A GENERALIZATION OF TSHEBYSHEV'S INEQUALITY TO TWO DIMENSIONS

By Z. W. Birnbaum, J. Raymond, and H. S. Zuckerman

University of Washington

1. Let $X_1, X_2, \cdots, X_n$ be independent random variables with expectations $E(X_j) = e_j$ and variances $\sigma^2(X_j) = \sigma_j^2$ for $j = 1, 2, \cdots, n$. The question may be asked: What is the upper bound for the probability $P\left(\sum_{j=1}^{n} \frac{(X_j - e_j)^2}{t_j^2} \geq 1\right)$ that the point $(X_1, X_2, \cdots, X_n)$ does not fall inside of the ellipsoid

$$\sum_{j=1}^{n} \frac{(X_j - e_j)^2}{t_j^2} = 1?$$

For $n = 1$ the answer to this question is given by Tshebyshev's inequality

$$P\left[\frac{(X - E(x))^2}{t^2} \geq 1\right] \leq \frac{\sigma^2(X)}{t^2} \tag{1.1}$$

which can not be improved without further assumptions. By a trivial generalization of the argument leading to (1.1) one can prove the inequality

$$P\left(\sum_{j=1}^{n} \frac{(X_j - e_j)^2}{t_j^2} \geq 1\right) \leq \sum_{j=1}^{n} \frac{\sigma_j^2}{t_j^2} \tag{1.2}$$

for any integer n. This inequality, however, can be improved for $n \geq 2$. In particular, for $n = 2$, the following theorem will be proved:

Theorem 1.1. *Let X and Y be independent random variables, with expectations $E(X) = X_0$, $E(Y) = Y_0$ and variances σ_X^2, σ_Y^2. Then, for any $s > 0$, $t > 0$* such that $\frac{\sigma_X^2}{s^2} \leq \frac{\sigma_Y^2}{t^2}$ *we have*

$$P\left[\frac{(X - X_0)^2}{s^2} + \frac{(Y - Y_0)^2}{t^2} \geq 1\right] \leq L(s, t) \tag{1.3}$$

where

$$L(s, t) = \begin{cases} 1 \quad \text{if } \frac{\sigma_X^2}{s^2} + \frac{\sigma_Y^2}{t^2} \geq 1 \\[2ex] \frac{\sigma_X^2}{s^2} + \frac{\sigma_Y^2}{t^2} - \frac{\sigma_X^2}{s^2} \cdot \dfrac{1 - \left(\frac{\sigma_X^2}{s^2} + \frac{\sigma_Y^2}{t^2}\right)}{1 - \frac{\sigma_X^2}{s^2}} \quad \textit{if} \\[2ex] \qquad \frac{\sigma_X^2}{s^2} + \frac{\sigma_Y^2}{t^2} \leq 1 \leq \frac{1}{2}\left(\frac{\sigma_X^2}{s^2} + \frac{2\sigma_Y^2}{t^2} + \sqrt{\frac{\sigma_X^4}{s^4} + \frac{4\sigma_Y^4}{t^4}}\right) \\[2ex] \frac{\sigma_X^2}{s^2} + \frac{\sigma_Y^2}{t^2} - \frac{\sigma_X^2\sigma_Y^2}{s^2\,t^2} \quad \textit{if } \frac{1}{2}\left(\frac{\sigma_X^2}{s^2} + \frac{2\sigma_Y^2}{t^2} + \sqrt{\frac{\sigma_X^4}{s^4} + \frac{4\sigma_Y^4}{t^4}}\right) \leq 1. \end{cases} \tag{1.4}$$

For any given σ_X^2, σ_Y^2, $s > 0, t > 0$ *such that* $\frac{\sigma_X^2}{s^2} \le \frac{\sigma_Y^2}{t^2}$ *there exist independent random variables X and Y with the variances* σ_X^2, σ_Y^2, *such that the equality sign is true in (1.3).*

This theorem is a special case of the more general statement:

