MULTIVARIATE CHANGEPOINT PROBLEM

Sivanandan Balakumar

Abstract. Procedures for detecting a changepoint in a sequence of N random p-vectors, when there is a location or a scale change are considered. An extension of such procedures for the case of simultaneous occurrence of location and scale changes is carried out. The asymptotic distributions of the proposed statistics under the null hypothesis, in two different changepoint models are obtained.

1. Introduction. A simple multivariate changepoint problem can be formulated as follows. Let $\mathbf{X_1}, \ldots, \mathbf{X_N}$ be a sequence of N independent random vectors of dimension p, where $\mathbf{X_i} = (X_{1i}, \ldots, X_{pi})^t$ for $i = 1, \ldots, N$. Let $F(\mathbf{X_i}, \mathcal{B}_\mathbf{i})$ be the continuous distribution function (cdf) of the random vector $\mathbf{X_i}$ and where $\mathcal{B}_\mathbf{i} = (\mathcal{B}_{1i}, \ldots, \mathcal{B}_{pi})^t$ for $i = 1, \ldots, N$ are parameters. The above sequence of random vectors is said to have a changepoint at time point n $(1 \leq n < N)$, if the random vectors $\mathbf{X_i}$ for $i = 1, \ldots, n$ have the cdf's $F(\mathbf{X_i}, \mathbf{0})$ and the random vectors $\mathbf{X_i}$ for $i = n+1, \ldots, N$ have the cdf's $F(\mathbf{X_i}, \mathcal{B})$ with $\mathbf{0} = (0, \ldots, 0)^t$ and $\mathcal{B} = (\mathcal{B}_1, \ldots, \mathcal{B}_p)^t$. The time point n may be called the single changepoint. The change may also occur smoothly over a period of time and the time point n at which the change begins to occur may be called the continuous changepoint.

Studies about changepoint problems in a multivariate setting are very rarely found in the literature. Sen and Srivastava [7] considered the problem of testing the hypothesis that the means of a sequence of N independent multivariate normal random variables are equal against the alternative that after an unknown time point r $(1 \leq r \leq N-1)$, the means have shifted.

In this paper we formulate two changepoint models, namely the single changepoint model (SC model) and the continuous changepoint model (CC model). Furthermore, we propose the appropriate test statistics for testing location, scale, and simultaneous location and scale changes in the above mentioned changepoint models. Section 2 contains the formulation of these models and the derivation of the appropriate test statistics. In Section 3, the asymptotic distributions of the proposed statistics are given and the concluding remarks are given in Section 4.

2. Multivariate Changepoint Models. In order to formulate the changepoint models, let $F[(1 + a_{1i})x_{1i} + b_{1i}, \ldots, (1 + a_{pi})x_{pi} + b_{pi}]$ for $i = 1, \ldots, N$ be the continuous cdf of the random vector $\mathbf{X_i}$, where a_{ji} and b_{ji} for $j = 1, \ldots, p$ and $i = 1, \ldots, N$ are the scale and location parameters, respectively.

Let $\mathbf{a_i} = (a_{1i}, \ldots, a_{pi})^t$, $\mathbf{b_i} = (b_{1i}, \ldots, b_{pi})^t$, $\mathbf{1} = (1, \ldots, 1)^t$, $\mathbf{0} = (0, \ldots, 0)^t$, $\mathbf{a} = (a_1, \ldots, a_p)^t$, and $\mathbf{b} = (b_1, \ldots, b_p)^t$. Now the SC model may be defined as follows.

$$\mathbf{a_i} = \begin{cases} \mathbf{0}, & 1 \le i \le n \\ \mathbf{a}, & n+1 \le i \le N, \end{cases} \qquad \mathbf{b_i} = \begin{cases} \mathbf{0}, & 1 \le i \le n \\ \mathbf{b}, & n+1 \le i \le N \end{cases} \tag{2.1}$$

with the hypotheses of interest $H_0 : \mathbf{a} = \mathbf{0}$ and $\mathbf{b} = \mathbf{0}$ and $H_a : \|\mathbf{a}\| > 0$ and $\|\mathbf{b}\| > 0$, where $\|\cdot\|$ denotes the Euclidean norm. The CC model takes the following form.

$$\mathbf{a_i} = \begin{cases} \mathbf{0}, & 1 \le i \le n \\ \mathbf{a}(i-n)/(N-n), & n+1 \le i \le N, \end{cases}$$

$$\mathbf{b_i} = \begin{cases} \mathbf{0}, & 1 \le i \le n \\ \mathbf{b}(i-n)/(N-n), & n+1 \le i \le N \end{cases} \tag{2.2}$$

with the hypotheses of interest $H_0 : \mathbf{a} = \mathbf{0}$ and $\mathbf{b} = \mathbf{0}$ and $H_a : \|\mathbf{a}\| > 0$ and $\|\mathbf{b}\| > 0$.

