超幾何関数と統計への応用

高山信毅 述
小山民雄 記

2013.9.2-2013.9.3

毎年１月に超幾何方程式研究会というのを神戸でやっていますが、今年から最低４年間、超幾何学校というのを夏に開こうということで、今回は、この第１回です。

早速、話題に移りたいですが、私の話では、こういう超幾何関数

$$\int_0^x t^{\frac{n}{2}-1}e^{-\frac{1}{2}}dt$$

を題材にして、我々が HGM (holonomic gradient method) と呼んでいるものの統計への応用を話す予定です。今年の朝ドラ [1] を見ていて北三陸編と東京編に分けるとスイッチできてよいと分かったので、これを参考にして、今日が代数編で明日が統計編になります。両方、かなり独立しています。

1 holonomic 系の積分イデアル

今日は代数編ということで、ホロノミック系の積分イデアルについて話します。ホロノミック系の積分イデアルを考えると、またホロノミック系になるという非常に基本的な定理があるのですが、それを証明するのが１日目の目標です。

1.1 Hilberlt 多項式

最初の４５分間で Hilbert 多項式について説明しましょう。まずは、簡単な話題から。F_k というのを

$$F_k = \{(x,y) \in \mathbf{N}_0^2 | x+y \leq k\}$$

[1] NHK 連続テレビ小説「あまちゃん」のこと

という格子点からなる集合だとしましょう (図 1、左、灰色の丸)。

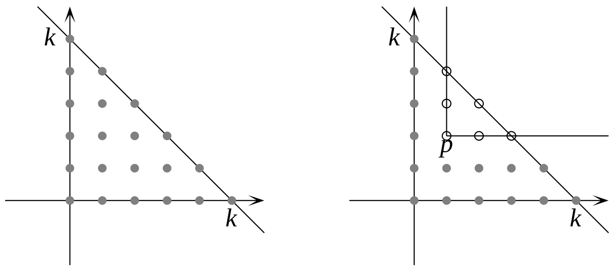

Figure 1: Hilbert 関数

このときに、F_k の要素の数を勘定すると、これは

$$
\begin{aligned}
\#F_k &= \binom{2+k}{2} = \frac{(k+2)(k+1)}{2} \\
&= \frac{1}{2!}k^2 + O(k).
\end{aligned}
$$

ここで、$O(k^m)$ は k の関数であって、$\left|\frac{O(k^m)}{k^m}\right|$ が $k \to \infty$ で有界であるものとします。

これは一番易しい例ですが、次の場合はどうなるでしょう？

$$F_k \backslash (p + \mathbf{N}_0^2).$$

図で描くと、これは図 1、右の図の灰色の丸の個数を勘定する問題になる訳ですね。これは、どうなるかというと、外側の大きい三角から内側の三角を引いてやればよろしいので、

$$
\begin{aligned}
\#\left(F_k \backslash (p + \mathbf{N}_0^2)\right) &= \binom{2+k}{2} - \binom{2+k-|p|}{2} \\
&= \frac{(k+2)(k+1)}{2!} - \frac{(k-|p|+2)(k-|p|+1)}{2!} \\
&= Ck + O(1) \\
&= O(k)
\end{aligned}
$$

という数になるわけです。ここで、$|p| = p_1 + p_2, p = (p_1, p_2)$.

一般的な問題としては、

$$\#\left(F_k\backslash\bigcup_{k=1}^{m}(P^{(i)}+\mathbf{N}_0^2)\right)$$

という個数の勘定します。図で描くと 図 2 の灰色の丸の個数を勘定しなさい

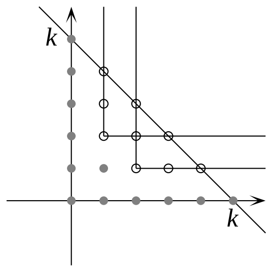

Figure 2: Hilbert 関数

という問題になる訳ですね。これは、今やった考え方が使えるはずです。全体の中から、これ引いてこれ引いて[2]、引き過ぎた分を足してやったりすると、この個数は二項係数の和と差で書ける訳です (k が十分大きいとき)。したがって、二項係数の部分は k の多項式で書けるので、個数は k の多項式として書けるということが推測できると思います。

さて以後、K と書いたら $\mathbf{Q}$, $\mathbf{C}$ のどちらかの体を表すことにして話を進めさせてください。

聴講者 B: $\mathbf{R}$(実数体) は使わないんですか？

実数体でも大丈夫です。

二変数の多項式環 $K[x_1,x_2]$ を考えます。この多項式環のイデアル I として、ある単項式 $x_1^{p_1}x_2^{p_2}$ で生成されるものを考えます。さっき F_k という記号を導入しましたが、ここで新たに新 F_k という記号を導入します。さっきの F_k とは違いますよ。新 F_k というのは、

$$F_k=\{\sum_{|\alpha|\le k}c_\alpha x_1^{\alpha_1}x_2^{\alpha_2}|c_\alpha\in K\}$$

[2]音声を字にするとこうなります。図から想像して下さい。

というものだとしましょう。ここで(講義記録音声で)絶対値と呼んでる $|\cdot|$ は、本物の絶対値じゃなくて、成分ごとの和をとった値 $\alpha_1+\alpha_2$ です。新 F_k は K-ベクトル空間になっているのて、この次元を勘定しなさいという問題を考えます。すると、以下の命題のように、さっきの二項係数の式と全く同じものが出てきます。