THEOREM 1.2. *Let W, Z be independent random variables such that*

$$P(W < 0) = P(Z < 0) = 0, \tag{1.4}$$

$$E(W) = \lambda, E(Z) = \mu, \tag{1.5}$$

$$\lambda \le \mu. \tag{1.6}$$

Then, for any $t > 0$, *we have*

$$P(W + Z \ge t) \le M(t) \tag{1.7}$$

where

$$M(t) = \begin{cases} 1 \quad \text{if} \quad t \le \lambda + \mu \\ \frac{\lambda + \mu}{t} - \frac{\lambda}{t} \cdot \frac{t - (\lambda + \mu)}{t - \lambda} = \frac{\mu}{t - \lambda} \\ \qquad \qquad \text{if} \quad \lambda + \mu \le t \le \frac{1}{2}(\lambda + 2\mu + \sqrt{\lambda^2 + 4\mu^2}) \\ \frac{\lambda + \mu}{t} - \frac{\lambda\mu}{t^2} \quad \text{if} \quad \frac{1}{2}(\lambda + 2\mu + \sqrt{\lambda^2 + 4\mu^2}) \le t. \end{cases} \tag{1.8}$$

For any given $\lambda > 0$, $\mu > 0$, $\lambda \le \mu$, *and* $t > 0$, *there exist independent variables W, Z such that (1.4) and (1.5) are fulfilled and that the equality sign is true in (1.7).*

Theorem 1.1 is obtained from Theorem 1.2 by writing

$$W = \frac{(X - X_0)^2}{s^2}, \qquad Z = \frac{(Y - Y_0)^2}{t^2}, \qquad t = 1.$$

2. Before proving Theorem 1.2 we shall derive two lemmas. The first of these lemmas deals with more than one variable. Since its proof for general m does not present any additional difficulties it will be stated and proven for any number $m \ge 1$ of variables, although in the proof of Theorem 1.2 it will be used only for $m = 1$.

LEMMA 1. *Let* $U, V_1, V_2, \cdots, V_m$ *be independent discrete random variables with only non-negative possible values, and let U have a probability distribution with the possible values* $0 \le U_1 \le U_2 \le \cdots \le U_n$ *and the probabilities* $P(U_i) = r_i$ *for* $i = 1, 2, \cdots, n$. *We consider any three possible values* U_j, U_k, U_l *of U such that*

$$0 \le U_j \le U_k \le U_l,$$

with the corresponding probabilities r_j, r_k, r_l. *Then, for any* $t > 0$, *there exists a random variable U′ with the same distribution as U except that the probabilities* r_j, r_k, r_l *of* U_j, U_k, U_l *are replaced by* r_j', r_k', r_l' *such that*

(2.1) $E(U') = E(U)$

(2.2) *one of* r'_j, r'_k, r'_l *is zero*

(2.3) $P(U' + V_1 + \cdots + V_m \geq t) \geq P(U + V_1 + \cdots + V_m \geq t)$.

PROOF: let r'_j, r'_k, r'_l be written

(2.4) $r'_j = r_j + \alpha\beta,\ r'_k = r_k - \beta,\ r'_l = r_l + (1 - \alpha)\beta.$

For any α, β we then have

$$r'_j + r'_k + r'_l = r_j + r_k + r_l .$$

Choosing

(2.5) $\alpha = (U_l - U_k)/(U_l - U_j)$

we obtain the equality

$$U_j r'_j + U_k r'_k + U_l r'_l = U_j r_j + U_k r_k + U_l r_l$$

so that (2.1) is true for any β.
We obviously have

$$P\left(U + \sum_{s=1}^{m} V_s \geq t\right) = \sum_{i=1}^{n} P(U = U_i)\cdot P\left(\sum_{s=1}^{m} V_s \geq t - U_i\right) \tag{2.6}$$
$$= \sum_{i=1}^{n} r_i P\left(\sum_{s=1}^{m} V_s \geq t - U_i\right).$$