Now we shall formulate the test statistics for both the changepoint models. Let R_N denote the $p \times N$ rank matrix corresponding to the observation vector $\mathbf{X} = (\mathbf{X_1}, \ldots, \mathbf{X_N})$ and let $\mathbf{R_i} = (R_{1i}, \ldots, R_{pi})^t$ be the rank vector of the ith observation vector $\mathbf{X_i}$. Thus,

$$R_N = \begin{pmatrix} R_{11} & \ldots & R_{1N} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ R_{p1} & \ldots & R_{pN} \end{pmatrix}_{p \times N} . \tag{2.3}$$

Let $a_{Nk}(R_{ji})$, $k = 1, 2$ denote the score functions corresponding to the rank R_{ji} for $j = 1, \ldots, p$ and $i = 1, \ldots, N$. Then the $p \times N$ matrix a_{Nk} of score functions is given by

$$a_{Nk} = \begin{pmatrix} a_{Nk}(R_{11}) & \ldots & a_{Nk}(R_{1N}) \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ a_{Nk}(R_{p1}) & \ldots & a_{Nk}(R_{pN}) \end{pmatrix}_{p \times N} , \tag{2.4}$$

where $k = 1, 2$. The matrix R_N in (2.3) has $(N!)^p$ realizations. The distribution of the rank vector $\mathbf{R_N} = (\mathbf{R_1}, \ldots, \mathbf{R_N})$ over the $(N!)^p$ realizations depend on the distribution of $\mathbf{X_i}$, $i = 1, \ldots, N$ even under the null hypothesis, since the p-variates in $\mathbf{X_i}$ are in general dependent. Therefore, we do not obtain a uniform distribution for the rank vector $\mathbf{R_N}$ in the multivariate case. To overcome this problem we consider the matrix $\mathbf{R*_N}$, which has the same columns as R_N but arranged so that the first row consists of the integers $1, \ldots, N$, in that order. The matrix $\mathbf{R*_N}$ has $(N!)^{p-1}$ realizations. The conditional distribution of $\mathbf{R_N}$ over the set of $N!$ possible realizations of the columns of the matrix $\mathbf{R*_N}$ is uniform under the null hypothesis. Furthermore, the matrix $\mathbf{R_N}$ is permutationally equivalent to the matrix $\mathbf{R*_N}$. Let $S(R*_N)$ be the set of matrices that are permutationally equivalent to $R*_N$. Then

$$P[\mathbf{R_N} = \mathbf{r_N} \mid S(R*_N), H_0] = 1/N!,$$

for all $\mathbf{r_N} \in S(R*_N)$. Thus, any statistic that depends explicitly on the elements of R_N gives a conditionally distribution free test. These results have been used by Duran and Mitchell [5] in formulating multisample multivariate nonparametric tests for simultaneous location/scale alternatives. For a complete treatment of conditionally distribution free tests, we refer the reader to Chatterjee and Sen [2, 3].

To derive a statistic for the SC model, define the vectors $\mathbf{S_1} = (S_{11}, \ldots, S_{1p})^t$, $\mathbf{S_2} = (S_{21}, \ldots, S_{2p})^t$, and $\mathbf{S_N} = (\mathbf{S_1}^t, \mathbf{S_2}^t)$, where

$$S_{kj} = \sum_{i=1}^{N} c_i a_{kj}(R_{ji}), \tag{2.5}$$

for $k = 1, 2$ and $j = 1, \ldots, p$ with $c_i = i - 1$. Note that the statistics S_{1j} and S_{2j} are used for location and scale testing, respectively in a univariate single changepoint problem by Balakumar [1]. In the multivariate situation, the statistic $\mathbf{S_1}$ may be used for location testing in the SC model given in (2.1) and likewise, $\mathbf{S_2}$ may be used for scale testing in the same model. The null hypothesis H_0 is rejected for large values of the statistics in either case.

