AN ABSTRACT OF THE DISSERTATION OF

Wichai Chattinnawat for the degree of Doctor of Philosophy in Industrial Engineering presented on September 5, 2003.

Title: Multivariate Control Charts for Nonconformities.

Abstract approved:

Redacted for Privacy

Robin C. Wurl

When the nonconformities are independent, a multivariate control chart for nonconformities called a demerit control chart using a distribution approximation technique called an Edgeworth Expansion, is proposed. For a demerit control chart, an exact control limit can be obtained in special cases, but not in general. A proposed demerit control chart uses an Edgeworth Expansion to approximate the distribution of the demerit statistic and to compute the demerit control limits. A simulation study shows that the proposed method yields reasonably accurate results in determining the distribution of the demerit statistic and hence the control limits, even for small sample sizes. The simulation also shows that the performances of the demerit control chart constructed using the proposed method is very close to the advertised for all sample sizes.

Since the demerit control chart statistic is a weighted sum of the nonconformities, naturally the performance of the demerit control chart will depend on the weights assigned to the nonconformities. The method of how to select weights that give the best performance for the demerit control chart has not yet been addressed in the literature. A methodology is proposed to select the weights for a one-sided demerit control chart with and upper control limit using an asymptotic technique. The asymptotic technique does not restrict the nature of the types and classification scheme for the nonconformities and provides an optimal and explicit solution for the weights.

In the case presented so far, we assumed that the nonconformities are independent. When the nonconformities are correlated, a multivariate Poisson lognormal probability distribution is used to model the nonconformities. This distribution is able to model both positive and negative correlations among the nonconformities. A different type of multivariate control chart for correlated nonconformities is proposed. The proposed control chart can be applied to nonconformities that have any multivariate distributions whether they be discrete or continuous or something that has characteristics of both, e.g., non-Poisson correlated random variables. The proposed method evaluates the deviation of the observed sample means from pre-defined targets in terms of the density function value of the sample means. The distribution of the control chart test statistic is derived using an approximation technique called a multivariate Edgeworth expansion. For small sample sizes, results show that the proposed control chart is robust to inaccuracies in assumptions about the distribution of the correlated nonconformities.

Multivariate Control Charts for Nonconformities

by
Wichai Chattinnawat

A DISSERTATION

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Presented September 5, 2003
Commencement June 2004

Doctor of Philosophy dissertation of Wichai Chattinnawat

presented on September 5, 2003.

APPROVED:

Redacted for Privacy

Major Professor, representing Industrial Engineering

Redacted for Privacy

Head of the Department of Industrial and Manufacturing Engineering

Redacted for Privacy

Dean of the Graduate School

I understand that my dissertation will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my dissertation to any reader upon request.

Redacted for Privacy

Wichai Chattinnawat, Author

ACKNOWLEDGEMENT

I would like to thanks my parents for giving me the opportunity to see, understand and to be a part of our educational-oriented way of life and finally be able to produce a supposedly piece of contribution. My wife Noy and my daughter Bambi have, and always, been my motivation for this dissertation and my works.

In the past years I have taken many classes in Industrial Engineering, Statistics and Mathematics. My most memorable learning experience is, in fact, from taking and sitting in the linear model classes of Dr. Justus F. Seely and Dr. David S. Birkes. I have seen a significant idea and method for developing and handling statistical works which have given me confidence and insight on how to handle my dissertation. This experience is the most valuable academic experience in my lifetime (even though I was unable to complete linear models I and II for a grade). Dr. Birkes had guided and assisted me as well as directed my work into a sensible direction and supported me with many suggestions, resources, and references for this dissertation.

My major professor Dr. Robin C. Wurl has been involved with this dissertation and has showed and taught me the most important concepts in conducting a research; and in determining and defining a practical contribution. I could not have finished this dissertation without her even though my English ability might have driven her crazy. Dr.Wurl has thoroughly, methodically and painstakingly directed my work, verified the results, and meticulously helped revise and write the dissertation. I am in debted to Dr.Wurl, Dr.Birkes and Dr.Seely for the substantial influence they had in this work.

Finally I would like to thank the department of Industrial and Manufacturing Engineering, Statistics and Mathematics at Oregon State University for providing me with the computing facilities and the library resources. I also would like to thank my committee for giving me a supportive environment.

CONTRIBUTION OF AUTHORS

Dr.David S. Birkes has guided the developments of the statistical analysis and gave a version of a proof for computing an optimal solution for designing a demerit weighting scheme for the one-sided demerit control chart paper and a proof that the adjusted chi-square statistic is invariant under affine transformation.

Dr. Robin C. Wurl contributed greatly in reviewing the work in this dissertation. She has carefully read, criticized, corrected, rewritten (where necessary) the papers given here into a professional format.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF TABLES (CONTINUED)

LIST OF APPENDICES

MULTIVARIATE CONTROL CHARTS FOR NONCONFORMITIES

GENERAL INTRODUCTION

Consider a batch process where the quality of the finished product is determined by the number and types of nonconformities on the product. For example, a weaving process for wire mesh from which rolls of finished wire are produced. Quality of the finished product is accessed by complete inspection of the finished product. The quantity, type, and severity of the nonconformities determine the overall quality of the product. The finished product is considered acceptable as long as the number of nonconformities of each type remains within preset limits.

We consider a multivariate control chart to monitor several types of nonconformities. When the nonconformity types are independent the demerit control chart can be used. The demerit control chart combines the nonconformities of different types into a single control chart statistic and hence a single control chart. The demerit control chart is more advantageous than separate control charts for each type of nonconformity because the number of control charts could grow quickly and may be difficult or impossible for an operator to manage. In addition, and possibly more importantly, the overall probability of obtaining a false alarm from one or more of the control charts is a direct and very undesirable result of adding more control charts.

The first contribution of this work is that we propose a control chart to simultaneously monitor several types of independent nonconformities as a generalized

demerit control chart that allows any arbitrary classification and weighting scheme to be used. A distribution approximation technique other than a traditional normal approximation is used to approximate the control chart statistic and produces a much more accurate control limit and hence better control chart properties than the conventional method.

The second contribution we propose is a methodology called an asymptotic technique to design the weights for a generalized demerit control chart. The technique provides an optimal and explicit solution for the weights that merely depends on the relative severities of the nonconformity types. The proposed technique assigns more relative weight to the more severe types of nonconformities.

The last contribution presented is an approach to statistically model correlated nonconformities and we propose two forms of a multivariate control chart that are in turn more robust and less dependent on the distribution for the nonconformities that is chosen. The proposed multivariate control chart statistically monitors changes in the mean of one or more of the nonconformities by judging the deviation of the observed sample means from known targets in terms of the density function of the sample means. The proposed method uses a distribution approximation technique to approximate the density function of the sample means and to empirically calculate a single sided control chart with an upper control limit.

In general, the exact control limit for the demerit control chart can be obtained for special cases only. The demerit control chart statistic is a weighted sum of the nonconformities and the performance of the demerit control chart depends on the

weights given to the nonconformities. Dodge and Torrey (1977) developed and presented a classification scheme of nonconformity types into four types, i.e., very serious, serious, moderately serious, and minor and the weighting scheme is fixed values for weights for the four nonconformity types. This method of classifying and weighting has since been adopted as standard practice. They defined the *number of demerits* as the weighted sum of the nonconformities of different types and based the control chart on the assumption that the number of demerits is approximately normally distributed and suggested 3-sigma control limits. Jones et al. (1999) observed that if the normality assumption is violated, the performance of the demerit control chart is severely compromised and, in response, developed an exact distribution for the demerit statistic using a numerical method to determine the distribution of the statistic and then compute the control limits. In particular, if the weights in the demerit function are not chosen as the ratio proposed by Jones et al. (1999), then the exact distribution for the demerit statistic cannot be obtained. Moreover the method of how to select weights that gives the best performance for the demerit control chart proposed by Dodge and Torrey (1977) or any other arbitrary weighting scheme has not yet been addressed in the literature.

In the first contribution, we present an integrated approach for dealing with the demerit control chart. First we propose a method of constructing the control limits of the demerit control chart with any weighting scheme. The proposed method uses a distribution approximating technique called an Edgeworth Expansion to approximate the distribution of the number of demerits and to compute the demerit control limits.

The proposed method removes the restriction on the weights and provides, in general, a more accurate estimation than the normal distribution. A simulation study shows that the proposed method yields reasonably accurate results in determining the distribution of the demerit statistic, the control limits and, estimating the properties of the control chart even for small sample sizes.

For the second contribution we propose a methodology to select the weights for the number of demerits. An asymptotic technique is used to find the weights and a one-sided demerit control chart is developed. The asymptotic technique does not restrict the nature of the types of nonconformities and can be applied to the weighting scheme suggested by Dodge and Torrey (1977) but also to any general weighting scheme. The asymptotic technique chooses weights for the demerit control chart that maximizes the probability of detecting an out-of-control condition and has several useful and advantageous properties. For example, the asymptotic technique provides an optimal and explicit solution for the weights and the solution does not depend on the magnitude of a shift or shifts in one or more of the nonconformity types but rather depends on the directions and the relative severities of the nonconformity types. The asymptotic technique assigns more weight to the more severe type of nonconformities.

The last contribution is a multivariate control chart for nonconformities, i.e., for nonconformities that are correlated. We present two forms of a multivariate control chart: the adjusted chi-square control chart if parameter values are known and the adjusted T^2 control chart if parameter values are unknown. Both control charts statistically monitor changes in the mean of one or more of the nonconformities. The

work on multivariate control charts for nonconformities has seldom been addressed in the literature. Patel (1973) proposed a statistical quality control method for a multivariate Poisson and multivariate binomial random variable. Patel (1973) assumes that the sample size is sufficiently large so that the Chi-square distribution is very nearly normally distributed and therefore the T^2 statistic can be applied to test for changes in the means of binomial or Poisson random variables; see Montgomery (2001) for the T^2 control chart. If the sample size is "small" or the test statistic is not distributed as multivariate normal, the performance of the control charts by Patel (1973) could be problematic; see Ajmani (1997). The performance of a multivariate control chart for nonconformities can be improved if the distribution of the test statistic were known or if a close approximation it could be found. We propose two forms of a multivariate control chart for correlated nonconformities that are more robust to the changes in the distribution of the nonconformities and to changes in the parameters of those distributions.

To develop the proposed multivariate control charts, a multivariate Poisson Lognormal distribution is used as a general model for the numbers of nonconformities. This distribution assumes the nonconformities are Poisson and also that the means of the Poisson random variables are also random variables and distributed as lognormal. The form of the test statistic was developed and a technique called a Multivariate Edgeworth Expansion was used to approximate the joint distribution of the sample means of the control chart statistic. A single sided control chart is presented, i.e., a single upper control limit is proposed for the multivariate control chart and the control

limit was obtained empirically. An advantage of the empirically obtained control limit is that it does not rely on large sample properties, that is, it is also accurate for small sample sizes especially if the sample size is one. Results show that for a single sample size the proposed control chart is very robust to changes in the distributional assumption of the correlated nonconformities.

A DEMERIT CONTROL CHART USING AN EDGEWORTH EXPANSION

ABSTRACT

In general, the exact distribution cannot be obtained for the statistic of a demerit control chart. A methodology is proposed that uses a technique called an Edgeworth Expansion to approximate the distribution of a demerit control chart statistic and hence to compute the control limits. A simulation study is performed to test the accuracy of the proposed technique. The results show that the proposed methodology yields more accurate results than the customary normal distribution approximation for small to large samples. For very large samples, the proposed technique performs equivalently with the normal distribution approximation. The propose methodology is more precise in estimating properties of the control chart. The results also show that the demerit control limits constructed using the proposed methodology have a probability of a false alarm very close to the desired for all sample sizes and an actual performance of the control chart is very closed to the advertised.

INTRODUCTION

Consider a process where the quality of the finished product is determined by both the number and type of nonconformities (*NCs*) that occur on the product. For example in a wire mesh weaving process there are different types of *NCs*, e.g., large gaps between adjacent wires or a broken wire in the mesh. The quality of a roll of finished wire mesh is determined by counting the number of *NCs* found for each type. The product is considered defective if the number of *NCs* in a sample exceeds the desired maximums set for each *NC* type. Suppose several product varieties of wire mesh need to be monitored using control charts. One alternative is to use a separate control chart for each *NC* type and for each product type. In this alternative, the number of control charts needed is potentially very large. For example, for five *NC* types and four product types, this scenario would need 20 separate control charts. A more reasonable alternative might be to use a demerit control chart that can simultaneously monitor multiple *NCs*. The number of control charts would be reduced to one chart per product variety. In the previous example, this would be four demerit control charts.

