No.

Date.

§「情報量」についての注意

確率分布$\mathbb{P}$に従うランダムなN文字の文章

↪ 情報量 $NH(\mathbb{P})$ bits　一文字の情報量

ランダムでなければ？

AABCAABCAABCAABC……

↓

はるかに圧縮可能（情報量が少ない）

完全にランダムな文章が最大の情報量をもつ。

与えられた文章がどこまで圧縮可能かはむずかしい問題。

例 9265358979323 38

$$H(\mathbb{P}) = -\sum_{i=1}^{\Omega} p_i \log_2 p_i$$

情報量が$H(\mathbb{P})$であることのラフな説明。

Ω種類の文字 かくりつ $p_1, p_2, \ldots, p_\Omega$で出現

N文字の文章。 $N \gg 1$.

N文字の文章 $N \gg 1$

N文字のうち 文字1は、ほぼ $N_1 = Np_1$ 回

2　$N_2 = Np_2$ 回

⋮

Ω　$N_\Omega = Np_\Omega$ 回

あらわれる。

$1, \ldots, \Omega$の文字がそれぞれ$N_1, N_2, \ldots, N_\Omega$回あらわれる文章の総数。

$$W = \frac{N!}{N_1! N_2! \cdots N_\Omega!}$$

（Stirlingの公式

$$n! = n(n-1)(n-2)\cdots 1 \sim \left(\frac{n}{e}\right)^n$$

$n \gg 1$のとき）

$$W \sim \left(\frac{N}{e}\right)^N \left(\frac{e}{N_1}\right)^{N_1} \left(\frac{e}{N_2}\right)^{N_2} \cdots \left(\frac{e}{N_\Omega}\right)^{N_\Omega}$$

$$= \left(\frac{e}{N_1}\right)^{N_1} \left(\frac{e}{N_2}\right)^{N_2} \cdots \left(\frac{e}{N_\Omega}\right)^{N_\Omega} \quad \left(\sum_{i=1}^{\Omega} N_i = N\right)$$

$$= p_1^{-Np_1} \cdot p_2^{-Np_2} \cdots p_\Omega^{-Np_\Omega}$$

ビットでかぞえる。

$$\log_2 W \sim \log_2 \{ p_1^{-Np_1} \cdots p_\Omega^{-Np_\Omega} \}$$

$$= -N \{ p_1 \log_2 p_1 + \cdots + p_\Omega \log_2 p_\Omega \}$$

$$= N \cdot H(\mathbb{P})$$