統計力学　Statistical Mechanics

第0章　はじめに

この「統計力学」は僕が学部時代にまとめていたものをネタに見やすいように TeX に起こした物です。読者の想定は自分自身で「普通」の大学生だと思っている方が基本的な対象です。僕自身あまり統計力学は計算が色々面倒で好きではありません。しかしながらこの分野が持つ物理における立場は承知していまので御座なりするつもりは毛頭ありません。統計力学は多くの物が集まってある物理系を作ったとき、その系の性質を現すための手法です。即ち、我々が普段目にするものの殆どはこの力学的な手法を用いています。また微視的な物理を記述する量子力学も統計力学を基本にして生まれたものです。尚更嫌いなどとは言ってられません。しかし... 初学者にはキツイ！統計力学が生まれる前は熱力学がこの役割をしていました。ですから統計力学は既にまとめられた分野であり、ある種集大成と呼べるような学問です。初学者は「当たり前」を学ぶ前にいるのですから、いきなり完成された物を見ても“？”が浮かぶだけで「そーいう物デスカ」で通り過ぎ単位をもらうための努力になってしまいがちです。
私なんぞは要領が悪いのでその努力すら出来ませんでした。案の定、酷い結果を頂戴し... え〜まぁそんなわけです。ですからここに書いてある事は１つのアドバイスとして見てもらいたいと思います。でもこれを書くに当たって真面目にやっていますので騙すつもりは全くありません。怪しいなと思ったら即座に有名図書を参考にして下さい。
話の流れとしては先ずは集団的な物理系を表す手法として「確率」を用いる事になります。そこで「確率」とは何ぞや？という所から入ります。特に心配なさそうな方はどんどんとばして下さい。あれっ？と思ったら帰ってくればいいですから。「エントロピー」は統計力学にせよ、熱力学にせよ大事な物理量であります。いきなり言われてもよく分からない物理量なので基本的な定義から進めます。「力学系の表現」は色々とあると思いますが所謂“ カノニカル ”な形式を見ていきましょう[1]。次はいよいよ統計力学っぽい「〜カノニカル集団」です。これ等は物理系の特徴を考えて、各々の記述方法を分類したものです。
この辺りまでをある程度詳しくやりたいと思います。

[1]Canonical は“ 正準 ”という意味で正しい基準という事です。物理的に正しい基準というのはどういった意味なのでしょう？そこら辺を話したいですね。

第1章　確率とエントロピー

この章で紹介する話は非常に当たり前の事です。が当前の事が曖昧のままの定義では後々の展開で分からなくなるのでちゃんとしておこう[1]。

1.1　単純事象、合成事象の確率表現

ある物事や事柄を総称として「事象」と呼ぶがこの事象は大きく分けて単純事象と合成事象に分けられる。例えばそれは

単純事象：ある実験を実行するにあたり起こり得る結果、状態がただ１つの場合
合成事象：単純事象の組み合わせで作られる事象

更に具体的に硬貨投げの例を挙げると

$\boxed{\text{一度だけ投げる}}$ (試行実験の内容)

$$\text{単純事象：}\begin{cases} (\text{表が出る}) \equiv A \\ (\text{裏が出る}) \equiv B \end{cases}$$

合成事象：？（一度の試行では定義できない？）

一度だけ硬貨を投げるという試行で現れる事象は上記の A, B の二つだけである。実験での事象を全確率 1 に割り振る。A, B は同程度現れると考えると各々の確率を $prob.(A) \equiv P_A = 1/2 = P_B$ と定義出来る。

$\boxed{\text{二度続けて投げる}}$

$$\text{単純事象}\ (Event)\text{：}\{E\} = \{AA, AB, BA, BB\}$$

$$\text{合成事象：例えば}\begin{cases} 1,\text{同じ面が二度続けて出る}\ (AA, BB) \\ 2,\text{異なる面が出る}\ (AB, BA) \\ 3,\text{少なくとも一度は表が出る}\ (AA, AB, BA) \\ \vdots \\ etc. \end{cases}$$

