Sankhyā: The Indian Journal of Statistics
1999, Volume 61, Series A, Pt. 2, pp. 286-291

THE MOMENT PROBLEM FOR THE STANDARD k-DIMENSIONAL SIMPLEX

By J.C. GUPTA
Indian Statistical Institute, Calcutta

SUMMARY. We give necessary and sufficient conditions for a multi-sequence of real constants to be the moment multi-sequence of a probability measure on the standard simplex in $\mathbb{R}^k$.

1. Historical Introduction

The problem of moments on S, a closed subset of $\mathbb{R}^k$, is as follows. Given a multi-sequence of real constants

$$\mu(\beta_1, \beta_2, \ldots, \beta_k),\ \beta_1, \beta_2, \ldots, \beta_k = 0, 1, 2, \ldots,$$

$$\mu(0, 0, \ldots, 0) = 1, \qquad \ldots (1.1)$$

one is interested in finding necessary and sufficient conditions on the multi-sequence so that there exists a probability measure P, supported on S, for which

$$\int_S x_1^{\beta_1} x_2^{\beta_2} \ldots x_k^{\beta_k} \quad dP = \mu(\beta_1, \beta_2, \ldots, \beta_k)$$

for all

$$\beta_1, \beta_2, \ldots, \beta_k = 0, 1, 2, \ldots \qquad \ldots (1.2)$$

We say that the moment problem on S is determined if any P, supported on S, is uniquely determined by its moment multi-sequence; otherwise we call it indeterminate. Clearly the moment problem on a compact set is determined. For the case $k = 1$ and $S = [0, \infty)$ the moment problem was posed and completely solved by Stieltjes (1894-95). The case $k = 1$ and $S = (-\infty, \infty)$ was studied and solved by Hamburger (1920-21). Hausdorff (1923) solved the problem for the unit interval of the real line and Haviland (1935-36) for rectangles in higher dimensions. For more details, see Shohat and Tamarkin (1943).

Paper received. June 1998.
AMS (1991) *subject classification.* 60E05, 30E05.
Key words and phrases. Moment problem, exchangeability.

2. Moment Problem on the Standard Simplex

We recall Hausdorff's solution to the moment problem on the 1-dimensional standard simplex, viz., $[0,1] \subset \mathbb{R}^1$. A sequence $\{\mu(n)\}_{n\geq 0}, \mu(0) = 1$, is called a completely monotone sequence if

$$(-1)^n \triangle^n \mu(k) \geq 0, \quad k, n = 0, 1, 2, \ldots \qquad \ldots (2.1)$$

where $\triangle\mu(k) := \mu(k+1) - \mu(k)$ and $\triangle^n$ stands for n applications of $\triangle$.

Theorem 2.1 (Hausdorff, 1923). *A sequence $\{\mu(n)\}_{n\geq 0}, \mu(0) = 1$, is the moment sequence of some probability measure on $[0,1]$ if and only if it is completely monotone.*

Hausdorff's proof exploits some properties of the Bernstein polynomials, see Shohat and Tamarkin (1943); also see Feller (1965).

We study the problem of moments on the standard k-dimensional simplex :

$$S_k = \{(x_1, x_2, \ldots, x_k) : \ x_i \geq 0 \ \forall i, \quad x_1 + x_2 + \ldots + x_k \leq 1\} \qquad \ldots (2.2)$$

We introduce the notion of a completely monotone multi-sequence.

Definition. The multi-sequence given in (1.1) is said to be *completely monotone* if

$$(-1)^{\beta_0} \triangle^{\beta_0} \mu(\beta_1, \beta_2, \ldots, \beta_k) \geq 0 \ \ \forall \beta_0, \beta_1, \beta_2, \ldots, \beta_k = 0, 1, 2 \ldots \qquad \ldots (2.3)$$

where

$$\triangle\mu(\beta_1, \beta_2, \ldots, \beta_k) := \mu(\beta_1 + 1, \beta_2, \ldots, \beta_k) + \mu(\beta_1, \beta_2 + 1, \ldots, \beta_k) + \\ \ldots + \mu(\beta_1, \beta_2, \ldots, \beta_k + 1) - \mu(\beta_1, \beta_2, \ldots, \beta_k). \qquad \ldots (2.4)$$

Theorem 2.2. *There exists a probability measure P on the standard k-simplex S_k such that*

$$\int_{S_k} x_1^{\beta_1} x_2^{\beta_2} \ldots x_k^{\beta_k} \ dP = \mu(\beta_1, \beta_2, \ldots, \beta_k) \qquad \ldots (2.5)$$

if and only if the multi-sequence $\mu(\beta_1, \beta_2, \ldots, \beta_k), \beta_1, \beta_2, \ldots, \beta_k = 0, 1, 2, \ldots$ is completely monotone.

