

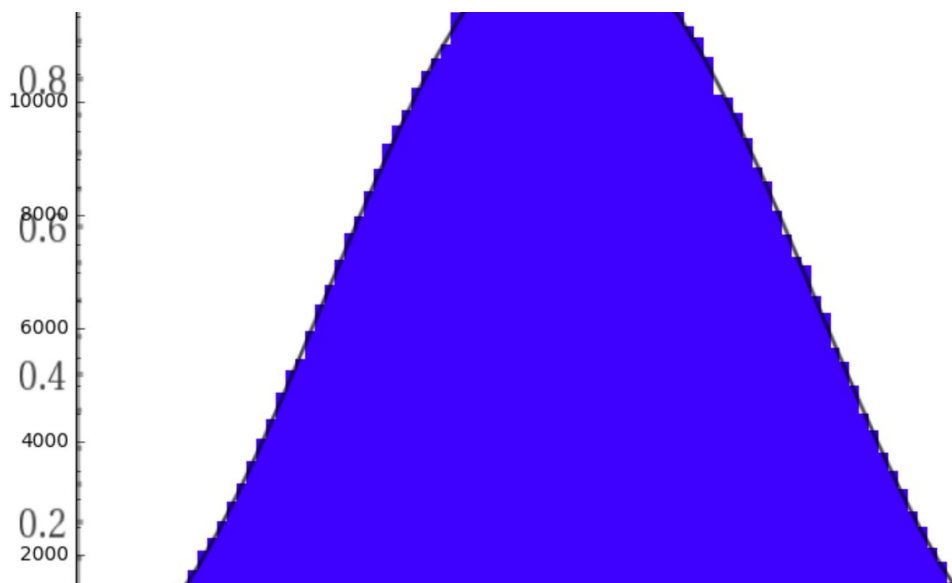


黒木玄 Gen Kuroki

@genkuroki

[mathtod.online/@iwaokimura/257...](https://mathtod.online/@iwaokimura/257...)

Sato-Tate 予想の例のグラフ。

 $y = \sin^2 x$  のグラフと重ねてみた。 [mathtod.online/media/fraB04cae...](https://mathtod.online/media/fraB04cae...)


2017年06月12日 15:42 · Web · 2 · ★ 4 · Webで開く



黒木玄 Gen Kuroki @genkuroki

Monday at 4:01pm

 $\sin^2$  分布は3次元球面(表面が3次元)上の一様分布を1次元の軸に射影したものに一致する。

その射影は  $A \in SU(2) \cong S^3$  をそのトレース  $\text{trace } A$  に移す写像と同一視でき、さらにこの写像は  $A \in SU(2)$  をその特性多項式  $\det(xE - A)$  に移す写像とも同一視でき、さらにさらにこれは  $A \in SU(2)$  をそれを含む共役類に移す写像とも同一視できる。

Sato-Tate予想は正標数の素体上での楕円曲線の有理点の個数がまるで  $SU(2) \cong S^2$  上の一様分布から来ているように見えるという予想。

佐藤幹夫さんがどのようにして  $\sin^2$  予想を発見したかについては次のリンク先が詳しくかつ面白い。

[www2.tsuda.ac.jp/suukeiken/mat...](http://www2.tsuda.ac.jp/suukeiken/mat...)

難波完爾さんが書いたものは面白いです。数学の世界が各分野ごとに狭く閉じていない雰囲気をもっと出してくれる。



黒木玄 Gen Kuroki @genkuroki

Monday at 4:21pm

より正確に言えば  $\sin^2$  分布の確率密度関数は

$$\frac{2}{\pi} \sin^2 \theta, \quad 0 \leq \theta \leq \pi$$

であり、これを  $x = \cos \theta$  と変数変換して得られる確率密度関数

$$\frac{2}{\pi} \sqrt{1-x^2}, \quad -1 \leq x \leq 1$$

を持つ確率分布が3次元単位球面  $x^2 + y^2 + z^2 + w^2 = 1$  上の一様分布を  $x$  軸に射影したものである。

一般に  $m$  次元単位球面上の一様分布の1次元への射影の確率密度関数は

$$Z_m^{-1}(1-x^2)^{(m-2)/2}$$

の形になる。ここで

$$Z_m = B(1/2, m/2) = 2^{m-1} B(m/2, m/2).$$

この公式は確率の総和が1になるという要請を2通りに計算すれば得られる。等号の2つ目のベータ関数の等式はLegendreのduplication formulaと同値。



