ISSN 0973-8975

J.Mech.Cont.& Math. Sci., Vol.-8, No.-2, January (2014) Pages 1175-1184

ON A PROBLEM OF MOMENTS

By

[1]Arvinda Banerjee and [2]Mihir B. Banerjee

[1]3A/143 Azad Nagar, Kanpur-2, U.P, India.

[2]C-36/45 Sector-K Aliganj, Lucknow, U.P., India.

Abstract

The necessary and sufficient conditions for a point (μ_1, μ_2) *in the* $\mu_1\mu_2-$*plane to be constituted of the first and second moment of a probability distribution have been established in the present paper. The main results are reported in Theorem 1 and Theorem 2.*

Keywords and phrases : probability distribution, first moment of a probability distribution, second moment of a probability distribution

বিমূর্ত সার (Bengali version of the Abstract)

(μ_1, μ_2) *বিন্দুর জন্য* $\mu_1\mu_2-$ *সমতলে সম্ভাব্যতা বন্টনের প্রথম এবং দ্বিতীয় ভ্রামক গঠিত হওয়ার জন্য প্রয়োজনীয় ও যথেষ্ট শর্তাবলীকে এই পত্রে প্রতিষ্ঠিত করা হয়েছে । উপপাদ্য -1 এবং উপপাদ্য - 2 -এ মূল ফলাফলগুলির বিবরণ পেশ করা হয়েছে ।*

1. INTRODUCTION

A problem concerning moments of a probability distribution will be investigated here. As the initial findings of this investigation were being written down, the authors came across a monograph titled "The problem of moments", by Shohat and Tamarkin [1] containing a description of the classical works of Heine, Tchebycheff, Markoff, Stieltjes, Hamburger, Nevanlina, Riesz, Carleman, Hausdorff and others, on the problem of moments. These, authors have investigated the problem exhaustively and in great generality, and that perhaps is the reason for the paucity of references to their work in the modern literature. The discovery of this monograph allows the present authors no scope for claiming originality for their results, and has dampened their motivation for further investigations. The results obtained by them (proved originally be Hausdorff) however are still worth reporting because of the original and elementary nature of the methods used, their geometrical flavour and the possibilities of generalization inherent in them.

2. MATHEMATICAL ANALYSIS

A continuous non-negative real valued function ϕ defined on [0,1] and satisfying $\int_0^1 \phi(x)dx = 1$ will be called a continuous probability distribution on [0,1]. The collection of all continuous probability distributions on [0,1] will be denoted by P.

A real valued function ϕ defined on [0,1] by $\phi(x) = \sum_{j=1}^{n} a_j \delta(x - x_j)$ and satisfying $\int_0^1 \phi(x)dx = 1$ will be called a discrete probability distribution on [0,1] where a_j and x_j are real numbers satisfying $0 \le a_j$ and $0 \le x_j \le 1$, and δ is the Dirac $\delta -$function. The collection of all discrete probability distributions will be denoted by P'.

For any $\phi \in P \cup P'$ and any positive integer n, let $\mu_n(\phi) = \int_0^1 x^n \phi(x)dx$. $\mu_n(\phi)$ will be called the nth moment of ϕ, as in [1].

Let $D = \{(\mu_1, \mu_2) \in R^2 \mid \mu_1^2 < \mu_2 < \mu_1\}$ as shown in Figure (1) and $\overline{D}$ its closure.

For any $\phi \in P \cup P'$ let $\mu(\phi) = (\mu_1(\phi), \mu_2(\phi)) \in R^2$. The main results to be proved here were first proved by Hausdorff [1923]. These are:

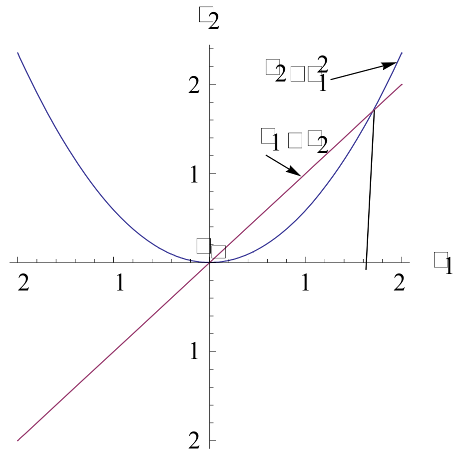

Figure 1

Theorem 1. $\mu(P') = \overline{D}$ i.e. a point $(\mu_1, \mu_2) \in R^2$ is equal to $\mu(\phi)$ for some discrete probability distribution ϕ if and only if $\mu_1^2 \le \mu_2 \le \mu_1$, and

