Some Optimization Problems Involving Moments of Discrete Random Variables

P.L. Sharma and R. Sharma
Department of Mathematics, H.P University, Shimla, (HP)
(Received 15 September, 2011, Accepted 20 October, 2011)

ABSTRACT : We use Lagrange multiplier method to give an alternative proof of the inequality involving moments of a discrete random variable. We also discuss an alternative proof of the inequality between arithmetic mean and variance of discrete uniform distributions.

Keywords: Moments, discrete distribution, variance, Lagrange multipliers.

I. INTRODUCTION

Let \( \{p_1, p_2, \ldots, p_n\} \) be the probability distribution with support \( \{x_1, x_2, \ldots, x_n\} \). The \( r \)th order moment \( \mu'_r \) is defined as

\[
\mu'_r = \sum_{i=1}^{n} p_i x_i^r
\]  

(1.1)

The inequalities between the moments of the discrete probability distributions have been studied extensively in literature. It is shown that the Lagrange and Kuhn Tucker methods are useful in investigating such inequalities, see [1-2]. The variance upper bounds are important in the field of theory of mathematical statistics. A number of important inequalities exist in literature, for more details see [3-10].

In the present paper, we first derive an inequality involving moments of discrete probability distributions (theorem 2.1, below). We then extend an inequality due to Muilwijk [10] and Mohr’s circle diagram in the theory of elasticity, (Lemma 2.2, below). It follows from Mohr’s circle diagram that the Muilwijk inequality is true for \( n = 3 \), we then show on using the similar analysis that the inequality must be true for \( n_i \) (Theorem 2.3, below) also see [11].

II. MAIN RESULTS

Theorem 2.1. Under the above notations:

\[
\mu'_3 \geq \mu'_1 \mu'_2
\]  

(2.1)

If \( x_i > 0, \; i = 1, 2, \ldots, n \), then

\[
\mu'_3 \geq \frac{H_2}{\mu_1}
\]  

(2.2)

Proof: We minimize the function

\[
f(x) = \sum_{i=1}^{n} p_i x_i^3
\]  

(2.3)

Subject to the constraints

\[
g_1(x) = \sum_{i=1}^{n} p_i x_i^2 - k_1
\]  

(2.4)

\[
g_2(x) = \sum_{i=2}^{n} p_i x_i^2 - k_2
\]  

(2.5)

The Lagrange function is

\[
I(x, \lambda) = \sum_{i=1}^{n} p_i x_i^3 - \lambda_1 \left( \sum_{i=1}^{n} p_i x_i^2 - k_1 \right) - \lambda_2 \left( \sum_{i=1}^{n} p_i x_i - k_2 \right)
\]  

(2.6)

The derivatives are

\[
\frac{\partial L}{\partial x_i} = (3x_i^2 - 2\lambda_1 x_i - \lambda_2) p_i
\]  

(2.7)

\[
\frac{\partial L}{\partial \lambda_1} = k_1 - \sum_{i=1}^{n} p_i x_i^2
\]  

(2.8)

and

\[
\frac{\partial L}{\partial \lambda_2} = k_2 - \sum_{i=1}^{n} p_i x_i
\]  

(2.9)

The solutions of these equations

\[
\frac{\partial L}{\partial x_i} = 0, \frac{\partial L}{\partial \lambda_2} = 0 \text{ and } \frac{\partial L}{\partial \lambda_2} = 0
\]  

(2.10)

give

\[
x_i = k_2
\]  

(2.11)

as \( \frac{\partial L}{\partial x_i} = 0 \) implies that all \( x_i \) are equal, \( i = 1, 2, \ldots, n \).

Also

\[
k_1 = k_2^2
\]  

(2.12)

For \( x_i \geq 0, \; i = 1, 2, \ldots, n \), the Hessian matrix

\[
\begin{bmatrix}
6p_1 x_1 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & 6p_n x_n
\end{bmatrix}
\]  

(2.13)

is positive definite, therefore the function is convex. So \( x_i = k_2 \) gives the minimum of \( f(x) \). Hence

\[
f(x) = \sum_{i=1}^{n} p_i x_i^3 \geq k_2^3
\]  

(2.14)

Since \( k_1 = k_2^2 \), therefore from (2.14), we have
Theorem 2.2. For real numbers \(x_1, x_2, \ldots, x_n\) we have
\[\mu_2 \geq (a + b)\mu_1' - ab\]  \hspace{1cm} (2.28)

and
\[\mu_2' \geq (x_{j-1} + x_j)\mu_1' - x_{j-1}x_j\]  \hspace{1cm} (2.29)

\(j = 2, 3, \ldots, n\).

