第2回： 今日の目標

- 情報理論の概略を説明できる
- 情報とは何かを説明できる
- ニュースバリューの要因を示せる
- 科学的に扱う情報を確率の概念で説明できる
- 情報量を数式で示し、その単位を記述できる
- 情報量の計算ができる

§１．４　情報理論

情報理論創始者

Claude E.Shannon（1916～2001、米）
“The Mathematical Theory of Communication”
University of Illinois Press, 1949
○情報の定量化
○通信のモデル
　情報源符号化と通信路符号化の概念

Publication

Norbert Wiener（1894～1964、米）
“Cybernetics or Control and Communication in the Animal and the Machine”
The M.I.T. Press and John Wiley & Sons, Inc., 1948
○システム理論
　情報、伝送、制御

情報伝達の効率化、高信頼化のための符号化理論

情報とは

情報を計るには　⇒　情報とは何か？

事象（事件）

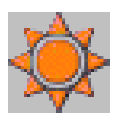
・猛暑日、連続記録更新
・介護労働者の離職率17%で最高
・国内観測史上最大のM9.0を記録
・
・明日は快晴
・

人への知らせ

情報

知らせる手段

・喋る（声）
・絵
・文字


晴


曇


雨

記号

情報＝事象を写し取った記号そのもの

ニュースバリュー

・かえるが蛇を飲み込んだ
・犬が猿の赤ちゃんを育てた
・バイク追突、宙舞って車屋根に着地

まれな事象
（確率が小さい）

人の主観

・福島原発汚染水／拡散防止策は国の責任
・中国、国交正常化40周年式典を中止 尖閣で対抗措置か
・大相撲秋場所　日馬富士、横綱昇進文句なし

ニュースバリューの要因＝確率的要因＋非確率的要因

完全事象系

事象	♥	♠	♦	♣	のカードを引く
確率	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	

互いに素な事象　：　$E_1, E_2, E_3, \cdots, E_n$

対応する確率　：　$p_i = P(E_i)$,

$$\Sigma\, p_i = 1$$

$\Rightarrow$集合$\{E_1, E_2, E_3, \cdots, E_n\}$を完全事象系という

$$\mathbf{E} = \begin{bmatrix} E_1, E_2, E_3, \cdots, E_n \\ \\ p_1, p_2, p_3, \cdots, p_n \end{bmatrix}$$

英文　$\left\{ \begin{matrix} a & b & c & \cdots & z & \_ & , & . \\ p_a & p_b & p_c & \cdots & p_z & p_\_ & p_, & p_. \end{matrix} \right\}$ ⇨ 情報

表 2.2　英語の文字の出現確率

文　字	確　率	文　字	確　率
スペース	1 0.1859	N	7 0.0574
A	4 0.0642	O	5 0.0632
B	0.0127	P	0.0152
C	0.0218	Q	0.0008
D	0.0317	R	9 0.0484
E	2 0.1031	S	8 0.0514
F	0.0208	T	3 0.0796
G	0.0152	U	0.0228
H	10 0.0467	V	0.0083
I	6 0.0575	W	0.0175
J	0.0008	X	0.0013
K	0.0049	Y	0.0164
L	0.0321	Z	0.0005
M	0.0198		

F. M. Reza 著，鶴見・大石訳「確率，情報，コード」(共立出版，1973)

アルファベット出現確率がキー配列に反映されているか？

表 2.3　英文における文字と単語の出現順位

順位	文字	2字組	3字組	単語	順位	文字	2字組	3字組	単語
1	E	TH	THE	THE	11	D	AT	HAT	FOR
2	T	IN	AND	OF	12	U	ON	HIS	IT
3	A	ER	ING	AND	13	C	HA	THA	WITH
4	O	RE	ION	TO	14	F	OU	ERE	AS
5	I	AN	ENT	A	15	M	IT	ARE	HIS
6	N	HE	TIO	IN	16	W	ES	ATE	ON
7	S	AR	FOR	THAT	17	Y	ST	RES	BE
8	R	EN	HER	IS	18	G	OR	ALL	AT
9	H	TI	TER	WAS	19	P	NT	VER	BY
10	L	TE	ATI	HE	20	B	HI	WAS	I

E. B. Montgomery : *Computer* Vol. 15. (IEEE, 1982), F. M. Reza 著, 鶴見・大石訳「確率, 情報, コード」(共立出版, 1973), 坂井利之著「情報基礎学」(コロナ社, 1982)