The demerit control chart was developed by Dodge and Torrey (1977). They defined the number of demerits as a weighted sum of the number of *NCs* of different types. The *NC*s were weighted by their importance or severity. They based the control chart on the assumption that the number of demerits is approximately normally distributed and suggested 3-sigma control limits. Jones et al. (1999) observed that if the normality assumption is violated, the performance of the demerit control chart is

severely compromised, i.e., the probability of a false alarm is considerably larger. In response, Jones et al. (1999) developed a methodology to find the exact distribution of the number of demerits using a numerical method. A requirement of the methodology is that the weights in the weighted sum are restricted to a specific ratio. It follows that if the weights in the demerit function cannot be chosen in the specified ratio then the exact distribution for the demerit statistic cannot be obtained.

An alternative methodology is proposed to approximate the distribution of the number of demerits that removes the restriction on the weights and that provides, in general, a more accurate estimation than the normal distribution. The proposed methodology uses a distribution approximating technique called an Edgeworth Expansion (*EE*) to approximate the distribution of the number of demerits. An *EE* uses relatively few computations in contrast to the iterative numerical method developed by Jones et al. (1999). Simulation studies are performed and the results show that an *EE* yields a very accurate approximation of the distribution of the demerit statistic and correspondingly accurate control limits. As a result of a more accurate distribution for the demerit statistic, the simulation study shows that the demerit control chart is able to perform considerably better than the demerit control chart developed by Dodge and Torrey (1977). Moreover, an *EE* works well, in our case, even for very small sample sizes whereas, as we would expect, the 3-sigma control limits by Dodge and Torrey (1977) results seriously in-accurate results. Better performance is measured by a lower rate of false alarms. Finally, if it is undesirable to

choose weights for the demerit function that conform to the ratio specified by Jones et al. (1999), then an *EE* seems the only clear alternative.

The rest of the paper is organized as follows. The next section defines the notation and assumptions for the demerit control chart and explains how to determine control limits for the demerit control chart using an *EE*. A numerical example is presented using data from a wire mesh weaving process. The following section explains the details and results of a simulation study used to compare the control limits obtained from the *EE* approximation and the normal approximation with the control limits obtained from the simulation for various sample sizes. The next section gives the results of how the simulation study was used to assess the performance of the proposed *EE* demerit control chart. The performance of the *EE* demerit control chart and the demerit control chart based on the normal approximation was compared to the performance of a control chart obtained from simulation. The last section provides the conclusions.

COMPUTING CONTROL LIMITS USING THE EDGEWORTH EXPANSION TECHNIQUE

Let the random variable (*rv*) X_{ij} represent the count of *NCs* for the i^{th} *NC* type in the j^{th} inspection unit in a random sample. Montgomery (1996) explains a sample in this context as an "area of opportunity" for the occurrence of *NCs* and it is typically a unit of length, area, or volume of product from the process. A sample consists of one or more inspection units, where an inspection unit is chosen for its "operational or data-collection simplicity" while the sample size is "chosen according to statistical

considerations" such as insuring that a sufficient number of *NC*s are observed per sample; see Montgomery (1996). Assume that the types of *NC*s and inspection units are independent and that X_{ij} are distributed *Poisson* with parameter $\lambda_i \ \forall\, j$.

Let ω_i be the weight assigned to the i^{th} *NC* type and then, for each inspection unit, define the number of demerits D_j as $D_j = \sum_{i=1}^{I} \omega_i X_{ij}$ where I is the total number of *NC* types. The mean and variance of D_j are $E\left[D_j\right] = \sum_{i=1}^{I} \omega_i \lambda_i$ and $V\left[D_j\right] = \sum_{i=1}^{I} \omega_i^2 \lambda_i$, respectively. Compute the average number of demerits per inspection unit as $U = \sum_{j=1}^{N} D_j / N$ where N is the number of inspection units per sample. We can write U in terms of X_{ij} as $U = \sum_{i=1}^{I} \sum_{j=1}^{N} \omega_i X_{ij} / N$ and compute the mean and standard deviation of U as

$$\mu_U = E[U] = \sum_{i=1}^{I} \omega_i \lambda_i \text{ and}$$

$$\sigma_U = \left(V[U]\right)^{1/2} = \left(\sum_{i=1}^{I} \omega_i^2 \lambda_i / N\right)^{1/2}. \tag{1}$$

If the distribution of U is known, the control limits for the demerit U are computed as follows: The centerline (*CL*) is determined by letting $CL = \mu_U$. The lower and upper control limits (*LCL*, *UCL*) are determined from the probability distribution of U, i.e., $P[U \le u]$. Select the probability of a false alarm, α, and find values for *LCL* and *UCL* so that sum of the area under the distribution of U to the left

of *LCL* and to the right of *UCL* equals α, i.e.,

$\alpha = P[U \le LCL] + (1 - P[U \le UCL])$. Let the areas under the distribution of *U* that are below and above the control limits be α_1 and α_2, respectively, such that $\alpha_1 + \alpha_2 = \alpha$, in this way α_1 need not equal α_2. Choose *LCL* and *UCL* such that $P[U \le LCL] = \alpha_1$ and $(1 - P[U < UCL]) = \alpha_2$. For the case where α is divided equally, i.e., $\alpha_1 = \alpha_2 = \alpha/2$, the control limits are computed as

$$LCL = \max\{u_1 : P[U \le u_1] \le \alpha/2\} \text{ and}$$
$$UCL = \min\{u_2 : P[U < u_2] \ge 1 - \alpha/2\}. \quad (2)$$

If there is no *LCL* that satisfies eqn. (2), then set *LCL*=0 and in order to preserve the same probability of a false alarm compute the *UCL* as

$UCL = \min\{u : P[U < u] \ge 1 - \alpha\}$.

Jones et al. (1999) proposed a method to obtain the exact cumulative distribution function (*CDF*) of *U*. If the weights are chosen such that every possible pair of weights conforms to the rule that, for all pairs of weights, the larger weight say ω_i is an integer multiple of the smaller weight ω_j, i.e.,

$(\omega_i/\omega_j) \in \{1, 2, ...\}$ $\forall\{i, j : i \ne j \text{ and } \omega_i \ge \omega_j\}$, then the exact *CDF* can be found. However, for a general weighting scheme the exact *CDF* cannot be obtained and therefore an approximation must be used. For "large" samples, the *CDF* is approximated well by a normal distribution but for "small" samples, the normal distribution approximation is inaccurate.

A more accurate approximation for the *CDF* of U can be obtained using an Edgeworth Expansion and the approximate *CDF* can be defined explicitly. An *EE* involves computations of third and fourth order standardized cumulant functions. Details about the *EE* approximation are provided in Appendix. The approximate *CDF* found by applying an *EE* is

$$F_U(u) = P[U \le u] = \Phi(z) - \phi(z)\left[\frac{\rho_3 h_2(z)}{6\sqrt{N}} + \frac{\rho_4 h_3(z)}{24N} + \frac{\rho_3^2 h_5(z)}{72N}\right] \tag{3}$$

where

$$z = \frac{u - \mu_U}{\sigma_U}, \tag{4}$$

and

$$\rho_3 = \sum_{i=1}^{I} \omega_i^3 \lambda_i \Big/ \left(\sum_{i=1}^{I} \omega_i^2 \lambda_i\right)^{3/2},$$
$$\rho_4 = \sum_{i=1}^{I} \omega_i^4 \lambda_i \Big/ \left(\sum_{i=1}^{I} \omega_i^2 \lambda_i\right)^{2}, \tag{5}$$

and where $h_2(z) = (z^2 - 1)$, $h_3(z) = (z^3 - 3z)$, $h_5(z) = (z^5 - 10z^3 + 15z)$. The functions $\Phi(z)$ and $\phi(z)$ are the *CDF* and the probability density function (*PDF*) of a standard normal *rv*, respectively.

If λ_i are unknown, estimates can be obtained from data. Say we have K samples each containing N_k inspection units from an in-control process. An estimate of λ_i can be computed as the grand average number of *NCs* of type i per inspection unit, i.e., as

$$\hat{\lambda}_i = \sum_{j=1}^{N}\sum_{k=1}^{K} x_{ijk} \Big/ \sum_{k=1}^{K} N_k \,. \qquad (6)$$

Notice that for the purposes of estimating λ_i the number of inspection units per sample N_k need not be equal. To estimate μ_U and σ_U, substitute $\hat{\lambda}_i$ from eqn. (6) into eqn. (1). Likewise, to estimate ρ_3 and ρ_4, substitute $\hat{\lambda}_i$ from eqn. (6) into eqn. (5). Finally, using the estimates $\hat{\mu}_U$ and $\hat{\sigma}_U$, values for $F_U(u)$ can be computed from eqn. (3).

Example – Computing control limits for the *EE* demerit control chart

Consider a wire weaving process at the Siam Wire Netting (SWN) factory in Thailand. The wire mesh product TMA-725 is to be monitored for multiple *NCs*. The mesh is produced in rolls where the width and length of each roll varies according to customer requirements, e.g., $1.5m \times 30m$, $1.25m \times 30m$ or, $1.25m \times 40m$. An inspection unit is chosen as one roll but since there are several roll dimensions, the most frequently produced roll size, i.e., $1.25m \times 40m = 50m^2$ is chosen as the inspection unit. The sample size to be statistically monitored is chosen to be $N = 25$ inspection units, which was chosen for the statistical property of having a sufficient number of *NCs* per sample. Five *NC* types are of interest. The names of the *NC*s have been coded as *1, 2, 3, 4 and 5* at the request of SWN.

Over a two-month period, data on the number of *NCs* for 36 samples of presumably acceptable quality TMA-725 were collected. Estimates for λ_i were

computed using eqn. (6). The estimates are $\hat{\lambda}_1 = 0.126$, $\hat{\lambda}_2 = 0.042$, $\hat{\lambda}_3 = 0.094$, $\hat{\lambda}_4 = 0.025$ and $\hat{\lambda}_5 = 0.051$. In the selection of weights, the relative importance of the *NC* types was not the overriding consideration. Since the relative importance of each *NC* type is considered fairly equal, it seems reasonable to choose weights as the reciprocal of the standard deviation, i.e., $\omega_i = 1/\sqrt{\hat{\lambda}_i}$. Choosing weights in this manner has a standardizing effect on the different types of *NCs*. The weights are computed as $\omega_1 = 2.817$, $\omega_2 = 4.879$, $\omega_3 = 3.262$, $\omega_4 = 6.326$ and $\omega_5 = 4.428$. It is apparent that these weights do not have a structure such that, for all pair of weights, the larger weight is an integer multiple of the smaller. Therefore, an *EE* is used to approximate the *CDF* of *U*. From eqns. (1) and (5) we get $\hat{\mu}_U = 1.25$, $\hat{\sigma}_U = 0.44$, $\hat{\rho}_3 = 1.94$ and $\hat{\rho}_4 = 4.07$. SWN is interested in monitoring the average daily production of TMA-725, which is, approximately 25 inspection units. Therefore, a sample consists of approximately $N = 25$ inspection units. From eqn. (3), the *CDF* of *U* is

$$F_U(u) = \Phi(z) - \phi(z)\left[\frac{1.94 * h_2(z)}{30} + \frac{4.07 * h_3(z)}{600} + \frac{3.77 * h_5(z)}{1800}\right] \qquad (7)$$

with $z = \dfrac{(u - 1.25)}{0.44}$. The centerline of the demerit *U* chart is $CL = 1.25$. If we choose $\alpha = 0.0027$ and divide α into two equal parts, i.e., $\alpha_1 = \alpha_2 = 0.00135$, then from eqn. (2), $LCL = \max\{u : F_U(u) \le 0.00135\} = 0.18$ and $UCL = \min\{u : F_U(u) \ge 1 - 0.00135\} = 2.81$.