For simultaneous location-scale testing in the SC model, we propose the statistic L_N given by

$$L_N = (\mathbf{S_N} - \mu_\mathbf{N})^t A_{N-1} (\mathbf{S_N} - \mu_\mathbf{N}),$$

where $\mu_\mathbf{N} = (\mu_1^t, \mu_2^t)$, $\mu_\mathbf{1} = (\mu_{11}, \ldots, \mu_{1p})^t$, $\mu_\mathbf{2} = (\mu_{21}, \ldots, \mu_{2p})^t$, $\mu_{kj} = E[S_{kj} \mid S(\mathbf{R*_N}), H_0] = [N(N-1)/2] \int_0^1 \phi_{kj}(u)\, du$, for $k = 1, 2$; $j = 1, \ldots, p$ in

which the score functions $a_{kj}(\cdot)$ for $k = 1, 2$ are given by the square integrable functions ϕ_k. The dispersion matrix A_N of dimension $2p \times 2p$ is given by

$$A_N = \begin{pmatrix} U & W \\ W & V \end{pmatrix},$$

with $U = (u_{rs})_{p\times p}$, $V = (v_{rs})_{p\times p}$, $W = (w_{rs})_{p\times p}$ for $r, s = 1, \ldots, p$ are $p \times p$ matrices with entries given by

$$u_{rs} = \mathrm{Cov}[S_{1r}, S_{1s} \mid S(\mathbf{R}*_{\mathbf{N}}), H_0] = t \int_0^1 [\phi_{1r}(u) - \overline{\phi}_{1r}][\phi_{1s}(u) - \overline{\phi}_{1s}]\, du,$$

$$w_{rs} = \mathrm{Cov}[S_{1r}, S_{2s} \mid S(\mathbf{R}*_{\mathbf{N}}), H_0] = t \int_0^1 [\phi_{1r}(u) - \overline{\phi}_{1r}][\phi_{2s}(u) - \overline{\phi}_{2s}]\, du,$$

and

$$v_{rs} = \mathrm{Cov}[S_{2r}, S_{2s} \mid S(\mathbf{R}*_{\mathbf{N}}), H_0] = t \int_0^1 [\phi_{2r}(u) - \overline{\phi}_{2r}][\phi_{2s}(u) - \overline{\phi}_{2s}]\, du,$$

where $t = N(N + 1)/12$. Now, by the choice of our score functions to give odd and even translation invariant statistics, we get $w_{rs} = 0$ for all r and s. Thus, the dispersion matrix A_N reduces to

$$A_N = \begin{pmatrix} U & 0 \\ 0 & W \end{pmatrix}.$$

Hence, the statistic L_N can be written as

$$L_N = (\mathbf{S_1} - \mu_\mathbf{1})^t U^{-1} (\mathbf{S_1} - \mu_\mathbf{1}) + (\mathbf{S_2} - \mu_\mathbf{2})^t V^{-1} (\mathbf{S_2} - \mu_\mathbf{2}). \tag{2.7}$$

A similar statistic to that of L_N for the CC model may be obtained by defining the vectors $\mathbf{T_1} = (T_{11}, \ldots, T_{1p})^t$, $\mathbf{T_2} = (T_{21}, \ldots, T_{2p})^t$ and $\mathbf{T_N} = (\mathbf{T_1}^t, \mathbf{T_2}^t)$, where

$$T_{kj} = \sum_{i=1}^{N} d_i a_{kj}(R_{ji}), \tag{2.8}$$

for $k = 1, 2$ and $j = 1, \ldots, p$ with $d_i = i(i-1)/2$. The statistics T_{1j} and T_{2j} are used for location and scale testing respectively, in a univariate continuous changepoint problem by Balakumar [1]. For location testing in a multivariate continuous changepoint problem the statistic $\mathbf{T_1}$ may be used and for scale testing the statistic $\mathbf{T_2}$ may be used. The simultaneous location-scale testing in the CC model given in (2.2) may be carried out by the statistic L_{N*} given by

$$\mathbf{L_{N*}} = (\mathbf{T_N} - \mathbf{v_N})^t B_N^{-1} (\mathbf{T_N} - \mathbf{v_N}),$$