The variable U' has the same possible values U_i as the variable U. Writing $P(U' = U_i) = r'_i$, for $i = 1, 2, \cdots, n$, we also have

$$P\left(U' + \sum_{s=1}^{m} V_s \geq t\right) = \sum_{i=1}^{n} r'_i P\left(\sum_{s=1}^{m} V_s \geq t - U_i\right). \tag{2.7}$$

From (2.6), (2.7), and (2.4) we obtain

$$P\left(U' + \sum_{s=1}^{m} V_s \geq t\right) - P\left(U + \sum_{s=1}^{m} V_s \geq t\right)$$
$$= \alpha\beta P\left(\sum_{s=1}^{m} V_s \geq t - U_j\right) - \beta P\left(\sum_{s=1}^{m} V_s \geq t - U_k\right) \tag{2.8}$$
$$+ (1 - \alpha)\beta P\left(\sum_{s=1}^{m} V_s \geq t - V_l\right).$$

For α determined by (2.5), the right-hand side of (2.8) is of the form $C\beta$, and will be positive if sign β = sign C. If sign C is positive, we choose $\beta = r_k$ and have, from (2.4), $r'_k = 0$, and, from (2.8), the inequality (2.3). If sign C is negative, we set $\beta = \text{Max}\left(-\frac{r_i}{\alpha}, -\frac{r_l}{1 - \alpha}\right)$ which leads to either $r'_j = 0$ or $r'_l = 0$, and again to (2.3). In both cases we have kept the probabilities r'_j, r'_k, r'_l non-negative as they should be.

LEMMA 2. *Let the discrete random variable U have only the two non-negative values $U_1 < U_2$, with the corresponding probabilities r_1, r_2, and let t be a given number such that*

$$E(U) < t < U_2. \tag{2.9}$$

Then there exists a number $\alpha \geq 0$ such that the random variable U' with the possible values

$$U_1' = U_1 + \alpha \tag{2.91}$$
$$U_2' = t$$

and the corresponding probabilities r_1, r_2, has the properties

$$0 \leq U_1' \leq U_2' \tag{2.92}$$

$$E(U') = E(U). \tag{2.93}$$

PROOF: to have (2.91) and (2.93) it is sufficient to choose

$$\alpha = \frac{r_2(U_2 - t)}{r_1}.$$

Then (2.92) is also fulfilled since, in view of (2.9), we have

$$U_1' = \frac{r_1 U_1 + r_2 U_2 - r_2 t}{r_1} = \frac{E(U) - r_2 t}{r_1} \leq \frac{t - r_2 t}{r_1} = t = U_2',$$

and obviously $\alpha \geq 0$ and hence $U_1' \geq U_1 \geq 0$.

3. Theorem 1 will first be proven under the assumption that W and Z are discrete random variables, each with a finite number of non-negative possible values. By repeatedly applying Lemma 1 with $m = 1$, $U = W$, $V_1 = Z$, we reduce the number of possible values of W which have non-zero probabilities to two, and denote those possible values by $W_1 \leq W_2$, and their probabilities by p_1 and $p_2 = 1 - p_1$. Then, applying Lemma 1 to the case $m = 1$, $U = Z$, $V_1 = W$, we similarly reduce the possible values of Z to the two non-negative values $Z_1 \leq Z_2$, and denote the corresponding probabilities by q_1 and $q_2 = 1 - q_1$. Throughout all these steps the expectations $E(W) = \lambda$ and $E(Z) = \mu$ remain unchanged, and $P(W + Z \geq t)$ is not decreased.

For $t \leq \lambda + \mu$, inequality (1.3) is obviously true, and equality is attained for W having the only possible value λ with probability 1 and Z having the only possible value μ with probability 1.