A SIMULATION STUDY TO COMPARE THE EE AND NORMAL APPROXIMATIONS

Comparison of Tail Probabilities

A simulation was conducted to assess the accuracy of the *EE* approximation for various sample sizes N =5, 10, 15, 20 and 25. For the simulation, an inspection unit of $50m^2$ was used. For instance with N=25, five values of *Poisson* random variates representing *NCs* of each type were generated for each inspection unit using the five estimates $\hat{\lambda}_i$ computed in the example. A total of (5)(25)=125 values of *NCs* were generated for a single sample. For each sample, the average number of demerits per inspection unit (U_{Sim}) was computed using the weights from the example. A total of 250,000 samples were generated and a total of 250,000 computed values of U_{Sim} were used to form the empirical distribution of U represented as $F_U^{Sim}(u)$. Since $F_U^{Sim}(u)$ is based on a very large number of samples, it's very close to an exact distribution of U and hereby assumed.

The *EE* approximation is much more accurate in approximating the *CDF* of U than the normal distribution not only for the specific control limit values but also in the tails of the distribution in general. Simulation results show that the *EE* demerit control chart consistently outperforms the demerit control chart with 3-sigma control limits in estimating the tails probabilities even when N is small. For moderate sample size of N equal 25 and 20, Figures 1 and 2 show that the plot of *CDF* of U calculated using an *EE* from eqn. (5) represented as $F_U^{EE}(u)$ falls very close to the $F_U^{Sim}(u)$. The

CDF using the normal distribution approximation, computed as $\Phi\left((u-1.25)/0.44\right)$ and denoted by $F_U^{Normal}(u)$, consistently over-estimates the values of $F_U^{Sim}(u)$ in both tails. These approximate *CDFs* of U were plotted against the standardized values of U, i.e., $(u-\hat{\mu}_U)/\hat{\sigma}_U$, for values ranging from (-3.00, -1.00), (1.00, 3.00) in Figure 1 and 2 respectively. The $F_U^{EE}(u)$ is more accurate in both extreme tails than the $F_U^{Normal}(u)$. For typical values of α, e.g., $\alpha=0.0027$, the control limits will lie in both extreme tails, therefore the accuracy in approximating the *CDF* of U in the extreme tails is very important in determining the control limits.

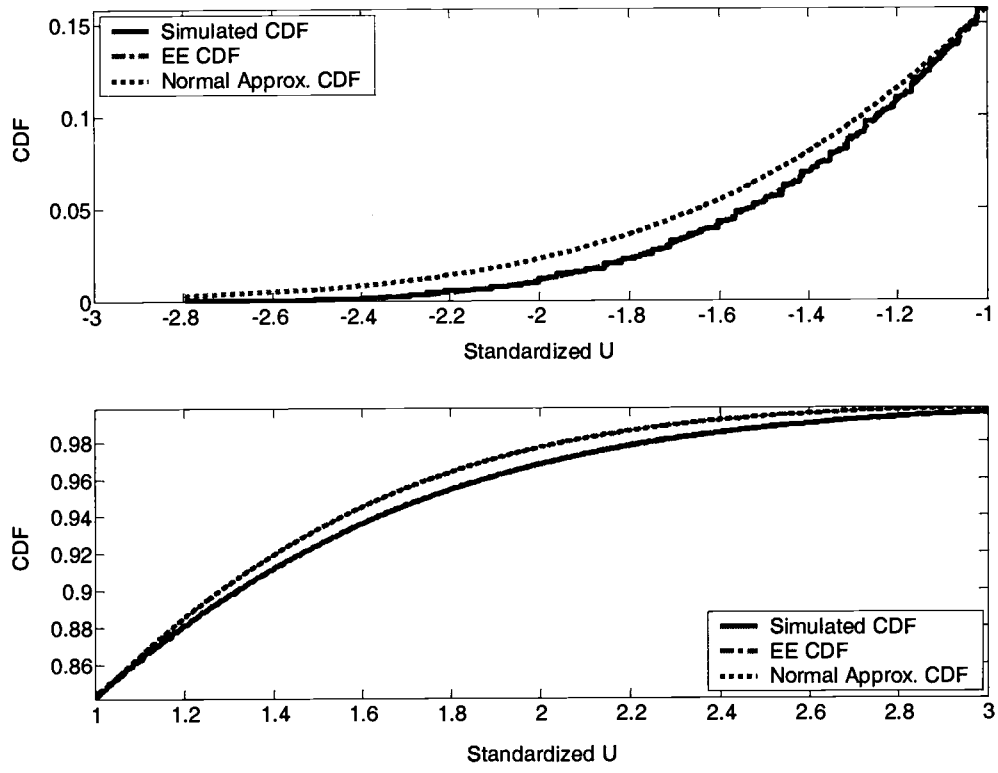

Figure 1: Plot of $F_U^{Sim}(u)$, $F_U^{EE}(u)$, and $F_U^{Normal}(u)$ vs. Standardized Values of U for N=25.

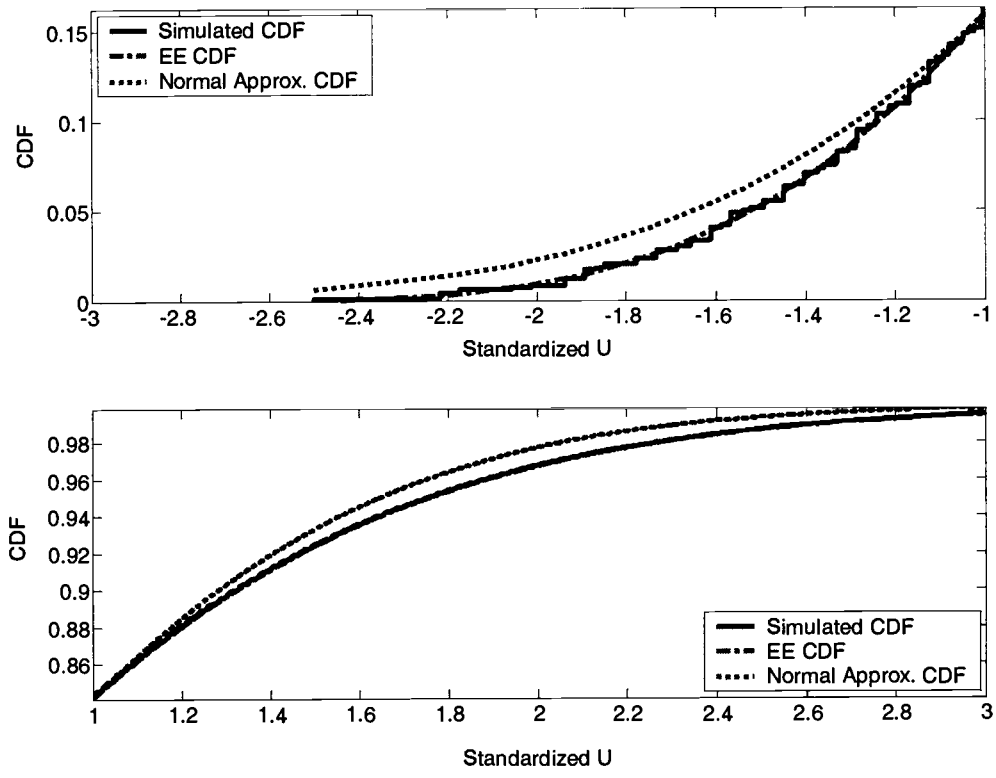

Figure 2: Plot of $F_U^{Sim}(u)$, $F_U^{EE}(u)$, and $F_U^{Normal}(u)$ vs. Standardized Values of U for N=20.

For moderate to small N, the $F_U^{EE}(u)$ is more accurate in approximating the tails probabilities, especially in the upper tail, than $F_U^{Normal}(u)$ even for small sample size such as N=5. The normal distribution approximation consistently over-estimates the upper tail probabilities. When N get smaller, the distribution of U is highly right skew. The minimum standardized value of U varies according to the sample sizes of 15, 10 and 5 are -2.16, -1.76, and -1.25 respectively. So, the approximate *CDFs* of U were plotted against the standardized values of U ranging from the minimum values to

zero instead and (1.00, 3.00) in figure 3 to 5. If a one-sided demerit control chart with only *UCL* is adopted, it's apparent that an *EE* will yield very precise *UCL* than the normal distribution approximation even for small sample size. These results indicate that the *EE* demerit control chart works well for both small and moderate sample size.

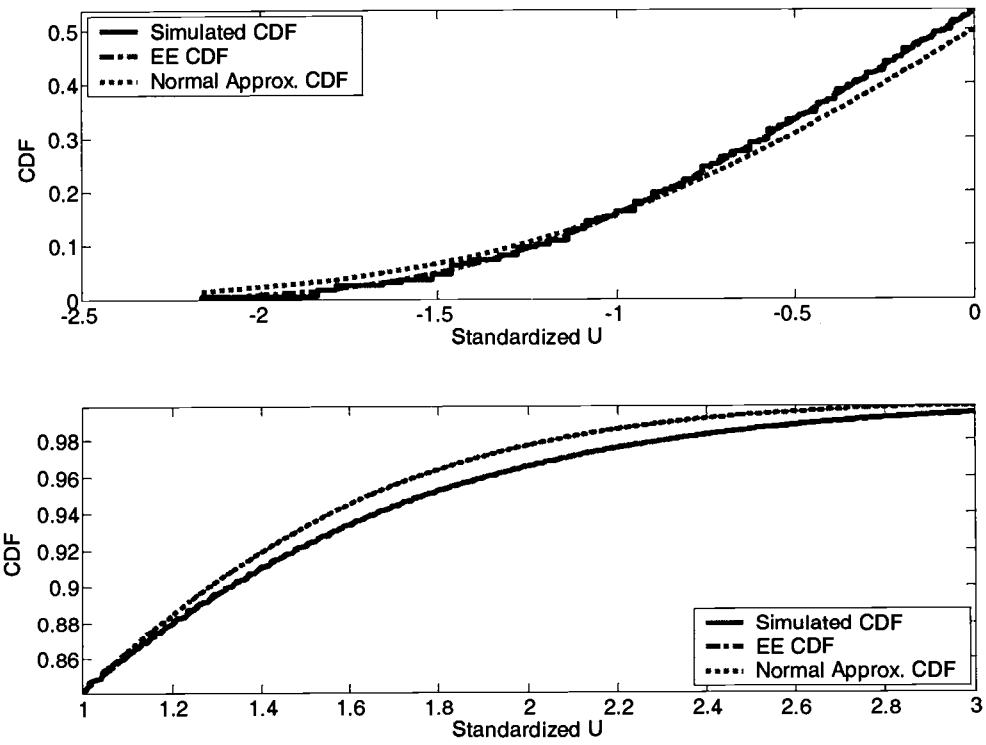

Figure 3: Plot of $F_U^{Sim}(u)$, $F_U^{EE}(u)$, and $F_U^{Normal}(u)$ vs Standardized Values of *U* for *N*=15.

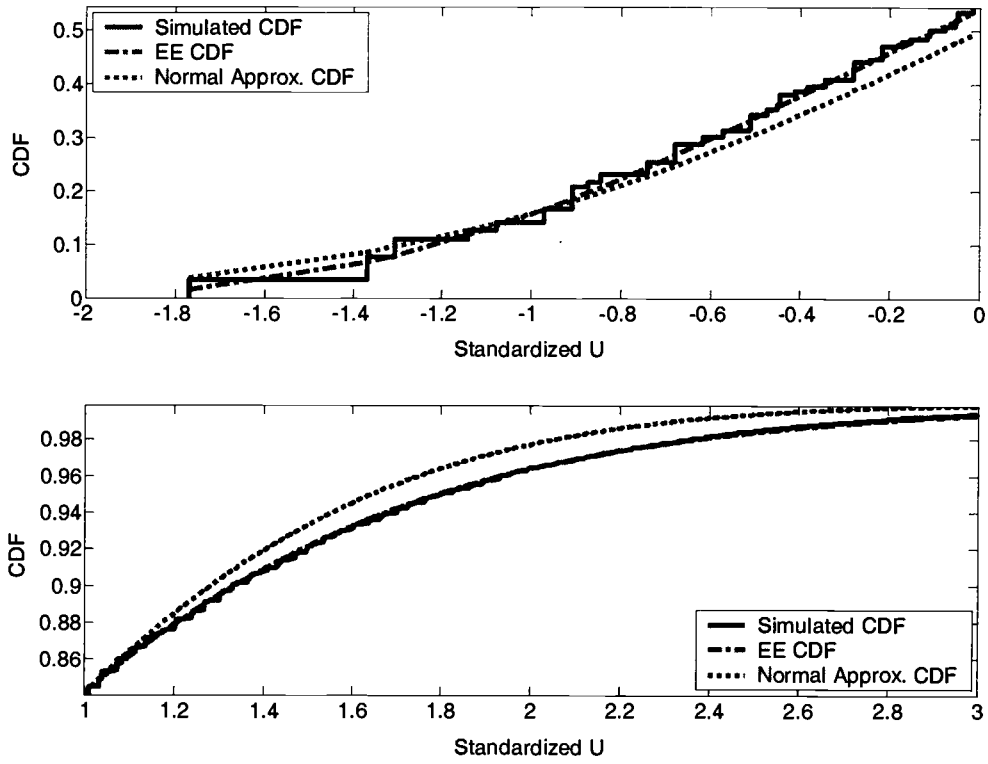

Figure 4: Plot of $F_U^{Sim}(u)$, $F_U^{EE}(u)$, and $F_U^{Normal}(u)$ vs Standardized Values of U for N=10.