where $\mathbf{v_N} = (\mathbf{v_1}^t, \mathbf{v_2}^t)$, $\mathbf{v_1} = (v_{11}, \ldots, v_{1p})^t$, $\mathbf{v_2} = (v_{21}, \ldots, v_{2p})^t$, $v_{kj} = E[T_{kj} \mid S(R_{N*}), H_0] = [N(N^2-1)/6] \int_0^1 \phi_{kj}\, du$ for $k = 1, 2$; $j = 1, \ldots, p$ and the dispersion matrix B_N of dimension $2p \times 2p$ is given by

$$B_N = \begin{pmatrix} Y & G \\ G & Z \end{pmatrix}, \tag{2.9}$$

with $Y = (y_{rs})_{p\times p}$, $G = (g_{rs})_{p\times p}$, $Z = (z_{rs})_{p\times p}$ for $r, s = 1, \ldots, p$ are $p \times p$ matrices with entries given by

$$y_{rs} = \mathrm{Cov}[T_{1r}, T_{1s} \mid S(R_{N*}), H_0] = m \int_0^1 [\phi_{1r}(u) - \overline{\phi}_{1r}][\phi_{1s}(u) - \overline{\phi}_{1s}]\, du,$$

$$g_{rs} = \mathrm{Cov}[T_{1r}, T_{2s} \mid S(R_{N*}), H_0] = m \int_0^1 [\phi_{1r}(u) - \overline{\phi}_{1r}][\phi_{2s}(u) - \overline{\phi}_{2s}]\, du,$$

and

$$z_{rs} = \mathrm{Cov}[T_{2r}, T_{2s} \mid S(R_{N*}), H_0] = m \int_0^1 [\phi_{2r}(u) - \overline{\phi}_{2r}][\phi_{2s}(u) - \overline{\phi}_{2s}]\, du,$$

where $m = N(N+1)(4N^2-1)/180$. As before, by the application of odd and even translation invariant score functions to the test statistics, the covariance component g_{rs} becomes zero and thus, the dispersion matrix B_N reduces to

$$B_N = \begin{pmatrix} Y & 0 \\ 0 & Z \end{pmatrix}.$$

Hence, $\mathbf{L_{N*}}$ becomes

$$\mathbf{L_{N*}} = (\mathbf{T_1} - \mathbf{v_1})^t Y^{-1} (\mathbf{T_1} - \mathbf{v_1}) + (\mathbf{T_2} - \mathbf{v_2})^t Z^{-1} (\mathbf{T_2} - \mathbf{v_2}) \tag{2.10}$$

and H_0 is rejected for large values of L_{N*}.

3. Asymptotic Distributions of the Statistics. In this section we shall obtain the asymptotic distribution of the statistics $\mathbf{S_1}, \mathbf{S_2}, \mathbf{T_1}, \mathbf{T_2}, \mathbf{L_N}$, and $\mathbf{L_{N*}}$ under the appropriate null hypothesis. To begin with, the statistic $\mathbf{S_1} = (S_{11}, \ldots, S_{1p})^t$, where S_{1j} for $j = 1, \ldots, p$ is as given by (2.5) is suitable for only location testing in the SC model given by (2.1). The hypotheses of interest are ${}^bH_0 : \mathbf{b} = \mathbf{0}$ and ${}^bH_a : \|\mathbf{b}\| > 0$. The following theorem gives the asymptotic distribution of the statistic $\mathbf{S_1}$.

Theorem 3.1. When ${}^bH_0$ holds, if the regression constants c_i's and the score functions a_{1j}'s for $i = 1, \ldots, N$; $j = 1, \ldots, p$ satisfy respectively, the following Noether's conditions

$$\lim_{N\to\infty} \frac{\max_{1\le i\le N}(c_i - \overline{c})^2}{\sum_{i=1}^N (c_i - \overline{c})^2} = 0 \quad \text{and} \quad \lim_{N\to\infty} \frac{\max_{1\le i\le N}(a_{1j} - \overline{a}_{1j})^2}{\sum_{i=1}^N (a_{1j} - \overline{a}_{1j})^2} = 0, \tag{3.1}$$

for each j, then the statistic $\mathbf{S_1}$ has an asymptotic p-variate normal distribution with mean vector $\mu_{\mathbf{1}} = (\mu_{11}, \ldots, \mu_{1p})^t$ and dispersion matrix U as given by (2.6).