観測する側が起こり得る単純事象を組み合わせて観測したい事象を合成する。または観測する事象が試行の段階ですでに複数の単純事象から出来るている事も当然あり得る。この様に考えれる事象を合成事象と呼ぶ。どちらにせよ対象とする事象が複数の事象から出来ていればそれは合成事象である。

[1] 様々な事柄についても当てはまる事だが「当然、前提」の定義が一番疎かにし易い部分かも知れない。

1.2　相関の有無とその確率

先述の合成事象を考える際に“ 二度 ”の実験を実行している。試行した実験の内容は全く等しいものであったが“ 一度目 ”と“ 二度目 ”という意味を付け加えると異なった実験と区別する必要がある。簡単に言えば“ 一度目 ”の結果が“ 二度目 ”に影響があるのか、無いのかという事である。一般的に合成事象が“ 同 ”実験の組み合わせで表現できる筈が無く、細かい事かも知れないが時間によってその実験の状況が変わってしまえば全く異なる実験と呼ばなければならない。実際こうでない実験を探す方が難しいであろう。二つの実験で現れる単純事象を用いて合成事象を作るとし、このとき元となる実験の名前を α, β とする。

1.2.1　相関の無い互いに独立な実験の確率表現

実験 α, β には各々i, j 個の単純事象があるとし、i 個の確率を $P_{i\alpha}$ とする。更に α の全単純事象を $\{E_{i\alpha}\}$ と表す。合成事象 $\{E_{i\alpha}E_{j\beta}\}$ のある事象の確率 $P_{i\alpha,j\beta}$ は単に

$$P_{i\alpha,j\beta} = P_{i\alpha} \cdot P_{j\beta} \tag{1.1}$$

と書く事が出来る。この表記はあまりにも単純であるがお互いに独立であるという重要な内容を含んでいるのである。微分方程式等を解く際に、変数分離が可能であるといった仮定を見た事があると思うが、事象が各々独立に起きて一つの合成事象を構成していると仮定する事に相当するのである。

1.2.2　相関のある実験の確率表現

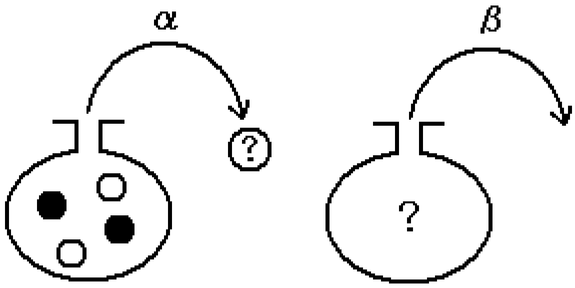
α, β に相関があるというのは α, β が互いに独立した実験ではなく α の事象によって β に発生する事象に変化が起き、結果 β における事象の確率が α の事象によって変化する事である。
従って先ほどの様に単に確率を掛け合わせるだけで合成事象の確率を表す事は出来ない。どちらかの確率に相関の影響を考慮せねばならない。
ここで非常に簡単な例ではあるが少し具体的に表記を示そう。

例

袋の中に入っている玉が全部で n 個とし、白が m 個黒が $n-m$ 個とする。この袋より玉を取り出す実験。右図に示す α, β は以下の内容を表す。

α：取り出した玉の色を確認し、戻さない。

β：α の後玉を取り出し色を確認する。



α, β で合成事象を考えたとすると β の結果はもちろん α の影響を受けるのでこの確率 $P_{i\alpha,j\beta}$ を

$$P_{i\alpha,j\beta} = P_{i\alpha} \cdot P_{i\alpha;j\beta} \tag{1.2}$$

の様に書くとする。もっと具体的に $P_{i\alpha,j\beta}$ を書くとしよう。
添え字の i, j は今、玉の色を表す。$i, j = 1$ を白、$i, j = 2$ を黒としよう。この様にすると

$$\alpha : P_{1\alpha} = \frac{m}{n} : 白が出た \ \ , \ \ P_{2\alpha} = \frac{n-m}{n} : 黒が出た$$

$$\beta \ : \begin{cases} P_{1\alpha;1\beta} = \frac{m-1}{n-1} : \alpha で白、\beta で白 \ \ , \ \ P_{2\alpha;1\beta} = \frac{m}{n-1} : \alpha で黒、\beta で白 \\ \\ P_{1\alpha;2\beta} = \frac{n-m}{n-1} : \alpha で白、\beta で黒 \ \ , \ \ P_{2\alpha;2\beta} = \frac{n-m-1}{n-1} : \alpha で黒、\beta で黒 \end{cases}$$