命題 1. *$I=\langle x^p\rangle$ $(x^p=x_1^{p_1}x_2^{p_2})$ の時、商線形空間 $F_k/(F_k\cap I)$ の K 上の次元は、次で与えられる。*

$$\dim_K F_k/(F_k\cap I)=\binom{k+2}{k}-\binom{k-|p|+2}{2}$$

Proof. まず、定義に戻って考えれば、$x^\alpha=x_1^{\alpha_1}x_2^{\alpha_2}\in I$ であるための必要十分条件は $\alpha\in p+\mathbf{N}_0^2$ で与えられます。これは、x^α が $F_k\cap I$ に入っているという条件は、ちょうど α が $p+\mathbf{N}_0^2$ と旧 F_k との共通部分に入っていることと同値であることを意味します。なので、さっきと全く同じ勘定をすることができて、命題の等式が言えます。つまり、$\beta\in F_k\backslash(p+\mathbf{N}_0^2)$ とすると、x^β が $F_k/(F_k\cap I)$ の K-ベクトル空間としての base であるということです。 □

一般にモノミアルで生成されるイデアル $I=\langle x^{p^{(1)}},\ldots,x^{p^{(m)}}\rangle$ に対しては、次が成り立ちます。

命題 2. *$\dim_K F_k/(F_k\cap I)$ は k が十分大で k の多項式となる。*

これは、次元の計算を格子点の数え上げに帰着することで、証明できます。怪訝な顔をしている人もいるので、例をやってみると、$K[x_1,x_2]$ のイデアル $I=\langle x_1x_2,x_1^2\rangle$ について、商ベクトル空間 $F_k/(F_k\cap I)$ の base は図 3 の

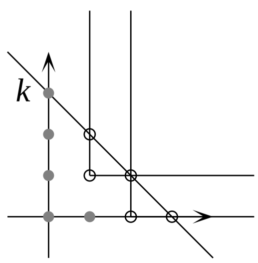

Figure 3: $I=\langle x_1x_2,x_1^2\rangle$

灰色の丸に対応するモノミアル $1,x_1,x_2,x_2^2,\ldots,x_2^k$ で与えられます。

この次元を表す多項式を I の Hilbert 多項式と呼び $H(k; I)$ と書きます。さらに、$H(k; I) = O(k^d)$ のとき、d を I の Krull 次元と呼びます。(Krull 次元 d は $V(I)$ の直感的次元と一致します。例えば $V(I)$ が manifold であれば d は manifold としての次元に一致します。)

話を微分作用素環に移しましょう。微分作用素環 $D_n = K\langle x_1, \ldots, x_n, \partial_1, \ldots, \partial_n\rangle$ とは、次の計算規則を持つ環です。

$$\begin{aligned} \partial_i x_j &= x_j \partial_i + \delta_{ij} \\ x_i x_j &= x_j x_i \\ \partial_i \partial_j &= \partial_j \partial_i \end{aligned}$$

もちろん、こんな風に定義してやって、結合法則とか分配法則とか成り立たせると、そのうち矛盾してくるかもしれないと不安になってきますが、それは大丈夫であることが、証明されています。本当に微分作用素環をきちんと定義するときには、多項式環を K ベクトル空間と見てやって、多項式環から多項式環への K 準同形が生成する代数で、多項式のかけ算と derivation で生成される部分代数として定義している訳ですね。これは問題を解くときに、非常に有用な考え方ですが [3]、今は一応こうゆうことにして、矛盾は生じないと信じてもらうことにします。

今、multi-index で $x^\alpha \partial^\beta = x_1^{\alpha_1} \cdots x_n^{\alpha_n} \partial_1^1 \cdots \partial_n^n$ と書かれた単項式の間の全順序を

$$\begin{aligned} &x^\alpha \partial^\beta > x^{\alpha'} \partial^{\beta'} \\ &\Leftrightarrow |\alpha| + |\beta| > |\alpha'| + |\beta'| \\ &\quad \text{or } (|\alpha| + |\beta| = |\alpha'| + |\beta'| \text{ かつ } (\alpha, \beta) - (\alpha', \beta') \text{ の最初の 0 でない成分が正}) \end{aligned}$$

で定めることにします。この順序を Graded lexicographic order と呼びます。微分作用素 $\ell = c_{\alpha\beta} x^\alpha \partial^\beta + (<$ についての低次項$)$ に対して多項式環 $K[x, \xi]$ の元

$$\mathrm{in}_<(\ell) = c_{\alpha\beta} x^\alpha \xi^\beta \in K[x, \xi]$$

を $<$ についての ℓ の initial と呼びます。微分作用素環 D_n の左イデアル I に対して、その (順序 $<$ についての) イニシャルイデアルを

$$\mathrm{in}_<(I) = \langle \mathrm{in}_<(\ell) | \ell \in I \rangle \subset K[x, \xi]$$

で定めます。$\mathrm{in}_<(I)$ はモノミアルイデアルです。左イデアル I の Hilbert 多項式を

$$H(k; I) = H(k; \mathrm{in}_<(I))$$

[3] 問 1.2 を参照

で定義しましょう。前半で、モノミアルイデアルに対する Hilbert 多項式を定義した訳ですが、これで、一般の左イデアルに対する Hilbert 多項式を定義します。こんなもの計算できるのと疑問に思う訳ですが、グレブナ基底を使うと Hilbert 多項式は計算できて、現在ではいろんな computer algebra のソフトウェアで hilbert なんとか... と入れると、すぐに計算してくれます (十分小さな問題なら)。

例 3. 一個くらい例をやっておきたいので、$I = \langle \partial_1, \ldots, \partial_n \rangle$ とすると、このイニシャルは、$\mathrm{in}_<(I) = \langle \xi_1, \ldots, \xi_n \rangle$ となります。これはグレブナ基底を使うことで示せます (この場合は使わなくても出来る)。$K[x,\xi]/\mathrm{in}_<(I) \cong K[x]$ なので、hilbert 多項式は $H(k;I) = \binom{k+n}{n} = O(k^n)$ と n 次式で書ける。

新たに $F_k = \{\sum_{|\alpha|+|\beta| \le k} c_{\alpha\beta} x^\alpha \partial^\beta\}$ をとってくると、次の定理が成り立ちます。