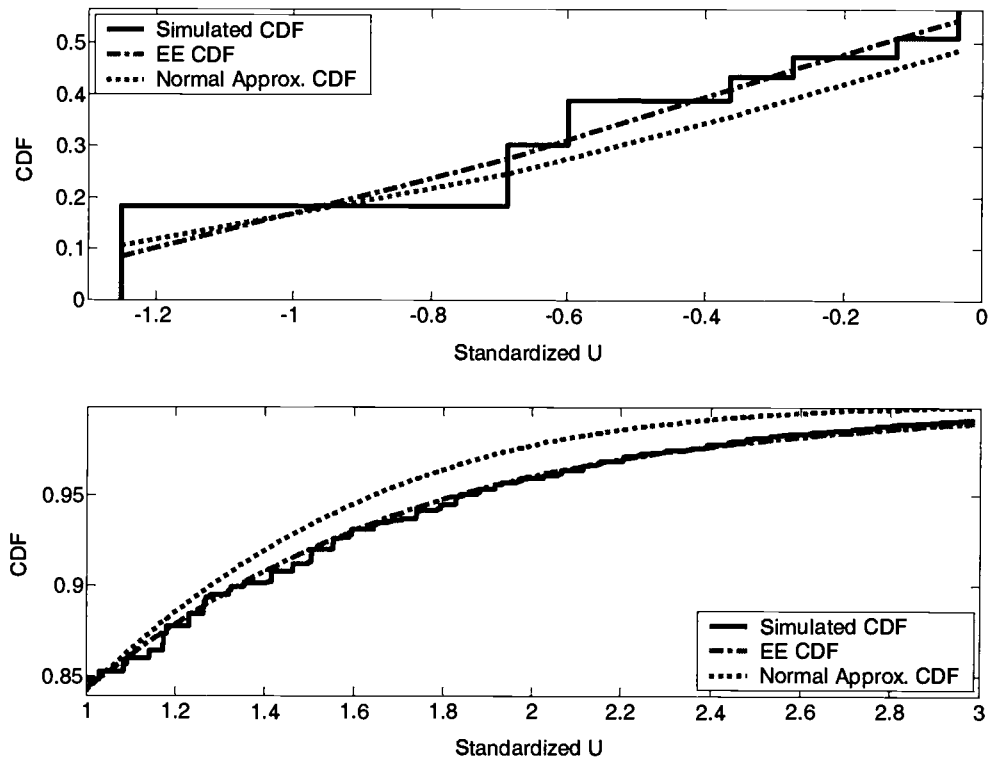

Figure 5: Plot of $F_U^{Sim}(u)$, $F_U^{EE}(u)$, and $F_U^{Normal}(u)$ vs Standardized Values of U for N=5.

Comparing control limits of the EE and the normal Approximations

The *EE* approximation is much more accurate in calculating the control limits than the normal distribution approximation. For comparison, Table 1 shows values for control limits obtained using $F_U^{Sim}(u)$, $F_U^{EE}(u)$, and $F_U^{Normal}(u)$ for N equal 5, 10, 15, 20 and 25. The empirical lower and upper control limits, denoted by LCL_{Sim} and UCL_{Sim}, were found by selecting the lower and upper percentiles from $F_U^{Sim}(u)$ that

correspond to the chosen probability of false alarm α. In the example, α was divided evenly between the lower and upper limits therefore the LCL_{Sim} is chosen as the largest number satisfying $F_U^{Sim}(u) = P[U_{Sim} \le u] \le \alpha/2$ and the UCL_{Sim} is chosen as the smallest number satisfying $F_U^{Sim}(u) = P[U_{Sim} < u] \ge 1-\alpha/2$. In other words, find the largest value U so that at most $(\alpha/2)\times 100\%$ of the U_{Sim} values are below LCL_{Sim}. Likewise, find the smallest value of U_{Sim} such that at least $(1-\alpha/2)\times 100\%$ of the U_{Sim} values are below UCL_{Sim}. If there is no LCL that satisfies the condition $F_U^{Sim}(u) = P[U_{Sim} \le u] \le \alpha/2$, then set LCL_{Sim}=0 and find UCL using α instead of $\alpha/2$. For the normal distribution 3-sigma limits were assumed, i.e., the limits were computed as $UCL_{Norm} = \hat{\mu}_U + 3\hat{\sigma}_U$ and $LCL_{Norm} = \max\{0, \hat{\mu}_U - 3\hat{\sigma}_U\}$. For each N, the values of LCL_{Sim}, UCL_{Sim}, LCL_{Norm}, UCL_{Norm} and the EE control limits were calculated. The absolute deviations of an EE and 3-sigma limits from the simulated ($|LCL - LCL_{Sim}|$, $|UCL - UCL_{Sim}|$) for all N were summarized in Table 2. When sample size is small to moderately small, i.e., between 5 and 15, there is no LCL that satisfies eqn. (2), then LCLs were set equal to 0 and UCLs were computed using $\alpha = 0.0027$.

Table 2 shows that the absolute error of the control limits for an EE, $|LCL - LCL_{Sim}|$ and $|UCL - UCL_{Sim}|$, were comparatively and consistently smaller than that of the 3-sigma limits for all sample sizes. For instance, with moderate N=25 where we expect the normal approximation to work reasonably well, the EE limits is

still much more accurate than the 3-sigma limits. For moderate sample size of 25, we expect the normal distribution approximation to work well and the control limits should be close to the actual. Even so, the absolute error for the normal distribution approximation is much larger with values 0.13 and 0.20 for the lower and upper limits respectively while the absolute error of the control limits for an *EE* is merely 0.05 for the lower limit and 0.02 for the upper limit. The absolute errors of the normal distribution approximation are consistently and approximately about 10 times larger than an *EE* at the upper limit. This affirms that the *EE* limits are uniformly more accurate than the 3-sigma limits. The more accurate the control limits, the more precise the performances of the demerit control chart. Therefore, the demerit control chart with the control limits computed using an *EE*, called the *EE* demerit control chart, has more precise performance than the demerit control chart using normal distribution approximation.

Table 1: Comparison of Control Limits.

N	Simulation		*EE* Approx.		Normal Dist. Approx. (3σ limits)	
	LCL	**UCL**	**LCL**	**UCL**	**LCL**	**UCL**
5	0	4.90	0	4.92	0	4.25
10	0	3.66	0	3.68	0	3.37
15	0	3.16	0	3.17	0	2.98
20	0.14	3.01	0.09	3.02	0	2.75
25	0.13	2.79	0.18	2.81	0	2.59

Table 2: Comparison of Absolute Errors of Control Limits.

N	*EE* Approx.		Normal Dist. Approx. (3σ limits)	
	$\lvert LCL - LCL_{Sim} \rvert$	$\lvert UCL - UCL_{Sim} \rvert$	$\lvert LCL - LCL_{Sim} \rvert$	$\lvert UCL - UCL_{Sim} \rvert$
5	0	0.02	0	0.65
10	0	0.02	0	0.29
15	0	0.01	0	0.18
20	0.05	0.01	0.14	0.26
25	0.05	0.02	0.13	0.20

ASSESSING THE PERFORMANCE OF THE EE DEMERIT CONTROL CHART

The performance of the *EE* demerit control chart can be determined by comparing it to the demerit control charts created from the simulation and the demerit control chart using normal distribution approximation. To compare the charts, the average run length (*ARL*) is used. The average run length for an in-control process (ARL_0) is related to the probability of a false alarm α by the relationship $ARL_0 = 1/\alpha$.

To assess ARL_0, an additional 250,000 samples of size 25 were randomly generated using the estimates $\hat{\lambda}_i$ and weights from the example and 250,000 values of U_{Sim} were computed for each *N*. The desired false alarm probability chosen for the *EE* demerit control chart in the example and throughout the simulation studies is $\alpha = 0.0027$ so that the extreme accuracy of the approximations in the tails can be investigated. Depending on the accuracy of the *EE* approximation, the actual *EE* false

alarm probability (α_{EE}) could be different. Assume that U_{Sim} values are representative of values for U that would be obtained from an in-control wire mesh weaving process. Then, the proportion of U_{Sim} that are outside the *EE* control limits in Table1 is an accurate estimate of the α_{EE} that we would experience when using the *EE* demerit control chart. In a similar fashion, the actual false alarm probability of the demerit control chart using the normal distribution approximation (α_{Norm}) can be determined in similar fashion. The corresponding *ARLs* were computed as $ARL_0^{Desired} = 1/\alpha$, $ARL_0^{EE} = 1/\alpha_{EE}$ and, $ARL_0^{Norm} = 1/\alpha_{Norm}$. Table 3 shows values of the ARL_0^{EE}, and ARL_0^{Norm} and the differences from $ARL_0^{Desired}$ =370.37 denoted as $|\Delta ARL_0|$, for each approximation.

The *EE* demerit control chart produces the actual false alarm rate that is closer to the desired than the demerit control chart using normal distribution approximation. From Table 3, the ARL_0^{EE} is different from $ARL_0^{Desired}$ about 9.40%, 14.60%, 4.01%, 15.19% and 12.90% for *N*=5, 10, 15, 20 and 25 respectively whereas the ARL_0^{Norm} is different from the desired about 69.40%, 56.17%, 47.43%, 52.33% and 37.79%. The large difference from $ARL_0^{Desired}$ of the demerit control chart using normal distribution approximation occurs such that the demerit control chart will falsely signal an out-of-control condition more frequent than the advertised.

Table 3: Comparison of ARL_0 and $|\Delta ARL_0|$.

N	EE Approx.		Normal Dist. Approx. (3σ limits)	
	ARL_0^{EE}	$\|\Delta ARL_0\|$	ARL_0^{Norm}	$\|\Delta ARL_0\|$
5	405.19	34.82	113.33	257.04
10	424.45	54.08	162.34	208.03
15	385.21	14.84	194.70	175.67
20	426.62	56.25	176.55	193.82
25	322.58	47.79	230.41	139.96

CONCLUSION AND DISCUSSION

Based on the simulations, an *EE* is more accurate than the normal distribution approximation for all sample sizes. The $F_U^{EE}(u)$ is closer to $F_U^{Sim}(u)$ than $F_U^{Normal}(u)$ especially in the extreme tails. The normal distribution approximation tends to over estimate the *CDF* of *U* than an *EE* does in both tails. Hence, the *EE* control limits are closer to the exact than those obtained using the normal distribution approximation.

The *EE* demerit control chart has actual performances in terms of *ARLs* very close to the advertised whereas the demerit control chart using normal distribution approximation has actual *ARLs* significantly departing from the advertised. The *EE* demerit control chart produces the actual α that is closer to the desired than the demerit control chart with 3-sigma control limits.

Consequently, an *EE* improves the performance of the demerit control chart. The normal approximation to the distribution of *U* is likely to results in *LCL* being

zero whereas an *EE* does not. With the downward shifts, the 3-sigma control limits limit will has very small probability of detecting the shift even for the large downward shift. The proposed *EE* demerit control chart has more detection area for the downward shift and thus more efficient.