For the remainder of the proof we assume $t > \lambda + \mu$. We then have

$$t > \lambda + \mu \geq \lambda + Z_1 \geq W_1 + Z_1.$$

If $W_2 > t$, we may replace it by $W_2 = t$ according to Lemma 2. Similarly, if $Z_2 > t$, we may replace it by $Z_2 = t$. The probability $P(W + Z \geq t)$ is not decreased in this process. We may thus assume, without loss of generality, that

$$W_2 \leq t, \qquad Z_2 \leq t.$$

The joint distribution of (W, Z) has now the possible values represented by the four points (W_1, Z_1), (W_1, Z_2), (W_2, Z_1), (W_2, Z_2). The coordinates of these four points and their probabilities fulfill the following conditions

$$0 \leq W_1 \leq \lambda \leq W_2 \leq t; \qquad 0 \leq Z_1 \leq \mu \leq Z_2 \leq t \tag{3.1}$$

$$p_1 + p_2 = q_1 + q_2 = 1 \tag{3.2}$$

$$p_1 W_1 + p_2 W_2 = \lambda, \qquad q_1 Z_1 + q_2 Z_2 = \mu. \tag{3.3}$$

In view of (3.1), the point (W_1, Z_1) always lies below the line $W + Z = t$. The other points may or may not lie below that line. Accordingly, we distinguish the cases listed in Table I. These clearly include all possible cases since (W_2, Z_2) can not be below the line $W + Z = t$ without all the other points being below that line.

In case V we have $P(W + Z \geq t) = 0$.

For the discussion of the remaining cases we note the following relationships which follow from (3.2) and (3.3).

TABLE I

Case	Points below line $W + Z = t$	Points not below line $W + Z = t$
I	(W_1, Z_1)	(W_2, Z_1), (W_1, Z_2), (W_2, Z_2)
II	(W_1, Z_1), (W_2, Z_1)	(W_1, Z_2), (W_2, Z_2)
III	(W_1, Z_1), (W_1, Z_2)	(W_2, Z_1), (W_2, Z_2)
IV	(W_1, Z_1), (W_2, Z_1), (W_1, Z_2)	(W_2, Z_2)
V	(W_1, Z_1), (W_2, Z_1), (W_1, Z_2), (W_2, Z_2)	none

$$p_1 = \frac{W_2 - \lambda}{W_2 - W_1}, \qquad p_2 = \frac{\lambda - W_1}{W_2 - W_1},$$

$$q_1 = \frac{Z_2 - \mu}{Z_2 - Z_1}, \qquad q_2 = \frac{\mu - Z_1}{Z_2 - Z_1}.$$

In case I we have

$$W_1 + Z_1 < t, \quad W_2 + Z_1 \geq t, \quad W_1 + Z_2 \geq t, \quad W_2 + Z_2 \geq t, \tag{3.41}$$

$$P = P(W + Z \geq t) = p_2 q_1 + p_1 q_2 + p_2 q_2 = 1 - p_1 q_1$$

$$= 1 - \frac{W_2 - \lambda}{W_2 - W_1} \cdot \frac{Z_2 - \mu}{Z_2 - Z_1}.$$

Since P is a decreasing function of W_1 and Z_1, we replace W_1 and Z_1 by the smallest values compatible with (3.41), namely $W_1 = t - Z_2$, $Z_1 = t - W_2$, and obtain

$$P \leq 1 - \frac{(W_2 - \lambda)(Z_2 - \mu)}{(W_2 + Z_2 - t)^2} = R(W_2, Z_2).$$

For fixed Z_2, $R(W_2, Z_2)$ has a minimum at $W_2 = Z_2 + 2\lambda - t$ and no other extremum, hence it assumes its maximum at one or both of the end-points of the interval for W_2 which, by (3.1) and (3.41), is

$$t - Z_1 \leq W_2 \leq t.$$

In view of (3.1) we also have $t - \mu \leq t - Z_1$, and hence

$$P \leq \text{Max}\,[R(t - \mu, Z_2), \quad R(t, Z_2)].$$

We find

$$R(t - \mu, Z_2) = 1 - \frac{t - \mu - \lambda}{Z_2 - \mu} \leq 1 - \frac{t - \mu - \lambda}{t - \mu} = \frac{\lambda}{t - \mu}$$

and

$$R(t, Z_2) = 1 - \frac{(t - \lambda)(Z_2 - \mu)}{Z_2^2} = R^{(1)}(Z_2).$$

This last expression has a minimum for $Z_2 = 2$ and no other extremum, hence it assumes its maximum at the ends of the interval for Z_2 which, by (3.41) and (3.1), is