Theorem 2. $\mu(P) = D$ i.e. a point $(\mu_1, \mu_2) \in R^2$ is equal to $\mu(\phi)$ for some continuous probability distribution ϕ if and only if $\mu_1^2 < \mu_2 < \mu_1$.

The rest of this section will be devoted to the proof of these two theorems.

Let ϕ_1 and ϕ_2 be both in P or both in P' and $0 \le x_j \le 1$. Let $\phi_t = (1-t)\phi_1 + t\phi_2$. It is clear that ϕ_t is a probability distribution of the same type as ϕ_1 and ϕ_2. Also $\mu(\phi_t) = (1-t)\mu(\phi_1) + t\mu(\phi_2)$. This gives

Lemma 1. Both $\mu(P)$ and $\mu(P')$ are convex subsets of R^2 i.e. they contain the line segment joining any two of their points.

We prove the following propositions.

Proposition 1: $\mu(P') \subset \overline{D}$ and $\mu(P) \subset D$ i.e. the conditions in Theorem 1 and Theorem 2 are necessary.

Proof: Let $\phi \in P \cup P'$ then

$$0 \le \int_0^1 [x - \mu_1(\phi)]^2 \phi(x)\, dx$$

$$= \mu_2(\phi) - \mu_1^2(\phi).$$

Further if $\phi \in P$ then the integrand is positive somewhere in [0,1] and so the inequality is strict.

On the other hand

$$0 \geq \int_0^1 x(x-1)\phi(x)\,dx$$

$$= \mu_2(\phi) - \mu_1(\phi).$$

Again, if $\phi \in P$ then the inequality is strict. The last argument is due to Banerjee and Shandil [2]. This completes the proof of Proposition 1.

Proof of Theorem 1. In view of Lemma 1, Proposition 1 and the obvious fact that D is the convex hull of (smallest convex set) containing its boundary points, it suffices to prove that every boundary point of D is in $\mu(P')$.

Let $\mu = (\mu_1, \mu_2)$ be a boundary point of D then either $\mu_2 = \mu_1^2$ or $\mu_2 = \mu_1$.

Suppose $\mu_2 = \mu_1^2$. For any $x' \in [0,1]$ let $\eta(x') = \int_0^1 \delta(x - x')\,dx$, so that $\eta(x') = 1$ if $x' \in (0,1)$ and $\eta(x') = \frac{1}{2}$ if $x' = 0$ and 1. Since μ is a boundary point of D it is clear that $0 \leq \mu_1 \leq 1$.

Let $\phi(x) = \eta^{-1}(\mu_1)\delta(x - \mu_1)$. Clearly then $\phi \in P'$ and $\mu(\phi) = \mu$.

On the other hand if $\mu_2 = \mu_1$, let $\phi(x) = 2[(1-\mu_1)\delta(x) + \mu_1\delta(x-1)]$. Clearly $\phi \in P'$ and $\mu(\phi) = \mu$. This argument is due to Kapur [3] and completes the Proof of Theorem 1.

Theorem 2 will now be proved by approximating discrete probability distributions by continuous ones, and crucial use will be made of the convexity of $\mu(P)$.

Lemma 2: $\mu(P)$ is dense in D i.e. given $\mu \in D$ there is a sequence $\phi_n \in P$ such that $\mu = \lim_{n \to \infty} \mu(\phi_n)$.

