数理解析研究所講究録
223巻 1974年 28-41

CHARACTERIZATION OF NORMAL DISTRIBUTION BY MEANS OF COLLISION, ENTROPY AND TIME REVERSAL

YOSHIAKI ITOH

THE INSTITUTE OF STATISTICAL MATHEMATICS

4-6-7 MINAMI-AZABU
MINATO-KU, TOKYO
JAPAN

Abstract

Using the language in kinetic theory of gases, the well-known theorems on characterization of normal distribution are considered. It is intended to clarify what the proof physically means through the concept of collision, entropy and time reversal.

§1. INTRODUCTION

Using the language in kinetic theory of gases, we will consider well-known theorems on characterizations of the normal distribution first proved by M. Kac and S. Bernstein extended by V. P. Skitovich, G. Darmois, S. G. Ghurye and I. Olkin and the others. (For example, see [1], [2], [4], [6], [9].)

An information-theoretic proof was given to the theorem in 1-dimensional case on the author's previous paper [3]. This time, it is intended to give extension to n-dimensional case and to clarify what the proof physically means through the concept of collision, entropy and time reversal. We have varying kind of entropies, whose use depends on situations we are dealing. (See [10].) For our problem we use "Linnik's entropy". (See [7], [8].) While in 1-dimensional case, there are several proofs known, I know in n-dimensional case, only one proof in which characteristic function is used. By the method we have developed here, both cases can be treated essentially in the same way. We must impose strict conditions, for example, the existence of variance, differentiability of densities, etc., although we can relax them by smoothing by the Gaussian distribution. Throughout this paper we do not write down those conditions explicitly, since our aim is to give the physical image to the theorem and its proof.

The theorems have a long history dating back to Maxwell's investigation. Although in recent works a characteristic function is the main tool for the proof, (See [9].), the origin of the theorems is in kinetic theory of gases. In this connection, the approach adopted in this paper is quite natural and justifiable.

§2. CASE OF KAC'S CARICATURE —— OUTLINE

First we discuss the case of Kac's caricature of a Maxwellian gas. Kac's caricature of a Maxwellian gas is a model for the motion of a molecule in a chaotic bath of like molecules, satisfying the followings.

i) The molecular velocities are 1-dimensional instead of 3.

ii) A collision is a 2-dimensional rotation:

$$X_1 \longrightarrow Y_1 = X_1\cos\theta + X_2\sin\theta$$
$$X_2 \longrightarrow Y_2 = -X_1\sin\theta + X_2\cos\theta,$$

where θ is uniformly distributed on $[0,2\pi)$, preserving the energy,

$$\tfrac{1}{2}X_1^2 + \tfrac{1}{2}X_2^2 = \tfrac{1}{2}Y_1^2 + \tfrac{1}{2}Y_2^2 \quad ,$$

but not preserving momentum.

iii) Colliding pairs are uniformly distributed, and each particle suffers on the average, one collision per unit time.

There are particles A_1, A_2, A_3..., and at time t the velocities of them are represented by the random variables X_1, X_2, X_3....

Theorem. If X_i and X_j are mutually independent and if the statistical independence are preserved in spite of the occurence of a collision, X_i, X_j, Y_i and Y_j are normally distributed.

Roughly speaking, we can realize the above theorem as the following. Because of the mutual independence between X_i and X_j, entropy of each particle increases by a collision. Consider the time reversal,

$$t \longrightarrow -t$$

$$Y_i \longrightarrow -Y_i \qquad Y_j \longrightarrow -Y_j$$

$$X_i \longrightarrow -X_i \qquad X_j \longrightarrow -X_j \ ,$$

that is, the particle A_i with the velocity $-Y_i$ and the particle A_j with $-Y_j$ collide with each other and change their velocities to $-X_i$ and $-X_j$.

By this time reversal, entropy of each particle also increases. This contradicts the former. So the increment of entropy must be zero on both time direction. Hence it is necessary X_i, X_j, Y_i and Y_j are normally distributed.