Before giving a formal proof we would like to discuss the main idea behind it. Let $E_k = \{0, 1, \ldots, k\}$ and let Q be an exchangeable probability on $E_k^\infty := E_k \times E_k \times \ldots$. By a theorem of de Finetti (1937) Q has the following representation :

$$Q(A) \ = \ \int_{\mathcal{P}} P^\infty(A) d \ \nu(P) \qquad \ldots (2.6)$$

for all Borel subsets A of E_k^∞. Here $P^\infty := P \times P \times \ldots$ is a product probability on E_k^∞, $\mathcal{P}$ is the class of all probabilities P on E_k and ν is a probability on $\mathcal{P}$ equipped with a suitable σ-field. We observe that a probability P on $E_k = \{0, 1, 2, \ldots, k\}$ can be identified with the element $(x_1, x_2, \ldots, x_k)$ of the simplex S_k where $x_i = P(\{i\}), i = 1, 2, \ldots, k$; of course $1 - x_1 - x_2 - \ldots - x_k = P(\{0\})$. The theorem of de Finetti establishes a one-one correspondence between the class of exchangeable probabilities on E_k^∞ and the class $\mathcal{P}$ of all probabilities on S_k. We exploit this correspondence in our proof; see the concluding paragraphs of this section.

PROOF. If $\mu(\beta_1, \beta_2, \ldots, \beta_k), \beta_1, \beta_2, \ldots, \beta_k = 0, 1, 2, \ldots$ is the multi-sequence of moments of a probability P on S_k, then

$$
\begin{aligned}
& (-1)^{\beta_0} \triangle^{\beta_0} \mu(\beta_1, \beta_2, \ldots, \beta_k) \\
= & \sum_{i_0 + i_1 + \ldots + i_k = \beta_0} \binom{\beta_0}{i_0, i_1, \ldots, i_k} (-1)^{i_1 + i_2 + \ldots + i_k} \mu(\beta_1 + i_1, \beta_2 + i_2, \ldots, \beta_k + i_k) \\
= & \int_{S_k} x_1^{\beta_1} x_2^{\beta_2} \ldots x_k^{\beta_k} \ (1 - x_1 - x_2 - \ldots - x_k)^{\beta_0} \ dP \\
\geq & \ 0
\end{aligned}
$$

for all $\beta_0, \beta_1, \ldots, \beta_k$ so that the multi-sequence is completely monotone.

Conversely, let $\mu(\beta_1, \beta_2, \ldots, \beta_k), \beta_1, \beta_2, \ldots, \beta_k = 0, 1, 2, \ldots, \mu(0, 0, \ldots, 0) = 1$ be a given completely monotone sequence. We define, for $\beta_0 + \beta_1 + \ldots \beta_k = n$,

$$q_n(\beta_0, \beta_1, \ldots \beta_k) := (-1)^{\beta_0} \triangle^{\beta_0} \mu(\beta_1, \beta_2, \ldots, \beta_k) \qquad \ldots (2.7)$$

and observe that

$$q_n(\beta_0, \beta_1, \ldots, \beta_k) \geq 0 \quad \forall \quad \beta_0, \beta_1, \ldots, \beta_k = 0, 1, 2, \ldots. \qquad \ldots (2.8)$$

For $\beta_0 \geq 1$ and $\beta_0 + \beta_1 + \ldots + \beta_k = n$, we define

$$
\begin{aligned}
\nabla q_n(\beta_0, \beta_1, \ldots, \beta_k) := & \ q_n(\beta_0, \beta_1, \ldots, \beta_k) + q_n(\beta_0 - 1, \beta_1 + 1, \ldots, \beta_k) \\
& + \ldots + q_n(\beta_0 - 1, \beta_1, \ldots, \beta_k + 1). \qquad \ldots (2.9)
\end{aligned}
$$

By (2.7) and (2.9) it easily follows that

$$\nabla q_n(\beta_0, \beta_1, \ldots, \beta_k) = q_{n-1}(\beta_0 - 1, \beta_1, \ldots, \beta_k) \qquad \ldots (2.10)$$

where $n = \beta_0 + \beta_1 + \ldots + \beta_k$.