黒木玄 Gen Kuroki @genkuroki

Monday at 4:41pm

$n-1$ 次元単位球面上の一様分布を1次元に射影して得られる分布の密度関数は

$$(1-x^2)^{n/2-3/2}, \quad -1 \leq x \leq 1$$

に比例する(定数倍を除いて一致)。ゆえに半径の二乗が  $n$  の球面上の一様分布を1次元に射影して得られる分布の密度関数は

$$\left(1 - \frac{1}{n/2} \frac{x^2}{2}\right)^{n/2-3.2}$$

に比例する。この  $n \rightarrow \infty$  での極限は

$$e^{-x^2/2}$$

である。標準正規分布の確率密度関数はこれに比例する。

半径の二乗が  $n$  の  $n-1$ 次元球面上の一様分布を1次元に射影して得られる分布の  $n \rightarrow \infty$  での極限は標準正規分布になる。

統計力学の教科書にあるMaxwell-Boltzmann分布の話は本質的に以上のような話である。

[genkuroki.github.io/documents/...](https://genkuroki.github.io/documents/...)  
の第9.8節を見よ。



黒木玄 Gen Kuroki @genkuroki

Monday at 5:14pm

各自由度ごとの運動エネルギーが

$$\frac{1}{2} v_i^2$$

と書けていて、自由度は  $n$  で運動エネルギーの総和は  $n/2$  であるとする。

$$\sum_{i=1}^n v_i^2 = n.$$

これは半径の二乗が  $n$  の  $n-1$ 次元球面の定義式に一致。

$v_i$  達はこの球面上で一様分布していると仮定する(等確率の原理とエネルギー保存則を仮定)。  $n$  が大きくなると、任意に選んだ  $v_i$  たちのうちの1つは標準正規分布に従うことを上で示した。まさにMaxwell-Boltzmann分布の話。

このタイプの正規分布の出現の仕方は中心極限定理とは異なる。この点を誤解してはいけない。

統計学を支える確率論における3つの基本定理は、大数の法則と中心極限定理とそしてSanovの定理である。Maxwell-Boltzmann則は中心極限定理ではなく、Sanovの定理から従う。

[genkuroki.github.io/documents/...](https://genkuroki.github.io/documents/...)



黒木玄 Gen Kuroki @genkuroki

Monday at 5:47pm

訂正：

$$\left(1 - \frac{1}{n/2} \frac{x^2}{2}\right)^{n/2-3.2}$$

は正しくはもちろん

$$\left(1 - \frac{1}{n/2} \frac{x^2}{2}\right)^{n/2-3/2}$$

です。こう書き直す前の式は

$$\left(1 - \frac{x^2}{n}\right)^{(n-3)/2}.$$



黒木玄 Gen Kuroki @genkuroki

Monday at 5:54pm

自由度→無限大の極限は「熱力学的極限」と呼ばれているようです。

何らかの自由度(次元)が極めて大きな熱浴があって、そこに注目したい系がどっぷりつかっているとき、注目したい系だけに実際に注目するとシンプルな良い性質が現れるというのは数学における一つの重要なパターンであるように思えます。

「熱浴」の概念を一般化するための中途半端な考察をツイッターに書いたことがありました。

[genkuroki.github.io/documents/...](https://genkuroki.github.io/documents/...)

[genkuroki.github.io/documents/...](https://genkuroki.github.io/documents/...)

[genkuroki.github.io/documents/...](https://genkuroki.github.io/documents/...)



黒木玄 Gen Kuroki @genkuroki

Monday at 6:09pm

話題をがんがん脱線させまくっていますが、「熱浴」っぽく見える数学的テクニックに「テンソルべきトリック」(tensor power trick)というのがあります。

[tricki.org/article/The\\_tensor\\_...](https://tricki.org/article/The_tensor_...)