§3. 3-DIMENSIONAL CASE —— OUTLINE

Boltzmann treated an assembly of particles considered as hard spheres of diameter δ. Consider the particles A_i and A_j whose velocities are $\vec{v}_i$ and $\vec{v}_j$. Center line at the time of collision is given by the unit vector $\vec{e}$. The collision is determined by $\vec{v}_i$, $\vec{v}_j$ and $\vec{e}$.

$\vec{v}_i$ and $\vec{v}_j$ change to

$$\vec{v}_i{}' = \vec{v}_i + (\vec{v}_i - \vec{v}_j, \vec{e})\vec{e}$$
$$\vec{v}_j{}' = \vec{v}_j - (\vec{v}_i - \vec{v}_j, \vec{e})\vec{e}$$

We say, "the collision is 3-dimensional," if the plane determined by $\vec{v}_i$ and $\vec{v}_j$ is different from the plane determined by $\vec{v}_i{}'$ and $\vec{v}_j{}'$.

Consider the particles A_i and A_j whose velocities are represented by the random variables $\vec{X}_i$ and $\vec{X}_j$. By a collision the velocities $\vec{X}_i$ and $\vec{X}_j$ change to $\vec{Y}_i$ and $\vec{Y}_j$. Using an orthogonal matrix T, we have the following relation.

$$\begin{pmatrix} \vec{Y}_i \\ \vec{Y}_j \end{pmatrix} = T \begin{pmatrix} \vec{X}_i \\ \vec{X}_j \end{pmatrix}.$$

Theorem. $\vec{X}_i$ and $\vec{X}_j$ are mutually independent random vectors. If the collision is 3-dimensional, and if the statistical independence are preserved in spite of the occurence of the collision, $\vec{X}_i$, $\vec{X}_j$, $\vec{Y}_i$ and $\vec{Y}_j$ have rotation invariant 3-dimensional normal densities.

In this case also, we can do the similar discussion as the case of Kac's caricature. By the collision entropy of each particle increases. Consider the time reversal. Particle A_i and A_j with their velocities $-\vec{Y}_i$ and $-\vec{Y}_j$ collide and change their velocities to $-\vec{X}_i$ and $-\vec{X}_j$, that is,

$$\begin{pmatrix} -\vec{X}_i \\ -\vec{X}_j \end{pmatrix} = T^{-1} \begin{pmatrix} -\vec{Y}_i \\ -\vec{Y}_j \end{pmatrix}.$$

By the time reversal the entropy of each particle also increases. This contradicts the case where the time is not reversed.

So the increment of entropy must be zero on both time direction. And we find the random vectors have rotation invariant 3-dimensional normal density.

§4. CASE OF KAC'S CARICATURE —— DETAIL

We introduce an "entropy" first used by Yu. V. Linnik to prove central limit theorem with Lindeberg condition, We name this, Linnik's entropy.

Consider one dimensional random variable X with continuous probability density p(x) and satisfying the conditions

$$\sup p(x)<\infty,\ E(X)=\int_{-\infty}^{\infty} xp(x)dx=0, \text{ and } D(X)=\int_{-\infty}^{\infty} x^2p(x)dx, \tag{1}$$

We put

$$I(X)=H(X)-\tfrac{1}{2}\log D(X) \tag{2}$$

following YU. V. Linnik where

$$H(X)=-\int_{-\infty}^{\infty} p(x)\log P(x)dx \tag{3}$$

I(X) has the following properties. (See [3],[5].)

Lemma 1. I(X) is invariant with respect to a homothetic transformation, i.e., for any α

$$I(\alpha X)=I(X). \tag{4}$$

Lemma 2. Let X and Y be mutually independent random variables with probability densities p(x) and q(y) and variances D(X) and D(Y) respectively. Then

$$I(X+\beta Y)-I(X)=\frac{1}{2}\beta^2D(Y)f(X)+o(\beta^2) \tag{5}$$

for sufficiently small β, where

$$f(X)=\int_{-\infty}^{\infty}\left(\frac{p'(x)}{p(x)}\right)^2 p(x)dx-\frac{1}{D(X)} \tag{6}$$

Lemma 3. Let f(X) be the value defined by (6), then we have

$$f(X) \geq 0$$

and

$$f(X) = 0$$

if and only if X is normally distributed.