It follows from (2.10) and (2.7) that, for all $\beta_0 = 0, 1, 2, \ldots$,

$$\mu(\beta_1, \beta_2, \ldots, \beta_k) = \bigtriangledown^{\beta_0} q_n(\beta_0, \beta_1, \ldots, \beta_k) \qquad \ldots (2.11)$$

where $n = \beta_0 + \beta_1 + \ldots + \beta_k$.

In particular we have

$$\sum_{\beta_0+\beta_1+\ldots+\beta_k=n} \binom{n}{\beta_0, \beta_1, \ldots, \beta_k} q_n(\beta_0, \beta_1, \ldots, \beta_k)$$
$$= \bigtriangledown^n q_n(n, 0, \ldots, 0) = \mu(0, 0, \ldots, 0) = 1. \qquad \ldots (2.12)$$

On the n-fold product E_k^n let Q_n be the symmetric measure assigning mass $q_n(\beta_0, \beta_1, \ldots, \beta_k)$ to each point $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ of E_k^n for which $\#\{i : \omega_i = j\} = \beta_j, \;\; j = 0, 1, \ldots, k$. By (2.8) and (2.12) Q_n is a probability. In an obvious way Q_n may be considered as a probability on the n-dimensional cylinder sets, defined in terms of the first n coordinates of E_k^∞ and, by (2.9) and (2.10), $Q_n, \; n = 1, 2, \ldots$ form a consistent set of probabilities on the finite-dimensional cylinders of E_k^∞. It then follows by, say, the Kolmogorov's consistency theorem that $Q_n, n = 1, 2, \ldots$ determine a unique probability Q on the Borel σ-field of E_k^∞; this Q is an exchangeable probability.

Let $X_n, \; n = 1, 2, \ldots$, be the coordinate variables on E_k^∞ and let

$$\xi_i : \{0, 1, 2, \ldots, k\} \longmapsto \{0, 1\} \text{ be the indicator of } \{i\},\; i = 0, 1, \ldots, k$$

and

$$\begin{aligned} S_n^i : &= \xi_i(X_1) + \xi_i(X_2) + \ldots + \xi_i(X_n) \\ &= \#\{j : X_j = i; j = 1, 2, \ldots, n\}. \end{aligned} \qquad \ldots (2.13)$$

Clearly,

$$Q\{S_n^0 = \beta_0, S_n^1 = \beta_1, \ldots, S_n^k = \beta_k\} = \binom{n}{\beta_0, \beta_1, \ldots, \beta_k} q_n(\beta_0, \beta_1, \ldots, \beta_k). \qquad \ldots (2.14)$$

Now choose and fix m and $\beta_1, \beta_2, \ldots, \beta_k$ such that $\beta_1 + \beta_2 + \ldots + \beta_k = m$. For all $n \geq m$, by (2.10), we have

$$\mu(\beta_1, \beta_2, \ldots, \beta_k) = \bigtriangledown^{n-m} q_n(n-m, \beta_1, \ldots, \beta_k)$$
$$= \sum_{i_0+i_1+\ldots+i_k=n-m} \binom{n-m}{i_0, i_1, \ldots, i_k} q_n(i_0, \beta_1 + i_1, \ldots, \beta_k + i_k)$$
$$= \sum_{i_0+i_1+\ldots+i_k=n-m} \binom{n-m}{i_0, i_1, \ldots, i_k} \frac{Q\{S_n^0 = i_0, S_n^1 = \beta_1 + i_1, \ldots, S_n^k = \beta_k + i_k\}}{\binom{n}{i_0, \beta_1 + i_1, \ldots, \beta_k + i_k}}$$

by (2.14)

$$= \frac{(n-m)!}{n!} \sum_{\alpha_0+\alpha_1+\ldots,\alpha_k=n} Q\{S_n^1=\alpha_1, S_n^2=\alpha_2,\ldots,S_n^k=\alpha_k\}.\prod_{i=1}^k(\alpha_i)_{\beta_i}$$

$$= \frac{n^m.(n-m)!}{n!} E_Q\left[\prod_{i=1}^k\left\{\left(\frac{S_n^i}{n}\right)\left(\frac{S_n^i-1}{n}\right)\cdots\left(\frac{S_n^i-\beta_i+1}{n}\right)\right\}\right]$$

$$= \lim_{n\to\infty} E_Q\left[\left(\frac{S_n^1}{n}\right)^{\beta_1}\left(\frac{S_n^2}{n}\right)^{\beta_2}\cdots\cdot\left(\frac{S_n^k}{n}\right)^{\beta_k}\right] \qquad \ldots(2.15)$$