定理 4. k が十分大で

$$\dim_K F_k/(F_k \cap I) = H(k;I) = H(k; \mathrm{in}_<(I))$$

この定理の証明は、グレブナ基底の理論を勉強すれば出来ますが、これをやりだすと話が他所へそれてしまうので、今日はやりません ([2] の 1.6 節, 6.6 節参照)。一言で、グレブナ基底は何をやっているのかというと、イデアル I についての情報を得たいときには、$\mathrm{in}_<(I)$ についての情報をとってきてやれば大体分かるよというのが、一言で言うとグレブナ基底の理論ですから、こういうふうにやることが出来る訳です。

表題にあったホロノミック系について、説明したいのですが、そのためには、もう一つ定理が必要になります。

定理 5 (Bernstein 不等式)**.** D_n とは異なる D_n の左イデアル I に対して、$H(k;I) = O(k^m)$ ならば $m \ge n$ となる。

こういう非常に不思議な定理が成り立ちます。多項式環のイデアルであれば、その Krull 次元は任意の非負整数をとれる訳ですね。1 でも 2 でも 3 でも 4 でもね。だけど、微分作用素環の 左イデアルに対する Krull 次元相当のものは必ず n 以上になる、それ以上は小さくなれないという非常に強い結果が成り立ちます。

さっきの、例 3 を見てもらうと、このイデアルの Krull 次元は n で lower bound を達成します。少し考えてもらえば分かるんですが、このイデアルにどんな微分作用素を足しても、イデアルは全体 D_n になってしまい、Krull 次

元を $n-1$ とか $n-2$ とかにはできません。絶対できないということを保証しているのが、Bernstein の不等式です。

この定理の証明も時間の関係で省略します。たとえばグレブナ道場 [2] の 6.8 節に証明が書いてあるので、気になる人は読んでみてください。

$m=n$ となる I を holonomic ideal と呼びます。また、商加群 $M=D_n/I$ は holonomic D_n-module と呼ばれる物の一つの例になっています。ホロノミックは定義が難しいですが、難しい分いろいろよい性質を満たしているという仕組みになっています。

これで前半一コマ目の 45 分が終わりましたが、いつも講義しているときはここで 5 分間休憩して私はのんびりするんですが、今日はのんびりせずに何か質問があれば質問してください。何かないですか？

> **聴講者 B**:あの、I を与えたときに holonomic かどうかというのは...

グレブナ基底を計算して initial 出しますね。initial 出したら、今度は Hilbert 関数を計算しますね。その計算できます。

> **聴講者 B**:例えば、イデアルの generator を与えたときに、holonomic かどうかを判定してくれるソフトウェアなんかはあるんですか？

あります、あります。Macaulay2 には `is.holonomic` という関数があります。それか Risa/asir にも `ns_twistedlog.holonomic` [4] という関数があります。

> **聴講者 B**: initial ideal の生成系というのは、イデアルの generator の initial だけでは不足していますよね。そういうのを自動的に増やして計算してくれる訳ですか？

そうですね。どこまで計算すれば十分であるかというのを、判定するのがグレブナ基底の理論です。

Holonomic の必要十分条件というのは、いっぱいあって、そのどれかをチェックすれば良い訳ですが、一般に超幾何系が与えられたとき、その holonomic 性の判定は難しいものも多いです。

他に何かないですか？

> **聴講者 C**: グレブナ基底じゃなくて、抱合基底の一般的な判定法はありますか？

---

[4] `import("ns_twistedlog.rr");` でパッケージを読み込むと使える.

involutive base の一般的な判定法もグレブナ基底かな。

聴講者 C: ブラケットで閉じていることを...

ブラケットで閉じているような系を作るときもグレブナ基底を用います。例えば、今のような順序でグレブナ基底を作ってやると、必ず blacket で閉じます。

今日証明したい代数編の main theorem はこれです。この定理はホロノミック系の理論の黎明期に得られた基本定理です。この定理に出てくる積分加群の構成をグレブナー基底で遂行するアルゴリズムを与えたのが大阿久氏でご本人の解説本 [4] にとてもわかりやすい解説が書いてあります。グレブナー道場 [2] 6 章には明日解説する HGM への応用とともにこの定理が解説してあります。ここでは Björk の本 [3] に掲載の証明法を紹介します。

定理 6. D_n/I が *holonomic* D_n*-module* ならば、積分加群 $D_n/(I+\partial_n D_n)$ は *holonomic* D_{n-1}*-module*

$M = D_n/I$,$M_k = F_k/(F_k \cap I)$ と置く。I の積分 module は $M/\partial_n M$ と同型になる。以下、$\partial = \partial_n$, $x = x_n$ とおく。

$\partial : M \to M$ が単射のときは、$\partial M_{k-1} \subset M_k$ なので、

$$\begin{aligned}\dim_k M_k/(\partial M \cap M_k) &\le \left(\frac{m}{n!}k^n + O(k^{n-1})\right) - \left(\frac{m}{n!}(k-1)^n + O(k^{n-1})\right) \\ &= O(k^{k-1})\end{aligned}$$

K ベクトル空間としての同型 $F_k/(F_k \cap (I + \partial D)) \cong M_k/(\partial M) \cap M_k$ が成り立つので OK.