よって例えば $P_{1\alpha,2\beta}$ は

$$P_{1\alpha,2\beta} = P_{1\alpha} \cdot P_{1\alpha;2\beta} = \frac{m}{n}\frac{n-m}{n-1}$$

といった様に書ける。最後に注意すべき事は「確率の規格化」である。確率の規格化とは実験の全ての事象を考えた事を示す。即ち、事象全ての確率を足し合わせた時に 1 となる事である。実験が単純事象であろうと合成事象であろうとも、物事を確率で表現する場合は少なくとも[2]全ての事象が分かっていると仮定しなければならない。これ等の仮定は次の様に表される。

$$\sum_i P_{i\alpha} \ = \ 1 \ \ , \ \ \sum_{i,j} P_{i\alpha,j\beta} \ = \ 1 \tag{1.3}$$

この様にする事を「確率の規格化」又は単に「規格化」と呼ぶ。

1.3　連続的確率事象

先ほどまで述べていた実験は全て「離散的確率事象」と分けられる物で i, j 等の離散数で事象の確率を分ける事にしていた。いわゆる物理量の殆どは連続的なものであるからこれに対応した事象の確率表現、即ち「連続的確率事象」の取り扱いを考えるのが自然である。

例 コロイド粒子の位置

水の中にコロイド粒子[3]を懸濁させると粒子はブラウン運動の挙動が現れる。この時の粒子の位置は確率的にしか決められないという。さてこの粒子の位置を確率で表すのはどうしたらよいだろうか。

この現象はその確率が連続的な x に依存している、「連続的確率事象」である。粒子が x' の位置にある確率を考える。x' に非常に近い位置 x がありそこから dx の間隔を取りその中に x' があるとする。この確率を

$$prob.(x < x' < x + dx) \equiv P(x')dx \tag{1.4}$$

と書く事にする。$P(x')dx$ が確率であり $P(x')$ は「確率密度」、「確率分布」と呼ぶ事が出来る。この定義において dx は不確定な範囲でこの間での事について我々はわからない。無限小の極限を考えるがそれは「不確定」を失くした事にはならない。またはこの間で劇的な変化は無いものとすると仮定している。

物理量が今、位置として x を変数としているとすると、この x を「確率変数」と呼ぶ。また確率は無次元である事から確率密度 $P(x)$ は確率変数 x の逆次元を持っている。

[2]自然界に起き得る現象を確率論で扱う場合、全ての事象が理解できた等とは到底言えない為である。

[3]コロイド粒子とは分子と比べて比較的大きな形を持つ粒子。例えば墨汁は何かが水に溶けるというのではなく、目に見える程度に細かい粒子が分散している状態になる。この時の粒子を指す。

1.4　連続的確率事象における相関の有無の表現

特に真新しい事は無いが連続的な確率についても表記を示しておく。

1.4.1　相関の無い場合

α, β の確率試行があるとき各々の確率変数を x, y とする。
互いに相関の無い α, β で合成事象を考えた時の確率 $P_{\alpha\beta}(x', y')$ は

$$\begin{aligned} prob.(x < x' < x + dx, y < y' < y + dy) &\equiv P_{\alpha\beta}(x', y')dxdy \\ &= P_\alpha(x')dx\ P_\beta(y')dy \end{aligned} \tag{1.5}$$

$$P_{\alpha\beta}(x, y) = P_\alpha(x)P_\beta(y) \tag{1.6}$$

と単に確率密度の掛け算で合成事象の確率密度を表現する事が出来よう。

1.4.2　相関の有る場合

互いに相関の有る α, β で合成事象を考えた時の確率 $P_{\alpha\beta}(x, y)$ は先ず先行する試行が α である場合

$$P_{\alpha\beta}(x, y)dxdy = P_\alpha(x)dx\ P_\beta(y \mid x)dy$$

$$P_{\alpha\beta}(x, y) = P_\alpha(x)\ P_\beta(y \mid x) = P_\alpha(x)\ P_\beta\left(y(x)\right) \tag{1.7}$$

