Alternative Hypothesis Tests for the Covariance Matrix Based on Eigenvalues and Multivariate Normality

Sueli Aparecida Mingoti
Letícia Pereira Pinto
Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG

Abstract

Three statistical tests for testing the covariance matrix in one population under multivariate normal assumption were developed and compared with the classical generalized variance test by means of Monte Carlo simulation. It is shown that the tests, which are based on the eigenvalues of the sample and the theoretical covariances matrices, presented better performance than the generalized variance particularly in cases which the parameters (determinant and traces) of the covariances matrices under the null and alternative hypothesis were similar. Since the three statistical tests are of simple implementation they can be considered as alternatives to be used in a general context.

Keywords: *Generalized variance, Covariance matrix, Eigenvalues, Monte Carlo, Hotelling, Hayter and Tsui.*

Introduction

Hypothesis tests for the vector mean and the covariance matrix are very common when sampling from the multivariate normal distribution. Most of the papers in the literature focus on the improvement of tests for the vector mean however, finding better tests for the covariance matrix is also important. Practical examples appear in several areas particularly in quality control when the quality of the process is evaluated by measuring p characteristics simultaneously ($p > 1$). In this case location and variability parameters of the process have to be kept in certain pre-determined values in order to satisfy the established quality requisites (Montgomery, 2001).

Some statistical tests have been proposed in the literature to test the covariance matrix (see Anderson, 1958; Costa and Machado, 2008; Yeh *et al.*, 2006, for examples), however, the generalized variance, which is based on the determinant of the sample covariance matrix, $|S|$, is still very popular and it has been used also in the quality control field (see Alt, 1985; Garcia-Diaz, 2007). It is known that the exact distribution of $|S|$ under the null hypothesis is related to the distribution of the product of p independent random variables with chi-square distributions (Anderson, 1958; Aparisi *et al.*, 1999). For simplicity, the normal distribution for $|S|$ under the null hypothesis, is usually used as an approximation to construct the critical region of the test in practical situations. However, as shown by Djauhari (2009) the normal approximation for $|S|$ just holds for large samples sizes. Another problem comes from the fact that

different matrices can present equal or similar determinants which contributes to the decreasing of the power of the generalized variance statistical test in detecting changes in the covariance matrix when they really take place. It is true that the determinant of the covariance matrix is equal to the product of its eigenvalues and that matrices with the same determinants not necessarily have similar eigenvalues structures (Timm, 2002). Therefore one possibility to improve the generalized variance test is to work with the eigenvalues of the covariance matrix instead of its determinant.

In this paper three statistical tests will be developed by using this approach. Two of them are adaptations of the T^2 Hotelling (1947) and the Hayter and Tsui (1994) statistical tests for population vector means and the third is based on the number of condition of the covariances matrices. The proposed tests will be compared to the generalized variance test by using Monte Carlo simulation. The power of the tests as well as the Average Run Length (Montgomery, 2001) will be estimated under several scenarios and different sample sizes. At the end of the paper an example of application of the tests in the quality control field is presented.

Generalized Variance Test

Let $X_1, X_2,...,X_n$, where $X_k = (X_{k1}, X_{k2},...,X_{kp})'$, $k=1,2,...,n$, be a random sample of size n from a p-variate normal distribution with mean vector $\mu = (\mu_1, \mu_2,...,\mu_p)'$ and covariance matrix Σ, a pxp positive definite matrix, where p is the number of random variables. Let S_{pxp} be the sample covariance matrix defined as

$$S_{pxp} = \frac{1}{n-1} \sum_{k=1}^{n} (X_k - \overline{X}_n)(X_k - \overline{X}_n)' \tag{1}$$

where $\overline{X}_n$ is the sample mean vector. Let the null and the alternative hypothesis be given as: $H_0: \Sigma = \Sigma_0$ and $H_1: \Sigma \neq \Sigma_0$ respectively, where Σ_0 is a pre-specified covariance matrix. For a significance level α, $0 < \alpha < 1$, the null hypothesis will be rejected for values of $|S| <$ LCL or $|S| >$ UCL, where LCL and UCL are defined as (2) and $|S|$ is the determinant of the sample covariance matrix.

$$LCL = max\left\{0, |\Sigma_0|\left(1 - z_{\alpha/2}\frac{\sqrt{b_2}}{b_1}\right)\right\} \quad , \quad UCL = |\Sigma_0|\left(1 + z_{\alpha/2}\frac{\sqrt{b_2}}{b_1}\right) \tag{2}$$

where

$$b_1 = \frac{1}{(n-1)^p} \prod_{i=1}^{p} (n-i) \; ; \; b_2 = \frac{1}{(n-1)^{2p}} \prod_{i=1}^{p} (n-i)\left\{\prod_{j=1}^{p} (n-j+2) - \prod_{j=1}^{p} (n-j)\right\},$$

$|\Sigma_0|$ is the determinant of the Σ_0 and $z_{\alpha/2}$ is the value from the standard normal distribution which area above is $\alpha/2$. The test statistic (2) is based on the fact that under the null hypothesis $E(|S|) = b_1|\Sigma_0|$ and $Var(|S|) = b_2|\Sigma_0|^2$ (Montgomery, 2001).

When the generalized variance test is used in quality control, for each sample of size n of the process the determinant of the sample covariance matrix as given in (1) is calculated. If the determinant $|S|$ belongs to the critical region of the test the process is

declared as out-of-control. Usually, the values of $|S|$ are plotted in a graph in sequential order of sample observations called control chart (Montgomery, 2001).

The generalized variance is an interesting measure since it transforms a *pxp* matrix into a scalar number. However, it fails in many situations since different matrices can have equal determinants as next example shows.

Example

Consider the two matrices given in (3). Both matrices have determinants equal 1 but their structures are completely different. The first represents a case of a bivariate normal distribution with no correlation between the two random variables while in the second the correlation is equal 0.37. However, their eigenvalues structures are completely different since for Σ_1 they are equal 1 and for Σ_2 they are equal 2.404 and 0.4159, respectively. This fact raises the idea of constructing hypothesis tests based on the respective eigenvalues of the covariances matrices separatedly instead of in the determinant form. In the next section three new tests based on this concept will be presented.

$$\Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \Sigma_2 = \begin{bmatrix} 2.32 & 0.4 \\ 0.4 & 0.5 \end{bmatrix} \tag{3}$$

Statistical Tests Based on Eigenvalues

In this section three statistical tests are presented. As described previously, for all tests it is assumed that $X_1, X_2, ..., X_n$, where $X_k = (X_{k1}, X_{k2}, ..., X_{kp})'$, $k=1,2,...,n$, is a random sample of size n from a p-variate normal distribution with mean vector $\mu = (\mu_1, \mu_2, ..., \mu_p)'$ and covariance matrix Σ, a *pxp* positive definite matrix, where p is the number of random variables.