Proof: By Lemma 2.2, the theorem is true for \(n = 3\). For \(n \geq 4\), we write
\[p_a + p_b + p_\gamma = 1 - \sum_{i=\alpha,\beta,\gamma}^n p_i\]  \hspace{1cm} (2.30)

and
\[x_\alpha^2 p_\alpha + x_\beta^2 p_\beta + x_\gamma^2 p_\gamma = \mu_2' - \sum_{i=\alpha,\beta,\gamma}^n p_i x_i^2\]  \hspace{1cm} (2.32)

The solution of the system of the linear equations (2.30), (2.31) and (2.32) can be written as
\[p_\alpha = \mu_2' - (x_\alpha + x_\beta)\mu_1' + x_\alpha x_\beta - \sum_{i=\alpha,\beta}^n p_i (x_i - x_\alpha)(x_i - x_\beta)\]  \hspace{1cm} (2.33)

\[p_\beta = \mu_2' - (x_\alpha + x_\beta)\mu_1' + x_\alpha x_\beta - \sum_{i=\alpha,\beta}^n p_i (x_i - x_\alpha)(x_i - x_\beta)\]  \hspace{1cm} (2.34)

and
\[p_\gamma = \mu_2' - (x_\alpha + x_\beta)\mu_1' + x_\alpha x_\beta - \sum_{i=\alpha,\beta}^n p_i (x_i - x_\alpha)(x_i - x_\beta)\]  \hspace{1cm} (2.35)

where \(\alpha, \beta\) and \(\gamma\) take values 1, 2, \ldots, \(n\) with \(\alpha \neq \beta \neq \gamma\).

Let \(\alpha = 1\) and \(\gamma = n\). From (2.34) we have
\[p_\beta = \mu_2' - (x_\alpha + x_n)\mu_1' + x_\alpha x_n - \sum_{i=\alpha}^n p_i (x_i - x_\alpha)(x_i - x_n)\]  \hspace{1cm} (2.36)

For \(x_1 \leq x_\beta \leq x_n\) we have \((x_\beta - x_\gamma)(x_\beta - x_n) \leq 0\). Also \(p_\beta \geq 0\), therefore it follows from (2.36) that
\[\mu_2' - (x_\alpha + x_n)\mu_1' + x_\alpha x_n \geq \sum_{i=\alpha}^n p_i (x_i - x_\alpha)(x_i - x_n)\]  \hspace{1cm} (2.37)

Since \((x_i - x_\alpha)(x_i - x_n) \leq 0\), for \(i = 1, 2, \ldots, n\) therefore the inequality (2.28) follows from (2.37).

We now consider the case when \(\alpha, \beta\) and \(\gamma\) take consecutive values. So \(x_\alpha \leq x_\beta \leq x_n\) and \((x_\gamma - x_\beta)(x_\gamma - x_n) \geq 0\). Since \(p_\gamma \geq 0\), therefore from (2.35) that
\[\mu_2' - (x_\alpha + x_n)\mu_1' + x_\alpha x_n \geq \sum_{i=\alpha}^{n-1} p_i (x_i - x_\alpha)(x_i - x_n)\]  \hspace{1cm} (2.38)

The inequality (2.29) therefore follows from (2.38), as \((x_i - x_\alpha)(x_i - x_n) \geq 0\), for \(i = 1, 2, \ldots, n - 1\).

Remark: If \(S^2\) be the variance of real numbers \(x_1, x_2, \ldots, x_n\), then \(\mu_2' = S^2 + \mu_1'^2\). The Muijik inequality, namely, \(S^2 \leq (b - \mu_1') \mu_1' - ab\), follows from the inequality (2.28).
Acknowledgements: Authors acknowledge the support of UGC-SAP.

REFERENCES