Proof. For every $x' \in [0,1]$, let $\{u_n(x,x')\}_{n=1}^{\infty}$ be a sequence of functions in P satisfying

$$\lim_{n\to\infty}\int_0^1 f(x)u_n(x,x')\,dx = \eta^{-1}(x')\int_0^1 f(x)\,\delta(x-x')\,dx = f(x')$$

for every continuous function f on [0,1]. It is well known that such a sequence exists, e.g.

$$u_n(x,x') = \left\{\int_0^1 v_n(x,x')\,dx\right\}^{-1}.v_n(x,x')$$

where

$$\begin{aligned} v_n(x,x') &= 0 \text{ if } 0 \le x \le x' - \frac{1}{n} \\ &= 2n^2\left(x - x' + \frac{1}{n}\right) \text{ if } x' - \frac{1}{n} \le x \le x' \\ &= 2n - 2n^2(x - x') \text{ if } x' \le x \le x' + \frac{1}{n} \\ &= 0 \text{ if } x' + \frac{1}{n} \le x \le 1. \end{aligned}$$

By **Theorem 1** there is a $\phi \in P'$ such that $\mu(\phi) = \mu$. Let $\phi(x) = \sum_{j=1}^{m} a_j \delta(x - x_j)$. Then

$$\int_0^1 \phi(x)\,dx = \sum_{j=1}^{m} a_j \eta(x_j) = 1.$$

Let $\phi_n(x) = \sum_{j=1}^{m} a_j \eta(x_j) u_n(x, x_j)$. Then $\int_0^1 \phi_n(x)\,dx = \sum_{j=1}^{m} a_j \eta(x_j) = 1$ because $u_n(x, x_j) \in P$. It follows that $\phi_n(x) \in P$ for every n.

Further for $k = 1$ and 2,

$$\lim_{n\to\infty}\mu(\phi_n) = \lim_{n\to\infty}\int_0^1 x^k \sum_{j=1}^m a_j \eta(x_j)\, u_n(x, x_j)\, dx$$

$$= \sum_{j=1}^m a_j \eta(x_j)\left\{\lim_{n\to\infty}\int_0^1 x^k u_n(x, x_j)\, dx\right\}$$

$$= \sum_{j=1}^m a_j x_j^k \eta(x_j)$$

$$= \mu_k(\phi).$$

$$= \mu_k.$$

It follows that $\lim_{n\to\infty}\mu(\phi_n) = \mu$ and this completes the proof of Lemma 2.

Proof of Theorem 2. In view of Proposition 1 it suffices to prove that $D \in \mu(P)$.

Let $\mu \in D$ be the point P as in Figure (2). Let $\alpha_1, \alpha_2, \alpha_3 \in D$ and represented by the points A_1, A_2 and A_3 respectively be such that the point P lies in the interior of the triangle $A_1A_2\, A_3$. Let $\in > 0$ be such that the ball of centre P and radius ε lies in the interior of triangle $A_1A_2A_3$. By Lemma 2 there are points A_1', A_2' and $A_3' \in \mu(P)$ such that each of the distances A_iA_i' is less than $\frac{\in}{4}$. It is easy to prove then that the point P lies in the interior of the triangle $A_1'A_2'A_3'$. Since $A_i' \in \mu(P)$ and $\mu(P)$ is convex by Lemma 1, it follows that $P \in \mu(P)$. This completes the proof of **Theorem 2**.

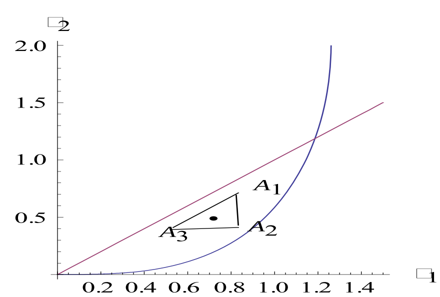

Figure 2

Some immediate consequences of these theorems are worth noticing. Let [a, b] be any finite interval. Let $\widetilde{P}$ and $\widetilde{P}'$ denote respectively the space of continuous and discrete probability distribution on [a, b]. Let $\tilde{\mu}_j$ be the jth moment function on $\tilde{P} \cup \tilde{P}'$ and let $\tilde{\mu}(\phi) = [\tilde{\mu}_1(\phi), \tilde{\mu}_2(\phi)] \in R^2$ for any $\phi \in \tilde{P} \cup \tilde{P}'$. Consider the linear homeomorphism $T : [a, b] \to [0,1]$ given by $x = T(y) = (b-a)^{-1}(y-a)$ so that its inverse is given by $y = T^{-1}(x) = a + (b-a)x$. T induces bijective mappings $P \to \tilde{P}'$ and $P' \to \tilde{P}'$, given by $\phi \leftrightarrow \tilde{\phi}$ where $\tilde{\phi}(y) = (b-a)^{-1}\phi[(b-a)^{-1}(y-a)]$ and $\phi(x) = (b-a)\,\tilde{\phi}[a + (b-a)x]$. Now