と表す事にする。一番最後の表記が分かり易いと思う。次に β が先行し、α が対応したとすると

$$P_{\alpha\beta}(x, y)dxdy = P_\beta(y)dy\ P_\alpha(x \mid y)dx$$

$$P_{\alpha\beta}(x, y) = P_\alpha(x \mid y)\ P_\beta(y) = P_\alpha\left(x(y)\right)\ P_\beta(y) \tag{1.8}$$

と表す事にする。またこの様な $P_\beta\left(y(x)\right), P_\alpha\left(x(y)\right)$ を条件付き確率密度と呼ぶ。

ここではあまり注意しなかったが、x, y の交換については注意してもらいたい。相関の無い場合はこのままで良いが、有る場合は少し厄介だと思われる。しかしながら量子論を含まない様な統計力学については変数の交換 $dxdy \to dydx$ は特に問題は無いだろう。

1.5　確率変数、確率空間、確率関数

ここでは今後の展開で出てくる用語の説明と出てきた用語をまとめておこう。
確率変数：一定の確率を割り振られた単純事象の名称。(観測する現象の要素)
離散的確率事象では $i, j, \cdots$ 等
連続的確率事象では $x, y, \cdots$ 等
確率空間：確率変数により張られる空間、確率変数の集合。(観測する現象が起こり得る空間、考える領域)
確率関数：確率変数による関数 (観測される現象)
例　離散的：サイコロの目 $(i = 1, 2, \cdots, 6) \times 100$ 円の配当金を表す関数 ￥$_j$(とか？[4])
連続的：ブラウン運動する粒子の重力場における位置エネルギー関数 $U(x)$

[4]適当な名前で自信が無い

1.6　平均値とモーメント

平均値はいいと思うが、ここで言うモーメントは「慣性〜」のそれとは違い統計での積率[5]を表すモーメントである。物理量を観測し最も頻度の高い現象を知るには平均値は重要である。また確率密度を異なる視点から見るのがモーメントである。これ等について見ていこう。

1.6.1　平均値

我々がある物理量に対し初めに観測し得るのは最も頻度が高い現象であろう。珍しい現象が起因となる物理量はなかなか観測し難いであろう。物理量の大体の性質を表すにはこの平均量が重要になってくる。即ち、確率関数 f の平均値 $\langle f\rangle$ を知る事である。確率関数、確率密度を用いて次の様に書く。

$$\langle f\rangle \equiv \sum_j f_j P_j \ \text{：離散}, \quad \langle f\rangle \equiv \int dx f(x)P(x) \ \text{：連続} \tag{1.9}$$

例 サイコロの目の期待値
今、サイコロに何の細工も無く $1,\cdots,6$ の目全て同確率で現れるとするとつまり

$$f_j = j = 1,2,\cdots,6 \quad , \quad P_j = \frac{1}{6} = const. \tag{1.10}$$

$$\langle f\rangle = \sum_j f_j P_j = \frac{1}{6}\sum_{j=1}^{6} j = \frac{1}{6}(1+2+3+4+5+6) = 3.5 \tag{1.11}$$

となる。勿論、物理現象がいつもこのサイコロの様に確率が全て同じという事はありえない。連続的確率事象ならば尚更である。

1.6.2　偏差、分散、モーメント

平均値と絡めて確率分布を考える事は非常に重要である。なぜならば我々はいつも観測した物理量の正確な確率分布を知っている筈が無いからである。むしろ観測した情報からその物理量に隠れている確率分布を見出し、物理量が何から構成されているか探る方が自然である[6]。
さて先ずは平均値からどれくらいズレているのかを示す量、偏差 σ は次の様に定義する[7]。

