

最大エントロピー分布の諸特性の検討

Study on Some Characteristics of Maximum Entropy Distribution

信州大学工学部 正 員 寒川 典昭

同 上 正 員 荒木 正夫

信州大学大学院 学生員 大原 基秀

同 上 学生員 森川 智

1. はじめに

最大エントロピー分布は, Pearson 系の分布あるいは Gram - Charlier の級数と同様一つの分布系と考えられる。^{1), 2)} また, それは, 1変数³⁾ から2変数⁴⁾, 多変数¹⁾, 条件つき分布²⁾へと拡張することができ, 水工計画上有効な分布群をなしている。我々は, いままでこのことを明らかにしてきたが, その間いくつかの検討課題が未解決のままであった。本稿では, その内, 原点にパルスをもつ最大エントロピー分布とその適用性, 高次モーメントまで制約条件とした場合の適用性, 複数の極値をもつ分布への対応, の3課題について検討する。これらは水文頻度分析上どこまで必要であるかという問題に議論を残すが, 最大エントロピー分布の特徴を把握するという見地から意義がある。また, 付録として, 上述の最大エントロピー分布のパラメーター同定の基礎知識として必要であるため, 黒田・横田⁵⁾が原点まわりのモーメントを制約条件とした離散的1変数最大エントロピー分布に対して行った, 推定された確率密度関数がエントロピーを最大にすること, 及び解の一意性の証明が, 多変数の任意関数を制約条件とした連続的多変数最大エントロピー分布についても同様に言えることを簡単に紹介する。

2. 諸特性の検討

(1) 原点にパルスをもつ分布

日降水量などのヒストグラムは原点に大きな頻度がみられる。このような問題に対処するため, 原点についてはパルスで確率を与え, その他の部分を最大エントロピー分布で適合することを考える。

a) 理論式

連続確率変数 $x_1, x_2, \dots, x_n$ に対する確率密度関数を $p(x_1, x_2, \dots, x_n)$ とすると, エントロピー $H(x_1, x_2, \dots, x_n)$ は

$$H(x_1, x_2, \dots, x_n) = - \int_0^\infty \dots \int_0^\infty \int_0^\infty p(x_1, x_2, \dots, x_n) \ln p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \quad (1)$$

また, 適切な任意関数を $g_r(x_1, x_2, \dots, x_n)$ ($r=1, 2, \dots, M$) とすると文献 1) の制約条件は次のように書き換えられる。

$$\int_0^\infty \dots \int_0^\infty \int_0^\infty p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = 1 - p(0, 0, \dots, 0) \equiv q \quad (2)$$

$$\int_0^\infty \dots \int_0^\infty \int_0^\infty g_r(x_1, x_2, \dots, x_n) p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n = E[g_r(x_1, x_2, \dots, x_n)], \quad r=1, 2, \dots, M \quad (3)$$

ここで, $E[\cdot]$ は期待値記号であり, q はパルス以外の確率を表わす。(2), (3)式の制約の下で (1)式を最大にする分布をラグランジュの未定乗数法で求めると

$$p(x_1, x_2, \dots, x_n) = \exp\left\{-\lambda_0 - \sum_{r=1}^M \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} \quad (4)$$

上式に含まれるパラメーター(ラグランジュ乗数) λ_r ($r=0, 1, \dots, M$) は次のように求められる。

(4) 式を (2)式に代入して変形すると

$$\exp(\lambda_0) = \frac{1}{q} \int_0^\infty \dots \int_0^\infty \int_0^\infty \exp\left\{-\sum_{r=1}^M \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} dx_1 dx_2 \dots dx_n \quad (5)$$

(4)式を(3)式に代入し、(5)式を用いて変形すると

$$\begin{aligned} & \int_0^\infty \dots \int_0^\infty \int_0^\infty g_r(x_1, x_2, \dots, x_n) \exp\left\{-\sum_{r=1}^M \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} dx_1 dx_2 \dots dx_n \\ &= \frac{E[g_r(x_1, x_2, \dots, x_n)]}{q} \int_0^\infty \dots \int_0^\infty \int_0^\infty \exp\left\{-\sum_{r=1}^M \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} dx_1 dx_2 \dots dx_n, \\ & \quad r=1, 2, \dots, M \end{aligned} \quad (6)$$