$\partial : M \to M$ が単射でないとき。

$$N = \{m \in M | \exists k \quad \partial^k m = 0 \quad \text{in } M\}$$

このとき、

(ア) N は左 D_n-module

(イ) $N \subset \partial M$

が成り立つとすると、$\bar{M} = M/N$ は左 D_n-module で $\partial : \bar{M} \to \bar{M}$ は単射となる。左 D_n-module としての同型 $M/\partial M \cong \bar{M}/\partial \bar{M}$ が成り立つので、単射の場合に帰着する。

(ア) の証明. 定義より N は左 $D_n\langle\partial_n\rangle$-module $m \in N$ に対して $x_n m \in N$ を示せばよいが、これは $\partial^k m = 0$ のとき

$$\begin{aligned}\partial^{k+1}(x_n m) &= x_n \partial^{k+1} m + (k+1)\partial^k m \\ &= 0\end{aligned}$$

より成り立つ。

(イ) の証明. $m \in N$ が $\partial^k m = 0$ を満たすとすると、

$$\partial^k x^k = x^k \partial^k + \sum_{i=1}^{k} \frac{1}{i!}\left(k(k-1)\cdots(k-i+1)\right)^2 x^{k-i}\partial^{k-i}$$

および

$$x^i \partial^i = (\partial x - 1)(\partial x - 2)\cdots(\partial x - i)$$

より、

$$\begin{aligned}\partial^k x^k m &= x^k \partial^k + \sum_{i=1}^{k} \frac{1}{i!}\left(k(k-1)\cdots(k-i+1)\right)^2 x^{k-i}\partial^{k-i} m \\ &= \partial(\dots)m + k!\sum_{i=1}^{k}\frac{(-1)^{k-i}k!}{i!(k-i)!}m\end{aligned}$$

したがって、$m = \partial(\dots)m \in \partial M$

最後に、$D_n/(I + \partial_n D_n)$ は左 D_{n-1}module として有限生成であることを示す。代表元から ∂_n^k は全部消せる。したがって、

$$\exists m_0 \forall m > m_0 x_n^m \equiv \sum_{i=0}^{m_0} \ell_i x_n^i, \quad \ell_i \in D_{n-1}$$

を示せばよい。

このとき、生成元は $1, x_n, x_n^2, \dots, x_n^{m_0}$ となる。$V_k = \{\sum_{\alpha_n - \beta_n \le k} c_{\alpha\beta} x^\alpha \partial^\beta\}$ とおく。これは kashiwara-malgrange filtration と呼ばれていて、M が holonomic のとき

$$\exists b(s) \in K[s] \exists \ell \in V_{-1} \ \text{ s.t. } \ b(\partial_n x_n) + \ell \in I$$

が成り立つことが知られています (簡単な可換環論を用いた証明は [6, Theorem 5.1.2, 5.1.3])。このような $b(s)$ で次数が最小のものを b 関数と呼ぶ。

$$\begin{aligned}x_n^m\left(b(\partial_n x_n) + \ell\right) &= b(\partial_n x_n - m)x_n^m + x_n^m \ell \\ &= \partial_n(\dots)x_n^m + b(-m)x_n^m + x_n^m \ell\end{aligned}$$

となるので、$b(-s) = 0$ の最大非負整数根 m_0 より大きい m に対しては $\bmod\ I + \partial_n D_n$ で x_n^m は x_n について m 次未満のもので書ける。

2　HGM, χ^2 分布を例として

二回目は、東京編で、一回目とは独立した話です (交わりはあるのですが)。今日は、holonomic gradient method を χ^2 分布を例として説明したいと思います。

> **聴講者 B**: (黒板に書かれた HGM という文字を見て) hypergeometric module なんて言うんじゃないのですか？

じゃないです。holonomic gradient method.

2.1　不完全 Γ 関数が統計学に出てくる理由

最初の４５分間は不完全 Γ 関数が統計学に出てくる理由を説明しましょう。実数上の関数 $T_n(x)$ を

$$T_n(x) = \begin{cases} x^{\frac{n}{2}-1}e^{-\frac{x}{2}}/N_T(n) & x > 0 \\ 0 & x \leq 0 \end{cases}$$

で定義します。ここで、$N_T(n)$ は、上の関数を積分したときに 1 になるための正規化定数で

$$N_T(n) = \int_0^\infty x^{\frac{n}{2}-1}e^{-\frac{x}{2}} = 2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)$$

で与えます。こんな風に定義してやると、関数 $T_n(x)$ は非負で等式 $\int_{-\infty}^\infty T_n(x)dx = 1$ が成り立つので、$T_n(x)$ は確率密度関数になります。関数 $T_n(x)$ を確率密度関数とする分布には、名前が付いていて χ^2 分布と呼ばれています。超幾何関数を勉強した人にとっては、これはガンマ関数に関係することが直に分かる訳ですね。

分布として χ^2 分布より、もっと有名で基本的なものには、密度関数が

$$e^{-\frac{(x-m)^2}{2\sigma^2}}/N, \quad N = \sqrt{2\pi\sigma^2}$$

で定義される平均 m、分散 σ^2 の正規分布 $N(m, \sigma^2)$ があります。正規分布というのは、確率統計において、一番基本的な分布ですが、χ^2 分布は、正規分布に従う独立な確率変数の 2 乗和が従う確率分布として現れます。

命題 7. $X_1,\ldots,X_n$ を $N(0,1)$ に従う独立な確率変数とすると、

$$Y=\sum_{i=1}^{n}X_i^2=(X_1,\ldots,X_n)\begin{pmatrix}X_1\\ \cdot\\ \cdot\\ \cdot\\ X_n\end{pmatrix}$$

は自由度 n の χ^2 分布に従う。

証明に入る前に、確率変数 (random variable) とは何かについて説明すると、ちゃんとした定義では、確率空間 $(\Omega,\mathcal{F},P)$ から $(\mathbf{R},\mathcal{B})$ への可測写像のことを確率変数と呼ぶのですが、直感的には乱数だと思ってください。例えば、確率変数を 1 個持ってきて、乱数を生成してヒストグラムを描いたら、密度関数のグラフと大体同じ形になるというのが「従う」の直感的な意味です。ちゃんとした定義は証明の中で述べます。

Proof. 証明は n に関する帰納法です ([8] の命題 7.5 (p.171) の証明)。まず、$n=1$ のときを考えましょう。確率変数 $Y=X^2$ が定数 $c\le 0$ より小さくなる確率は

$$P(Y<c)=P(-\sqrt{c}<X<\sqrt{c})$$

と書けますが、確率変数 X が $N(0,1)$ に従うことの定義は

$$P(a\le X\le b)=\int_a^b\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx\quad(a\le b)$$