マルコフ情報源　＝　文脈依存型情報源 ⟷ 文脈自由型情報源

単純マルコフ過程 : $P(X_t=\text{'H'}|X_{t-1}=\text{'T'})=p_{TH}$

疑似英文

二重マルコフ過程 : $P(X_t=\text{'E'}|X_{t-2}=\text{'T'}, X_{t-1}=\text{'H'})=p_{THE}$

独立な事象

A = { ♥ ♠ ♦ ♣ }

B = { ⚀ ⚁ ⚂ ⚃ ⚄ ⚅ }

p(♥) = 1/4

p(⚀) = 1/6

p(♥ ∩ ⚀) = p(♥) · p(⚀) $= \frac{1}{4} \cdot \frac{1}{6}$

$= \frac{1}{24}$

情報量

i (E) : 事象E関する情報量

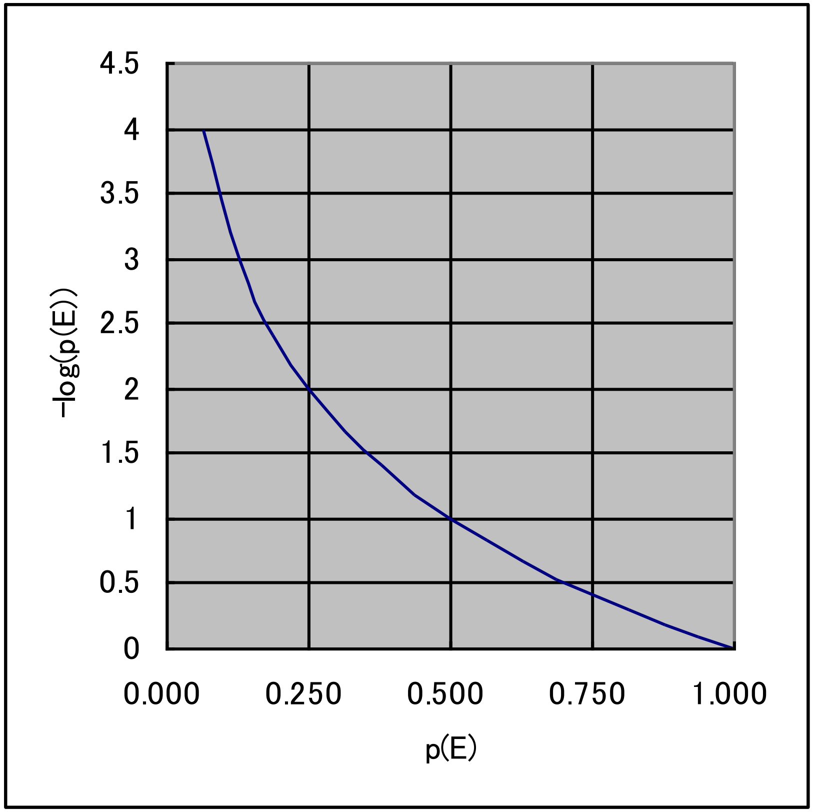
条件（1） i (E) ≧0

（2） i (E∩F)=i (E)+i (F)

（3） p (E)=1/2 のとき、 i (E)=1

$$\mathbf{E} = \begin{bmatrix} \text{yes} & \text{no} \\ 1/2 & 1/2 \end{bmatrix}$$

情報の単位



条件を満たす関数

$$i(E) = -\log_2 p(E) \quad [\text{bit}]$$

（ビット＝Binary digit (unit) ）

（１） $0 \leqq p(E) \leqq 1$

（２） $p(E \cap F) = p(E) \cdot p(F)$

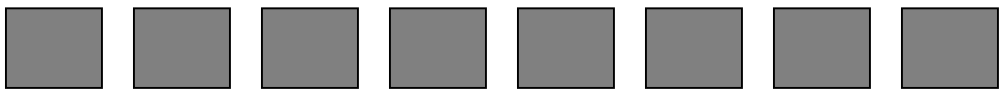
（３） $-\log_2 (1/2) = 1$

例： “コインをn回投げてn回とも表が出る”事象E

$$p(E) = p(1回目表)\cdot p(2回目表)\cdot\cdot\cdot p(n回目表) = \left(\frac{1}{2}\right)^n$$

$$i(E) = -\log_2\left(\frac{1}{2}\right)^n = n\ [\mathrm{bit}]$$

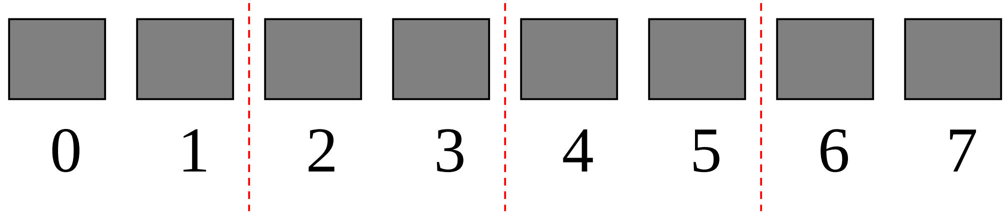
例:中が見えない8つの箱の1つにダイヤが入っている



状態　0　1　2　3　4　5　6　7

ダイヤの箱を当てる確率: P(アタリ)＝1/8

$$情報量:\ i(アタリ) = -\log_2\left(\frac{1}{8}\right) = -\log_2\left(\frac{1}{2}\right)^3 = 3[\mathrm{bit}]$$



状態　0　1　2　3　4　5　6　7

質問：半分に分けて右に入っていますか？

答：1回目	2	3	状態
no	no	no	0
no	no	yes	1
no	yes	no	2
no	yes	yes	3
yes	no	no	4
yes	no	yes	5
yes	yes	no	6
yes	yes	yes	7

状態数８ $\Rightarrow$ 2進数3桁で表現

0から999の状態 $\Rightarrow$ $\log_2 1000 = \log_{10} 1000 / \log_{10} 2 = 9.966 \fallingdotseq 10$ bit

2進数10桁

2進数16桁　$2^{16}=65536$

■　演習

（１）128本のくじがある。その1本が当たりくじであるとき、当たりくじを教えてもらって得られる情報量はいくらか。

（２）2進数12桁で表される数はいくらか。

（３）明日は雨の降る確率が75%という予報である、神様に“明日は晴れ”と教えてもらったときの情報量はいくらか。ただし、天気は晴れか雨しかないものとする。

（４）2つのサイコロを投げて、ぞろ目になるときの情報量はいくらか。ただし、$\log_2 3=1.58$とする。

情報科学概論のトップへ
明治薬科大学のホームへ