(6)式が λ_r ($r=1, 2, \dots, M$) を求める方程式である。具体的には、文献1)の $E[\cdot]$ を $E[\cdot]/q$ に置き換えれば、そこで述べているパラメーター同定法はすべて使える。(6)式で求まった λ_r ($r=1, 2, \dots, M$)を用いると λ_0 は(5)式、すなわち次式で求めることができる。

$$\lambda_0 = -\ln q + \ln \left[\int_0^\infty \dots \int_0^\infty \int_0^\infty \exp\left\{-\sum_{r=1}^M \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} dx_1 dx_2 \dots dx_n \right] \quad (7)$$

b) 実測データに対する適用

解析資料は長野県の藪原(1938年～1983年, 9月15日)と福島(1938年～1983年, 6月全部)の日降水量とする。Fig.1は藪原について、Fig.2は福島について、ヒストグラムに対するモーメントを制約条件とした1変数最大エントロピー分布の適合度をパルスを与えた場合と与えない場合で比較したものである。両図より

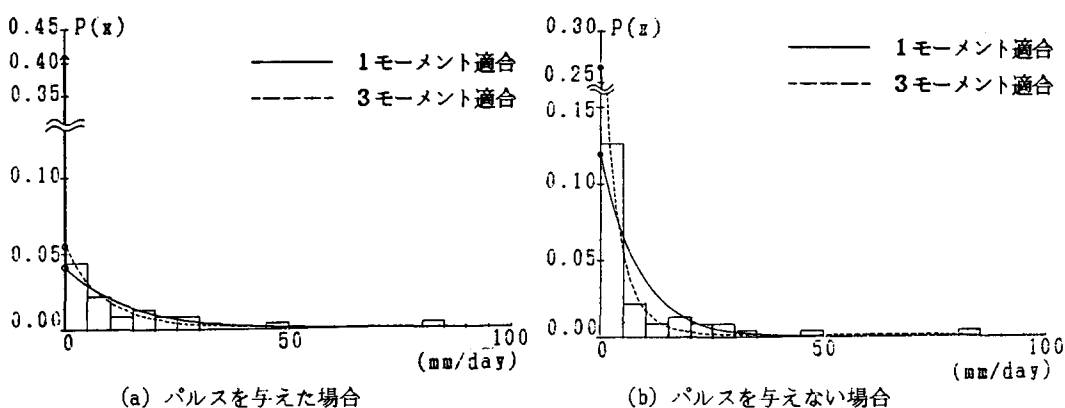


Fig.1 藪原日降水量のヒストグラムと最大エントロピー分布

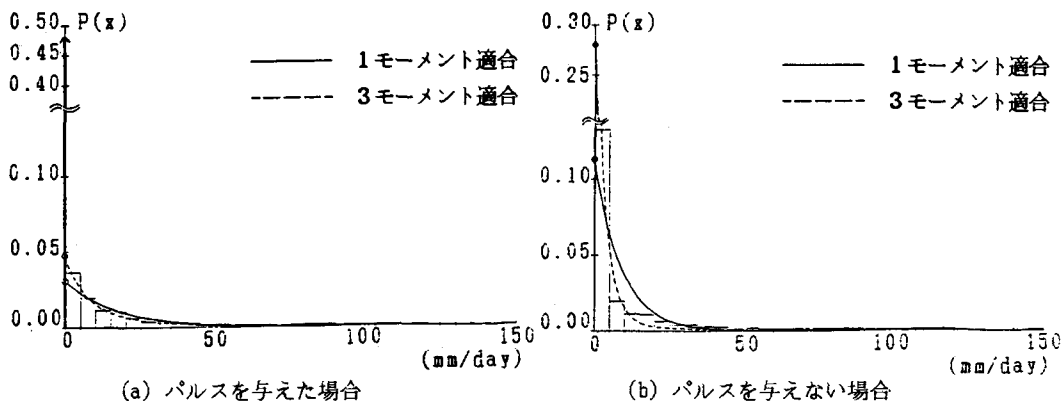


Fig.2 福島日降水量のヒストグラムと最大エントロピー分布

パルスを与えた場合適合度の改善がみられ、その有効性が読み取れる。

(2) 高次モーメントまで用いた適用性

我々が水文頻度解析をする上で知りたいのは母集団の確率密度関数であるが、それが未知であるため限られたデータからいかにうまく母集団を推定するかということに多くの問題が集約される。それを大別すれば確率密度関数の選択の問題とパラメーター推定の問題になる。後者の問題の1つに高次モーメントにおけるサンプリングエラーがある。