According to the spectral decomposition theorem (Timm, 2002) any covariance matrix Σ_{pxp} can be expressed in terms of its eigenvalues λ_i and the corresponding normalized eigenvectors e_i, $i=1,2,...,p$, as given in (4)

$$\Sigma_{pxp} = \sum_{i=1}^{p} \lambda_i e_i e_i' \tag{4}$$

Let $\hat{\lambda}_i$, $i = 1, 2, ..., p$, be the eigenvalues of the sample covariance matrix S_{pxp}. When the random vector X has a p-variate normal distribution the eigenvalues $\hat{\lambda}_i$ are asymptotically independent and for each i, $\hat{\lambda}_i$ has a normal distribution with mean equal to the corresponding parameter λ_i and variance $2\lambda_i^2 / (n-1)$ (Timm, 2002). This result can be used in the construction of hypothesis tests for the covariance matrix Σ_{pxp} as shown next.

Hotelling T² Test Adapted for the Eigenvalues of the Covariance Matrix

Let the null and the alternative hypothesis be given as $H_0: \Sigma = \Sigma_0$ and $H_1: \Sigma \neq \Sigma_0$, respectively. Let $\lambda_0 = (\lambda_1, \lambda_2, ..., \lambda_p)'$, $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p$, be the vector with the eigenvalues of the pre-specified matrix Σ_0, and $\hat{\lambda} = (\hat{\lambda}_1, \hat{\lambda}_2, ..., \hat{\lambda}_p)'$ be the vector with

the eigenvalues of the sample covariance matrix S_{pxp}, $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \ldots \geq \hat{\lambda}_p$. The T^2 statistic given in (5) is proposed to test the null hypothesis.

$$T^2 = (\hat{\lambda} - \lambda_0)' \Sigma_{\lambda_0}^{-1} (\hat{\lambda} - \lambda_0) = \sum_{i=1}^{p} \frac{(n-1)}{2\lambda_i^2} (\lambda_i - \hat{\lambda}_i)^2 \tag{5}$$

where

$$\Sigma_{\lambda_0} = \begin{bmatrix} 2\lambda_1^2 / (n-1) & 0 & \ldots & 0 \\ 0 & 2\lambda_2^2 / (n-1) & 0 & 0 \\ \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & 2\lambda_p^2 / (n-1) \end{bmatrix}$$

is the covariance matrix of the vector $\hat{\lambda} = \left(\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_p\right)'$ calculated considering $\Sigma = \Sigma_0$. Under the null hypothesis T^2 has approximately a chi-square distribution with p degrees of freedom (χ_p^2). For a given significance level α, the null hypothesis will be rejected for observed values of T^2 larger than the critical value c, where $P[\chi_p^2 > c] = \alpha$, $0 < \alpha < 1$.

The test statistic given in (5) is an adaptation of the Hotelling T^2 test statistic for the mean vector of one population (Hotelling, 1947). Due to the fact that the normal distribution of $\hat{\lambda}_i$ is only asymptotic, the chi-square distribution of (5) is also an approximation. The exact joint distribution of the eigenvalues $\hat{\lambda}_i$ can be found in Anderson (1958) but is complex. Due to these facts, in this paper the exact distribution of the test statistic given in (5) will be generated by Monte Carlo simulation and the performance of the test will be also evaluated under this situation.

Hayter and Tsui Test Adapted for the Eigenvalues of the Covariance Matrix

In 1994, Hayter and Tsui proposed a hypothesis test for the vector mean of one population. Its advantage to the Hotelling T^2 statistical test was the facility in finding out automatically the variables whose means were different than the null hypothesis postulated values. However, as shown in Hayter and Tsui (1994) neither of the tests were uniformly more powerful. In this section we will present an adaptation of Hayter and Tsui´s test to the situation where the objective is to test the covariance matrix. Let the null and the alternative hypothesis be defined as before. The M test statistic proposed by Hayter and Tsui (1994) and adapted in this paper for the eigenvalues takes the form given in (6).

$$M = max\left\{ \left|Y_j\right|, j = 1,2,\ldots,p \right\} \tag{6}$$

where

$$Y_j = \frac{\hat{\lambda}_j - \lambda_j}{\lambda_j \sqrt{2/(n-1)}}.$$

Due to the multivariate normality of the random vector X, the random variables Y_j are independent and asymptotically normal. Therefore, under the null hypothesis the test statistic M is distributed as the maximum of the absolute values of p independent random variables with standard normal distribution. For a significance level α, $0 < \alpha < 1$, the null hypothesis is rejected for observed values of M larger than the constant called $C_{R\alpha}$, such that $P[M > C_{R,\alpha}] = \alpha$. The $C_{R\alpha}$ value is obtained by using a simulation of samples from a p-variate normal distribution with zero mean vector and covariance matrix equal the correlation matrix (P_{pxp}) of the random vector $\hat{\lambda} = \left(\hat{\lambda}_1, \hat{\lambda}_2, ..., \hat{\lambda}_p\right)'$ which under the null hypothesis is the identity matrix. The basic steps of the simulation algorithm (see Hayter and Tsui, 1994) are given as follows:

- *Step 1.* Generate a large number N of vectors of observations from a p-variate normal distribution with zero mean vector and covariance matrix I_{pxp}. The generated vectors are denoted by $Z^1, Z^2, ..., Z^N$.

- *Step 2.* Calculate the statistic M for each of the generated vectors $Z_i = \left(Z_1^i, Z_2^i, ..., Z_p^i\right)'$ from step 1, i.e., for every $i = 1, 2, ..., N$, calculate the value of the statistics
$$M^i = \max\left\{| Z_j^i |, j = 1, 2, ..., p\right\}.$$

- *Step 3.* From the empirical distribution obtained from the sample ($M^1, M^2, ..., M^N$) find the value corresponding to the percentil of order $(1 - \alpha)$ and use this value as the critical constant $C_{R\alpha}$, $0 < \alpha < 1$.

The values of Y_i larger than $C_{R\alpha}$ identify the eigenvalues of the matrix Σ_0 which were significantly different than the corresponding eigenvalues of the sample covariance matrix S_{pxp}.