[terrytao.wordpress.com/2007/09...](https://terrytao.wordpress.com/2007/09...)

数学用語で「トリック」と言えばだましのトリックという意味ではないです。

テンソルべきトリックとは

$$\alpha \leq \beta$$

を示すために任意の  $k$  について

$$\alpha^k \leq M\beta^k$$

を示す方法のことです。両辺の  $k$  乗根を取って  $k \rightarrow \infty$  とすると、 $M^{1/k} \rightarrow 1$  より  $\alpha \leq \beta$  が得られる。



黒木玄 Gen Kuroki @genkuroki

Monday at 7:04pm

テンソルべきトリックはWeil予想の証明でも使われています。

任意の  $d$  次元の  $X$  に関する

$$|\alpha| \leq q^{d/2}$$

を証明するために、任意の  $d$  次元の  $X$  に関して証明できる弱い条件

$$|\alpha| \leq q^{(d+1)/2}$$

から、 $nd$  次元の  $X^n$  に関する

$$|\alpha^n| \leq q^{(nd+1)/2}$$

を出して、

$$|\alpha| \leq q^{(d+1/n)/2}$$

の  $n \rightarrow \infty$  での極限で

$$|\alpha| \leq q^{d/2}$$

示す方針。そして、Poincare双対性より、逆向きの不等式も得られて

$$|\alpha| = q^{d/2}$$

が得られる。



黒木玄 Gen Kuroki @genkuroki

Monday at 7:25pm

テンソルベキトリックは以下のようにも使われる。

Mcmillanの不等式:  $r$  種類の文字  $s_1, \dots, s_r$  の  $b$  種類の文字による可逆な符号化で  $s_i$  の符号化( $b$  種類の文字の文字列になる)の長さが  $l_i$  のとき

$$a := \sum_{i=1}^r b^{-l_i} \leq 1.$$

証明: 正の整数  $n$  に対して、 $l_{i_1} + \dots + l_{i_n} = k$  を満たす  $s_{i_1} \dots s_{i_n}$  全体の個数を  $N_k$  と書くと、符号化が可逆なこと(単射性)より

$$N_k \leq b^k.$$

$l = \max\{l_1, \dots, l_r\}$  とおくと

$$\begin{aligned} a^n &= \sum_{k=0}^{nl} N_k b^{-k} \\ &\leq \sum_{k=0}^{nl} b^k b^{-k} = nl. \end{aligned}$$

$a^n \leq nl$  の両辺の  $n$  乗根を取って  $n \rightarrow \infty$  とすると  $a \leq 1$  が得られる. q.e.d.



黒木玄 Gen Kuroki @genkuroki

Monday at 7:53pm

このMcmillanの不等式から、Shannon情報量が可逆な符号化の平均符号長の下界になっていることが出ます。ファイルの中に文字  $s_i$  が確率  $p_i$  で現われるとすると、そのファイルのShannon情報量(Shannonエントロピーと言うことも多い)は

$$\sum_{i=1}^r p_i (-\log_b p_i)$$

と定義されます。一般に  $a_i \geq 0, \sum_{i=1}^r a_i \leq 1$  のとき、Gibbsの情報不等式より、

$$\sum_{i=1}^r p_i (-\log a_i) \geq \sum_{i=1}^r p_i (-\log p_i).$$

Mcmillanの不等式より、 $a_i = b^{-l_i}$  に関してこの不等式が使えるので、平均符号長について、

$$\begin{aligned} \sum_{i=1}^r p_i l_i &= \sum_{i=1}^r p_i (-\log_b a_i) \\ &\geq \sum_{i=1}^r p_i (-\log p_i). \end{aligned}$$



黒木玄 Gen Kuroki @genkuroki

Monday at 8:17pm

専門外過ぎるせいで、正直まだよくわかっていないことは、Sanovの定理と平均符号長に関する話の関係。

Sanovの定理はもろに「熱浴」の数学そのものだし、平均符号長の不等式を証明するときに使ったMcmillanの不等式の証明には「熱浴」っぽいテンソルベキトリックを使っているので、その辺の関係をうまく言えたらいいのだが。



黒木玄 Gen Kuroki @genkuroki

Monday at 9:07pm

あと「熱浴」ネタで面白そうなのでもっと理解したいのがPoincare予想のPerelmanさんによる証明における熱浴の利用。

[terrytao.wordpress.com/2008/04...](http://terrytao.wordpress.com/2008/04...)