The *EE* approximation yields reasonable accurate result in determining the performance of the control chart when the exact distribution is not obtainable. As in the example, we can see that it is very common and natural to come up with a weighting scheme in which the technique provided by the Jones et al. (1999) technique is not applicable. When the exact distribution is not obtainable, determining the performance of the demerit control chart must rely on some approximation. Even if it does, evaluating the performance of the demerit control chart using exact distribution requires an extensive tabulation among possible $\boldsymbol{\lambda}^*$, μ_U^*. An *EE* requires a simple numerical computation to obtain the control limits comparing to the technique by Jones et al. (1999). Therefore, the proposed demerit control chart using *EE* technique is recommended because of its simplicity.

Lastly an *EE* would be a fast, easy and very effective tool in determining the best set of $\boldsymbol{\omega}$. To choose the best weight set, one might have to determine the performance of the demerit control chart across not only $\boldsymbol{\lambda}$ but also $\boldsymbol{\omega}$. Even if we restrict ourselves to a class of weight such that the specific requirement is satisfied, the technique by Jones et al. (1999) could result in an exponentially intensive tabulation for the *CDF* of U. An *EE* would require comparatively small fraction of time to do so. Besides, this proposed technique allows us to search for weights in a wider class

with less effort. If one wants to search for the best weights regarding to some criteria, the *EE* approximation presented in this paper will definitely better suite this type of problem and therefore is recommended.

APPENDIX: DEVELOPMENT OF THE EDGEWORTH EXPANSION

The cumulant generating function (*CGF*) of a *rv* Y is defined as $\log E(e^{tY})$ provided that the expectation of e^{tY}, denoted as $E(e^{tY})$, exist for some t; see Stuart and Ord (1994). For example, the *CGF* of a Poisson X_{ij} is $K_{X_{ij}}(t) = \lambda_i \exp(t-1)$. The *CGF* of D_j can be derived as

$$\log E(e^{tD_j}) = \log\left[E\left(\exp\left(t * \sum_{i=1}^{I} \omega_i X_{ij}\right)\right)\right]$$

$$= \log\left[\prod_{i=1}^{I} E\left(\exp\left((t\omega_i)X_{ij}\right)\right)\right]$$

$$= \sum_{i=1}^{I} \log\left[E\left(\exp\left((t\omega_i)X_{ij}\right)\right)\right]$$

$$= \sum_{i=1}^{I} K_{X_{ij}}(t\omega_i)$$

$$= \sum_{i=1}^{I} \lambda_i \exp(t\omega_i - 1)$$

Without loss of generality, assume that a sample consists of N inspection units. The average number of demerits per inspection unit, $U = \sum_{j=1}^{N} D_j / N$ can also be viewed as the sample mean of independent and identically distributed (*iid*) *rvs*, D_j. The standardized *rv* $G = \dfrac{U - \mu_U}{\sigma_U}$ has a mean zero and a unit variance. The cumulant

generating function of G is $K_G(t) = \log\left[E\left(\exp\left(t(U-\mu_U)/\sigma_U\right)\right)\right]$ which is equal to

$\log\left[E(\exp(tU/\sigma_U))\right] - \left[(t\mu_U)/\sigma_U\right]$. Note that

$\log\left[E(\exp(tU/\sigma_U))\right] = \log\left[E\left(\exp\left[t*\left(\sum_{j=1}^{N} D_j / N\right)\Big/\sigma_U\right]\right)\right]$. Since D_j is assumed

independent, the expectation of the product is the product of the expectation, i.e.,

$E\left(\exp\left[t*\left(\sum_{j=1}^{N} D_j / N\right)\Big/\sigma_U\right]\right) = \prod_{j=1}^{N} E\left(\exp\left[\left(t/(N\sigma_U)\right)*D_j\right]\right)$. Then the *CGF* of G is

$$K_G(t) = N*\log\left[E\left(\exp\left(tD_j/(N\sigma_U)\right)\right)\right] - \left[(t\mu_U)/\sigma_U\right]$$

$$= N*K_{D_j}\left(\frac{t}{N\sigma_U}\right) - \left[(t\mu_U)/\sigma_U\right] \qquad \text{(A1)}$$

Using results from Kolassa (1997), it can be shown that an *EE* to the *CDF* of G is

$$F_G(z) \approx \Phi(z) - \phi(z)\left[\frac{\rho_3 h_2(z)}{6\sqrt{N}} + \frac{\rho_4 h_3(z)}{24N} + \frac{\rho_3^2 h_5(z)}{72N}\right] \qquad \text{(A2)}$$

where $\rho_3 = \dfrac{\omega_1^3\lambda_1 + \omega_2^3\lambda_2 + \ldots + \omega_I^3\lambda_I}{\left(\omega_1^2\lambda_1 + \omega_2^2\lambda_2 + \ldots + \omega_I^2\lambda_I\right)^{3/2}}$ and $\rho_4 = \dfrac{\omega_1^4\lambda_1 + \omega_2^4\lambda_2 + \ldots + \omega_I^4\lambda_I}{\left(\omega_1^2\lambda_1 + \omega_2^2\lambda_2 + \ldots + \omega_I^2\lambda_I\right)^2}$ are the

standardized third and fourth cumulant respectively, $h_2(z) = (z^2 - 1)$, $h_3(z) = (z^3 - 3z)$,

$h_5(z) = (z^5 - 10z^3 + 15z)$, $\Phi(z)$, $\phi(z)$ are *CDF*, *PDF* of $N(0,1)$. See Kolassa (1997)

for a theoretical discussion on the Edgeworth Expansion (*EE*) and its properties.

Thus, an *EE* for the *CDF* of U is $F_U(u) = P[U \le u] = P\left[G \le \frac{u-\mu_U}{\sigma_U}\right] \approx F_G\left(\frac{u-\mu_U}{\sigma_U}\right)$

where F_Z was defined as in eqn. (A2).

Kolassa (1997) mention that the error of an *EE* in approximating the discrete distribution can be claimed to be uniformly as $O(1/\sqrt{N})$ according to a jump of size $O(1/\sqrt{N})$ at the discontinuity points of any discrete distribution. See Vaart (1998) for the definition of "big" and "little O" notation. Since X_{ij} takes on discrete values, the *CDF* of U is a step function with a discontinuity at each discrete value of U. The size of the steps decreases as N increases. There is no approximation for the *CDF* of a discrete *rv* that will out-govern the rate at which the size of the steps decrease at the discontinuity points as N increases, i.e., no approximation has an error rate less than the decreasing rate of the step size. However, the accuracy of the approximation to the *CDF* of a discrete *rv* can be improved at non-discontinuity points and an *EE* uniformly yields at least as good or better of an approximation than the normal distribution for all N; see Kolassa (1997).

REFERENCES

Dodge, Harold F. and Torrey, M. N. (1977), "A check inspection and demerit rating plan (Reprint from Industrial Quality Control)", Journal of Quality Technology, 9, 146-153.

Jones, Allison L., Woodall, William H., and Conerly, Michael D. (1999), "Exact properties of demerit control charts", Journal of Quality Technology, 31, 207-216.

Kolassa, John Edward (1997), "Series approximation methods in statistics", Springer-Verlag Inc., New York.

Montgomery, Douglas. C. (1996), "Introduction to Statistical Quality Control", 3rd edition, John Wiley & Sons, New York.

Stuart, Alan , and Ord, J. Keith (1994), "Kendall's Advanced Theory of Statistics. Volume 1. Distribution theory (Sixth edition)", Edward Arnold Publishers Ltd., London.

Vaart, van der A. W. (1998), "Asymptotic statistics", Cambridge University Press, Cambridge, UK.

AN ASYMPTOTIC TECHNIQUE FOR DESIGNING A DEMERIT WEIGHTING SCHEME FOR A ONE-SIDED DEMERIT CONTROL CHART

ABSTRACT

A methodology to design the demerit weights for a demerit control chart with general classification scheme is proposed. The proposed method assigns the optimal weights to the number of demerits such that the demerit control chart has the maximum probability of detecting an increase in the mean numbers of nonconformities. The optimal weights depend on the directions of the shifts in the mean numbers of nonconformities rather than the magnitudes of the shifts. The proposed method assigns more relative weight to more severe types of nonconformities. The solution is explicit and is easily computed.

INTRODUCTION

Consider a process where the quality of the finished product is determined by the number and type of nonconformities (*NC*) that occur on the product. The product is considered "defective" only if the numbers of *NCs* for one or more *NC* types exceeds predefined limits. Dodge and Torrey (1977) developed a demerit control chart to monitor such a process. They classified the *NCs* into four types, i.e., very serious, serious, moderately serious, and minor and used fixed values for weights for those four *NC* types. This method of classifying and weighting the types of *NCs* has since been adopted as standard practice for demerit control charts; see Montgomery

(1996). The performance of the demerit control chart depends on the weights given to the *NCs*. The method of how to select weights that give the best performance for the demerit control chart presented by Dodge and Torrey (1977) and then other classification schemes has not yet been addressed in the literature.

A methodology of selecting values of weights for a one-sided demerit control chart (*OSDCC*) using a method called an asymptotic technique is proposed. The asymptotic technique does not restrict the nature of the types of *NCs* and can be applied not only to the classification scheme suggested by Dodge and Torrey (1977) but also to a general scheme. The asymptotic technique chooses weights for the demerit control chart that maximize the probability of detecting an out-of-control condition. The asymptotic technique has many useful and desirable properties. For example, the asymptotic technique provides an optimal and explicit solution for the weights and the solution does not depend on the absolute magnitude of a shift or shifts in one or more of the *NC* types but rather depends on the directions. The asymptotic technique always assigns more weight to the more severe type of *NC*.

The rest of the paper is organized as follows. Notation and assumptions and the procedure for determining a control limit for the *OSDCC* is presented in the first section. The following section gives the development of the asymptotic technique and the procedure for obtaining the weights. This section also presents several properties of the asymptotic technique. In the next section, a numerical example from an actual wire mesh weaving process is given. The last section is the conclusion.

DEVELOPMENT OF THE ASYMPTOTIC TECHNIQUE FOR OSDCC

Consider a *OSDCC* that is used to simultaneously monitor an increase in *NCs* of several types. For example in a wire mesh weaving process there are different types of *NCs*, e.g., large gaps between adjacent wires or a broken wire in the mesh. The demerit control chart is constructed as follows. First, determine an inspection unit, e.g., in wire mesh weaving we let an inspection unit equal fifty square feet. At equal time intervals, take a sample consisting of N inspection units, e.g., a finished roll of wire mesh with a total area of 500 square feet is inspected and the number of *NCs* is counted. Let the random variable (*rv*) X_{ij} represent the count of *NCs* for the i^{th} *NC* type in the j^{th} inspection unit in the random sample. Assume that the types of *NCs* and inspection units are independent and that X_{ij} are distributed *Poisson* with parameter λ_i $\forall\, j$. Let ω_i be the weight assigned to the i^{th} *NC* type. For each inspection unit, define the demerit value, D_j, as a weighted sum of the number of *NCs* of different types, i.e., let $D_j = \sum_{i=1}^{I} \omega_i X_{ij}$ where I is the number of *NC* types. The mean and variance of D_j are $E\left[D_j\right] = \sum_{i=1}^{I} \omega_i \lambda_i$ and $V\left[D_j\right] = \sum_{i=1}^{I} \omega_i^2 \lambda_i$, respectively. Next, compute an average demerit value per inspection unit as $U = \sum_{j=1}^{N} D_j / N$ where N is the number of inspection units per sample. The mean and standard deviation of U are $\mu_U = \sum_{i=1}^{I} \omega_i \lambda_i$ and $\sigma_U = \left(\sum_{i=1}^{I} \omega_i^2 \lambda_i / N \right)^{1/2}$ respectively.

The control limits for the *OSDCC* are determined by letting the centerline equal μ_U. The upper control limit (*UCL*) is determined from the probability distribution of U, i.e., $P[U \le u]$. Select the probability of a false alarm, α, and find values for the *UCL* so that the area under the distribution of U to the right of the *UCL* equals α, i.e., $\alpha = P[U \ge UCL]$. For instance, if the normal distribution is used to approximate the distribution of U, the *UCL* becomes $UCL = \mu_U + z_\alpha \sigma_U$ where z_α is the $100(1-\alpha)^{th}$ percentile of the standard normal distribution.

