な一様性をもたらす性質のものと、累積することによって情報として働き新たな秩序を作り出す 2 種類が存在する。前者の偶然性はエントロ

⁶ 竹内啓：偶然とは何か, 岩波新書, 2010.

ピーを増大させ一様性という安定状態を作り出そうとするが、後者は偶然に作り出された新しい秩序が情報システムによって捉えられることにより安定維持されることになる。

すなわち竹内は、この自然界には無秩序の安定性と秩序ある安定性があり、共に自然界の法則として成立すると言っている。確かに、定義式では、熱力学的エントロピーと情報理論的エントロピーでは、表現形式は同じ（対数表示）であるが、符号は逆向きである。熱力学エントロピーを気体分子の存在状態数で定義したのに対し、情報エントロピーは事象の発生確率の逆数で表しており、発生確率が小さい事象ほど情報量は大きく定義され、発生確率が小さいものほどある秩序を作り出し易いことを意味している。そのこと（秩序を作り出す発生確率の小さな事象）が「情報」の本質と言えよう。しかし皮肉なことに、発生確率が小さな事象ほど「情報」として珍重されるので、その世界に安定的に存在する力は大きく、「情報量」として増大していく性質があるのであろう。自然界において情報量が増大する性質をうまく表現し、自然界の熱力学的エントロピーと同意義的扱いを可能にしたシャノンの先見性に脱帽である。

4. 本講のまとめ

今回の学修テーマは「情報とは何か」である。2章において、情報には以下の2つのモデルがあることを指摘した。

- ・情報の意味・価値に関する構造モデル：R. Ackoff(1989)の DIKW ヒエラルキー
- ・情報の形式に関する確率モデル：C. Shannon(1948)の情報理論

通常我々が使う「情報」は、その情報が持つ {意味・価値} に意義を見いだす構造モデルである。しかし、情報工学が扱う「情報（シャノンが定義した情報）」とは、情報の持つ意味ではなく、その情報、例えば文章の場合は、その文章を構成しているシンボルを伝えるのに必要な情報単位（bit 数）が情報量として重要な量ということである。

その意味を理解してもらった教材として、演歌歌詞と JPOP 歌詞の情報量比較（エントロピー比較）をしてもらった。エントロピー（情報量＝情報の不確定性）は圧倒的に JPOP のほうが大きいことに気がつくであろう。JPOP の歌詞（シンボル）を情報として流す場合、必要な bit 数は演歌歌詞よりも多くなると言うことである。そこに主観の入る余地はない。しかし、我々は歌詞を情報としてみた場合、そのメッセージ性に意義を見いだす。演歌と JPOP の世界観がどれだけ伝わるか、またどれだけ多くの人に共感を持ってもらえるか、そこに歌詞の持つ「情報」の意義を見いだす。演歌の世界観は人と人との間に交わされる「情（なさけ）」であり、その情けの深さをキーフレーズ（たとえば、「おふくろさん」「あなた」というキーフレーズ）の繰り返しで表現している。そのメッセージ性は極めて強く、情報伝達という理論から見るなら、繰り返しによりその出現確率が上がりエントロピーは小さくなり、情報量としては小さくて済む。一方、JPOP が伝えたいメッセージは「恋人に対する切ない思い」で、演歌に通じる面があるが、その表現方法として繰り返しを使うのではなく、

アルファベットやカタカナなど乾燥したイメージのキーフレーズを多用することで、相手との距離感を演出してその切なさ表現している。執着性がない分、多くのキーフレーズが出現し、エントロピーは大きくなり必要伝送量も多くなる。

相手に思いを伝える「情報」として、あなたはどちらに軍配を上げるだろうか。

この問題を通して、エントロピーを比較することにどのような意味があるか理解できたであろうか。エントロピー（平均情報量）が大きいほどそのテキスト（情報源）の持つ情報不確定性は大きいことを意味し、情報量が多いことを意味する。誤解を与えずに他者に情報を渡す行為（これをコミュニケーションという）において、曖昧さを持つ情報を曖昧さを排除して伝えるのにどれだけの時間（話す量）を要するかという問題を、そのテキストを送送するのに必要なビット数に変換しコンピュータ処理の言語に言い換えた訳であり、シャノンがコンピュータ情報処理の技術発展に果たした役割は極めて大きいことが分かるであろう。

最後に、以下の点について各自考えてもらいたい。

シャノンの情報理論に基づけば、情報量は確率（不確定性）の関数であることが理解できたであろう。ここでもう一度考えてみよう。シャノンが扱おうとした不確定性とは、なんだろうか。工学的に確率論を扱うことの意味と意義について考察せよ。特に、自分の研究領域において有効と思われる確率論の応用展開方法について、議論せよ。

次回は、斉藤の例題による、確率論の優位性について解説する。