Proof. As

$$\lim_{N\to\infty} \left\{ \frac{\max_{1\le i\le N}(c_i - \overline{c})^2}{\sum_{i=1}^N (c_i - \overline{c})^2} \right\} = \lim_{N\to\infty} \frac{[(N-1)/2]^2}{[N(N^2-1)/12]^2} = 0$$

for each j, the statistic S_{1j} has an asymptotic normal distribution with mean

$$\mu_{1j} = [N(N-1)/2] \int_0^1 \phi_{ij}(u)\ du$$

and variance

$$\sigma_{1j}^2 = [N(N+1)/12] \int_0^1 [\phi_{1j}(u) - \overline{\phi}_{1j}]^2\ du \tag{3.2}$$

according to Theorem V.I.5a of Hajek and Sidak [6]. Thus, the mean of the vector $\mathbf{S_1} = (S_{11}, \ldots, S_{1p})$ is given by $\boldsymbol{\mu_1} = (\mu_{11}, \ldots, \mu_{1p})$, where μ_{1j} for $j = 1, \ldots, p$ is given in (3.2). Furthermore, since

$$\text{Cov}\ [S_{1r}, S_{1s}] = [N(N+1)/12] \int_0^1 [\phi_{1r}(u) - \overline{\phi}_{1r}][\phi_{1s}(u) - \overline{\phi}_{1s}]\ du$$

the dispersion matrix of $\mathbf{S_1}$ is as given by (2.6). To show the multinormality of $\mathbf{S_1}$, it suffices to prove the normality for any linear combination of S_{1j}'s for $j = 1, \ldots, p$ according to a theorem by Cramer [4]. For any h_j for $j = 1, \ldots, p$ consider the linear combination

$$\sum_{j=1}^{p} h_j S_{1j} = \sum_{j=1}^{p} h_j \left\{ \sum_{i=1}^{N} c_i a_{1j}(R_{ji}) \right\} = \sum_{i=1}^{N} c_i \left\{ \sum_{j=1}^{p} h_j a_{1j}(R_{ji}) \right\}$$

and let

$$b_i = \sum_{j=1}^{p} h_j a_{1j}(R_{ji}).$$

Since

$$\frac{\max_{1\le i\le N}(b_i-\overline{b})^2}{\sum_{i=1}^{N}(b_i-\overline{b})^2}=\frac{\max_{1\le i\le N}\left\{\sum_{j=1}^{p}h_j[a_{1j}(R_{ji})-\overline{a}_{1j}]\right\}^2}{\sum_{i=1}^{N}\left\{\sum_{j=1}^{p}h_j[a_{1j}(R_{ji})-\overline{a}_{1j}]\right\}^2}$$

$$\le\frac{p\max_{1\le i\le N}h_{12}[a_{11}(R_{1i})-\overline{a}_{11}]^2}{\sum_{i=1}^{N}h_{12}[a_{11}(R_{1i})-\overline{a}_{11}]^2}+\cdots+\frac{p\max_{1\le i\le N}h_{p^2}[a_{1p}(R_{pi})-\overline{a}_{1p}]^2}{\sum_{i=1}^{N}h_{p^2}[a_{1p}(R_{pi})-\overline{a}_{1p}]^2}$$

by condition (3.1) of the theorem, each term on the right hand side of the last inequality above will approach zero as N approaches ∞. Hence, the theorem is proved.

For scale testing in the SC model the hypotheses of interest are ${}^aH_0 : \mathbf{a} = \mathbf{0}$ and ${}^aH_a : \|\mathbf{a}\| > 0$ and the required test statistic is $\mathbf{S_2}$. The asymptotic distribution of $\mathbf{S_2}$ under the null hypothesis is given in the following theorem.

Theorem 3.2. When ${}^aH_0$ holds and if the regression constants c_i's for $i = 1,\ldots,N$ and the score functions a_{2j}'s for $j = 1,\ldots,p$ satisfy the Noether's conditions, then the statistic $\mathbf{S_2}$ has an asymptotic p-variate normal distribution with mean vector $\mu_{\mathbf{2}} = (\mu_{21},\ldots,\mu_{2p})^t$ and dispersion matrix V as given by (2.6).