だったので、変数変換 $y=x^2$ を用いると

$$\begin{aligned}P(Y<c)&=\int_{-\sqrt{c}}^{\sqrt{c}}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx=2\int_0^{\sqrt{c}}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx\\&=\int_0^c\frac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{1}{2}y}dy=\int_0^c T_1(y)dy\end{aligned}$$

となって、Y が自由度 1 の χ^2 分布に従うことが言えました。

次に、n まで定理が成り立っているとしましょう。次の定理を使います。

定理 8. 確率変数 X,Y が独立で、X と Y の分布の密度関数がそれぞれ $f(x),g(y)$ で与えられたとする。このとき $X+Y$ の分布の密度関数は $f*g(x)=\int_{-\infty}^{\infty}f(x-t)g(t)dt$ で与えられる。

この定理は、分布の定義に戻って計算すれば、簡単に得られます。そうすると計算するべきことは $T_n * T_1$ で、これが T_{n+1} なること言えば良い訳です。この計算は簡単なベータ関数の計算になります。実際にやってみると、

$$(T_n * T_1)(x) = \frac{1}{N_T(n)N_T(1)} e^{-\frac{x}{2}} \int_0^x (x-t)^{\frac{n}{2}-1} t^{-\frac{1}{2}} dt$$

となり (T_n と T_1 の台に注意せよ)、変数変換 $t = xs$ を用いると右辺の積分は

$$x^{\frac{n-1}{2}} \int_0^1 (1-s)^{\frac{n}{2}-1} s^{-\frac{1}{2}} dt = x^{\frac{n-1}{2}} \frac{\Gamma(\frac{n}{2})\Gamma(\frac{1}{2})}{\Gamma\left(\frac{n+1}{2}\right)}$$

となるので証明が完了する訳です。 □

これを多変量の正規分布でやる、すなわち確率変数の値域が $\mathbf{R}^d$ の場合に、これと同じような構成をやると所謂 matrix hypergeometric ${}_1F_1$ という、zonal 多項式がモノミアルの代りになる超幾何関数が自然に出てきます。

次に、早わかり検定論です。みなさんはまだ健康診断で赤信号が出るような年じゃないですが、例えば、「尿酸値の値が高くてね」という時にお肉を食べると尿酸値の値が高くなるという結果があるんですが、そういうのは統計的な検定論の結果なんです。

例えば、「コインを 3 回投げたら 3 回表が出た」という状況を考えます。このときに帰無仮説 (否定したい仮説) “このコインは公平である。すなわち、表が出る確率も裏が出る確率も $\frac{1}{2}$” を検定するという問題を考えます。このとき、統計的な仮説検定では、三回表がでるのと同じくらいかそれ以上珍しいことの起こることの確率を求めよ。こういうことをします。“これ以上珍しいこと” とというのが数学的じゃないので、問題や実験に応じて、いろいろ決めると私は理解しているのですが、そんなんでいいんですか聴講者 A 君。

> **聴講者 A** そうですね。どういう量を基準に検定するかは、検定統計量といい, その場その場に応じて経験的に決まっています。

この場合、コインが出るパターンを書いてみると (1 が表, 0 が裏)

0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1

の 8 通りあります。帰無仮説が正しいとすると、

$$P(\{\text{ 表が 2 回、裏が 1 回出る }\}) = P(\{\text{ 表が 1 回、裏が 2 回出る }\}) = \frac{3}{8}$$
$$P(\{\text{ 表が 3 回、裏が 0 回出る }\}) = P(\{\text{ 表が 0 回、裏が 3 回出る }\}) = \frac{1}{8}$$

と成ります。この場合は、「表が 3 回出る」ことと「裏が 3 回出る」ことを同じくらい珍しいと思いたいです (異論があるかも知れませんが)。これらが起こる確率のことを統計では P 値と読んでいます。すると、この値は $\frac{1}{8}+\frac{1}{8}=\frac{1}{4}=25\%$ と成ります。統計では伝統的に P 値が 5% または 1% 以下のときに、帰無仮説を捨てます。この場合には P 値が 25% なので、帰無仮説は否定できないとなります。否定できないからといって、肯定できる訳ではなくて、もっと実験が必要と成ります。つまり、こんな結果が出たからといって、まだ不公平だと結論は出来ないですよということです。これが統計的検定の一番簡単な例です。

ちょっと人工的な例題をやります。

人工的例題: ソフトウェア A には正規分布 $N(0, 0.25)$ に従う乱数を生成する関数 `rk()` があると宣伝している。命題 7 を用いて、上の宣伝を検定せよ。

命題 7 によれば, $\sum_{k=1}^{n} \texttt{rk()}^2/(0.25)^2$ は自由度 n の χ^2 分布に従うので以下の手順で検定する.

1. n 回 `rd()` を呼び出して値の 2 乗和 y を求める。
2. $\frac{y}{(0.25)^2} = \bar{y}$ とおく。
3. $\bar{y}$ がどれくらい珍しいか？ $\int_{\bar{y}}^{\infty} T_n(x)$ を計算して判断。5%以下なら帰無仮説を棄却。

例 9. 一様分布 (平均 0、分散 $(0.25)^2$) を `rd()` とする。$n = 50$ まで大きく取ると、P 値が 5% 以下になって帰無仮説が棄却される場合もある。ただし試行を繰り返すと P 値が 5% 以下にならない場合もある。n をもっと大きくとると (たとえば $n = 500$) 確実に棄却されることとなる。ここに掲載するのは統計ソフトウエア R による P 値の計算である。

Listing 1: p-value

```
> N<-50;
> a<-runif(N)-0.5;
> b<-sum(a^2)/(0.25)^2;
> p<-1-pchisq(b,df=N);
> p
```

```
[1] 0.2052216
#上をもう一度実行して
[1] 0.01388657
#上をもう一度実行して
[1] 0.02053797
#累積確率密度関数のグラフを描く.
> curve(pchisq(x,df=50),from=0,to=100);
```