$$\mu_1(\phi) = \int_0^1 x\phi(x)\,dx$$

$$= \int_a^b (b-a)^{-1}(y-a)^{-1}\phi(y)\,dy$$

$$= (b-a)^{-1}[\tilde{\mu}_1(\tilde{\phi}) - a]$$

and

$$\mu_2(\phi) = \int_0^1 x^2\phi(x)\,dx$$

$$= (b-a)^{-2}[\tilde{\mu}_2(\tilde{\phi}) - 2a\tilde{\mu}_1(\tilde{\phi}) + a^2].$$

Therefore

$$\mu_1^2(\phi) \le \mu_2(\phi) \Leftrightarrow \tilde{\mu}_1^2(\tilde{\phi}) \le \tilde{\mu}_2(\tilde{\phi})$$

and

$$\mu_1^2(\phi) < \mu_2(\phi) \Leftrightarrow \tilde{\mu}_1^2(\tilde{\phi}) < \tilde{\mu}_2(\tilde{\phi}).$$

Similarly

$$\mu_2(\phi) \le \mu_1(\phi) \Leftrightarrow \tilde{\mu}_2(\tilde{\phi}) - 2a\tilde{\mu}_1(\tilde{\phi}) + a^2 \le (b-a)\{\tilde{\mu}_1(\tilde{\phi}) - a\}$$

$$\Leftrightarrow \tilde{\mu}_2(\tilde{\phi}) \le (b+a)\tilde{\mu}_1(\tilde{\phi}) - ab$$

and

$$\mu_2(\phi) < \mu_1(\phi) \Leftrightarrow \tilde{\mu}_2(\tilde{\phi}) < (b+a)\tilde{\mu}_1(\tilde{\phi}) - ab.$$

Stated geometrically this means that for any $\phi \in P \cup P'$,

$$\mu(\phi) \in D \Leftrightarrow \tilde{\mu}_1(\tilde{\phi}) \in \tilde{D}$$

and

$$\mu(\phi) \in \overline{D} \Leftrightarrow \tilde{\mu}_1(\tilde{\phi}) \in \overline{\tilde{D}}$$

where $\tilde{D} = \{\tilde{\mu} = (\tilde{\mu}_1, \tilde{\mu}_2) \in R^2 \mid \tilde{\mu}_1 < \tilde{\mu}_2^2 < (b-a)\,\tilde{\mu}_1 - ab\}$ is the parabolic sector determined by the two points (a, a^2) and (b, b^2) on the parabola $\tilde{\mu}_2 = \tilde{\mu}_1^2$ (Figure (3)), and $\overline{\tilde{D}}$ its closure.

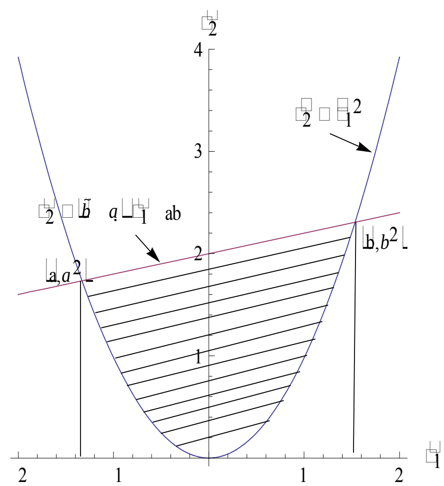

Figure 3

As immediate corollaries to the above arguments one gets

Corollary 1. The necessary and sufficient condition for a point $(\tilde{\mu}_1, \tilde{\mu}_2) \in R^2$ to be $\tilde{\mu}(\phi)$ for some ~ discrete probability distribution $\tilde{\phi}$ on [a, b] is $\tilde{\mu}_1^2 \le \tilde{\mu}_2 \le (b+a)\tilde{\mu}_1 - ab$ and