We name $-\int p \log p$, Boltzmann's entropy. It was used by Boltzmann to prove H theorem. For the $p(x,t)$ which satisfies a kind of Boltzmann equation (7), we can prove H theorem, that is, $\frac{\partial}{\partial t} - \int_{-\infty}^{\infty} p(x,t) \log p(x,t)\,dx \geq 0$

(7) $$\frac{\partial}{\partial t} p(x,t) = \int_{-\infty}^{\infty}\int_{0}^{2\pi} p(x\cos\theta + y\sin\theta, t)\,p(-x\sin\theta + y\cos\theta, t) - p(x,t)\,p(y,t)\,dy\,\frac{d\theta}{2\pi}$$

Consider the diffusion equation,

(8) $$\frac{\partial}{\partial t} p(x,t) = \frac{1}{2} \frac{\partial^2}{\partial x^2} p(x,t),$$

and use Linnik's entropy,

$$-\int_{-\infty}^{\infty} p(x,t) \log p(x,t)\,dx - \tfrac{1}{2}\log t,$$

we can prove H theorem also in this case. It is interesting we can use Linnik's entropy for a problem of kinetic theory of gases in which usually Boltzmann's entropy is used.

(9) Let $Y_1(\theta) = X_1 \cos\theta + X_2 \sin\theta$

$Y_2(\theta) = -X_1 \sin\theta + X_2 \cos\theta$

and let $X_1 \perp\!\!\!\perp X_2$, (that is X_1 and X_2 are mutually independent), $Y_1(\theta_1) \perp\!\!\!\perp Y_2(\theta_1)$, $Y_1(\theta_2) \perp\!\!\!\perp Y_2(\theta_2)$, we can decompose a collision which preserves independence to two collisions which preserve independence.

Collision 1. A_1 with X_1 and A_2 with X_2 collide and change their velocities to $Y_1(\theta_1)$ and $Y_2(\theta_1)$.

Collision 2. A_1 with $Y_1(\theta_1)$ and A_2 with $Y_2(\theta_1)$ collide and change their velocities to $Y_1(\theta_2)$ and $Y_2(\theta_2)$.

Theorem 1. The random variables X_1 and X_2 are mutually independent. If θ is sufficiently small and if $Y_1(\theta)$ and $Y_2(\theta)$ are mutually independent, X_1 and X_2 are normally distributed.

Proof. The outline of proof has already been shown in §2. By the collision X_1 and X_2 change to

$$
\begin{aligned}
&Y_1(\theta)=X_1\cos\theta+X_2\sin\theta \\
&Y_2(\theta)=-X_1\sin\theta+X_2\cos\theta, \text{ respectively.}
\end{aligned} \tag{10}
$$

Using Lemma 1 and Lemma 2, we see

$$
\begin{aligned}
&I(Y_1(\theta))-I(X_1)=I(Y_1(\theta)/\cos\theta)-I(X_1)=I(X_1+X_2\tan\theta)-I(X_1) \\
&=\tfrac{1}{2}\tan^2\theta D(X_2)f(X_1)+o(\theta^2) \\
&I(Y_2(\theta))-I(X_2)=I(Y_2(\theta)/\cos\theta)-I(X_2)=I(X_2-X_1\tan\theta)-I(X_2) \\
&=\tfrac{1}{2}\tan^2\theta D(X_1)f(X_2)+o(\theta^2).
\end{aligned} \tag{11}
$$

So the entropy of each particle increases since $f(X_1)\geqq 0$ and $f(X_2)\geqq 0$ by Lemma 3.

Consider the time reversal,

A_1 with $-Y_1(\theta)$ and A_2 with $-Y_2(\theta)$ collide and change to

$$
\begin{aligned}
&-X_1=-Y_1(\theta)\cos\theta+Y_2(\theta)\sin\theta \\
&-X_2=-Y_1(\theta)\sin\ -Y_2(\theta)\cos\theta, \text{ respectively.}
\end{aligned} \tag{12}
$$