$$\sigma \equiv \mid f(x) - \langle f(x)\rangle \mid \tag{1.12}$$

特に後々モーメントという量を考えるので都合上 $f(x)\equiv x$ で話を進める。一般的に上記の偏差は正しい。更に正負を考えない様に二乗の量の平均 (x 各々に確率分布が掛かっている) を考える。これは

$$\begin{aligned} \langle\ \left(x-\langle x\rangle\right)^2\ \rangle &= \langle\ x^2 - 2x\langle x\rangle + \langle x\rangle^2\ \rangle = \langle\ x^2\ \rangle - 2\langle x\langle x\rangle\rangle + \langle\langle x\rangle^2\rangle \\ &= \langle\ x^2\ \rangle - \langle x\rangle^2 \equiv \tilde{V} \end{aligned} \tag{1.13}$$

と定義され $\tilde{V}$ を分散と呼ぶ[8]。

[5] ここで意味は分からなくとも区別してね、という事
[6] 理論先行し過ぎるとこの方向を往々にして無視する傾向が現れる。現実 (観測量) に基づいた正しさが科学の本質であり、これからの逸脱は本来ありえない。しかしながら物理を理解する上で数学の力を借り、一時的に無視する場合もあるのは仕方が無い。
自分の信じた、作った理論を大事にするのは分かるが、現実からの逸脱は虚構でしかない。この様な症状を Pygmalion 症候群などと呼んでいるが、気をつけよう。
[7] 少し違いますが所謂、標準偏差とかいうやつです。
[8] ちょっと定義が違うが σ^2 を分散とも言う。

偏差や分散等を知る事で確率分布のどの程度広がっているかが分かる。更にここで次の量を定義する。

$$M_1 \equiv \langle x \rangle \quad , \quad M_2 \equiv \langle x^2 \rangle \quad , \quad M_n \equiv \langle x^n \rangle \tag{1.14}$$

これ等はモーメントと呼ばれる量で **確率変数** の冪乗の期待値で定義される。M_1 は一次のモーメント M_n はｎ次のモーメントと呼ぶ。確率分布 $P(x)$ を用いて表せば

$$M_n = \langle j^n \rangle = \sum_j j^n P_j \quad , \quad M_n = \langle x^n \rangle = \int x^n P(x) dx \tag{1.15}$$

先の内容に照らし合わせると M_1 を知る事は確率分布の中心を知る事に相当する。
もう少し詳しくモーメントを知るために特性関数と呼ばれる物を持ち出す。

特性関数

特性関数 $\Phi(\xi)$ は確率分布 $P(x)$ の Fourier 変換として定義される関数で、離散的な場合と連続的な場合でそれぞれ

$$\Phi_\xi = \sum_j e^{i\xi j} P_j \quad , \quad \Phi(\xi) = \int e^{i\xi x} P(x) dx \tag{1.16}$$

と表される。連続的な場合について注目し exp 部分を Taylor 展開すると

$$\Phi(\xi) = \int e^{i\xi x} P(x) dx = \int \sum_{n=0}^{\infty} \frac{(i\xi x)^n}{n\ !} P(x) dx = \sum_{n=0}^{\infty} \frac{i\xi}{n\ !} \int x^n P(x) dx = \sum_{n=0}^{\infty} \frac{i\xi}{n\ !}\ M_n \tag{1.17}$$

と出来る。つまり全ての M_n が分かれば $\Phi(\xi)$ が分かり、Fourier 逆変換をすれば確率分布 $P(x)$ の形状が分かるのである。M_n は実際の観測される量と密接に関係しているので観測情報と物理量の裏に隠れている理論を見出す手がかりとなるのである。離散的な場合も同様である。

1.7　エントロピー (Entropy)

ある確率実験の結果を予想する際、その確率分布が偏っていると予想しやすくなるだろう。しかしながら逆に平坦な分布であると予想しにくくなる。簡潔な言葉を用いれば、前者は不確定性が小さく、後者は大きいと言える。この様な不確定性を表す物理量エントロピー[9] を定義していこう。