Generally, the properties and performance of the *OSDCC* depend heavily on the weights, $\boldsymbol{\omega} = [\omega_1, ..., \omega_I]$ in the demerit function $D_j = \sum_{i=1}^{I} \omega_i X_{ij}$, therefore it is reasonable to select weights that will optimize the performance of the *OSDCC*. Performance is measured as a trade-off between the probability of a false alarm α and the probability of detecting an out-of-control condition $(1-\beta)$. Assume that the probability of a false alarm is selected and fixed, then the performance of the *OSDCC* is measured in terms of $(1-\beta)$. The goal is then to find the weights, $\boldsymbol{\omega}$, that maximize the probability of detection, $(1-\beta)$.

A change in μ_U can occur from a change in the mean of one or more *NC* types, i.e., a change in any $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, ..., \lambda_I]$. Without loss of generality, represent the shifts by a vector $\boldsymbol{\delta} = [\delta_1, \delta_2, ..., \delta_I]$. Unfortunately, there does not exist a unique solution for $\boldsymbol{\omega}$ such that the *OSDCC* will have the maximum probability of detecting

all possible shift vectors $\boldsymbol{\delta}$. For instance, the weight $\boldsymbol{\omega} = [1, 0, ..., 0]$ results in the maximum probability of detection for $\boldsymbol{\delta}$ of the form $[\delta_1, 0, ..., 0]$ whereas the weight $\boldsymbol{\omega} = [0, 1, ..., 0]$ result in that if $\boldsymbol{\delta}$ is of the form $[0, \delta_2, ..., 0]$.

An asymptotic technique presented here determines the weights for detecting a special class of shifts for the *OSDCC*. First, the probability of detecting a shift certainly depends on $P[U \leq u]$ which in turn depends on I and $\boldsymbol{\lambda}$. The exact distribution of U cannot be obtained for all sets of $\boldsymbol{\omega}$ but can be approximated well by the normal distribution if N is large enough. Jones et al. (1999) developed a numerical method to obtain the exact distribution for $P[U \leq u]$ but it requires that, $\boldsymbol{\omega}$ have specific ratios, element-wise, which does not allow for a comprehensive selection of $\boldsymbol{\omega}$. The proposed asymptotic technique determines weights by using the approximate distribution of U when N is large. Define the limiting probability as the approximate probability $P[U \leq u]$ in the limit as $N \to \infty$.

The limiting probability for an arbitrary shift will approach one, regardless of its magnitude; see Appendix 1 for verification. So, it is not informative to consider just any arbitrary shift. To avoid this difficulty, consider a class of $\boldsymbol{\delta}$ such that $\lim_{N \to \infty} \boldsymbol{\delta} = 0$ element wise. Let $\boldsymbol{\delta}$ be represented in the form $\boldsymbol{\delta} = \left[k_1/\sqrt{N}, ..., k_I/\sqrt{N}\right]$ where $\mathbf{k} = [k_1, ..., k_I]$ are fixed constants. When N is large, the probability of detection converges to a constant less than one; see Appendix 2 for details. Therefore, the asymptotic technique takes advantage of the limiting probability of detection and

determines a set of weights, assuming N is reasonably large. The derivation of the asymptotic technique invokes the properties of the Central Limit Theorem. The asymptotic technique finds the set of $\boldsymbol{\omega}$ that maximizes the limiting probability of detecting the shift of type $\mathbf{k}/\sqrt{N}$. The maximization problem is represented as

$$\underset{\boldsymbol{\omega}}{Max} \sum_{i=1}^{I} \omega_i k_i \Big/ \sqrt{\sum_{i=1}^{I} \omega_i^2 \lambda_i} \ , \tag{1}$$

or equivalently

$$\underset{\mathbf{r}>0:\sum r_i=1}{Max} \sum_{i=1}^{I} \left(k_i/\sqrt{\lambda_i}\right)\sqrt{r_i} \tag{2}$$

where

$$r_i = \omega_i^2 \lambda_i \Big/ \left(\omega_1^2 \lambda_1 + \omega_2^2 \lambda_2 + ... + \omega_I^2 \lambda_I\right). \tag{3}$$

The interpretation of r_i is as the proportion or contribution of the variance of the i^{th} *NC* type to the total variance of D_j; see Appendix 3 for the derivation of eqn. (2). Let the optimal solution to eqn. (2) be denoted $\mathbf{r}^* = (r_1^*, ..., r_I^*)$.

In what follows, some properties of the asymptotic technique are given.

Property 1: The solution to eqn. (2) depends only on $\left(k_1/\sqrt{\lambda_1}, ..., k_I/\sqrt{\lambda_I}\right)$ and has both a unique and explicit optimal solution $\mathbf{r}^* = \left[r_1^*, r_2^*, \ldots, r_I^*\right]$. The solution is

$$r_i^* = \left(k_i/\sqrt{\lambda_i}\right)^2 \Big/ \left(\sum_{i=1}^{I}\left(k_i/\sqrt{\lambda_i}\right)^2\right) \text{ for } i = 1,...,I; \tag{4}$$

see Appendix 4 for the proof.

Property 2: For a given solution $\mathbf{r}^*$ the solution for $\boldsymbol{\omega}$ is determined uniquely up to a scalar multiple of

$$\boldsymbol{\omega} = (k_1/\lambda_1,.....,k_I/\lambda_I); \tag{5}$$

see Appendix 5 for the proof.

Property 3: The asymptotic technique computes larger relative values for r_i for larger values of $k_i/\sqrt{\lambda_i}$. If the values $k_i/\sqrt{\lambda_i}$ are all equal to a positive constant, viz., if $k_i/\sqrt{\lambda_i} = C$ where $C > 0 \ \forall i$ then $r_i = 1/I \ \ \forall i$; see Appendix 6 for the proof.

Property 4: The asymptotic technique assigns non-zero weights only to *NC* types with non-zero $k_i/\sqrt{\lambda_i}$; see Appendix 7 for the proof.

COMPUTING THE WEIGHTS FOR A OSDCC

The asymptotic technique obtains values for $\mathbf{r}^*$ and $\boldsymbol{\omega}$ from the values $\left(k_1/\sqrt{\lambda_1},.....,k_I/\sqrt{\lambda_I}\right)$ and $(k_1/\lambda_1,.....,k_I/\lambda_I)$ which are simply the shift size weighted by the standard deviation and variance of *NC*, viz., the *NC* type with a larger variance receives a lower weight for the same magnitude shift. Values for $[\lambda_1,\lambda_2,\ldots,\lambda_I]$ are either known or can be estimated using data from an in-control process. The values

$[k_1, k_2, \ldots, k_I]$ are the direction of the shifts that we are interested in detecting. The simplest method for specifying $\mathbf{k}$ is to let $k_1 = k_2 = \cdots = k_I = C$ where C is a positive constant and then the *NC* type with the smallest variance will have the largest comparative loss and correspondingly will have the highest relative weight. Assigning weights reciprocally to variance is a common method in many statistical techniques, e.g., weighted least squares regression.

The asymptotic technique possesses several appealing properties. Firstly, there always exists a unique and explicit solution for $\mathbf{r}$ and explicit solution for $\boldsymbol{\omega}$. Secondly, from Property 4, larger values of r_i (and so ω_i) are assigned to larger values of $k_i/\sqrt{\lambda_i}$. For each *NC* type, the technique considers the direction of the absolute shift k_i relative to the standard deviation $\sqrt{\lambda_i}$ or equivalently the variance. The *NC* type with a relatively higher value for the weighted shift size is more important than a *NC* type with a lower weighted shift size and is therefore assigned a larger weight. Simply stated, larger weights $\boldsymbol{\omega}$ are given to the type of *NC* that matter most. If the values of $k_i/\sqrt{\lambda_i}$ are all equal to a positive non-zero constant, this means that the *NC* types are equally important. Then, as we would expect, the asymptotic technique will produce equal weights $\mathbf{r}$. Thirdly, the asymptotic technique only assigns non-zero weights to types of *NC* where $k_i/\lambda_i > 0$. The condition $k_i/\lambda_i = 0$ holds *iff* $k_i = 0$ since λ_i is strictly positive. Since, an out-of-control condition can occur from a change in the mean of one or more *NC* types, the type of *NC* that has $k_i/\sqrt{\lambda_i} = 0$ means that the

i^{th} type of *NC* does not contribute to the out-of-control condition. Therefore, to maximize the power of detecting the shift for *NC* types with non-zero $k_i/\sqrt{\lambda_i}$ or k_i/λ_i, we must assign a zero weight to the type of *NC* that has no effect on the quality of the product.

Example – computing the Weights for a OSDCC

Consider a weaving process for wire mesh at the Siam Wire Netting (SWN) factory in Thailand. The overall product quality for wire mesh is accessed by the quantity, type, and severity of *NC* found on a finished roll. More than 20 types of *NC* are to be inspected during the inspection process. SWN has initially planned to adopt a control chart to statistically monitor the mean number of *NC* for the product type TMA-725 and has plans to monitor other product types in the near future. Since the there are several types of products each with several types of *NC*, to avoid an unmanageable number of control charts, a single *OSDCC* is to be used for each product. For TMA-725, a Pareto chart was constructed for the different types of *NC*. The Pareto chart in Figure 1 shows that the first 5 types of *NC*, coded at the request of SWN as *1, 2, 3, 4 and 5,* accounts for about 73 % of the total number of *NC* found on the product. Therefore, it was decided that these five *NC* types would be monitored using the *OSDCC*.

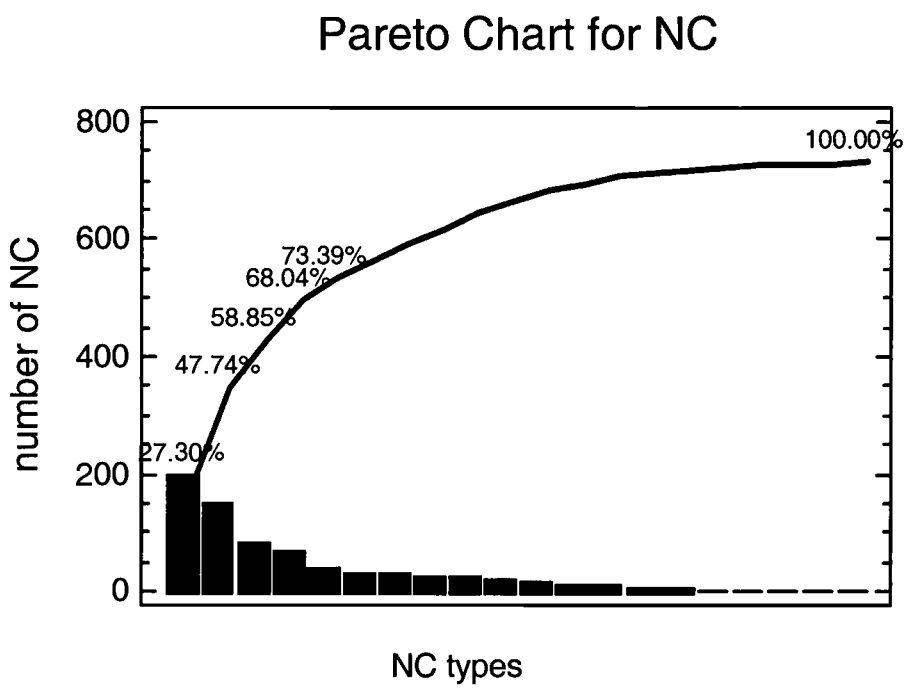

Figure 6: Pareto Chart for the Number of *NC* for Wire Mesh TMA-725.

It was decided that the values for CL_i would be determined by defining values for $k_i/\sqrt{\lambda_i}$. Estimates for λ_i were computed using the data from 1,697 rolls of good product that was collected from several different looms over a two month period. In total, thirty-six samples were collected, each sample containing a different number of rolls of TMA-725. Let X_{ijk} be the count of *NC* for the i^{th} *NC* type in the j^{th} inspection unit in k^{th} sample. Let N_k be the number of inspection units in the k^{th} sample. Then, estimate λ_i as the average number of *NC* of type *i* per inspection unit, i.e.,

$\hat{\lambda}_i = \sum_{j=1}^{N}\sum_{k=1}^{K} X_{ijk} \Big/ \sum_{k=1}^{K} N_k$. The estimates for $\boldsymbol{\lambda}$ were calculated as $\hat{\lambda}_1 = 0.126$, $\hat{\lambda}_2 = 0.042$, $\hat{\lambda}_3 = 0.094$, $\hat{\lambda}_4 = 0.025$ and $\hat{\lambda}_5 = 0.051$.