なお、この例題は命題 7 のような定理をどのように検定等に使うかの精神の説明のための例題で、実際の乱数の検定ではもっとスマートな方法を用いています。たとえば D.Knuth, 準数値算法, 乱数を参照。

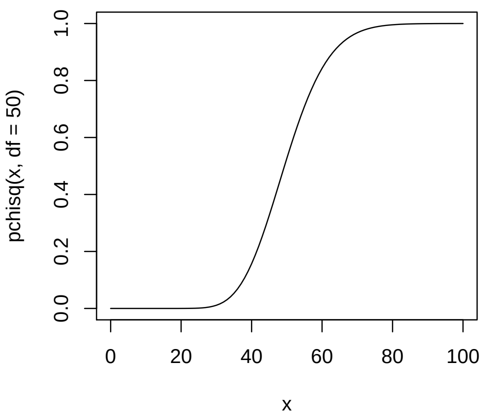

Figure 4: χ^2 分布の累積分布密度関数 `pchisq`, $\int_0^x T_n(t)dt$

2.2 HGM による P 値の計算

そしたら次の話題に行きましょう。関数

$$\gamma(x) = \int_0^x t^{\frac{n}{2}-1} e^{-\frac{t}{2}} dt$$

の数値計算を考えます。上の人工的な例題の場合では $1-\gamma(\bar{y})/\gamma(+\infty)$ が P-値になるのでこの数値計算は統計で重要です。

区間 $[0,x]$ の積分を扱うのは、何かと不便なので、関数 $\gamma(x)$ を Heaviside 関数を用いて

$$\int_{-\infty}^{\infty} H(x-t)H(t)t^{\frac{n}{2}-1}e^{-\frac{t}{2}}dt$$

の形に書きます。ここで、Heaviside 関数というのは、

$$H(t) = \begin{cases} 1 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

で定義される関数です。

今からやりたいのは、この積分表示を元にして、$\gamma(x)$ の満たす超幾何方程式を求めることです (やっと超幾何という言葉が出てきました)。昨日やった積分加群を使って考えることにします。単に微分方程式を求めるだけでなく、微分方程式を満たす理由を込みで説明しましょう。まず、超関数の基本的な関係式を復習すると

$$\partial_t H(t) = \delta(t)$$
$$t\delta(t) = 0$$

という公式がありました。ここで、$\delta(t)$ は delta 関数です。これらの関係式を組み合わせると $t\partial_t H(t) = 0$ が得られます。従って、$f(t,x) = H(x-t)H(t)t^{\frac{n}{2}-1}e^{-\frac{t}{2}}$ と置くと, 積の微分公式から

$$\partial_t f(t,x) = H'(t)H(x-t)g(t) - H(t)H'(x-t)g(t) + H(t)H(x-t)g'(t)$$

となってくれるわけです。delta 関数の性質に注意すれば、この両辺に $(x-t)t$ を掛けると右辺の第一項と第二項は消えてしまって、

$$(x-t)t\partial_t f(t,x) = H(t)H(x-t)g'(t)$$

が得られます。$g'(t)$ は計算すると有理式と $g(t)$ の積で書けることが分かるので、以上を整理すると、微分作用素

$$\ell_1 := (x-t)\left(t\partial_t - \left(\frac{n}{2}-1\right) + \frac{1}{2}t\right)$$
$$\ell_2 := (x-t)\partial_x$$

が関数 $f(t,x)$ を消去することが言えた訳です。今、微分作用素環 $D = K\langle x,t,\partial_x,\partial_t\rangle$ の左イデアル $I = D\ell_1 + D\ell_2$ を考えると、I に属する微分作用素は f を零化します (つまり $\ell \in I$ なら $\ell \bullet f = 0$ となる)。

さて今、$(I+\partial_t D)\cap K\langle x,\partial_x\rangle \ni \ell \neq 0$ とします。すると、

$$\ell = P_1 + \partial_t P_2 (P_1 \in I, P_2 \in D) \tag{1}$$

と書けるので、

$$\begin{aligned}\ell \bullet \int_{-\infty}^{\infty} f(x,t)dt &= \int_{-\infty}^{\infty} P_1 f(x,t)dt + \int_{-\infty}^{\infty} \partial_t P_2 f(x,t)dt \\ &= [P_2 f]_{-\infty}^{\infty} = 0\end{aligned}$$

となって、$\ell \bullet \gamma(x) = 0$ が求める常微分方程式 (O. D. E.) であることが分かります。ここで漸く、北三陸編じゃないや (笑)、昨日の代数編と繋がってですね、以下のような議論をします。

命題 10. *D/I はホロノミック*

Proof. グレブナー基底の考え方を使います。$x > t > \partial_x > \partial_t$ であるような graded lexicographic order で考えます. このとき $\mathrm{in}_<(\ell_1) = 2xt\xi_t$, $\mathrm{in}_<(\ell_2) = x\xi_x$ です。ここで $\xi_t = \mathrm{in}_<(\partial_t)$, $\xi_x = \mathrm{in}_<(\partial_x)$ とおきました. $\langle xt\xi_t, x\xi_x\rangle$ の Hilbert 多項式の次数は 3 です (計算練習)。調べるモノミアルが足りないので, $\partial_x\ell_1 - (2t\partial_t)\ell_2 - (t-n+2)\ell_2$ を計算すると, $\underline{2t\partial_x} + 2t\partial_t + t - n + 2$ となり, $\langle xt\xi_t, x\xi_x, t\xi_x\rangle$ の Hilbert 多項式の次数は 2 です (計算練習)。今 I で零化される超関数が存在するので $I = D$ とはなりません。したがって, Bernstein の不等式より D/I は ホロノミックです。 □