しかし、そのモーメントも母集団に対する何らかの情報を有しており、最大エントロピー分布を利用する上で何次までのモーメントを用いるか、あるいはモーメントのもつ誤差をいかにおさえるかということについての検討は重要と思われる。しかし、ここでは、この分布の究極の特性を調べるため、あえて高次モーメントまで採用して、ヒストグラムに対するこの分布の適合度を考察する。

Table 1 は (1) のパルスをもつ場合の藪原の日降水量について同定した1変数最大エントロピー分布のラグランジュ乗数であり、Fig.3 はヒストグラムと5～8モーメント適合との対応を示したものである。推定

Table 1 藪原日降水量のラグランジュ乗数

ラグランジュ乗数	λ_0	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8
1 モーメント 適合	1.06545	0.58700							
3 モーメント 適合	0.73332	1.05764	-0.09686	0.00413					
5 モーメント 適合	0.30432	3.09804	-1.87683	0.52801	-0.05974	0.00236			
6 モーメント 適合	0.49847	1.79896	-0.20297	-0.25893	0.10310	-0.01275	0.00051		
7 モーメント 適合	0.23162	4.15565	-4.38585	2.55461	-0.78534	0.12825	-0.01035	0.00032	
8 モーメント 適合	0.84459	-2.69593	12.30450	-13.31000	6.58568	-1.70536	0.23766	-0.01678	0.00047

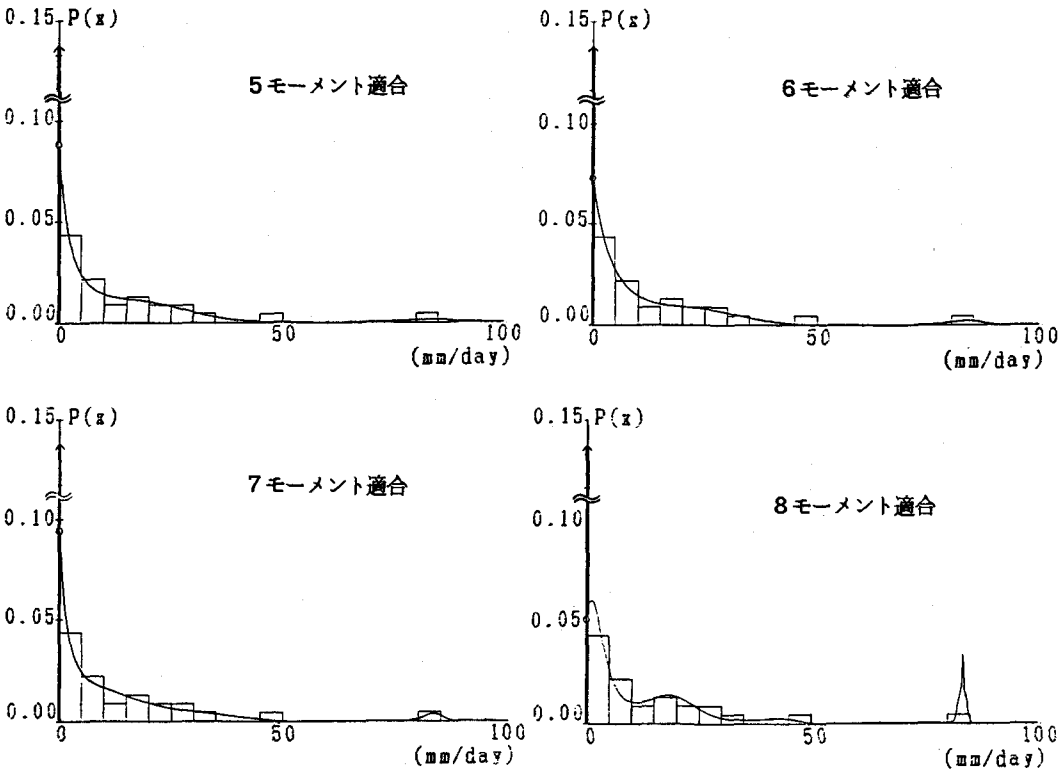


Fig.3 藪原日降水量の高次モーメント適合

された分布は、採用するモーメントの個数が増加するに伴い、ヒストグラムの細部にわたってよく記述できることがわかる。ここで提示していない他の計算例について同様なことが言えている。