Hypothesis Test based on the Number of Condition of the Covariance Matrix

Another measure related to the structure of the covariance matrix is the number of condition which is used to evaluate the singularity of the matrix, i.e., to check if the matrix is badly conditioned or not. Generally speaking, the number of condition of a matrix A is defined as $\kappa(A) = \|A\|.\|inv(A)\|$, being $\|.\|$ any pre-specified norm and $inv(.)$ the inverse matrix. When the quadratic norm is chosen and the matrix A is equal its transpose ($A = A'$), the number of condition of A is the ratio between its largest and its smallest eigenvalues as given in (7)

$$\kappa(A) = \frac{\lambda_{max}}{\lambda_{min}} \tag{7}$$

The matrix A is considered badly conditioned if the number of condition is large. For some authors $\kappa(A) > 20$ is considered large enough (Greene, 1997). An interesting fact is that matrices with similar determinants may have different number of conditions. To illustrate this let´s consider again the matrices Σ_1 and Σ_2 given in (3). They both have determinants equal 1 but the number of conditions are different since $\kappa(\Sigma_1) = 1$ and $\kappa(\Sigma_2) = 5.7788$. Therefore, the number of condition can be used as an alternative to differentiate matrices with similar determinants but different components structures. Let $H_0: \Sigma = \Sigma_0$ and $H_1: \Sigma \neq \Sigma_0$, be the null and the alternative hypothesis. Let the test statistic $\kappa(S_{pxp})$ be the number of condition of the sample covariance matrix

S_{pxp} as define in (1). The null hypothesis will be rejected for $\kappa(S_{pxp}) < c_1$ or $\kappa(S_{pxp}) > c_2$, where the constants c_1 and c_2 are obtained from the distribution of the test statistic $\kappa(S_{pxp})$ under H_0 and they are such that

$$P_{H_0}[\kappa\,(S_{pxp}) < c_1] + P[\kappa\,(S_{pxp}) > c_2] = \alpha \quad , \quad 0 < \alpha < 1.$$

The distribution of the test statistic $\kappa(.)$ under the null hypothesis is generated by Monte Carlo simulation and the critical region of the test is found according to the pre-specified significance level (α) of the test. In this paper the constants c_1 and c_2 are such that

$$P_{H_0}[\kappa\,(S_{pxp}) < c_1] = P[\kappa\,(S_{pxp}) > c_2] = \frac{\alpha}{2}$$

Comparing the Eigenvalues and the Generalized Variance Tests

In this section the results of a Monte Carlo study are presented. A total of 10.000 random samples of sizes $n = 5, 10, 25, 50$ and 100 were generated from a multivariate normal distribution under the null ($H_0: \Sigma = \Sigma_0$) and the alternative ($H_1: \Sigma = \Sigma_1 \neq \Sigma_0$) hypothesis being H_0 tested for each sample and each test using 0.05 as the nominal significance level. The proportion of rejection was then calculated given an estimate of the probability of type I error of the test when data were generated under the null hypothesis and an estimate of the power of the test when data were generated under the alternative hypothesis. This procedure was repeated $k=50$ times under the null and the alternative hypothesis and at the end, average estimates of the probability of type I error and the power of the test were obtained by taking the average over all 50 repetitions for each test, respectively. For the generalized variance and the T^2 eigenvalues tests the probability of type I error and the power were determined by using the asymptotic distribution of the test statistic as well as by the exact distribution obtained by Monte Carlo simulation through a generation of 50.000 random samples under the null hypothesis. Without loss of generality the study was performed considering one particular structure for matrix Σ_0 for $p = 2,3$, assuming $\mu=0$ under H_0 and H_1.

Under the statistical point of view, for a fixed significance level, the hypothesis test which presents the highest power values is the best. However, in the quality control field it is usual to analyse the performance of the test by observing the ARL (Average Length Run) value since in quality control samples of size n are observed sequentially at certain units of time (frequency of sampling). The ARL is defined as the average number of samples of size n taken until one indicates that the process is in the out-of-control condition (i.e. the value of the test statistic falls into the critical region of the test, or in other words, it falls outside the control limits). The ARL is given by the inverse of the probability (q) that the value of the test statistic falls in the critical region of the test, i.e., exceeds the control limits. When the process is under control q is the value of the probability of type I error and when the process is out-of- control q is the power of the test. Usually, when designing a control chart a small value of q, under control, is chosen such as 0.0027 for example, in order to decrease the rate of false alarms (a false alarm occurs when the process is declared out-of-control by the control chart when in fact it is not).Then, the test (or control chart) that detects faster

the true out-of-control situation of the process is preferred taking into consideration economical and operational issues also.

For the study presented in this paper the value of q for the process under control is 0.05 which is not a value usually chosen to build a control chart. Independent of that, without loss of generality, the ARL out-of-control estimates of the tests were also estimated with the purpose of compared them on the light of the quality control point of view.

Simulated Models

The simulated models under the alternative hypothesis were chosen to make it possible to evaluate whether the statistical tests based on eigenvalues were able to detect small differences from Σ_0 since it is already known that for larger samples and larger changes in Σ_0 the generalized variance test has good performance. Tables 1 and 2 present the simulated models for $p = 2$ and $p = 3$ with the respective determinants,

Table 1. Simulated models – $p = 2$.

Cases	Covariance Matrices	$\|\Sigma\|$	$tr(\Sigma)$	$\kappa(\Sigma)$	λ_1, λ_2	TR	ER	CR
null $\rho = 0.37$	$\begin{bmatrix} 2.32 & 0.40 \\ 0.40 & 0.50 \end{bmatrix}$	1	2.82	5.7802	2.4040 0.4159	--	--	--
1 $\rho = 0.60$	$\begin{bmatrix} 2.32 & 0.65 \\ 0.65 & 0.50 \end{bmatrix}$	0.7375	2.82	8.6575	2.5280 0.2920	1	1.05 0.7	1.5
2 $\rho = 0$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	1	2	1	1 1	1.41	0.42 2.40	0.17
3 $\rho = 0.66$	$\begin{bmatrix} 2.32 & 0.90 \\ 0.90 & 0.80 \end{bmatrix}$	1.0460	3.12	7.1675	2.7380 0.3820	1.11	1.14 0.92	1.24
4 $\rho = 0.28$	$\begin{bmatrix} 2.32 & 0.30 \\ 0.30 & 0.50 \end{bmatrix}$	1.0700	2.82	5.2417	2.3682 0.4518	1	0.99 1.09	0.91
5 $\rho = 0.38$	$\begin{bmatrix} 2.32 & 0.50 \\ 0.50 & 0.75 \end{bmatrix}$	1.4900	3.07	4.0828	2.4660 0.6040	1.09	1.03 1.45	0.71
6 $\rho = 0.37$	$\begin{bmatrix} 2.32 & 0.57 \\ 0.57 & 1.00 \end{bmatrix}$	1.9900	3.32	3.2280	2.5340 0.7850	1.18	1.05 1.89	0.56
7 $\rho = 0.37$	$\begin{bmatrix} 2.32 & 0.80 \\ 0.80 & 2.00 \end{bmatrix}$	4.000	4.32	2.2135	2.975 1.344	1.53	1.24 3.23	0.38

TR , ER and CR: respectives ratios between the traces, eingenvalues and the number of conditions of the covariances matrices under H_1 and H_0. Case null denotes the covariance model under the null hypothesis; ρ is the correlation coefficient between the two variables.

traces, eigenvalues and the number of conditions of the covariances matrices under H_0 and H_1. The ratios between traces, eigenvalues (only for $p = 2$), determinants and the number of conditions from the covariances matrices under H_0 and H_1 are also presented.