Proof. The proof is similar to that of Theorem 3.1.

In a similar manner we are able to determine the asymptotic distributions of the test statistics $\mathbf{T_k}$, $k = 1, 2$ that is used for detecting location and scale respectively, in the CC model. For location testing in the CC model the required hypotheses are ${}^{b^*}H_0 : \mathbf{b} = \mathbf{0}$ and ${}^{b^*}H_a : \|\mathbf{b}\| > 0$ and for scale testing the required hypotheses are ${}^{a^*}H_0 : \mathbf{a} = \mathbf{0}$ and ${}^{a^*}H_0 : \|\mathbf{a}\| > 0$. The following theorems give the asymptotic distribution of the statistics $\mathbf{T_k}$, $k = 1, 2$. Since the proofs are similar to that of Theorem 3.1 we omit the proofs here and state the theorems.

Theorem 3.3. When ${}^{b^*}H_0$ holds, if the regression constants d_i's for $i = 1,\ldots,N$ as given in (2.8) and the score functions a_{1j}'s for $j = 1,\ldots,p$ satisfy the Noether's condition then the statistic $\mathbf{T_1}$ has an asymptotic p-variate normal distribution with mean vector $\mathbf{v_1} = (v_{11},\ldots,v_{1p})^t$ and dispersion matrix Y as given by (2.9).

Theorem 3.4. When ${}^{a^*}H_0$ holds, if the regression constants d_i's for $i = 1,\ldots,N$ as given in (2.8) and the score functions a_{2j}'s for $j = 1,\ldots,p$ satisfy

the Noether's condition then the statistic $\mathbf{T_2}$ has an asymptotic p-variate normal distribution with mean vector $\mathbf{v_2} = (v_{21}, \ldots, v_{2p})^t$ and dispersion matrix Z as given by (2.9).

By virtue of the preceding theorems the asymptotic distributions of the statistics $\mathbf{L_N}$ and $\mathbf{L_{N^*}}$ can be deduced. The results are given in the following theorems. The proofs of the theorems are very straightforward and we omit them here.

Theorem 3.5. Under the conditions of Theorem 3.1, the statistic $\mathbf{L_N}$, given by (2.7) has a distribution asymptotically converging in probability to a chi-square distribution with $2p$ degrees of freedom.

Theorem 3.6. Under the conditions of Theorem 3.2, the statistic $\mathbf{L_{N^*}}$, given by (2.10) has a distribution asymptotically converging in probability to a chi-square distribution with $2p$ degrees of freedom.

4. Concluding Remarks. In this paper we have presented some techniques for testing simultaneous location-scale changes in a multivariate changepoint problem. The underlying probability distribution of the random vectors $\mathbf{X_1}, \ldots, \mathbf{X_N}$ can be any multivariate distribution. The efficiency of these methods and the asymptotic distribution of the test statistics under a suitable alternative hypothesis will be investigated in the future.

References

1. S. Balakumar, "Changepoint Detection Using Nonparametric Procedures," *Missouri Journal of Mathematical Sciences,* 9 (1997), 178–183.
2. S. K. Chatterjee and P. K. Sen, "Nonparametric Tests for the Bivariate Two-Sample Location Problem," *Cal. Stat. Assoc. Bulletin,* 13 (1964), 18–58.
3. S. K. Chatterjee and P. K. Sen, "Some Nonparametric Tests for the Two-Sample Association Problem," *Cal. Stat. Assoc. Bulletin,* 14 (1965), 14–35.
4. H. Cramer, *Mathematical Methods of Statistics,* Princeton University Press, Princeton, New Jersey, 1946.
5. B. S. Duran and C. C. Mitchell, "A Class of Multisample Multivariate Nonparametric Tests for Location-Scale Changes," *Communications in Statistics. Theory & Method,* 18 (1988), 67–105.

6. T. Hajek and Z. Sidak, *Theory of Rank Tests,* Academic Press, New York, 1967.

7. A. K. Sen and M. S. Srivastava, "On Multivariate Tests for Detecting Change in Mean," *Sankhya,* 35 (1973), 173–186.

Sivanandan Balakumar
Department of Natural Sciences and Mathematics
Lincoln University
Jefferson City, MO 65102-0029
email: Balakuma@lincolnu.edu