1.7.1　エントロピーの性質

先に述べたような性質、不確定性を表す量をエントロピーと呼び σ で表す。確率実験 α、単純事象が k 個ある時のエントロピーを $\sigma_\alpha(k)$ と書くとしよう。しかしエントロピーの具体的な形はさっぱり分からない。そこでエントロピーの性質より、要請される形を考える事にする。

1, $\sigma_\alpha(k)$ は k について単調増加である。
k の数が増える、事象の数が増加する、確率空間が広がる、これ等の事が起こると勿論、結果は予想しにくくなるわけであるから、エントロピーは増加する性質を見せる筈である。従って

$$\frac{\partial \sigma_\alpha(k)}{\partial k} \geq 0 \tag{1.18}$$

少なくともこの性質を持つ関数形をしている。

2, k が１つ (only) しか無い場合、$\sigma_\alpha(k)$ はゼロである。
事象が１つしか起こりえないとき不確定性は当然無いということ。即ち

$$\sigma_\alpha(k : \mathrm{only}) = 0 \tag{1.19}$$

3, α とは独立な実験 β があり、そのエントロピーが $\sigma_\beta(n)$ とするとき α, β の合成実験によるエントロピー $\sigma_{\alpha\beta}(kn)$ は $\sigma_\alpha(k)$ と $\sigma_\beta(n)$ の和で表される。

$$\sigma_{\alpha\beta}(kn) = \sigma_\alpha(k) + \sigma_\beta(n) \tag{1.20}$$

これら三つの性質を満たす関数形は対数関数しかないだろう。よってエントロピーは

$$\sigma_\alpha(k) = const.\ \log_a(k) \tag{1.21}$$

の形をしている。const. が Boltzmann 定数 k_B の時、熱力学的なエントロピーになる。熱力学的なエントロピーは少し異なった概念を含む事に注意する。この概念は熱力学的な立場からであるからここでは触れないでおく。

三番目の性質はエントロピーの相加性と呼ばれるものであるが、示している様に単なる和で表現される事は不可解である。もともと熱力学で現れた物理量であるから実際のところ結構天下り的で統計力学的には確率を絡めた話にしないとよく分からない。例を見ていく中で理解されると思う。

[9] この単語は著者にとって天敵であった。熱力学を学んでいた時、悲しいくらい意味が取れなくて魔法の言葉くらいにしか思ってなかった。しかし統計力学でこの定義を見たとき非常に合点のいく物理量で興味深いと思った。読者の皆様にもそんな気持ちになったらなぁと思う。感想はこれくらいで...

1.7.2　一般的な確率分布に対するエントロピー

先述のエントロピーの形は確率事象 k 個で一様な確率分布を持つものに対応している。いきなりそう言われても分からんので具体的な例を考えながらエントロピーの形をより明らかにしていこう。
今エントロピーを次の様に書くとしよう。

$$\sigma_\alpha(k) = \log_a(k) \tag{1.22}$$

この対数関数の底である a はエントロピーの単位として選択される確率試行の事象数を表す。具体的に言えば $a = 2$ は硬貨の裏表での確率試行の不確定性との比較で考えるよ、という事。

例 カードを数枚から 1 枚を選択する。(どのカードも等確率でめくられる。)
カード 2 枚から 1 枚を選ぶ。$a = 2$ でこのエントロピーを考えると

$$\sigma(2) = \log_2(2) = 1 \tag{1.23}$$

これは即ち、この確率試行が硬貨の裏表を決めるのと同じくらいの不確定性があるという事。

カード 4 枚から 1 枚を選ぶ。$a = 2$ でこのエントロピーを考えると

$$\sigma(4) = \log_2(4) = 2 \tag{1.24}$$

これは硬貨の裏表を決め方の倍の不確定性を持つという事

勿論、カード 3 枚から 1 枚を選ぶとしても

$$\sigma(3) = \log_2(3) = 1.585\ldots \tag{1.25}$$

と不確定性を表す事が出来る。
更に一般化するために底を $a = e$ とし自然対数にする。比較が硬貨から指数関数になっただけである[10]。次に一般的な確率分布に対するエントロピーについて考える事にする。
一般に離散的確率試行におけるエントロピーは以下の式で表される。

$$\sigma_\alpha[P_i] \equiv -\sum_{i=1}^{k} P_i \ \mathrm{ln} P_i \tag{1.26}$$

$P_i = 1/k$ に注意して (1.26) が (1.22) を満たす事を示すと

$$\sigma_\alpha[P_i] \equiv -\sum_{i=1}^{k} \frac{1}{k} \ \mathrm{ln}\frac{1}{k} = \mathrm{ln}\ k \tag{1.27}$$