For $p = 2$, models 1 and 4 have the same variance structure as Σ_0 but different correlations; models 5, 6 and 7 preserve the correlation structure of Σ_0 but with changes in the variances; models 2 and 3 have different variances and correlations. For $p = 3$, all models have the same variances structures as Σ_0 except model 6. The covariances matrices of models 1,2,3 present similar determinant as Σ_0 but completely different

Table 2. Simulated models - $p = 3$ variables.

Cases	Covariance matrices	$|\Sigma|$	$\kappa(\Sigma)$	$\lambda_i, i = 1, 2, 3$	DR	TR	CR
null	$\begin{bmatrix} 1 & 0.6 & 0.6 \\ 0.6 & 1 & 0.8 \\ 0.6 & 0.8 & 1 \end{bmatrix}$	0.22	11.69	2.34 0.46 0.20	-	-	-
1	$\begin{bmatrix} 1 & 0.3 & 0.2 \\ 0.3 & 1 & -0.8 \\ 0.2 & -0.8 & 1 \end{bmatrix}$	0.22	16.09	1.76 1.13 0.11	1.00	1.00	1.38
2	$\begin{bmatrix} 1 & 0.3 & 0.2 \\ 0.3 & 1 & 0.8 \\ 0.2 & 0.8 & 1 \end{bmatrix}$	0.33	10.02	1.93 0.87 0.19	1.50	1.00	0.86
3	$\begin{bmatrix} 1 & 0.0 & 0.0 \\ 0.0 & 1 & 0.8 \\ 0.0 & 0.8 & 1 \end{bmatrix}$	0.36	9.00	1.80 1.00 0.20	1.64	1.00	0.77
4	$\begin{bmatrix} 1 & 0.3 & 0.3 \\ 0.3 & 1 & 0.4 \\ 0.3 & 0.4 & 1 \end{bmatrix}$	0.73	2.78	1.67 0.73 0.60	3.32	1.00	0.24
5	$\begin{bmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{bmatrix}$	0.50	4.00	2.00 0.50 0.50	2.27	1.00	0.34
6	$\begin{bmatrix} 1 & 0.6 & 0.6 \\ 0.6 & 4 & 0.8 \\ 0.6 & 0.8 & 1 \end{bmatrix}$	2.14	11.10	4.36 1.25 0.39	9.73	2.00	0.95

DR, TR, and CR: respectives ratios between the determinants, traces and the number of conditions of the covariance matrices under H_1 and H_0; Case null denotes the covariance model under the null hypothesis.

covariances (correlations) structures and eigenvalues. The determinants of models 4, 5, 6 are more different than the respective value of Σ_0. An interesting case is model 6 since it presents the largest difference in the determinant compared to Σ_0, different eigenvalues but similar number of conditions. The ratio of the determinants for the bivariate models, ranged from 0.7375 (case 1) to 4 (case 7). For the trivariate models the ratio ranged from 1 (case 1) to 9.73 (case 6).

Results and Discussion

The averages proportion of rejection of H_0 for each test discussed in this paper are shown in Tables 3 and 4, for $p = 2$, under the null and the alternative hypothesis models. The ratios (*PR*) between the estimated powers of the eigenvalues tests and the generalized variance test are presented in Table 5. When the exact distributions were used to build the critical region of the tests, the estimates of the probability of type I error were 0.05 as expected. However, for all sample sizes the estimates of the probability of type I error from the generalized variance test based on the normal distribution were smaller than the pre-specified significance level of the test. In the quality control context it means that the amount of false alarms of the control chart will be larger than the expected under the null hypothesis. The difference was about 0.03 for $n = 5$, 0.02 for $n = 10,25$ and 0.01 for larger samples ($n = 50, 100$). This reinforces the fact that the normal approximation for the test statistic of the generalized variance test should be avoided for small samples and used with some care for samples sizes which are usually considered large such as $n = 100$. For $n = 5$, the estimate of the probability of type I error was also smaller than 0.05 for T^2 test when the chi-square distribution was used as a reference distribution to build the critical region of the test (estimate = 0.04). Due to these facts, in this study the comparison of the generalized variance with the three other statistical tests will be performed by considering the exact distribution only. For the T^2 test both distributions (asymptotic and exact) will be used in the comparisons except for $n = 5$ for which only the exact will be considered.

In general, the statistical tests based on the eigenvalues performed better than the generalized variance test considering the exact distribution (*EGV*) resulting in more similar values only for situations where the determinants of Σ_0 and Σ_1 matrices were very different (*DR* ≥ 1.49 ; cases 5, 6 and 7). However, even for these cases the powers estimates of the generalized variance test were smaller than the values obtained for the T^2(*ET2*) and Hayter and Tsui (*HT*) tests except for $n = 5$ where *EGV* presented similar performance than these two tests. For $n \geq 10$, the estimated power gains of *ET2* and *HT* tests (see *PR* values in Table 5) over the *EGV* test in case 5 ranged from 18 to 46%; for case 6 from 6 ($n = 100$) to 40% ($n = 25$). For case 7 there was gain only for $n = 10$ (10 to 14%) being the *EGV* similar to *ET2* and *HT* for the other samples sizes what was expected since the determinant of matrix Σ_1 in case 7 is 4 times larger than the determinant of Σ_0, a situation that favours the performance of the *EGV* test. The condition number test (*CN*) presented smaller power estimates than *EGV* for cases 6 and 7 and more similar values for case 5.

The generalized variance test presented very poor performance for cases 2, 3 and 4 since the power estimates were around 0.05 for all n. This can be explained by the proximity of the determinants values of Σ_0 and Σ_1 although the matrices have complete different correlation structures. Not even for $n = 100$ the *EGV* was able to

Table 3. Probability of type I error and power estimates of the tests - p = 2, n = 5, 10, 25.