By the same argument,

$$
\begin{aligned}
&I(-X_1)-I(-Y_1(\theta))=I(X_1)-I(Y_1(\theta))= \\
&\qquad\qquad \tfrac{1}{2}\tan^2\theta D(Y_2(\theta))f(Y_1(\theta))+o(\theta^2) \\
&I(-X_2)-I(-Y_2(\theta))=I(X_2)-I(Y_2(\theta))= \\
&\qquad\qquad \tfrac{1}{2}\tan^2\theta D(Y_1(\theta))f(Y_2(\theta))+o(\theta^2)
\end{aligned} \tag{13}
$$

So the entropy increases also since

$$f(Y_1(\theta))\geqq 0 \text{ and } f(Y_2(\theta))\geqq 0 \text{ by Lemma 3.}$$

This contradicts the case where the time is not reversed, unless $f(X_1)$, $f(Y_1(\theta))$, $f(X_2)$, and $f(Y_2(\theta))$ are zero. So X_1 and X_2 are normally distributed.

Theorem 2. If $X_1 \perp\!\!\!\perp X_2$, $Y_1(\theta_1) \perp\!\!\!\perp Y_2(\theta_1)$, $Y_1(\theta_2) \perp\!\!\!\perp Y_2(\theta_2)$, and if $\theta_2 - \theta_1$ is sufficiently small, X_1 and X_2 are normally distributed.

Proof. In this case we can decompose a collision to two collisions which preserve independence.

Collision 1.

X_1 and X_2 collide and change to $Y_1(\theta_1)$ and $Y_2(\theta_1)$.

Collision 2.

$Y_1(\theta_1)$ and $Y_2(\theta_1)$ collide and change to $Y_1(\theta_2)$ and $Y_2(\theta_2)$. The relation between $(Y_1(\theta_1), Y_2(\theta_1))$ and $(Y_1(\theta_2), Y_2(\theta_2))$ is the following.

$$\begin{aligned} Y_1(\theta_2) &= Y_1(\theta_1)\cos(\theta_2-\theta_1)+Y_2(\theta_1)\sin(\theta_2-\theta_1) \\ Y_2(\theta_2) &= -Y_1(\theta_1)\sin(\theta_2-\theta_1)+Y_2(\theta_1)\cos(\theta_2-\theta_1). \end{aligned} \tag{14}$$

Applying Theorem 1 to Collision 2, we find $Y_1(\theta_1)$ and $Y_2(\theta_1)$ are normally distributed. So X_1 and X_2 are normally distributed.

§5. 3-DIMENSIONAL CASE —— DETAIL

Yu. V. Linnik extended his result of one dimensional case to a class of n dimensional random vectors $\vec{X}$ which have a probability density $p(\vec{x})$ subjected to the conditions sup $p(\vec{x})$; $E\vec{X}=\vec{o}$; $B_{\vec{X}}= E\vec{X}\,\vec{X}^T > 0$ and finite. For the class he defined following information functional.

$$I(\vec{X}) = -\int_{R^n} p(\vec{x}) \log p(\vec{x})\,dx_1 dx_2 \ldots dx_n - \frac{1}{2}\log(\det B_{\vec{X}}) \quad . \tag{15}$$

Yu. V. Linnik proved the following Lemmas.

Lemma 1. If A is n dimensional non-singular matrix, $\vec{Y} = A\vec{X}$, then

$$I(\vec{Y}) = I(\vec{X}) \quad . \tag{16}$$

Lemma 2. Let $\vec{X}$ and $\vec{Y}$ be mutually independent random vectors and β is a matrix, then

(17) $$I(\vec{X}+\beta\vec{Y}) = \frac{1}{2}\sum_{i,j=1}^{n} B_{\beta\vec{Y}\,ij}\left[\, I_{\vec{X}\,ij} - B_{\vec{X}\,ij}^{-1}\,\right] + o(\mathrm{Tr}B_{\beta\vec{Y}})$$

where $$I_{\vec{X}\,ij} = \frac{\partial^2}{\partial x_i \partial x_j} \log p(\vec{x})\; p(\vec{x})dx, \ldots dx_n \;.$$

Lemma 3

(18) $$I_{\vec{X}\,ij} - B_{\vec{X}\,ij} \geqq 0 \quad .$$

Equality holds if and only if $\vec{X}$ is normally distributed.