さて $K\langle x,\partial_x\rangle/(I+\partial_t D)\cap K\langle x,\partial_x\rangle$ は定理 6 の積分加群の部分加群となるので, ホロノミックです (詳しくは道場 [2] 6.7 節から 6.10 節を参照)。したがって $(I+\partial_t D)\cap K\langle x,\partial_x\rangle$ (これを積分イデアルと呼びます) は 0 で生成されるイデアルではありません。よって (1) のように書ける ℓ が存在することが言えます。これは $\gamma(x)$ の満たす微分方程式です。

実際、今の場合に Risa/Asir (コラム Risa/Asir とは、を参照) で積分イデアルを計算すると $J = \langle 2x\partial_x^2 + (x-n+2)\partial_x\rangle$ となります。

Listing 2: 積分イデアル

```
import("nk_restriction.rr");
L=[(x-t)*(t*dt-(n/2-1)+t/2),
   (x-t)*dx];
G=nk_restriction.integration_ideal(L,[t,x],[dt,dx],[1,0]);
```

最後に holonomic gradient method (HGM) について説明しましょう。$\gamma(x)$ や正規化定数 $N_T(n) = \gamma(+\infty)$ の数値計算は前節で説明したように応用上重

Risa/Asir とは

Risa/Asir の利点の一つは簡潔で安定した計算代数環境が無償で提供されていることおよび C 言語に似たユーザ言語が提供されており、本格的にソフト開発の勉強をしたいときには必ずマスターしておきたい C 言語の入門にもなるという点です。Risa/Asir や C 言語のプログラミングの入門や数学ソフトウエアの開発については私の講義録のページ [10] およびここからリンクされているビデオ、参考文献などを参照してください。

要です。応用に出てくる類似の問題の多くはパラメータ付の定積分の数値計算である場合が多く、この積分の数値計算を以下のような手順で行うことを我々は HGM と呼んでいます。

H. G. M. .
Step 1. 積分の満たす holonomic 系 J を計算する。
Step 2. J を Pfaffian へ
Step 3. 級数展開、数値積分などを初期値として **Step 2** の Pfaffian を数値解析して、広い範囲の積分値を求める。

この三つのステップを HGM の三段算法と呼びます。Step 2 に出てくる Pfaffian は

$$\frac{\partial Q}{\partial x_i} = P_i Q, \quad i = 1, \ldots, n$$

と書ける方程式系です (道場 [2] 6.2 節参照)。$n = 1$ の時は連立一階の常微分方程式にほかなりません。Holonomic 系は必ず Pfaffian に書換えることが可能です。Pfaffian は $\mathrm{grad}(Q) = (P_1 Q, \ldots, P_n Q)$ とも書けて、この関係式を用いて Q を計算するので我々はこの方法を holonomic gradient method と呼んでいます。

さて各ステップを $\gamma(x)$ の例で説明しましょう。Step 1 はすでに考察済みです。この例では理論的考察と計算機による計算を組み合わせましたが、問題によって理論的考察のみで導いたり、または計算機による計算のみで導いたりします。Step 2 では、$(2x\partial_x^2 + (x-n+2)\partial_x)\,\gamma(x) = 0$ なので、

$$\partial_x \begin{pmatrix} \gamma(x) \\ \partial_x \gamma(x) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & -\frac{1}{2x}(x-n+2) \end{pmatrix} \begin{pmatrix} \gamma(x) \\ \partial_x \gamma(x) \end{pmatrix}$$

Step 3. $\gamma(x)$ の冪級数展開を求めると、

$$\begin{aligned}\int_0^x t^{\frac{n}{2}-1}e^{-\frac{t}{2}}dt &= \sum_{k=0}^{\infty}\int_0^x t^{\frac{n}{2}-1}\frac{1}{k!}\left(-\frac{t}{2}\right)^k dt \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}\left(-\frac{1}{2}\right)^k \int_0^x t^{\frac{n}{2}-1+k}dt \\ &= \sum_{k=0}^{\infty}\frac{1}{k!}\left(-\frac{1}{2}\right)^k \frac{x^{\frac{n}{2}+k}}{\frac{n}{2}+k} \\ &= \frac{n}{2}x^{\frac{n}{2}}{}_1F_1\left(\frac{n}{2};\frac{n}{2}+1;-\frac{x}{2}\right)\end{aligned}$$

が得られます。差分法の初期値には、十分小さな x に対して、この冪級数で計算した近似値を用いることが出来ます。しかしながら、n が大きくなってくると、単純に差分法を適用するだけでは、精度が得られなくなってきます。このような場合には、$\gamma(x) = e^{-\frac{x}{2}}x^{\frac{n}{2}}h(x)$ と置いて、$h(x)$ に対する方程式 $\theta\left(\theta+\frac{n}{2}\right) - \frac{x}{2}(\theta+1)$ を数値解析するなどの工夫が必要になります。大きなパラメータを持つ Pfaffian の解の漸近展開などの理論も将来的には数値解析に応用されるのではないかと思ってます。