For $\mathbf{k}$, information about the desired direction of the shift to be detected was not known, so it was decided to let $k_1 = k_2 = k_3 = k_4 = k_5 = 1$. Therefore, the severity of the *NC* type was determined by the magnitude of its corresponding variance estimate, i.e., by $\hat{\lambda}_i$. From the estimates $\hat{\lambda}_i$, all five values for $k_i/\sqrt{\lambda_i}$ are computed as $[2.81, 4.87, 3.26, 6.32, 4.42]$. The *NC* type with smallest variance is assigned the largest comparative loss value and therefore has the largest relative weight.

From eqn. (4), $\mathbf{r}^*$ is computed as $\mathbf{r}^* = [0.07,\ 0.23,\ 0.10,\ 0.39,\ 0.19]$ and eqn. (5) the $\boldsymbol{\omega}$ is easily calculated to be $\boldsymbol{\omega} = C[7.936\ \ 23.809\ \ 10.638\ \ 40\ \ 19.607]$ for some $C > 0$.

From the example we can see that the *NC* type with a relatively higher value for the weighted shift size is more important and is assigned a larger weight. For instance, since $k_1 = k_2 = k_3 = k_4 = k_5 = 1$, then an increase in one unit of the mean for *NC* type 4 is considered 2.24 times more important than that for *NC* of type 1, i.e., $k_1/\sqrt{\hat{\lambda}_1} = 1/\sqrt{0.126} = 2.81$ and $k_4/\sqrt{\hat{\lambda}_4} = 6.32$ and $\left(k_4/\sqrt{\hat{\lambda}_4}\right)\Big/\left(k_1/\sqrt{\hat{\lambda}_1}\right) = 6.32/2.81 = 2.24$. The impact of increasing the mean of *NC* type 4 one unit is considered comparatively 2.24 times more critical than that of type 1. As a result, the asymptotic technique has assigned a larger weight to *NC* type 4

than to *NC* type 1, i.e. $\omega_1 = 4.04$ and $\omega_4 = 20.4$. Then, for the same magnitude of increase in the means of type 1 and 4, the *OSDCC* detects the increase in the mean for *NC* of type 4 quicker than type 1. This implies that the loss or cost from increasing the mean for *NC* of type 4 is comparatively higher than that of type 1. Since $k_1 = k_2 = k_3 = k_4 = k_5 = 1$ in the Example, the asymptotic technique will give the same results for any positive values of *k*. Hence the asymptotic technique assign the weights that depend on the directions of the shifts instead of the magnitudes.

CONCLUSION

The major characteristics of the asymptotic technique are that the asymptotic technique provides an explicit solution in terms of $\mathbf{r}$ and $\boldsymbol{\omega}$ which do not depend on the shift size chosen but depend on the direction of the shift. This constitutes one of the critical properties of the asymptotic technique. If the values $[k_1, k_2, \ldots, k_I]$ are set equally, the asymptotic technique first adjusts the out-of-control direction $[k_1, k_2, \ldots, k_I]$ according to the *NC* variances and assigns weights in the same order (as is common in many statistical techniques) by assigning weights reciprocally to the standard deviation. Since the asymptotic technique has these important and practical properties, it is a logical method for determining the weights for the *OSDCC*.

APPENDICES

Appendix 1: Limiting Probability for an Arbitrary Shift

Given the non-negative shift vector $\boldsymbol{\delta}$, the mean of the individual *NC* becomes $\boldsymbol{\lambda}^* = [\lambda_1 + \delta_1, \ldots, \lambda_I + \delta_I] = \boldsymbol{\lambda} + \boldsymbol{\delta}$ and the mean and standard deviation of U is $\mu_U^* = \mu_U + \sum_{i=1}^{I} \omega_i \delta_i$ and $\sigma_U^* = \sqrt{\sigma_U^2 + \sum_{i=1}^{I} \omega_i^2 \delta_i / N}$, respectively. The probability that the *OSDCC* detects the shift $(1-\beta)$ is computed as $P[U \geq UCL | \boldsymbol{\delta}] = P[(U - \mu_U^*)/\sigma_U^* \geq (UCL - \mu_U^*)/\sigma_U^*]$. First, re-write the UCL as $\mu_U + c\sigma_U$ with $c = (UCL - \mu_U)/\sigma_U$. Then, $(UCL - \mu_U^*)/\sigma_U^*$ can be written as $\frac{1}{\sigma_U^*}\left(c\sigma_U - \sum_{i=1}^{I} \omega_i \delta_i\right) = c\sqrt{\frac{\sigma_U^2}{\sigma_U^2 + \left(\sum_{i=1}^{I} \omega_i^2 \delta_i / N\right)}} - \left(\frac{1}{\sigma_U^*}\right) * \sum_{i=1}^{I} \omega_i \delta_i$. The term $\left(\frac{1}{\sigma_U^*}\right)$ gets large as N increases whereas the term $\left(\frac{\sigma_U}{\sigma_U^*}\right) = \sqrt{\frac{\sigma_U^2}{\sigma_U^2 + \left(\sum_{i=1}^{I} \omega_i^2 \delta_i / N\right)}}$ approaches a constant while the term $\left(\frac{1}{\sigma_U^*}\right) * \sum_{i=1}^{I} \omega_i \delta_i$ approaches ∞. Then, $\frac{1}{\sigma_U^*}\left(c\sigma_U - \sum_{i=1}^{I} \omega_i \delta_i\right)$ will approach $-\infty$ as N gets large. Let $G = (U - \mu_U^*)/\sigma_U^*$ be an arbitrary standard *rv* with zero mean and unit variance. The, $P[U \geq UCL]$ is simply $P[G \geq -\infty]$ which is one.

Appendix 2: Limiting Probability of Detection Converges to a Constant Less Than One

Suppose λ^* is of the form $\lambda^* = \lambda + \mathbf{k}/\sqrt{N}$. Then, as $N \to \infty$, λ^* approaches λ, i.e., $\lim_{N\to\infty} \lambda^* = \lambda$, and σ_U^* converges to σ_U. Given λ^*, the mean and standard deviation of U become $\mu_U^* = \mu_U + \sum_{i=1}^{I} \omega_i k_i \Big/ \sqrt{N}$ and $\sigma_U^* = \sqrt{\sigma_U^2 + \sum_{i=1}^{I} \omega_i^2 k_i \Big/ \left(N\sqrt{N}\right)}$, respectively. Then,

$$(1-\beta) = P[U \geq UCL] = P\left[(U - \mu_U^*)/\sigma_U^* \geq \frac{1}{\sigma_U^*}\left(c\sigma_U - \sum_{i=1}^{I} \omega_i k_i \Big/ \sqrt{N} \right) \right].$$ The term

$$\left(\frac{\sigma_U}{\sigma_U^*}\right) = \sqrt{\frac{\sigma_U^2}{\sigma_U^2 + \left(\sum_{i=1}^{I} \omega_i^2 k_i \Big/ \left(N\sqrt{N}\right) \right)}}$$ approaches one whereas $\left(\frac{1}{\sigma_U^*}\right) \sum_{i=1}^{I} \omega_i k_i \Big/ \sqrt{N}$ which is equal to $\frac{\sum_{i=1}^{I} \omega_i k_i \Big/ \sqrt{N}}{\sqrt{\sum_{i=1}^{I} \omega_i^2 \lambda_i \Big/ N + \sum_{i=1}^{I} \omega_i^2 k_i \Big/ \left(N\sqrt{N}\right)}} = \frac{\sum_{i=1}^{I} \omega_i k_i}{\sqrt{\sum_{i=1}^{I} \omega_i^2 \lambda_i + \sum_{i=1}^{I} \omega_i^2 k_i \Big/ \sqrt{N}}}$ will approach the constant $\frac{\sum_{i=1}^{I} \omega_i k_i}{\sqrt{\sum_{i=1}^{I} \omega_i^2 \lambda_i}}$. Then, the

$$P[U \geq UCL] = P\left[G \geq \left(c - \sum_{i=1}^{I} \omega_i k_i \Big/ \sqrt{\sum_{i=1}^{I} \omega_i \lambda_i^2} \right) \right] < 1.$$

Appendix 3: Derivation of The Maximization Problem

To incorporate the notion of sequences depending on N, let U_N, $\mu_N(\boldsymbol{\lambda})$, $\sigma_N(\boldsymbol{\lambda})$ and UCL_N denote the statistic U, the mean of U, the standard deviation of U, and the UCL for the $OSDCC$ based on N inspection units assuming no shift, i.e., for an in-control process with parameter $\boldsymbol{\lambda}$. When the process is in-control, the UCL_N can be expressed as $\mu_N(\boldsymbol{\lambda})+c_N(\boldsymbol{\lambda})\sigma_N(\boldsymbol{\lambda})$ with $c_N(\boldsymbol{\lambda})=\left(UCL_N-\mu_N(\boldsymbol{\lambda})\right)/\sigma_N(\boldsymbol{\lambda})$.

In the presence of a mean shift of the form $\mathbf{k}/\sqrt{N}$, the parameter for the out-of-control process becomes $\boldsymbol{\lambda}^*=\boldsymbol{\lambda}+\mathbf{k}/\sqrt{N}$. Denote the shifted mean and standard deviation of U as $\mu_N(\boldsymbol{\lambda}^*)$ and $\sigma_N\left(\boldsymbol{\lambda}^*\right)$, respectively. The limiting probability of detecting the shift is $P\left[U_N \geq UCL_N \middle| \boldsymbol{\lambda}^*\right]=P\left[G \geq \left(UCL_N-\mu_N\left(\boldsymbol{\lambda}^*\right)\right)/\sigma_N\left(\boldsymbol{\lambda}^*\right)\right]$ where $G=\left(U_N-\mu_N\left(\boldsymbol{\lambda}^*\right)\right)/\sigma_N\left(\boldsymbol{\lambda}^*\right)$. First, expand $\mu_N(\boldsymbol{\lambda}^*)$ using a Taylor series about $\boldsymbol{\lambda}$ to get $\mu_N\left(\boldsymbol{\lambda}^*\right)=\mu_N\left(\boldsymbol{\lambda}\right)+\left[\mathbf{k}/\sqrt{N}\right]\nabla\mu_N\left(\boldsymbol{\lambda}\right)$ where $\nabla\mu_N\left(\boldsymbol{\lambda}\right)$ is the gradient column vector $\left[\frac{\partial\mu_N\left(\boldsymbol{\lambda}\right)}{\partial\lambda_1} \quad \cdots \quad \frac{\partial\mu_N\left(\boldsymbol{\lambda}\right)}{\partial\lambda_I}\right]^T$ and since $\nabla^m\mu_N\left(\boldsymbol{\lambda}\right)=\left[\frac{\partial^m\mu_N\left(\boldsymbol{\lambda}\right)}{\partial\lambda_1^m} \quad \cdots \quad \frac{\partial^m\mu_N\left(\boldsymbol{\lambda}\right)}{\partial\lambda_I^m}\right]^T=0$ for $m\geq 2$. Note that $\frac{\partial\mu_N\left(\boldsymbol{\lambda}\right)}{\partial\lambda_i}=\frac{\delta}{\delta\lambda_i}\left(\sum_{i=1}^{I}\omega_i\lambda_i\right)=\omega_i,\ \forall i$, therefore $\nabla\mu_N\left(\boldsymbol{\lambda}\right)$ is simply equal to the weight vector, i.e., $\nabla\mu_N\left(\boldsymbol{\lambda}\right)=\boldsymbol{\omega}^T$. Next, substituting for the expressions $\mu_N\left(\boldsymbol{\lambda}^*\right)$ and UCL_N, the probability of detecting the out-of-control parameter can then be expressed as