$$t - W_1 \leq Z_2 \leq t.$$

From (3.1) we also have $t - \lambda \leq t - W_1$; and thus

$$R(t, W_2) \leq \text{Max}\,[R^{(1)}(t - \lambda), \quad R^{(1)}(t)] = \text{Max}\left[\frac{\mu}{t - \lambda}, \quad \frac{\lambda + \mu}{t} - \frac{\lambda\mu}{t^2}\right].$$

Finally, we obtain

$$P \leq \text{Max}\left[\frac{\lambda}{t - \mu}, \quad \frac{\mu}{t - \lambda}, \quad \frac{\lambda + \mu}{t} - \frac{\lambda\mu}{t_2}\right].$$

Each of the values $P = \frac{\lambda}{t - \mu}, \frac{\mu}{t - \lambda}, \frac{\lambda + \mu}{t} - \frac{\lambda\mu}{t^2}$ can be attained in case I, as is shown by the probability distributions

(3.42)
$$W_1 = 0, \quad W_2 = t - \mu, \quad Z_1 = \mu, \quad Z_2 = t,$$
$$p_1 = 1 - \frac{\lambda}{t - \mu}, \quad p_2 = \frac{\lambda}{t - \mu}, \quad q_1 = 1, \quad q_2 = 0;$$

(3.43)
$$W_1 = \lambda, \quad W_2 = t, \quad Z_1 = 0, \quad Z_2 = t - \lambda,$$
$$p_1 = 1, \quad p_2 = 0, \quad q_1 = 1 - \frac{\mu}{t - \lambda}, \quad q_2 = \frac{\mu}{t - \lambda};$$

(3.44)
$$W_1 = 0, \quad W_2 = t, \quad Z_1 = 0, \quad Z_2 = t,$$
$$p_1 = 1 - \frac{\lambda}{t}, \quad p_2 = \frac{\lambda}{t}, \quad q_1 = 1 - \frac{\mu}{t}, \quad q_2 = \frac{\mu}{t}.$$

In case II we have

$$W_1 + Z_1 < t, \qquad W_2 + Z_1 < t, \qquad W_1 + Z_2 \geq t, \qquad W_2 + Z_2 \geq t,$$

(3.51)
$$P = P(W + Z \geq t) = p_1q_2 + p_2q_2 = q_2 = \frac{\mu - Z_1}{Z_2 - Z_1}.$$

This is a decreasing function of Z_1 as well as of Z_2 and hence takes its maximum for the smallest values of Z_1 and Z_2 compatible with (3.1) and (3.5), that is for $Z_1 = 0, Z_2 = t - \lambda$. We thus obtain

$$P \leq \frac{\mu}{t - \lambda}.$$

This upper bound can be attained in case II, as may be seen from the distribution

$$W_1 = \lambda, \qquad W_2 = \lambda, \qquad Z_1 = 0, \qquad Z_2 = t - \lambda,$$

(3.52)
$$p_1 = \tfrac{1}{2}, \qquad p_2 = \tfrac{1}{2}, \qquad q_1 = 1 - \frac{\mu}{t - \lambda}, \qquad q_2 = \frac{\mu}{t - \lambda}.$$

Case III is symmetrical with case II and leads to the inequality

$$P \leq \frac{\lambda}{t - \mu}.$$

In case IV we have

$$W_1 + Z_1 < t, \qquad W_2 + Z_1 < t, \qquad W_1 + Z_2 < t, \qquad W_2 + Z_2 \geq t,$$

(3.61)
$$P = P(W + Z \geq t) = p_2 q_2 = \frac{(\lambda - W_1)(\mu - Z_1)}{(W_2 - W_1)(Z_2 - Z_1)}.$$

The right hand side is a decreasing function of each of the variables W_1, W_2, Z_1, Z_2, and hence is increased by chosing for these variables the smallest values compatible with (3.61), i.e.