にあるTaoさんによる解説ではPerelmanさんが使ったケースの類似として「Laplace方程式+熱浴」から「熱方程式」を出す話を扱っています。

その話はMaxwell-Boltzmann則の話の拡張になっています。正規分布の確率密度関数は分散を時間変数とする熱方程式の基本解になっています。

その意味での時間変数はMaxwell-Boltzmann則では温度に対応しています。

しかし、Laplace方程式との関係は統計力学の教科書でも見たことがありません。

自由度→無限大でのLaplace方程式と自由度有限での熱方程式の関係について書かれている統計力学の文献があれば教えて欲しいです。



黒木玄 Gen Kuroki @genkuroki

yesterday at 9:44am

「熱欲」話の続き

物理的には熱欲はでかければ何でも良いと思うのですが、そういう完全な一般論を考えるのではなく、以上で挙げた例に共通の条件(雰囲気)について考えてみる。

(1) Weil予想では有限体上の多様体  $X$  (注目する系)をその  $n$  個の直積  $X^n$  (熱欲)に対角的に埋め込み、 $n$  を大きくする。熱欲に埋め込むメリットは、素数定理に対応する荒い不等式から、Riemann予想に対応する精密な不等式を出せること。

Weil予想は色々難しいのでもっと易しい例を考える。続く



黒木玄 Gen Kuroki @genkuroki

yesterday at 9:59am

(2) 確率変数  $X$  (注目する系)を独立同分布確率変数列  $X_1, X_2, \dots, X_n$  (熱欲)の中の任意の1つとして埋め込むことは統計学における最も基本的な設定です。

物理における熱欲にズバリ対応するのは  $X_1, X_2, \dots, X_n$  について

$$\frac{1}{n} \sum_{i=1}^n h(X_i) \underset{\geq}{\overset{\leq}{\approx}} c$$

の型の制限のもとでの条件付き確率分布を考えることです。

物理ではエネルギー保存則があるので、状態  $x$  のエネルギーを  $h(x)$  としてエネルギーの総和が「一定」という条件を課しますが、不等式に条件を緩めた方が数学的には扱い易いです。

田崎晴明さんの統計力学の教科書でも不等式で制限をかけて議論をクリアにしています。



黒木玄 Gen Kuroki @genkuroki

yesterday at 10:11am

独立同分布確率変数列の1つとして元の確率変数(注目する系)を埋め込んで、独立同分布確率変数列の方で

$$\frac{1}{n} \sum_{i=1}^n h(X_i) \underset{\geq}{\overset{\leq}{\approx}} c$$

の型の制限による条件付き確率分布を考えて、 $n$  を大きくして、さらに注目する系だけに注目するようにすると、いわゆるカノニカル分布が得られます。

注目する系  $X$  の元々の確率密度関数が  $\varphi(x)$  ならば条件付き確率分布から得られる確率分布の確率密度関数は

$$Z(\beta)^{-1} e^{-\beta h(x)} \varphi(x),$$

$$Z(\beta) = \int e^{-\beta h(x)} \varphi(x) dx$$

になることを示せます。逆温度  $\beta$  は  $c$  から決まります。(Sanovの定理またはCramerの定理の帰結)



黒木玄 Gen Kuroki @genkuroki

yesterday at 10:27am

(3) さらに易しい例。(2)の例の特別な場合。Maxwell-Boltzmann則。統計力学の教科書に本質的に載っている。

$\mathbb{R}^n = \{(x_1, \dots, x_n)\}$  における原点を中心とする半径  $\sqrt{nt}$  の球面  $\sqrt{nt} S^{n-1}$  を考え、その上の一様分布の  $x_1$  軸への射影を考えましょう。その計算はせいぜい大学2年生レベル。