Corollary 2. The necessary and sufficient condition for a point $(\tilde{\mu}_1, \tilde{\mu}_2) \in R^2$ to be $\tilde{\mu}(\phi)$ for some ~ continuous probability distribution $\tilde{\phi}$ on [a, b] is $\tilde{\mu}_1^2 < \tilde{\mu}_2 < (b+a)\tilde{\mu}_1 - ab$ and

Corollary 3. If $a \le c < d \le b$ and $\tilde{\phi}$ a discrete (respectively continuous) probability distribution on [a, b] such that $\tilde{\phi}(x) = 0$ for every $x \notin [c, d]$ then $\tilde{\mu}(\tilde{\phi})$ is in the closed (respectively open) parabolic sector of the parabola determined by the point (c, c^2) and (d, d^2)

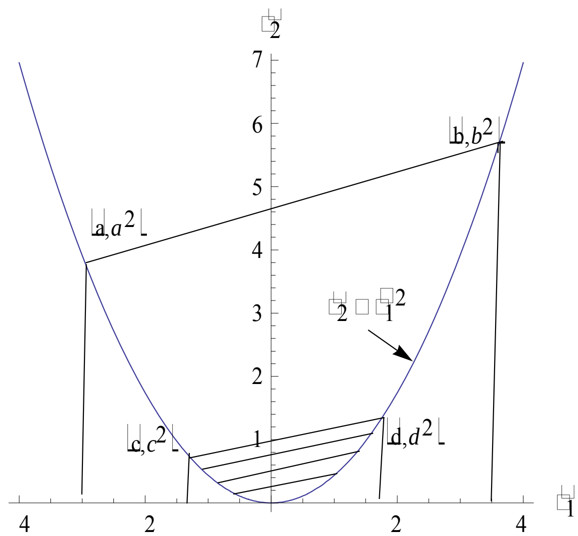

Figure 4

Proof. This follows immediately from Corollary 1 and Corollary 2 because $\tilde{\phi}$ when restricted to [c, d] defines a probability distribution in [c, d], of the same type as $\tilde{\phi}$ with the same moments.

Corollary 4. If $a \le c < d \le b$ and $\tilde{\phi}$ a probability distribution on [a, b] such that $\tilde{\phi}(x) = 0$ for every $x \in [c, d]$ then $\tilde{\mu}(\tilde{\phi})$ is not in the open parabolic sector of the parabola $\tilde{\mu}_2 = \tilde{\mu}_1^2$ determined by the point (c, c^2) and (d, d^2).

Proof: If either a=c or b=d and then the result follows from Corollary 3. Let $a < c < d < b$. Let $\phi_1(x) = \tilde{\phi}(x)$ if $x \in [a, c]$ and $\phi_1(x) = 0$ otherwise and let

$\phi_2(x) = \tilde{\phi}(x)$ if $x \in [d, b]$ and $\phi_2(x) = 0$ otherwise. By Corollary 3 we may assume ϕ_1T0 ϕ_2 T0 and. Clearly $\tilde{\phi}(x) = \phi_1(x) + \phi_2(x)$.

Let $\alpha_j = \int_a^b \phi_j(x)dx$ for j = 1 and 2. Clearly $\alpha_1, \alpha_2 > 0$ and $\alpha_1 + \alpha_2 = 1$. Let $\psi_j(x) = \alpha_j^{-1}\phi_j(x)$. It follows then that ψ_1 and ψ_2 are probability distributions on [a, b] and so is $\psi = (1 - \alpha_2)\psi_1 + \alpha_2\psi_2$. By Lemma 1 $\tilde{\mu}(\psi)$ lies on the line segment joining $\tilde{\mu}(\psi_1)$ and $\tilde{\mu}(\psi_2)$. By Corollary 3 $\tilde{\mu}(\psi_j)$ for j = 1, 2 lie in the closed unbounded sector of the parabola determined by the points (c, c^2) and (d, d^2). The corollary now follows from the obvious fact that this sector is convex.

REFERENCES

1) Shohat, J.A. and Tamarkin, J. D. The Problem of Moments, American Mathematical Society,1963.

2) Banerjee, M.B. and Shandil, R.G. A theorem of mean and standard deviation of a statistical variate, Ganita, 46,21-23.

3) Kapur, J.N. On inequality between moments of probability distribution, Ganita, 47,37-41.