(3.62)
$$W_1 = Z_1 = 0, \qquad W_2 + Z_2 = t$$

for which we obtain

$$P \leq \frac{\lambda}{W_2} \frac{\mu}{t - W_2} = R^{(2)}(W_2).$$

Since $R^{(2)}(W_2)$ has a minimum at $W_2 = \dfrac{t}{2}$ and no other extremum, it attains its largest value at one of the end points of the interval for W_2 which, by (3.1), (3.61) and (3.62), is

$$\lambda \leq W_2 \leq t - \mu.$$

This leads to

$$P \leq \text{Max}\,[R^{(2)}(\lambda), \qquad R^{(2)}(t - \mu)] = \text{Max}\left[\frac{\mu}{t - \lambda}, \qquad \frac{\lambda}{t - \mu}\right].$$

The upper bounds $\frac{\mu}{t-\lambda}$, $\frac{\lambda}{t-\mu}$, respectively, are attained in case IV for the probability distribution

$$W_1 = 0, \quad W_2 = \lambda, \quad Z_1 = 0, \quad Z_2 = t - \lambda,$$

$$(3.63) \qquad p_1 = 0, \quad p_2 = 1, \quad q_1 = 1 - \frac{\mu}{t-\lambda}, \quad q_2 = \frac{\mu}{t-\lambda}$$

and

$$W_1 = 0, \quad W_2 = t - \mu, \quad Z_1 = 0, \quad Z_2 = \mu,$$

$$p_1 = 1 - \frac{\lambda}{t-\mu}, \quad p_2 = \frac{\lambda}{t-\mu}, \quad q_1 = 0, \quad q_2 = 1.$$

From the preceding discussion we conclude that $P = P(W + Z \geq t)$ always fulfills the inequality

$$P \leq \text{Max}\left[\frac{\lambda}{t-\mu}, \quad \frac{\mu}{t-\lambda}, \quad \frac{\lambda+\mu}{t} - \frac{\lambda\mu}{t^2}\right] = U(t)$$

for $t \geq \lambda + \mu$. Since we have assumed $\lambda \leq \mu$, we have $\frac{\lambda}{t-\mu} \leq \frac{\mu}{t-\lambda}$ for $t \geq \lambda + \mu$, and therefore

$$U(t) = \text{Max}\left[\frac{\mu}{t-\lambda}, \quad \frac{\lambda+\mu}{t} - \frac{\lambda\mu}{t^2}\right] \quad \text{for} \quad t \geq \lambda + \mu.$$

It is easily verified that

$$\frac{\mu}{t-\lambda} \leq \frac{\lambda+\mu}{t} - \frac{\lambda\mu}{t^2} \quad \text{for} \quad \lambda + \mu \leq t \leq \tfrac{1}{2}\,(\lambda + 2\mu + \sqrt{\lambda^2 + 4\mu^2})$$

and

$$\frac{\mu}{t-\lambda} \leq \frac{\lambda+\mu}{t} - \frac{\lambda\mu}{t^2} \quad \text{for} \quad \tfrac{1}{2}\cdot(\lambda + 2\mu + \sqrt{\lambda^2 + 4\mu^2}) \leq t$$

so that we have $U(t) = M(t)$ as defined in (1.8). For given λ, μ and any $t \geq \lambda + \mu$, the equality $P = \frac{\mu}{t-\lambda}$ is fulfilled for the distributions (3.43), (3.52) and (3.63), while the equality $P = \frac{\lambda+\mu}{t} - \frac{\lambda\mu}{t^2}$ is true for the distribution (3.44).

This completes the proof of Theorem 1.2 for discrete random variables. If W and Z are independent random variables with the cumulative probability functions $P(W \leq w) = F(w)$ and $P(Z \leq z) = G(z)$, then each of these cumulative probability functions can be uniformly approximated by a step function with a finite number of steps, that is by the cumulative probability function of a discrete random variable with a finite number of possible values. Since for such variables Theorem 1.2 is proven, it also is true for the general random variables W and Z.