実際に計算すると、 $x_1$  軸上の分布の密度関数は

$$\left(1 - \frac{x^2}{nt}\right)^{n/2-3/2}$$

の定数倍になることがわかります。これは  $n$  が大きいとき

$$e^{-x^2/(2t)}$$

で近似されるので、 $n \rightarrow \infty$  での極限で  $x_1$  軸上の確率分布の密度関数は

$$\frac{1}{\sqrt{2\pi t}} e^{-x^2/(2t)}$$

になります。平均0、分散  $t$  の正規分布。この密度関数は熱方程式

$$u_t = \frac{1}{2} u_{xx}$$

を満たし、 $t \rightarrow +0$  でデルタ関数に収束する。



黒木玄 Gen Kuroki @genkuroki

yesterday at 11:00am

(4) さらにさらに易しい例。Mcmillanの不等式。

$r$  種類の文字  $s_1, \dots, s_r$  の各々に  $b$  種類の文字  $t_1, \dots, t_b$  を並べた文字列  $w_1, \dots, w_r$  を対応させることを符号化と呼びます。以下、 $w_i$  の文字列としての長さを  $l_i$  と書きます。

長さ  $n$  の文字列  $s_{i_1} \dots s_{i_n}$  は長さ  $k = l_{i_1} + \dots + l_{i_n}$  の文字列  $w_{i_1} \dots w_{i_n}$  に符号化されると考えます。

符号化が可逆であるとは  $s_{i_1} \dots s_{i_n}$  を  $w_{i_1} \dots w_{i_n}$  に移す写像が単射になることだと定義します。

続く



黒木玄 Gen Kuroki @genkuroki

yesterday at 11:03am

Mcmillanの不等式: 可逆な符号化について

$$a := \sum_{i=1}^r b^{-l_i} \leq 1.$$

証明:  $l_i$  の最大値を  $l$  と書くと、長さ  $n$  の文字列  $s_{i_1} \dots s_{i_n}$  の全体は  $b$  種類の文字を並べた長さ  $nl$  以下の文字列全体の中に単射で符号化される。

ゆえに、長さ  $k$  の文字列に符号化される  $s_{i_1} \dots s_{i_n}$  全体の個数  $N_k$  は  $b$  種類の文字を並べた長さ  $k$  の文字列全体の個数  $b^k$  以下になる。

したがって

$$a^n = \sum_{i=1}^{nl} N_k b^{-k}$$

$$\leq \sum_{i=1}^{nl} b^k b^{-k} = nl.$$

これより  $n \rightarrow \infty$  で

$$a \leq (nl)^{1/n} = e^{(1/n)\log(nl)} \rightarrow 1.$$

これで  $a \leq 1$  が示された。q.e.d.



黒木玄 Gen Kuroki @genkuroki

yesterday at 11:13am

証明のポイントは1文字の符号化の長さ  $l_i$  に関するMcmillanの不等式について直接考えるのではなく、長さ  $n$  の文字列  $s_{i_1} \cdots s_{i_n}$  の符号化に関する可逆性(単射性)から得られる自明な不等式を考えることです。

各文字  $s_i$  達(注目する系)の符号化は長さ  $n$  の文字列全体(熱欲)の符号化の中に埋め込まれており、長さ  $n$  の文字列の符号化の単射性による制限は各文字の符号化の制限(Mcmillanの不等式)を導くわけです。

Mcmillanの不等式の証明は、「注目する系」を「熱欲」に埋め込んで考えることとテンソルべきトリックの関係の最もわかり易い例になっていると思います。



黒木玄 Gen Kuroki @genkuroki

yesterday at 11:20am

$a^n$  の計算は以下の通り、 $a = \sum_i b^{-l_i}$  より

$$a^n = \sum_{i_1, \dots, i_n} b^{-(l_1 + \dots + l_n)}$$

$$= \sum_k \left( \sum_{l_1 + \dots + l_n = k} 1 \right) b^{-k}$$

$$= \sum_k N_k b^{-k}.$$

この手の計算はよくします。



黒木玄 Gen Kuroki @genkuroki  
一般に

yesterday at 12:36pm

$$p_i \geq 0, \quad \sum_{i=1}^r p_i = 1,$$

$$a_i \geq 0, \quad \sum_{i=1}^r a_i \leq 1$$

のとき

$$\sum_{i=1}^r p_i (-\log a_i) \geq \sum_{i=1}^r p_i (-\log p_i).$$

これはGibbsの情報不等式と呼ばれています。

Gibbsの情報不等式については、証明できなくても  $r = 2$  の場合に固定された  $p_1, p_2$  について  $a_1, a_2$  の函数としてのグラフをコンピューターで描いて納得するのも良いと思います。証明できる人もグラフは見た方が良いでしょう。