	Average rejection of the null hypothesis																		
	n=5						n = 10						n = 25					
Case	GV	EGV	T2	ET2	HT	CN	GV	EGV	T2	ET2	HT	CN	GV	EGV	T2	ET2	HT	CN
null	0.02	0.05	0.04	0.05	0.05	0.05	0.03	0.05	0.05	0.05	0.05	0.05	0.03	0.05	0.05	0.05	0.05	0.05
1	0.01	0.05	0.05	0.05	0.05	0.06	0.01	0.06	0.05	0.05	0.04	0.08	0.01	0.10	0.09	0.05	0.08	0.25
2	0.02	0.05	0.04	0.05	0.05	0.10	0.03	0.05	0.16	0.17	0.17	0.33	0.03	0.05	0.75	0.77	0.66	0.97
3	0.03	0.05	0.07	0.08	0.07	0.05	0.04	0.05	0.07	0.08	0.07	0.06	0.05	0.05	0.09	0.09	0.09	0.14
4	0.03	0.05	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.06	0.05	0.05	0.05	0.06	0.06	0.06	0.06	0.11
5	0.06	0.08	0.08	0.09	0.08	0.06	0.12	0.11	0.12	0.13	0.13	0.07	0.22	0.18	0.25	0.26	0.26	0.21
6	0.11	0.13	0.12	0.14	0.13	0.06	0.23	0.20	0.25	0.27	0.26	0.10	0.48	0.42	0.58	0.59	0.59	0.40
7	0.29	0.31	0.30	0.32	0.30	0.08	0.60	0.57	0.63	0.65	0.64	0.19	0.94	0.92	0.97	0.97	0.97	0.72

GV and EGV: the generalized variance tests with normal approximation and the exact distributions respectively; T2 and ET2: T2 Hotelling adapted for eigenvalues tests with chi-square and the exact distributions respectively; CN: condition number test. Case null denotes the covariance model under the null hypothesis.

detect any difference between these two matrices (power estimate ≤ 0.07). For case 2 the power estimates of the *ET2* and *HT* tests ranged from 0.05 ($n = 5$) to 1 ($n = 50,100$) being equal to 0.77 and 0.66 respectively for $n = 25$; the best test for case 2 however, was the *CN* (power estimates from 0.10 ($n = 5$) to 1 ($n = 50,100$) being equal 0.97 for $n = 25$. By Table 5 it can be seen that the power gains obtained by using *CN* instead of *EGV* for case 2 were very large (*PR* estimates ranged from 6.6 to 20).

For case 3, although the power estimates of *ET2*, *HT* and *CN* tests were not very large, they were larger than the power estimates of the *EGV* for all sample sizes. The ratio (*PR*) between the estimated powers of these three tests compared to *EGV* ranged from 1.2 to 4.67 being *CN* the best test for this case (estimated powers between 0.05 to 0.28).

For case 4, which is also a difficult model to be detected by the tests since the structure of the matrix Σ_1 is very similar than the structure of Σ_0, the estimated power values for all tests were low being *CN* the test more capable of detecting the difference between the two covariance matrices for $n \geq 25$ (power estimates ranged from 0.11 to 0.14).

In all cases, except for $n = 5$, the T^2 test using the chi-square distribution presented similar power estimates than T^2 with the exact distribution. This is an indication that for $n \geq 10$ the asymptotic distribution for the T^2 test statistic could be used instead of the exact distribution to build the critical region of the test. Under the practical point of view this is a good point in favour of using T^2 test regarding to its competitors *HT* and *CN*.

It is important to point out that for $n = 5$ the majority of the power estimates of all tests were small (around 0.05 except for cases 6 and 7). This is due partially by the nature of the simulated models since some parameters of the Σ_0 and Σ_1 matrices

Table 4. Probability of type I error and power estimates of the tests - $p = 2$, $n = 50, 100$.

	Average rejection of the null hypothesis											
	n = 50						n = 100					
Case	GV	EGV	T2	ET2	HT	CN	GV	EGV	T2	ET2	HT	CN
null	0.04	0.05	0.05	0.05	0.04	0.05	0.04	0.05	0.05	0.05	0.04	0.05
1	0.06	0.16	0.21	0.22	0.19	0.40	0.20	0.30	0.49	0.50	0.48	0.64
2	0.04	0.05	1.00	1.00	0.99	1.00	0.04	0.05	1.00	1.00	1.00	1.00
3	0.05	0.05	0.12	0.12	0.11	0.19	0.06	0.06	0.18	0.18	0.16	0.28
4	0.06	0.06	0.07	0.07	0.07	0.12	0.07	0.07	0.08	0.09	0.08	0.14
5	0.36	0.30	0.43	0.44	0.44	0.32	0.57	0.53	0.71	0.71	0.70	0.53
6	0.74	0.68	0.85	0.86	0.86	0.64	0.94	0.93	0.99	0.99	0.99	0.89
7	1.00	1.00	1.00	1.00	1.00	0.95	1.00	1.00	1.00	1.00	1.00	1.00

GV and *EGV*: the generalized variance tests with normal approximation and the exact distributions; *T2* and *ET2*: T^2 Hotelling adapted for eigenvalues tests with chi-square and the exact distributions; *CN*: condition number test. Case null denotes the covariance model under the null hypothesis.

Table 5. Ratios (PR) between the estimated power of the eingenvalues and the generalized variance (EGV) tests.

p = 2

Case	n = 10				n = 25				n = 50				n = 100			
	T2	ET2	HT	CN	T2	ET2	HT	CN	T2	ET2	HT	CN	T2	ET2	HT	CN
1	0.83	0.83	0.67	1.33	0.90	1.00	0.80	2.50	1.31	1.37	1.19	2.50	1.63	1.67	1.60	2.13
2	3.20	3.40	3.40	6.60	15.00	15.40	13.20	19.40	20.00	20.00	19.80	20.00	20.00	20.00	20.00	20.00
3	1.40	1.60	1.40	1.20	1.80	1.80	1.80	2.80	2.40	2.40	2.20	3.80	3.00	3.00	2.67	4.67
4	1.00	1.20	1.00	1.00	1.00	1.00	1.00	1.83	1.17	1.16	1.17	2.00	1.14	1.29	1.14	2.00
5	1.09	1.18	1.18	0.64	1.39	1.44	1.44	1.17	1.43	1.46	1.46	1.07	1.34	1.34	1.32	1.00
6	1.25	1.35	1.30	0.50	1.38	1.40	1.40	0.95	1.25	1.26	1.26	0.94	1.06	1.06	1.06	0.96
7	1.10	1.14	1.12	0.33	1.05	1.05	1.05	0.78	1.00	1.00	1.00	0.95	1.00	1.00	1.00	1.00