4. An attempt to extend the method used in proving Theorem 1.2 to more than two variables leads to arguments of a prohibitive length. It is possible, however, to obtain corollaries of Theorems 1.1 and 1.2 which lead to an improvement of inequality (1.2) for n variables.

COROLLARY 2.1 *Let $X_1, X_2, \cdots, X_n$ be independent random variables with expectations $E(X_j) = e_j$ and variances $\sigma^2(X_j) = \sigma_j^2$. Then, for any $t_j \geq 0$, $j = 1, 2, \cdots, n$, and any m such that*

$$\Sigma_1 = \sum_{j=1}^{m} \frac{\sigma_j^2}{t_j^2} \leq \sum_{j=m+1}^{n} \frac{\sigma_j^2}{t_j^2} = \Sigma_2,$$

we have the inequality

$$P\left(\sum_{j=1}^{n} \frac{(X_j - e_j)^2}{t_j^2} \geq 1\right) \leq \begin{cases} 1 \quad \text{if} \quad 1 \leq \sum_{j=1}^{n} \frac{\sigma_j^2}{t_j^2} = \Sigma_1 + \Sigma_2 \\ \sum_{j=1}^{n} \frac{\sigma_j^2}{t_j^2} - \Sigma_1 \cdot \frac{t - (\Sigma_1 + \Sigma_2)}{t - \Sigma_1} \\ \qquad \text{if} \quad \Sigma_1 + \Sigma_2 \leq t \leq \frac{1}{2}\left[\Sigma_1 + 2\Sigma_2 + \sqrt{\Sigma_1^2 + 4\Sigma_2^2}\right] \\ \sum_{j=1}^{n} \frac{\sigma_j^2}{t_j^2} - \Sigma_1 \cdot \Sigma_2 \quad \text{if} \quad \frac{1}{2}\left[\Sigma_1 + 2\Sigma_2 + \sqrt{\Sigma_1^2 + 4\Sigma_2^2}\right] \leq t. \end{cases}$$

This corollary is a special case of the following corollary to Theorem 1.2

COROLLARY 2.2. *Let $W_1, W_2, \cdots, W_n$ be independent random variables such that $P(W_j < 0) = 0$ for $j = 1, 2, \cdots, n$, and let m be any integer such that*

$$\sum_{j=1}^{m} E(W_j) = \lambda, \qquad \sum_{j=m+1}^{n} E(W_j) = \mu, \qquad \lambda \leq \mu.$$

Then, for any $t \geq 0$, we have

$$P\left(\sum_{j=1}^{n} W_j \geq t\right) \leq M(t)$$

where $M(t)$ is defined by (1.8).

This corollary follows immediately from Theorem 1.2 by writing

$$W = \sum_{j=1}^{m} W_j, \qquad Z = \sum_{j=m+1}^{n} W_j.$$

To obtain Corollary 2.1, one only has to write in Corollary 2.2

$$W_j = \frac{(X_j - e_j)^2}{t_j^2}.$$

If some additional assumptions are made on the expectations $E(W_j)$ or on the variances σ_j^2, the upper bounds in Corollaries 2.1 and 2.2 may be minimized by proper choice of m or of the t_j. For example, if all the variances are equal

$$\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2 = \sigma^2$$

and n is even, one obtains the inequality

$$P\left[\sum_{j=1}^{n} (X_j - e_j)^2 \geq t^2\right] \leq \begin{cases} 1 & \text{if} \quad t^2 \leq n\sigma^2 \\ 1 - \dfrac{t^2 - n\sigma^2}{t^2 - \dfrac{n}{2}\sigma^2} & \text{if } n\sigma^2 \leq t^2 \leq \dfrac{3 + \sqrt{5}}{4} n\sigma^2 \\ \dfrac{n\sigma^2}{t^2}\left(1 - \dfrac{1}{4}\dfrac{n\sigma^2}{t^2}\right) & \text{if} \quad \dfrac{3 + \sqrt{5}}{4} n\sigma^2 \leq t^2. \end{cases}$$