Under Q let ν_n be the law of $(\frac{S_n^1}{n},\frac{S_n^2}{n},\ldots,\frac{S_n^k}{n})$ on the simplex S_k. Then, for all $\beta_1,\beta_2,\ldots,\beta_k=0,1,\ldots$, by (2.15),

$$\int_{S_k} x_1^{\beta_1}x_2^{\beta_2}\ldots x_k^{\beta_k} \quad d\nu_n \longrightarrow \mu(\beta_1,\beta_2,\ldots,\beta_k).$$

It follows from the compactness of S_k that there exists a probability measure $\tilde{\nu}$ on S_k such that ν_n weakly converges to $\tilde{\nu}$ and

$$\int_{S_k} x_1^{\beta_1}x_2^{\beta_2}\ldots x_k^{\beta_k} \quad d\tilde{\nu} = \mu(\beta_1,\beta_2,\ldots,\beta_k) \qquad \ldots(2.16)$$

for $\beta_1,\beta_2,\ldots,\beta_k=0,1,2,\ldots$.

This completes the proof. □

To complete the story we identify the probability $\tilde{\nu}$ in (2.16) with the 'mixing' probability ν in the de Finetti representation of the exchangeable probability Q as given in (2.6).

For each $n=1,2,\ldots$, let $\mathcal{S}_n$ be the σ-field of those Borel sets of E_k^∞ which are invariant under any permutation of the first n coordinates. Further let $\mathcal{S}$ be the σ-field of symmetric Borel sets of E_k^∞. Then $\mathcal{S}_n\downarrow\mathcal{S}$ and, by the reverse martingale convergence theorem,

$$\frac{S_n^j}{n} = E_Q(\xi_j(X_1)||\mathcal{S}_n) \overrightarrow{a.s.} E_Q(\xi_j(X_1)||\mathcal{S}) = Q(X_1=j||\mathcal{S}). \qquad \ldots(2.17)$$

By de Finetti's theorem $X_1,X_2,\ldots$ are conditionally i.i.d. given $\mathcal{S}$, and consequently, for $1\le j\le k$,

$$(\frac{S_n^j}{n})^{\beta_j} \overrightarrow{a.s.} \{Q(X_1=j||\mathcal{S}).Q(X_2=j||\mathcal{S})\ldots Q(X_{\beta_j}=j||\mathcal{S})$$

$$= Q(X_1=X_2=\ldots=X_{\beta_j}=j||\mathcal{S}).$$

Thus

$$E_Q[\prod_{j=1}^{k}(\frac{S_n^j}{n})^{\beta_j}] \to \quad Q\{X_i = j, \quad \sum_{m=1}^{j-1}\beta_m \quad < i \leq \sum_{m=1}^{j}\beta_m; \quad j = 1, 2, \ldots, k\}$$

$$= \int_{\mathcal{P}} \prod_{1}^{k}(P(\{j\}))^{\beta_j} \quad d\nu(P)$$

$$= \int_{S_k} x_1^{\beta_1} x_2^{\beta_2} \ldots x_k^{\beta_k} \quad d\tilde{\nu}.$$

Acknowledgements. The author acknowledges fruitful mathematical conversations with B.V. Rao. The author thanks the Indian Statistical Institute, both at Calcutta and Delhi, for the facilities extended to him.

References

Stieltjes, T.J. (1894-95). Recherches sur les fractions continue. *Annales de la Faculté des Sciences de Toulouse,* **8** 1-122; **9** 5-47.

Hamburger, H. (1920-21). Uber eine Erweiterung der Stieltjes'schen Momentenproblems. *Math. Ann.* **81** 235-319; **82** 120-164; **82** 168-187.

Hausdorff, F. (1923). Momentprobleme für ein endliches Intervall, *Math. Zeit.,* **16** 220-248.

Haviland, E.K. (1935). On the momentum problem for distributions in more than one dimension. *Amer. Jour. Math.,* **57** 562-568.

Haviland, E.K. (1936). On the momentum problem for distribution functions in more than one dimension. II. *Amer. Jour. Math.,* **58** 164-168.

De Finetti, B. (1937). La prévision : ses lois logiques, ses sources subjectives, *Ann. Instt. Henri Poincaré*, **7** 1-68.

Shohat, J.A. and Tamarkin, J.D. (1943). *The Problem of Moments*, Math. Surveys **1**, American Mathematical Society, New York, N.Y..

Feller, W. (1965). *An Introduction to Probability Theory and Its Applications* Vol. II, John Wiley, New York, N.Y.

J.C. Gupta
32, Mirdha Tola
Budaun 243601
India