と満足はしている。具体的に (1.26) の形になる事を見せよう。
前の例で 4 枚のカードから 1 枚を選択する実験を見た。この時 k は事象の数であり、連続的な確率を持つものに移行できない。そこで離散、連続のどちらにも適用できそうな確率分布を用いてこのエントロピーを表そうというのである。先ずは簡単のために $a = 2$ として $k \to P_i(= 1/k)$ でエントロピーを考える。

$(k = 4)$

$$\sigma_\alpha[P_i] = \log_2 \frac{1}{4} = -2$$

[10]大きな変化であると思いますがどんなものと比較してもいいので基準は分かり易い、計算しやすいものにしましょう。

という事で符号を変えないように「-」をつけて

$$\sigma_\alpha[P_i] = -\log_2\frac{1}{4} = 2$$

と、答えは一緒になったが意味がおかしい。これは 1/4 の確率試行の一部を考えたに過ぎない。4 回の試行を全て行って初めてこの実験のエントロピーになるのだから全て足し合わせる。

$$\sigma_\alpha[P_i] = -\log_2\frac{1}{4} - \log_2\frac{1}{4} - \log_2\frac{1}{4} - \log_2\frac{1}{4} = 8$$

これは規格化がされていない事が原因で起きている。
それぞれの事象に規格化された各々の確率をかけてやれば

$$\sigma_\alpha[P_i] = 4\left(-\frac{1}{4}\log_2\frac{1}{4}\right) = 2$$

となる。以上の議論を一般的に形にすると

$$\sigma_\alpha[P_i] \equiv -\sum_{i=1}^{k} P_i\ \ln P_i$$

の形になるのである。同様にして考えれば連続的確率分布に対するエントロピーは

$$\sigma[P(x)] \equiv -\int dx P(x)\ \ln P(x) \tag{1.28}$$

と書ける。
さてここまできてこれらの形の「エントロピー」が先の性質を全て満たしているのか確かめてみる。
1, k について単調増加である。 $\rightarrow k$ の増加は P_i を小さくするので明らかに満足する。
2, k が１つ (only) しか無い場合、$\sigma_\alpha(k)$ はゼロである。 $\rightarrow P_i = 1$ の事であるから明らかに満足する。
3, 互いに独立な確率試行 α, β によって考える合成実験 $\alpha\beta$ のエントロピーは α, β 各々のエントロピーの和で表される。

$$\begin{aligned}\sigma_{\alpha\beta} &= -\sum_i^k\sum_j^n P_{i\alpha,j\beta}\ \ln P_{i\alpha,j\beta} = -\sum_i^k\sum_j^n P_{i\alpha}P_{j\beta}\ \ (\ln P_{i\alpha} + \ln P_{j\beta}) \\ &= -\sum_i^k\sum_j^n P_{i\alpha}P_{j\beta}\ \ln P_{i\alpha} - \sum_i^k\sum_j^n P_{i\alpha}P_{j\beta}\ \ln P_{j\beta} = -\sum_i^k P_{i\alpha}\ \ln P_{i\alpha} - \sum_j^n P_{j\beta}\ \ln P_{j\beta} \\ &= \sigma_\alpha + \sigma_\beta \end{aligned} \tag{1.29}$$