(3) 複数の極値をもつ分布への対応

ここでも、(2)で述べたことと同様の目的で、複数の極値をもつ分布への対応について検討する。

a) 理論式

一般的な多変数最大エントロピー分布は文献1)より次式となる。

$$p(x_1, x_2, \dots, x_n) = \exp\left\{-1 - \sum_{r=1}^M \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} \quad (8)$$

ここで、 λ_r はラグランジュ乗数、 $g_r(\cdot)$ は適切な任意関数であり、ここでは $g_r(\cdot)=1$ を含むものとする。(8)式を x_1 で偏微分して 0 とおくと

$$\frac{\partial p(x_1, x_2, \dots, x_n)}{\partial x_1} = \exp\left\{-1 - \sum_{r=1}^M \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} \left\{-\sum_{r=1}^M \lambda_r \frac{\partial g_r(x_1, x_2, \dots, x_n)}{\partial x_1}\right\} = 0 \quad (9)$$

ここで、 $\exp\left\{-1 - \sum_{r=1}^M \lambda_r g_r(x_1, x_2, \dots, x_n)\right\} \neq 0$ より

$$\sum_{r=1}^M \lambda_r \frac{\partial g_r(x_1, x_2, \dots, x_n)}{\partial x_1} = 0 \quad (10)$$

上式は、 $x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n$ を固定したときの x_i 方向の極値を与える。いま制約条件として $g_r(x) = x^r$ を採用するなら、文献5)にも述べられているように s 個の極値をもつヒストグラムを記述するためには $r \geq s+1$ とする必要がある。

b) ヒストグラムへの適用

Fig. 4 のように人工的にヒストグラムを作成し、これに対してモーメントを制約条件とした1変数最大エントロピー分布の適用を試みた。Table 2はそのとき得られたラグランジュ乗数を、Fig. 4は1～5モーメント適合で得られた分布曲線を示している。ヒストグラムは4個の極値をもつため、5モーメント適合はよくヒストグラムを記述していることがわかる。

3. まとめと今後の課題

計算例が少ないので断定することはできないが、得られた結果をまとめると以下ようになる。

1) 原点に大きな確率をもつヒストグラムに対しては、パルスをもつ最大エントロピー分布を用いれば適合度が改善され、この分布は日降水量等の頻度解析に有効である。

2) 高次モーメント適合では、採用するモーメントの個数の増加に伴いヒストグラムの細部にわたってよく記述できる。ここで用いた例では8モーメント適合でほぼヒストグラムと合致している。

3) (10)式は極値を与える式であり、ここで用いた4個の極値をもつヒストグラムへの適用例では5モーメント適合でほぼヒストグラムに合致している。

今後の課題は、本稿で取り上げた問題に対する計算例を増やすとともに、

- 1) 高次モーメントのもつ誤差をいかに押さえるか
 - 2) 水文量ごとにいかに最適な $g_r(\cdot)$ を見つけるか
- 等である。

参 考 文 献

- 1) 寒川典昭・荒木正夫・佐藤健次：多変数最大エントロピー分布とその基礎特性に関する研究，土木学会論文集，No.375 / II-6, pp.89 ~ 98, 1985 年11月。

Table 2. 人工的なヒストグラムのラグランジュ乗数

ラグランジュ乗数	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6
1 モーメント 適合	-0.000020	1.000010				
2 モーメント 適合	1.142890	-1.648760	0.982830			
3 モーメント 適合	0.768556	0.071836	-0.759612	0.473096		
4 モーメント 適合	0.747348	0.224962	-1.016750	0.624380	-0.028642	
5 モーメント 適合	-0.139149	11.656400	-32.819100	34.308900	-15.094500	2.39487