p = 3

Case	n = 10				n = 25				n = 50				n = 100			
	T2	ET2	HT	CN	T2	ET2	HT	CN	T2	ET2	HT	CN	T2	ET2	HT	CN
1	6.00	6.80	5.80	1.20	17.20	17.60	15.80	2.40	20.00	20.00	20.00	4.00	20.00	20.00	20.00	7.40
2	2.00	2.33	2.11	0.55	3.71	3.86	3.71	0.43	3.61	3.65	3.61	0.35	2.53	2.54	2.54	0.31
3	2.30	2.60	2.40	0.50	3.53	3.63	3.53	0.42	2.85	2.88	2.85	0.39	1.82	1.82	1.82	0.45
4	1.00	1.13	1.00	1.06	1.30	1.32	1.27	1.26	1.07	1.08	1.07	1.07	1.00	1.00	1.00	1.00
5	0.68	0.84	0.68	0.95	1.34	1.39	1.34	1.29	1.37	1.37	1.37	1.31	1.09	1.09	1.09	1.09
6	0.99	1.03	0.97	0.07	1.00	1.00	0.99	0.05	1.00	1.00	1.00	0.05	1.00	1.00	1.00	0.06

T2 and ET2: T² Hotelling adapted for eigenvalues tests with chi-square and the exact distributions respectively; CN: condition number test.

were very similar, and also by the sample size which statistically speaking, is too small to estimate the population covariance matrix properly.

Table 6 presents the averages of the estimated proportions of rejection of H_0 for each test, for $p = 3$. Only the results from sample sizes $n \geq 10$ were considered since the proper estimation of the covariance matrix when $p = 3$ requires n larger than 6. The ratios (*PR*) between the estimated powers of the eigenvalues tests and the generalized variance test are presented in Table 5. When the exact distributions were used to build the critical region of the tests the estimates of the type I error were 0.05 as expected. Similar as seen for $p = 2$, the probability of type I error estimates for the generalized variance test with the normal distribution (*GV*) were smaller than 0.05 even for $n = 50$ (value around 0.03) being equal 0.04 only for $n = 100$. For these two respective sample sizes the *GV* estimates for all cases were very similar to the values of the exact test *EGV*. Due to these facts the results from *GV* will not be shown in the tables used as a support for the discussion presented in this section.

The results from Table 6 show that the statistical tests T^2 and *HT* based on eigenvalues presented larger power estimates than the generalized variance test (except for case 5, $n = 10$, where they were similar) in some of the cases and they were able to detect the differences from Σ_0 to Σ_1 even for case 1, which the determinants of both matrices are equal. For this particular case the *EGV* test failed in detecting the differences even for larger samples sizes (estimated power = 0.05; $n = 100$) whereas the other two tests presented power estimates around 0.30 (for $n = 10$) and 1 (for $n \geq 25$). The performance of the *EGV* test improved in cases 4, 5 and 6, resulting in similar values than T^2 and *HT* for case 6 which was expected since the determinant of Σ_1 is 9.73 larger than the respective value of Σ_0. For case 4 the *EGV* test also presented similar performance than *HT* for $n = 10$ but lower power estimates for $n \geq 25$. By Table 5 it can be seen that for $n \geq 10$ the gains in the power of the test by using the T^2 and *HT* were large for cases 1,2 and 3 (*PR* ranged from for cases 1.82 to 20). For cases 4 and 5 the *PR* ratio ranged from 1.09 to 1.39 (for $n \geq 25$) being around 1 for $n = 10$ and close to 1 for case 6.

The number of condition test (*CN*) performs better than *EGV* for cases 1, 4 and 5 particularly for case 1 (*PR* ranged from 1.2 to 7.4) although the power estimates were not very expressive for this case (ranged from 0.06 to 0.37). For cases 2, 3 and 6 the *CN* test power estimates were lower than *EGV* specially for case 6 (estimated power around 0.05 for all sample sizes) which can be explained by the high similarity of the number of condition of both matrices Σ_0 and Σ_1.

For $p = 3$, the T^2 test based on the chi-square distribution (T^2) had similar performance than the T^2 based on the exact distribution (*EGV*).

The ARL estimates out-of-control are shown in Figure 1 (for $p = 2$) and Figure 2 (for $p = 3$). The results indicated that the tests based on the eigenvalues could be an alternative to the generalized variance test in the quality control field since in the majority of the cases discussed in this paper these tests presented ARL values similar or lower than *EGV*, i.e., they were able to identify the true out-of-control condition faster than *EGV* even for smaller samples sizes and situations where the differences between the covariance matrices under the null and the alternative hypothesis were not very large.

Table 6. Probability of type I error and power estimates of the tests - p = 3, n = 10, 25,50,100.

		Average rejection of the null hypothesis																		
	n = 10					n = 25					n = 50					n = 100				
Case	EGV	T2	ET2	HT	CN	EGV	T2	ET2	HT	CN	EGV	T2	ET2	HT	CN	EGV	T2	ET2	HT	CN
null	0.05	0.04	0.05	0.04	0.05	0.05	0.05	0.05	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
1	0.05	0.30	0.34	0.29	0.06	0.05	0.86	0.88	0.79	0.12	0.05	1.00	1.00	1.00	0.20	0.05	1.00	1.00	1.00	0.37
2	0.09	0.18	0.21	0.19	0.05	0.14	0.52	0.54	0.52	0.06	0.23	0.83	0.84	0.83	0.08	0.39	0.99	0.99	0.99	0.12
3	0.10	0.23	0.26	0.24	0.05	0.19	0.67	0.69	0.67	0.08	0.33	0.94	0.95	0.94	0.13	0.55	1.00	1.00	1.00	0.25
4	0.32	0.32	0.36	0.32	0.34	0.69	0.90	0.91	0.88	0.87	0.93	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
5	0.19	0.13	0.16	0.13	0.18	0.41	0.55	0.57	0.55	0.53	0.68	0.93	0.93	0.93	0.89	0.92	1.00	1.00	1.00	1.00
6	0.75	0.74	0.77	0.73	0.05	0.99	0.99	0.99	0.98	0.05	1.00	1.00	1.00	1.00	0.05	1.00	1.00	1.00	1.00	0.06

EGV: the generalized variance test with the exact distribution; T2 and ET2: T^2 Hotelling adapted for eigenvalues tests with chi-square and the exact distributions respectively; CN: condition number test. Case null denotes the covariance model under the null hypothesis.

Under the quality control point of view the T^2 test has some advantage on Hayter and Tsui since for n ≥ 10 the chi-square distribution may be used to build the critical region of the test instead of the exact distribution, on the contrary of Hayter and Tsui´s which requires a simulation procedure to determine the constant $C_{R\alpha'}$ for any sample size n. Therefore, in practice the T^2 test is more feasible and for this reason it would be more appropriated in the quality control area even considering that the simulation procedure necessary to determine the constant $C_{R\alpha}$ is very simple and computationally fast.