次の例は $\gamma(X)/N_T(N)$ を HGM で計算する Risa/Asir のプログラムです。

Listing 3: p-value

```
import("names.rr")$
import("taka_runge_kutta.rr")$
Glib_math_coordinate=1$
extern Aig$
/* Df : 自由度
   N : approx degree
*/
def poch(A,N) {
  R=1;
  for (I=0; I<N; I++) {
    R = R*(A+I);
  }
  return R;
}
def igs__(X,Df,N) {
  S = 1;
  for (K=1; K<=N; K++) {
    S += eval((X*1/2)^K)/poch(Df/2+1,K);
  }
```

```
  return S/(Df/2);
}
def igs_(X,Df,N) { return eval(igs__(X,Df,N)*exp(0)); }
def igs(X,Df,N) {    return igs_(X,Df,N)*eval(X^(Df/2)*exp(-X/2)); }
/* diff(igs), 級数で gamma(X) を計算. */
def igs1__(X,Df,N) {
  S = 0;
  for (K=1; K<=N; K++) {
    S += (1/2)*K*eval((X*1/2)^(K-1))/poch(Df/2+1,K);
  }
  return S/(Df/2);
}
def igs1_(X,Df,N) { return eval(igs1__(X,Df,N)*exp(0)); }
def ig_(X,Df) { /* h(X) の計算 */
  Step=0.01;
  X0=1;  N=40; /* 級数近似の次数 */
  Iv=[igs_(X0,Df,N),igs1_(X0,Df,N)];
  Eq=[[0,1],[1/(2*x),(-1/x)*(Df/2+1-x/2)]];
  /* Runge-Kutta 法で微分方程式の解を近似計算 */
  A=tk_rk.runge_kutta_4_linear(Eq,x,[],X0,Iv,X,Step);
  return A;
}
/* 自由度 Df で gamma(X) を計算. */
def ig(X,Df) {
  extern Aig;
  Aig=ig_(X,Df);
  A=[];
  for (I=0; I<length(Aig); I++) {
    V=Aig[I]; X=V[0];
    A = cons([X,V[1]*eval(exp(-X/2)*X^(Df/2))],A);
  }
  return reverse(A);
}
/* 正規化定数 N_T(Df) */
def nc(Df) {    return(pari(gamma,Df/2)*eval(2^(Df/2))); }
end$
```

$\gamma(70)$ の値および $\gamma(70)/N_T(50)$ の計算は以下のようになります。

```
[1893] load("evalig3.rr");
[1980] A=ig(70,50)$
[1981] A[0];
[70,2.01448037926921063 6 E31]
[1982] A[0][1]/nc(50);
0.9676258902264072996
```

本日は χ^2 分布を例として HGM の手法を解説しました。もちろん χ^2 分布の場合は別の手法による計算の方が断然うまくいきます。たとえば数値計算の有名な教科書 Numerical Recipes を読んで見て下さい。しかしながら、HGM によりはじめて現実的時間内に数値計算が可能になった統計分布の例も沢山あります。それらの研究については [11] をご覧下さい。HGM の超幾何関数論から見た面白さは数値計算と統計への応用という視点を持つことにより今までにない形のある意味新鮮な数学的問題を提供可能だという点だと思います。

3 練習問題

問 3.1. [10] $F_k = \{(m_1, m_2) \in \mathbf{N}_0^2 \,|\, m_1 + m_2 \leq k\}$ と置く. $p, q \in \mathbf{N}_0^2$ の時, $F_k \setminus (p + \mathbf{N}_0^2) \cup (q + \mathbf{N}_0^2)$ の要素の数を k の多項式として表せ (k が十分大きいとき). n 変数の時に同様な数えあげをすると?

問 3.2. [10] 微分作用素環 $\mathbf{Q}\langle x, \partial \rangle$ において次の等式を示せ.

1. $x^k \partial^k = \theta(\theta - 1) \cdots (\theta - k + 1)$, ここで $\theta = x\partial$.
2. $b(\theta) x^k = x^k b(\theta + k)$. ここで $b(\theta)$ は変数 $\theta = x\partial$ の一変数多項式.
3. $\partial^k x^k = x^k \partial^k + \sum_{i=1}^{k} \frac{1}{i!}(k(k-1) \cdots (k-i+1))^2 x^{k-i} \partial^{k-i}$.

問 3.3. [10] $e^{xt - t^n}$ の満す微分方程式系の t についての積分イデアルを計算せよ.

問 3.4. [15] $\int_0^{700} T_{500}(t) dt$ の値を適当なプログラム言語で近似計算せよ.

問 3.5. [20]

1. 不完全ガンマ関数を合流型超幾何級数 ${}_1F_1(a, c; x)$ で表せ.
2. この級数の隣接関係式 (a, c についての漸化式) を用いて不完全ガンマ関数の連分数表示を一つ求めよ.

問 3.6. [45] (研究課題) 多次元正規分布に対する χ^2 分布相当のものを考察すると行列引数の超幾何関数 ${}_1F_1$ が得られる. 古典的 ${}_1F_1$ のモノミアルを zonal 多項式にすると級数展開が, 積分領域を positive definite symmetric matrix のある集合にすると積分表示が得られる. この多変数関数に現代的な超幾何関数論の種々の手法を適用して数値計算に有用な新しい公式達を導きだせ. 参考:

1. A.G.Constantine, Some Non-Central Distribution Problems in Multivariate Analysis, The Annals of Mathematical Statistics 34 (1963), 1270–1285.

2. H.Hashiguchi, Y.Numata, N.Takayama, A.Takemura, The holonomic gradient method for the distribution function of the largest root of a Wishart matrix, Journal of Multivariate Analysis, 117, (2013) 296-312.

References

[1] D.Cox, J.Little, D.O'Shea, Ideals, Varieties, and Algorithms, Springer. 日本語訳もあり. グレブナー基底の基礎を知るにはこの本の 1 章を読んでから次のグレブナー道場の 1 章を読む, のが一つの方法.

[2] JST CREST 日比チーム, グレブナー道場, 共立出版.

[3] Björk, Rings of Differential Operators. この本の一章が Weyl 代数の初歩.

[4] 大阿久, D 加群と計算数学, 朝倉書店.

[5] T.Kimura, Hypergeometric Functions of Two Variables. ガウスの超幾何関数を含む多変数超幾何関数の入門書. ネットで検索して下さい.

[6] M.Saito, B.Sturmfels, N.Takayama, Gröbner Deformations of Hypergeometric Differential Equations, Springer.

[7] 竹村, 統計, 第 2 版, 共立出版. R システムを用いた統計の入門書.

[8] 小針, 確率・統計入門, 岩波書店. 上記の本に書いてない証明等がきちんと書いてある.

[9] `http://www.math.kobe-u.ac.jp/HOME/taka/2013/hgs-2013`

[10] `http://www.math.kobe-u.ac.jp/HOME/taka/2014/keisan-1/ref.html`

[11] `http://www.math.kobe-u.ac.jp/OpenXM/Math/hgm/ref-hgm.html`