Theorem 3.

i) A_1 with the velocity $\vec{X}_1$ and A_2 with $\vec{X}_2$ collide and change their velocities to $\vec{Y}_1(T)$ and $\vec{Y}_2(T)$,

where

(19) $$\begin{pmatrix} \vec{Y}_1(T) \\ \vec{Y}_2(T) \end{pmatrix} = T \begin{pmatrix} \vec{X}_1 \\ \vec{X}_2 \end{pmatrix}$$

and T is a 6-dimensional orthogonal matrix.

ii) $\|T-E\|$ is sufficiently small with respect to Frobenius norm, where E is 6-dimensional unit matrix.

iii) The collision is 3-dimensional.

If $\vec{X}_1$, $\vec{X}_2$, $\vec{Y}_1(T)$ and $\vec{Y}_2(T)$ satisfy the above conditions, $\vec{X}_1$ and $\vec{X}_2$ have rotation invariant normal densities.

Proof. T is represented as the following

(20)
$$T = \begin{pmatrix} \alpha_{11}(T) & \alpha_{12}(T) \\ \alpha_{21}(T) & \alpha_{22}(T) \end{pmatrix}$$

for 3-dimensional matrices $\alpha_{kj}(T)$.

(19) is equivalent to

(21)
$$\begin{pmatrix} \vec{X}_1 \\ \vec{X}_2 \end{pmatrix} = T^{-1} \begin{pmatrix} \vec{Y}_1(T) \\ \vec{Y}_2(T) \end{pmatrix}$$

where

$$T^{-1} = \begin{pmatrix} \beta_{11}(T) & \beta_{12}(T) \\ \beta_{21}(T) & \beta_{22}(T) \end{pmatrix}$$

for 3-dimensional matrix $\beta_{kj}(T)$.

By a collision $\vec{X}_1$ changes to

(22)
$$\vec{Y}_1(T) = \alpha_{11}(T)\vec{X}_1 + \alpha_{12}(T)\vec{X}_2 \quad .$$

Applying Lemma 1 and Lemma 2 to (22), we have the following relation

(23)
$$I(\vec{Y}_1(T)) - I(\vec{X}_1) = I(\alpha_{11}^{-1}(T)\vec{Y}_1(T)) - I(\vec{X}_1)$$
$$= \frac{1}{2} \sum_{j,k=1}^{3} B_{\alpha_{11}^{-1}(T)\alpha_{12}(T)\vec{X}_2 \, jk} \left[I_{\vec{X}_1 \, jk} - B^{-1}_{\vec{X}_1 \, jk} \right)$$
$$+ 0 \left(\mathrm{Tr} B_{\alpha_{11}^{-1}(T)\alpha_{12}(T)\vec{X}_2}\right) \quad .$$

So the entropy increases by Lemma 3.

Consider the time reversal. By a collision $-\vec{Y}_1(T)$ changes to

(24)
$$-\vec{X}_1 = -\beta_{11}(T)\vec{Y}_1(T) - \beta_{12}(T)\vec{Y}_2(T) \quad .$$

By Lemma 1 and Lemma 2,

$$(25)\qquad I(-\vec{X}_1) - I(-\vec{Y}_1(T)) = I(\vec{X}_1) - I(\vec{Y}_1(T))$$

$$= \frac{1}{2}\sum_{j,k=1}^{3} B_{\beta_{11}^{-1}(T)\beta_{12}(T)\vec{Y}_2(T)\ jk}\left[I_{\vec{Y}_1(T)\ jk} - B^{-1}_{\vec{Y}_1(T)\ jk}\right]$$

$$+ o(\mathrm{Tr}B_{\beta_{11}^{-1}(T)\beta_{12}(T)\vec{Y}_2(T)})$$

So the entropy increases by Lemma 3. This contradicts the case where the time is not reversed, unless

$I_{\vec{X}_1\ jk} - B^{-1}_{\vec{X}_1\ jk}$ and

$I_{\vec{Y}_1(T)\ jk} - B^{-1}_{\vec{Y}_1(T)\ jk}$ are zero for every jk.

Taking accpunt the condition "the collision is 3-dimensional", we find $\vec{X}_1$ and $\vec{X}_2$ have rotation invariant 3-dimensional normal densities.