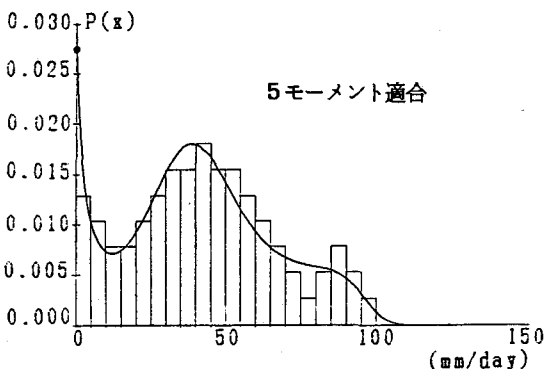
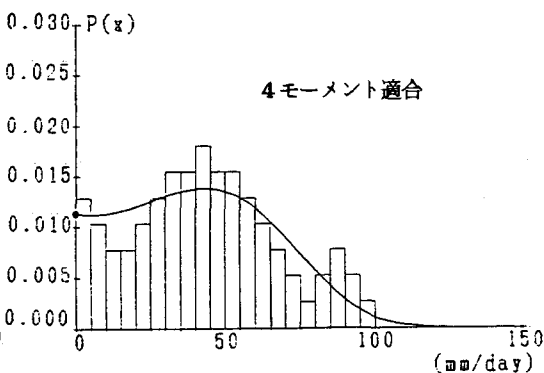
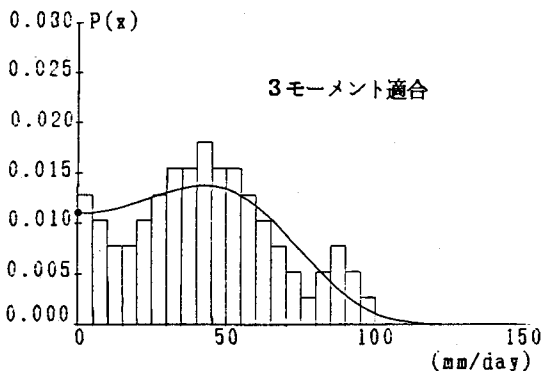
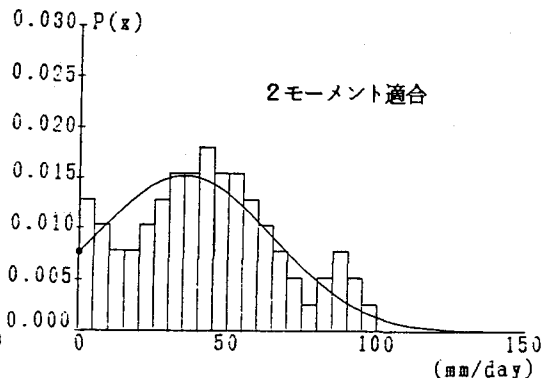
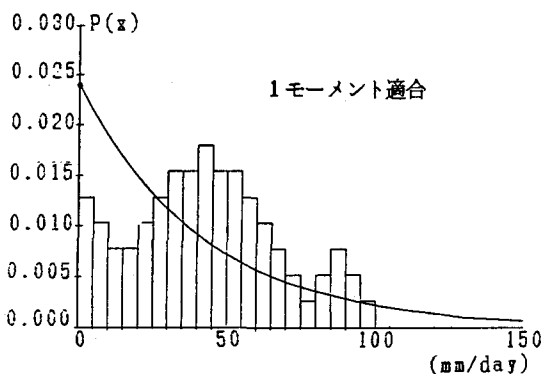


Fig.4 人工的なヒストグラムと最大エントロピー分布

- 2) Sogawa, N., Araki, M. and Imai, T.:
Studies on Multivariate Conditional
Maximum Entropy Distribution and
its Characteristics, Journal of
Hydroscience and Hydraulic
Engineering, Vol.4, No.1, April,
1986, pp.79~97.
- 3) 寒川典昭・荒木正夫: 水文事象の頻度分析へ
のME P導入について, 土木学会論文報告集,
No.335, pp.89 ~ 95, 1983年7 月.
- 4) 寒川典昭・荒木正夫・寺島 彰: 2変数ME
P分布とその特性に関する研究, 水理講演会論文集, No.28, pp.397 ~ 402, 1984年2 月.
- 5) 黒田義輝・横田勝弘: 最大エントロピー法による確率密度関数の推定法, 計測自動制御学会論文集,
Vol.18, No.1, pp.15 ~ 22, 1982年 1月.