Example of Application

In Montgomery (2001) an example was presented which two quality characteristics, the tensile strength (X_1) and the diameter (X_2), were measure in $k = 20$ samples of $n = 10$ fiber specimens. The statistical analysis of the data presented in Montgomery had established the following vector μ_0 and the matrix Σ_0 as the parameters of the process in the under control condition:

$$\mu_0 = (115.59, 1.06)' \quad \Sigma_0 = \begin{bmatrix} 1.23 & 0.79 \\ 0.79 & 0.83 \end{bmatrix}$$
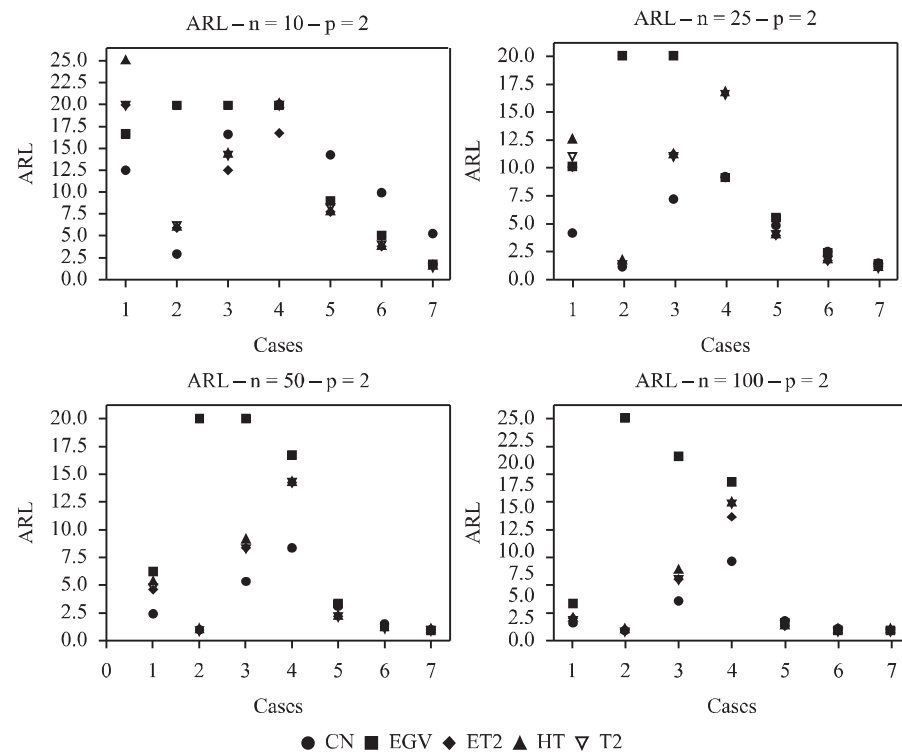

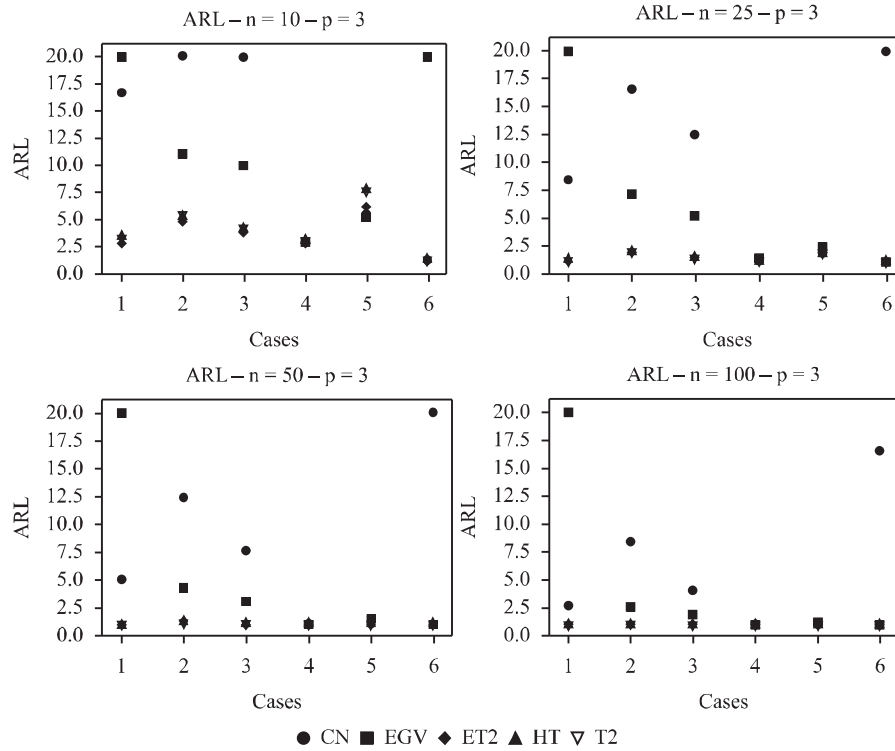**Figure 1.** ARL out-of-control estimates – n = 10,25,50,100 – p = 2.

Figure 2. ARL out-of-control estimates – n = 10,25,50,100 – p = 3.

The determinant of Σ_0 is equal 0.3968, the eigenvalues are λ_1 = 1.8449 and λ_2 = 0.2151, the trace is 2.06 and the number of condition is 8.5769. The correlation between X_1 and X_2 is equal 0.782.

To illustrate the statistical tests discussed in this paper 5 samples of size n = 10 were generated from a multivariate normal distribution with mean vector μ_0 being the first sample from the process under control and the remaining by introducing special causes in the process which affected: (i) the standard deviation of the diameter of the fiber only (an increase of 1.6) with no change in the correlation structure (sample 2); (ii) the standard deviation of the tensile strength of the fiber (an increase of 2) decreasing the correlation between variables (sample 3); (iii) the standard deviations of both variables (an increase of 1.6 for the first variable and 2 for the second) increasing the correlation between variables (sample 4); (iv) the standard deviations of both variables (an increase of 2.5 for the first variable and 3 for the second) decreasing the correlation between variables (sample 5). The respective sample covariances matrices are given by

$$S_1 = \begin{bmatrix} 1.13 & 0.87 \\ 0.87 & 1.04 \end{bmatrix} S_2 = \begin{bmatrix} 1.28 & 0.95 \\ 0.95 & 1.81 \end{bmatrix} S_3 = \begin{bmatrix} 4.26 & 0.25 \\ 0.25 & 0.73 \end{bmatrix}$$

$$S_4 = \begin{bmatrix} 2.80 & 2.69 \\ 2.69 & 3.00 \end{bmatrix} S_5 = \begin{bmatrix} 6.21 & 0.52 \\ 0.52 & 5.17 \end{bmatrix}$$

Table 7. Example – Observed values of the tests statistics - $p = 2$ - $n = 10$.