Let
$$\begin{pmatrix}\vec{Y}_1(T_i)\\ \vec{Y}_2(T_i)\end{pmatrix} = T_i\begin{pmatrix}\vec{X}_1\\ \vec{X}_2\end{pmatrix}$$

and let $\vec{X}_1 \perp\!\!\!\perp \vec{X}_2$, $\vec{Y}_1(T_1) \perp\!\!\!\perp \vec{Y}_2(T_1)$, $\vec{Y}_1(T_2) \perp\!\!\!\perp \vec{Y}_2(T_2)$, we can decompose a collision which preserves independence to two collisions which preserve independence.

Collision 1. A_1 with $\vec{X}_1$ and A_2 with $\vec{X}_2$ collide and change their velocities to $\vec{Y}_1(T_1)$ and $\vec{Y}_2(T_1)$.

Collision 2. A_1 with $\vec{Y}_1(T_1)$ and A_2 with $\vec{Y}_2(T_1)$ collide and change their velocities to $\vec{Y}_1(T_2)$ and $\vec{Y}_2(T_2)$. The relation between $\vec{Y}_i(T_1)$ and $\vec{Y}_i(T_2)$ is represented as the following,

$$\begin{pmatrix} \vec{Y}_1(T_2) \\ \vec{Y}_2(T_2) \end{pmatrix} = T_2 T_1^{-1} \begin{pmatrix} \vec{Y}_1(T_1) \\ \vec{Y}_2(T_1) \end{pmatrix} .$$

So we can extend theorem 2 to 3-dimensional case.

In the above discussion, "3-dimensional" is not essential we can prove the case of "n-dimensional" by the same argument.

Strictly speaking, condition ii) in Theorem 3 must be stated as the following.

Theorem 3'.

i) A_1 with the velocity $\vec{X}_1$ and A_2 collide and change their velocities to $\vec{Y}_1(T_i)$ and $\vec{Y}_2(T_i)$,

where

$$\begin{pmatrix} \vec{Y}_1(T_i) \\ \vec{Y}_2(T_i) \end{pmatrix} = T_i \begin{pmatrix} \vec{X}_1 \\ \vec{X}_2 \end{pmatrix} \tag{19}$$

and T_i is a 6-dimensional orthogonal matrix.

ii) T_i $(i=1,2,3,\ldots .)$ converges to E with respect to Frobenius norm, where E is 6-dimensional unit matrix.

iii) The collision is 3-dimensional.

If $\vec{X}_1$, $\vec{X}_2$, $\vec{Y}_1(T_i)$ and $\vec{Y}_2(T_i)$ satisfy the above conditions, $\vec{X}_1$ and $\vec{X}_2$ have rotation invariant normal densities.

Acknowledgement

I am much indepted to Dr. R. Shimizu and Mr. R. Shibata for their valuable advices.

References

[1] Feller, W. (1966). An introduction to probability theory and its applications. Vol. II., John Wiley & Sons, Inc. New York. London, Sydney.

[2] Ghurye, S. G. and Olkin, I. (1962). A characterization of the multivariate normal distribution. Ann. Math. Statist., 33 533-541.

[3] Itoh, Y. (1970). The information theoretic proof of Kac's theorem. Proc. Japan Acad. Vol. 46, 283-286.

[4] Kac, M. (1939). On a characterization of normal distribution. Amer. J. Math., 61.

[5] Kac, M. (1959). Probability and related topics in physical sciences. New York.

[6] Kagan, A. M., Linnik, Yu. V., Rao, C. R. (1973). Characterization problems in mathematical statistics. John Wiley & Sons.

[7] Linnik, Yu. V. (1959). An information theoretic proof of the central limit theorem with the Lindeberg condition. Theory of Probability and its applications, 4, 288-299.

[8] Linnik, Yu. V. (1959). On certain connections of the information theory of C. Shannon and R. Fisher with the theory of symmetrization of random vectors. Transactions of the Second Prague Conference, 313-327.

[9] Lucacs, E. and Laha, R. G. (1964), Applications of characteristic functions. Hafner Publishing Co., New York.

[10] Moran, P. (1961). Entropy, Markov processes, and Boltzmann's H-theorem. Proc. Cambridge philos. Soc. 57, 833-842.