付 録

黒田・横田⁵⁾は、原点まわりのモーメントを制約条件とした離散的1変数最大エントロピー分布について、推定された確率密度関数がエントロピーを最大にしていること、及びその分布のパラメーターの一意性を証明した。ここでは、我々が議論してきた任意関数を制約条件とする連続的多変数最大エントロピー分布についても同様なことが言えることを簡単に紹介し、パラメーター同定の基礎知識にする。

(1) 推定された確率密度関数がエントロピーを最大にすること

推定された $p(x_1, x_2, \dots, x_n)$ と $g_r(x_1, x_2, \dots, x_n)$ の期待値が等しい確率密度関数を $p'(x_1, x_2, \dots, x_n)$ とする。すなわち

$$\begin{aligned} & \int \dots \int g_r(x_1, x_2, \dots, x_n) p'(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \\ &= \int \dots \int g_r(x_1, x_2, \dots, x_n) p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \end{aligned} \quad (11)$$

一般に

$$\begin{aligned} & - \int \dots \int p'(x_1, x_2, \dots, x_n) \ln p'(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \\ & \leq - \int \dots \int p'(x_1, x_2, \dots, x_n) \ln p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \end{aligned} \quad (12)$$

(12)式の右辺は確率密度関数が具備すべき条件と(8),(11)式より次のように変形される。

$$\begin{aligned} (12) \text{式の右辺} &= - \int \dots \int p'(x_1, x_2, \dots, x_n) \ln \left\{ \exp \left\{ -1 - \sum_{r=1}^M \lambda_r g_r(x_1, x_2, \dots, x_n) \right\} \right\} dx_1 dx_2 \dots dx_n \\ &= 1 + \sum_{r=1}^M \lambda_r \int \dots \int g_r(x_1, x_2, \dots, x_n) p'(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \\ &= 1 + \sum_{r=1}^M \lambda_r \int \dots \int g_r(x_1, x_2, \dots, x_n) p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \\ &= - \int \dots \int p(x_1, x_2, \dots, x_n) \ln p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \end{aligned} \quad (13)$$

ここで、 $p'(x_1, x_2, \dots, x_n)$ は(11)式を満足する任意の確率密度関数であるから、推定された $p(x_1, x_2, \dots, x_n)$ はエントロピーを最大にすることがわかる。

(2) 解の一意性

$\lambda_r (r=1, 2, \dots, M)$ とは別のもう1組の解を $\lambda'_r (r=1, 2, \dots, M)$ とすると最大エントロピー分布は

$$p(x_1, x_2, \dots, x_n) = \exp \left\{ -1 - \sum_{r=1}^M \lambda'_r g_r(x_1, x_2, \dots, x_n) \right\} \quad (14)$$

上式のエントロピー $H'(x_1, x_2, \dots, x_n)$ は

$$\begin{aligned} H'(x_1, x_2, \dots, x_n) &= - \int \dots \int \left\{ -1 - \sum_{r=1}^M \lambda'_r g_r(x_1, x_2, \dots, x_n) \right\} \exp \left\{ -1 - \sum_{r=1}^M \lambda'_r g_r(x_1, x_2, \dots, x_n) \right\} dx_1 dx_2 \dots dx_n \\ &= 1 + \sum_{r=1}^M \lambda'_r E[g_r(x_1, x_2, \dots, x_n)] \end{aligned} \quad (15)$$

同様に $\lambda_r (r=1, 2, \dots, M)$ を解とする最大エントロピー分布のエントロピー $H(x_1, x_2, \dots, x_n)$ は

$$H(x_1, x_2, \dots, x_n) = 1 + \sum_{r=1}^M \lambda_r E[g_r(x_1, x_2, \dots, x_n)] \quad (16)$$

条件より $H(x_1, x_2, \dots, x_n)$ と $H'(x_1, x_2, \dots, x_n)$ は等しい。すなわち、

$$\sum_{r=1}^M \lambda'_r E[g_r(x_1, x_2, \dots, x_n)] = \sum_{r=1}^M \lambda_r E[g_r(x_1, x_2, \dots, x_n)] \quad (17)$$

より

$$\lambda'_r = \lambda_r, r=1, 2, \dots, M \quad (18)$$

したがって、我々が今まで数値計算により同定した解以外の解は存在しないことがわかる。