Test	Sample 1 $r_{12} = 0.80$	Sample 2 $r_{12} = 0.62$	Sample 3 $r_{12} = 0.14$	Sample 4 $r_{12} = 0.93$	Sample 5 $r_{12} = 0.09$
EGV	0.418	1.414	3.047	1.163	31.835
T2	0.016	12.111	31.882	18.566	2212.97
ET2	0.016	12.111	31.882	18.566	2212.97
HT	0.127	3.389	4.905	4.308	46.746
CN	9.149	4.530	6.004	26.866	1.297

Rejection Regions: EGV <0.024 or > 1.802; T2>11.829; HT>3.209; CN < 1.387 or > 108.119. r_{12}: sample correlation between variables; $\alpha = 0.0027$.

The corresponding eigenvalues ($\hat{\lambda}_1, \hat{\lambda}_2$) of the sample covariances matrices are: 1.9562 and 0.2138 (sample 1); 2.5313 and 0.5587 (sample 2); 4.2776 and 0.7124 (sample 3); 5.5919 and 0.2081 (sample 4); 6.4254 and 4.9546 (sample 5).

For each sample the null hypothesis $H_0: \Sigma = \Sigma_0$, was tested against $H_1: \Sigma \neq \Sigma_0$ at the significance level $\alpha = 0.0027$, by the generalized variance implemented with the exact distribution and the tests based on the eigenvalues discussed in this paper. The results are given in Table 7 with the corresponding rejection limits for H_0. Under the null hypothesis the standard deviations of the first and second eigenvalues ($\hat{\lambda}_1, \hat{\lambda}_2$) are given respectively by 0.8697 and 0.1014.

All tests did not reject the null hypothesis for sample 1 and rejected for sample 5. This was expected since sample 1 came from a process under control and sample 5 came from a process with large variation in both variables affecting the covariance matrix strongly (the determinant of the sample covariance matrix S_5 is about 80 times larger than Σ_0). The condition number (*CN*) failed in rejecting the null hypothesis in samples 2,3 and 4 whereas the generalized variance failed in samples 2 and 4. On the contrary, Hayter and Tsui and T^2 (with the exact and asymptotic distributions) statistical tests were able to detect all the changes performed in the parameters of the process since they rejected the null hypothesis for samples 2 to 5. Therefore, they were more appropriated than *EGV* and *CN* tests in this example.

Final Remarks

The results presented in this paper showed that the adaptation of T^2 Hotelling and Hayter and Tsui statistical tests based on the eigenvalues of the covariances matrices were more powerful than the generalized variance test, except for larger changes in the determinant of the covariance matrix postulated in the null hypothesis, (Σ_0), cases in which the tests presented similar power estimates. It is important to point out that the T^2 Hotelling and Hayter and Tsui statistical tests were able to detect small changes in the parameters of Σ_0, on the contrary of the generalized variance test. The test based on the number of condition (*CN*) did not perform well only in situations where Σ_0 and Σ_1 had different eigenvalues structures but similar number of conditions.

Considering the results from the simulation study the T^2 Hotelling and Hayter and Tsui statistical tests are better alternatives than the generalized variance test for the

covariance matrix under multivariate normal distribution, since they presented similar or larger power than the respective test. In terms of quality control Hotelling T^2 test has some advantage since for $n \geq 10$ it can be implemented by using the chi-square as a reference distribution to build the critical region of the test while Hayter and Tsui and the number of condition tests depend upon a simulation procedure to find it, being the simulation required for Hayter and Tsui simpler than the simulation required for the number of condition test.

Acknowledgments

References

Alt, F. (1985) Multivariate Quality Control, in: Kotz, S. and Johnson, NL. (Eds), Encyclopedia of Statistical Sciences. New York: John Wiley, pp. 110-122.

Anderson, T.W. (1958) An Introduction to Multivariate Statistical Analysis. New York: John Wiley.

Aparisi, F.; Jabaioyes, J. and Carrion, A. (1999) Statistical Properties of the |s| Multivariate Control Chart. Communication in Statistics - Theory and Methods, Vol. 28, No. 11, pp. 2671-2686.

Costa A.F.B. and Machado, M.A.G. (2008) A New Multivariate Control Chart for Monitoring the Covariance Matrix of Bivariate Processes. Communication in Statistics - Simulation and Computation, Vol. 37, No. 7, pp. 1453-1465.

Djauhari, M.A. (2009) Asymptotic Distribution of Sample Covariance Determinant. Matematika, Vol. 25, No. 1, pp. 79-85.

Garcia-Diaz, J.C. (2007) The Effective Variance Control Chart for Monitoring the Dispersion Process with Missing Data. European Journal of Industrial Engineering, Vol. 1, No. 1, pp. 40-55. http://dx.doi.org/10.1504/EJIE.2007.012653

Greene, W.H. (1997) Econometric Analysis. Upper Saddle River: Prentice Hall.

Hayter, A.J. and Tsui, K.L. (1994) Identification and Quantification in Multivariate Quality Control Problems. Journal of Quality Technology, Vol. 26, No. 3, pp. 197-208.

Hotelling, H. (1947) Multivariate Quality Control, in: Eisenhart, M.; Hastay, W. and Wallis, A. (Eds.), Techniques of Statistical Analysis. New York: MacGraw-Hill, pp. 111-184.

Montgomery, D.C. (2001) Introduction to Statistical Quality Control. New York: John Wiley.

Timm, N.H. (2002) Applied Multivariate Analysis. New York: Spring Verlag.

Yeh, A.B.; Lin, D.K.J. and McGrath, R.N. (2006) Multivariate Control Charts for Monitoring Covariance Matrix: A Review. Quality Technology et. Quantatitative Management, Vol. 3, No. 4, pp. 415-436.

Biography

Sueli Aparecida Mingoti is an Associate Professor at the Statistics Departament of Federal University of Minas Gerais (UFMG) located at Minas Gerais, Brazil. Her research fields are Multivariate and Industrial Statistics.

Contact: sueliam@est.ufmg.br

Letícia Pereira Pinto is doctor student at the Statistics Department of Federal University of Minas Gerais (UFMG). Her research fields are Quality Control and Spatial Statistics.

Contact: leticiap@dcc.ufmg.br

Article Info