$\sum_i^k P_{i\alpha} = \sum_j^n P_{j\beta} = 1$ ：規格化条件

確かに α, β のエントロピーの和で書けた。では α, β に相関があった場合どの様にかけるか。

$$\begin{aligned}\sigma_{\alpha\beta} &= -\sum_i^k\sum_j^n P_{i\alpha,j\beta}\ \ln P_{i\alpha,j\beta} = -\sum_i^k\sum_j^n P_{i\alpha}P_{i\alpha;j\beta}\ \ (\ln P_{i\alpha} + \ln P_{i\alpha;j\beta}) \\ &= -\sum_i^k P_{i\alpha}\ \ln P_{i\alpha} - \sum_i^k P_{i\alpha}\sum_j^n P_{i\alpha;j\beta}\ \ln P_{i\alpha;j\beta}\ \equiv\ \sigma_\alpha + \sigma_\beta^{(\alpha)} \end{aligned} \tag{1.30}$$

とこの様に書く事は出来る。$\sigma_\beta^{(\alpha)}$ を条件付エントロピーと呼ぶが、なんら具体的な形は見えない。
少しエントロピーの性質を見るために次の例題を考えよう。

(例題) 最大のエントロピー
規格化条件

$$\sum_i^k P_i = 1$$

よりエントロピー $\sigma[P_i]$

$$\sigma[P_i] = -\sum_{i=1}^k P_i \ \mathrm{ln} P_i$$

が最大になる確率分布 P_i を求めよ。

(解答)
求める確率分布を $\bar{P}_i$ とし、エントロピー σ の変分 $\delta\sigma$ を考える。

$$\begin{aligned}
\delta\sigma &= \sigma(\bar{P}_i + \delta P_i) - \sigma(\bar{P}_i) \\
&= -\sum_i^k (\bar{P}_i + \delta P_i) \ \mathrm{ln}(\bar{P}_i + \delta P_i) + \sum_i^k \bar{P}_i \ \mathrm{ln}\bar{P}_i \\
&= -\sum_i^k \bar{P}_i \ \mathrm{ln}(\bar{P}_i + \delta P_i) - \sum_i^k \delta P_i \ \mathrm{ln}(\bar{P}_i + \delta P_i) + \sum_i^k \bar{P}_i \ \mathrm{ln}\bar{P}_i \\
& \left(\begin{array}{l} \mathrm{ln}(\bar{P}_i + \delta P_i) = \mathrm{ln}\bar{P}_i \left(1 + \frac{\delta P_i}{\bar{P}_i}\right) = \mathrm{ln}\bar{P}_i + \mathrm{ln}\left(1 + \frac{\delta P_i}{\bar{P}_i}\right) \simeq \mathrm{ln}\bar{P}_i + \frac{\delta P_i}{\bar{P}_i} \\ \sum_i^k \delta P_i = 0 \quad \Rightarrow \quad -\sum_i^{k-1} \delta P_i = \delta P_k \quad , \quad \delta P_i^2 = 0 \text{ とする。} \end{array} \right) \\
&\simeq -\sum_i^k \bar{P}_i \ \left(\mathrm{ln}\bar{P}_i + \frac{\delta P_i}{\bar{P}_i}\right) - \sum_i^k \delta P_i \ \left(\mathrm{ln}\bar{P}_i + \frac{\delta P_i}{\bar{P}_i}\right) + \sum_i^k \bar{P}_i \ \mathrm{ln}\bar{P}_i \\
&= -\sum_i^k \delta P_i \left(1 + \mathrm{ln}\bar{P}_i\right) \\
&= -\delta P_k \left(1 + \mathrm{ln}\bar{P}_k\right) - \sum_i^{k-1} \delta P_i \left(1 + \mathrm{ln}\bar{P}_i\right) = \sum_i^{k-1} \delta P_i \left(1 + \mathrm{ln}\bar{P}_k\right) - \sum_i^{k-1} \delta P_i \left(1 + \mathrm{ln}\bar{P}_i\right) \\
&= -\sum_i^{k-1} \delta P_i \mathrm{ln} \ \frac{\bar{P}_i}{\bar{P}_k}
\end{aligned} \tag{1.31}$$

となる。
$\delta\sigma = 0$ より任意の δP_i で成立するのは $\bar{P}_i = \bar{P}_k$ を満たせばよい。これは事象全てが同確率で現れる事を意味するから、エントロピーが最大になる時、確率分布は一様分布である。