$a_i = b^{-l_i}$  に関するMcmillanの不等式とGibbsの情報不等式より、

$$\sum_{i=1}^r p_i l_i \geq \sum_{i=1}^r p_i (-\log_b p_i).$$



左辺は文字  $s_i$  が現われる確率が  $p_i$  のときの1文字あたりの平均符号長であり、右辺は Shannon情報量です。これはファイルの可逆圧縮の限界を示す不等式。



黒木玄 Gen Kuroki @genkuroki

yesterday at 3:02pm

訂正：ずーっと上の方で  $SU(2) \cong S^2$  と書いた部分があるが、もちろん正しくは  $SU(2) \cong S^3$ .

具体的には

$$\begin{bmatrix} z & -\bar{w} \\ w & \bar{z} \end{bmatrix} \in SU(2)$$

と

$$(a, b, c, d) \in S^3$$

が

$$z = a + ib, \quad w = c + id$$

によって一対一に対応している。



黒木玄 Gen Kuroki @genkuroki

yesterday at 3:04pm

誤植訂正版

$a^n$  の計算は以下の通り、 $a = \sum_i b^{-l_i}$  より

$$\begin{aligned} a^n &= \sum_{i_1, \dots, i_n} b^{-(l_{i_1} + \dots + l_{i_n})} \\ &= \sum_k \left( \sum_{l_{i_1} + \dots + l_{i_n} = k} 1 \right) b^{-k} \\ &= \sum_k N_k b^{-k}. \end{aligned}$$

この手の計算はよくします。



黒木玄 Gen Kuroki @genkuroki

yesterday at 3:20pm

例えば、以下  $p$  は素数を動くことにすると、

$$\frac{1}{1 - p^{-s}} = \sum_{k_p=0}^{\infty} p^{-k_p s}$$

について

$$\begin{aligned} &\prod_p \frac{1}{1 - p^{-s}} \\ &= \sum_{k_2, k_3, \dots} \left( \prod_p p^{k_p} \right)^{-s} \\ &= \sum_{n=1}^{\infty} \left( \sum_{\prod_p p^{k_p} = n} 1 \right) n^{-s} \\ &= \sum_{n=1}^{\infty} n^{-s}. \end{aligned}$$

最後の等号で素因数分解の一意存在を使った。



黒木玄 Gen Kuroki @genkuroki

yesterday at 3:31pm

他の例:  $n$  の分割数を  $p(n)$  と書く。例えば



$$\begin{aligned}
 4 &= 1 + 1 + 1 + 1 \\
 &= 1 + 1 + 2 \\
 &= 2 + 2 \\
 &= 1 + 3 \\
 &= 4
 \end{aligned}$$

の5通りなので、 $p(4) = 5$  です。 $n$  の分割は  $n = k_1 + 2k_2 + 3k_3 + \cdots$  (有限和) となる非負の整数達  $k_i$  で決まります。ゆえに

$$\frac{1}{1 - q^i} = \sum_{k_i=0}^{\infty} q^{ik_i}$$

について

$$\begin{aligned}
 &\prod_{i=1}^{\infty} \frac{1}{1 - q^i} \\
 &= \sum_{k_1, k_2, \dots \geq 0} q^{k_1 + 2k_2 + 3k_3 + \cdots} \\
 &= \sum_{n=0}^{\infty} \left( \sum_{k_1 + 2k_2 + 3k_3 + \cdots = n} 1 \right) q^n \\
 &= \sum_{n=0}^{\infty} p(n) q^n.
 \end{aligned}$$