$P\left[U_N \geq UCL_N \middle| \lambda^*\right] = P\left[G \geq \left(UCL_N - \mu_N(\lambda^*)\right)/\sigma_N(\lambda^*)\right]$. Note that

$$UCL_N - \mu_N(\lambda^*) = \left[\mu_N(\lambda) + c_N(\lambda)\sigma_N(\lambda)\right] - \left\{\mu_N(\lambda) + \left[\mathbf{k}/\sqrt{N}\right]\nabla\mu_N(\lambda)\right\}$$
$$= c_N(\lambda)\sigma_N(\lambda) - \left[\mathbf{k}/\sqrt{N}\right]\nabla\mu_N(\lambda).$$

Therefore,

$$P\left[U_N \geq UCL_N \middle| \lambda^*\right] = P\left[G \geq \left(c_N(\lambda) * \sigma_N(\lambda) - \left[\mathbf{k}/\sqrt{N}\right]\nabla\mu_N(\lambda)\right)/\sigma_N(\lambda^*)\right]$$
$$= P\left[G \geq c_N(\lambda) * \left(\frac{\sigma_N(\lambda)}{\sigma_N(\lambda^*)}\right) - \frac{\left[\mathbf{k}/\sqrt{N}\right]\nabla\mu_N(\lambda)}{\left(\frac{\sigma_N(\lambda^*)}{\sigma_N(\lambda)}\right) * \sigma_N(\lambda)}\right].$$

By the Central Limit Theorem, $U_N = \sum_{j=1}^{N} D_j / N$ converges to a normal *rv* with mean $\mu_N(\lambda)$ and standard deviation $\sigma_N(\lambda)$. So, the UCL_N converges to an upper limit calculated from a normal distribution, i.e., $c_N(\lambda)$ converges to z_α where z_α is a $100(1-\alpha)^{th}$ percentile of a standard normal distribution. Also, the *rv* G converges to a standard normal distribution. Note that

$\sigma_N(\lambda^*) = \sqrt{\sigma_N^2(\lambda) + \sum_{i=1}^{I} \omega_i^2 k_i / \left(N\sqrt{N}\right)}$ converges to $\sigma_N(\lambda)$. So, the ratio $\sigma_N(\lambda)/\sigma_N(\lambda^*)$ as well as its reciprocal converge to 1. Finally, the probability of detecting the shift reduces to $P\left[Z \geq z_\alpha - \left\{\left[\mathbf{k}/\sqrt{N}\right]\nabla\mu_N(\lambda)/\sigma_N(\lambda)\right\}\right]$.

The Asymptotic technique finds the set of $\boldsymbol{\omega}$ that maximize the limiting probability of detection which is equivalent to maximizing $\left[\mathbf{k}/\sqrt{N}\right]\nabla\mu_N(\lambda)/\sigma_N(\lambda)$.

Maximizing $\left[\mathbf{k}/\sqrt{N}\right]\nabla\mu_N(\lambda)/\sigma_N(\lambda)$ is the same as maximizing

$\sum_{i=1}^{I}\omega_i k_i \Big/ \sqrt{\sum_{i=1}^{I}\omega_i^2\lambda_i}$, since $\left[\mathbf{k}/\sqrt{N}\right]\nabla\mu_N(\lambda)/\sigma_N(\lambda) = \left[\mathbf{k}/\sqrt{N}\right]\boldsymbol{\omega}^{\mathrm{T}}/\sigma_N(\lambda)$ which is equal to $\sum_{i=1}^{I}\omega_i\left(k_i/\sqrt{N}\right)\Big/\sqrt{\sum_{i=1}^{I}\omega_i^2\lambda_i/N} = \sum_{i=1}^{I}\omega_i k_i \Big/ \sqrt{\sum_{i=1}^{I}\omega_i^2\lambda_i}$. Letting $r_i = \omega_i^2\lambda_i/\left(\omega_1^2\lambda_1+\omega_2^2\lambda_2+\ldots+\omega_I^2\lambda_I\right)$, we can rewrite $\sum_{i=1}^{I}\omega_i k_i \Big/ \sqrt{\sum_{i=1}^{I}\omega_i^2\lambda_i}$ as $\sum_{i=1}^{I}\left(k_i/\sqrt{\lambda_i}\right)*\sqrt{\omega_i^2\lambda_i/\sum_{i=1}^{I}\omega_i^2\lambda_i} = \sum_{i=1}^{I}\left(k_i/\sqrt{\lambda_i}\right)\sqrt{r_i}$.

Appendix 4: Proof of Property 1

To show that a unique optimal solution exists, notice that the function $f(\mathbf{r}) := \sum_{i=1}^{I}\left(k_i/\sqrt{\lambda_i}\right)\sqrt{r_i}$ is a linear combination of the function $g_i(\mathbf{r}) := \left(k_i/\sqrt{\lambda_i}\right)\sqrt{r_i}$ for $i = 1,\ldots,I$. The second derivative of $g_i(\mathbf{r})$, $\left[\frac{\partial^2 g_i(r_i)}{\partial r_i^2}\right]$ is $\left(-\frac{1}{4}\left(k_i/\sqrt{\lambda_i}\right)r_i^{-3/2}\right)\ \forall i$ and is negative for all $\mathbf{r}$. Therefore, $f(\mathbf{r})$ is a concave function since it is a linear combination of only concave functions. Since $f(\mathbf{r})$ is the concave for all values of $\mathbf{r}$, it's also concave on a restricted region $\sum_{i=1}^{I} r_i = 1$. Hence the function $f(\mathbf{r}>0): \sum_{i=1}^{I} r_i = 1$ possess a global unique optimal solution $\mathbf{r}^*$.

The explicit solution to eqn. (4) is found by rewriting eqn. (2) as

$\underset{\mathbf{r}>0}{Max}\left\{\sum_{i=1}^{I-1}\left(\left(k_i/\sqrt{\lambda_i}\right)*\sqrt{r_i}\right)+\left(k_I/\sqrt{\lambda_I}\right)\sqrt{1-\sum_{i=1}^{I-1}r_i}\right\}$ with a gradient vector of

$\left[\ldots\left(0.5*\left(k_i/\sqrt{\lambda_i}\right)/\sqrt{r_1}\right)-0.5\left(k_I/\sqrt{\lambda_I}\right)/\sqrt{1-\sum_{i=1}^{I-1}r_i}\ldots\right]$. The optimal solution is found by setting the gradient vector equal to the zero vector and solving the simultaneous system of $I-1$ equations, i.e., solve an equation

$r_i=\left[\left(k_i/\sqrt{\lambda_i}\right)^2/\left(k_I/\sqrt{\lambda_I}\right)^2\right]\left(1-\sum_{i=1}^{I-1}r_i\right)=\left[\left(k_i/\sqrt{\lambda_i}\right)^2/\left(k_I/\sqrt{\lambda_I}\right)^2\right]r_I$ for $i=1,\ldots,I-1$.

Since $\sum_{i=1}^{I}r_i=1$ this implies $r_I+\sum_{i=1}^{I-1}\left[\left(k_i/\sqrt{\lambda_i}\right)^2/\left(k_I/\sqrt{\lambda_I}\right)^2\right]r_I=1$ which in turn yields a solution of $r_I=\left(k_I/\sqrt{\lambda_I}\right)^2/\sum_{i=1}^{I}\left(k_i/\sqrt{\lambda_i}\right)^2$. Hence the optimal solution

$\mathbf{r}^*=\left[r_1^*,r_2^*,\ldots,r_I^*\right]$ is $r_i^*=\left(k_i/\sqrt{\lambda_i}\right)^2/\left(\sum_{i=1}^{I}\left(k_i/\sqrt{\lambda_i}\right)^2\right)$ for $i=1,\ldots,I$.

Appendix 5: Proof of Property 2

Recall that $r_i=\dfrac{\omega_i^2\lambda_i}{\omega_1^2\lambda_1+\omega_2^2\lambda_2+\ldots+\omega_I^2\lambda_I}$ where $r_i=\left(k_i/\sqrt{\lambda_i}\right)^2/\left(\sum_{i=1}^{I}\left(k_i/\sqrt{\lambda_i}\right)^2\right)$.

Next rewrite ω_i^2 as $\omega_i^2=\left(k_i/\lambda_i\right)^2\left(\omega_1^2\lambda_1+\omega_2^2\lambda_2+\ldots+\omega_I^2\lambda_I\right)/\left(\sum_{i=1}^{I}\left(k_i/\sqrt{\lambda_i}\right)^2\right)$. Since both $\omega_1^2\lambda_1+\omega_2^2\lambda_2+\ldots+\omega_I^2\lambda_I$ and $\sum_{i=1}^{I}\left(k_i/\sqrt{\lambda_i}\right)^2$ are constant, we can rewrite ω_i as

$\omega_i = (k_i/\lambda_i) * \text{constant}$. So there exists a unique solution for $\boldsymbol{\omega}$ up to a scalar multiple of $\boldsymbol{\omega} = [k_1/\lambda_1 \ldots k_I/\lambda_I]$.

Appendix 6: Proof of Property 3

Denote the ordered $\left(k_i/\sqrt{\lambda_i}\right)$ as $A_{(1)}, A_{(2)}, \ldots, A_{(I)}$ where $A_{(1)}$ is the smallest and $A_{(I)}$ is the largest then eqn. (4) can be rewritten as $r^*_{(i)} = A^2_{(i)} \Big/ \left(\sum_{i=1}^{I} A^2_{(i)} \right)$. It is clear that the order of r_i depend on and is the same as the order of A_i. If for $i = 1, 2, \ldots, I$ the $k_i/\sqrt{\lambda_i}$ are equal then eqn. (4) becomes $r_i = 1/I \ \ \forall i$.

Appendix 7: Proof of Property 4

Consider the case of a shift in only the mean *NC* of type *i* with $\mathbf{k} = [0, \ldots, k_i, \ldots, 0]$. A uniformly most powerful *OSDCC* to detect such a class of shifts is of the form $\boldsymbol{\omega} = [0, \ldots, \omega_i, \ldots, 0]$ for some $\omega_i \neq 0$. See Lehman (1997) for the definition of the most powerful statistical hypothesis test. If $\mathbf{k} = (0, \ldots, k_i, \ldots, 0)$ then eqn. (1) reduces to $\underset{\boldsymbol{\omega}}{Max} \ \omega_i k_i \Big/ \sqrt{\sum_{i=1}^{I} \omega_i^2 \lambda_i} = \underset{\mathbf{r}: \sum r_i = 1}{Max} \left(k_i/\sqrt{\lambda_i}\right)\sqrt{r_i}$ which is maximized *iff* $r_i = 1$ and then asymptotic technique results in a solution of the form $\boldsymbol{\omega} = [0, \ldots, \omega_i, \ldots, 0]$.

REFERENCES

Dodge, Harold F. and Torrey, M. N. (1977), "A check inspection and demerit rating plan (Reprint from Industrial Quality Control)", Journal of Quality Technology, 9, 146-153.

Jones, Allison L., Woodall, William H., and Conerly, Michael D. (1999), "Exact properties of demerit control charts", Journal of Quality Technology, 31, 207-216.

Lehmann, E. L. (1997), Testing Statistical Hypotheses, 2nd edition, Springer.

Montgomery, Douglas. C. (1996), "Introduction to Statistical Quality Control", 3rd edition, John Wiley & Sons, New York.

MULTIVARIATE CONTROL CHART FOR NONCONFORMITIES

ABSTRACT

A multivariate control chart is developed to monitor the quality of a product that is based on the number of different types of nonconformities. The nonconformities are correlated and assumed to be distributed with a multivariate Poisson distribution. Since multivariate Poisson probability distributions cannot model negative correlations, the distribution used in this work is the multivariate Poisson lognormal probability distribution. This distribution is able to model both the positive and negative correlations that exist among the nonconformities.

A different type of multivariate control chart for correlated nonconformities is proposed. The technique developed for the proposed control chart can be applied to nonconformities that have any multivariate distribution, discrete or continuous or other correlated random variables. The proposed method evaluates the deviation of the observed sample means from pre-defined targets in terms of its density function value. The distribution of the control chart test statistic is derived using an approximation technique called a multivariate Edgeworth expansion. For small sample sizes, the results show that the proposed control chart is very robust to inaccuracies in the assumption of the distribution for the correlated nonconformities.

INTRODUCTION

Consider a process where the quality of the finished product is determined by both the number and type of nonconformities (*NCs*) that occur on the product. For example in a wire mesh weaving process there are different types of *NCs*, e.g., large gaps between adjacent wires or a broken wire in the mesh. The quality of a roll of finished wire mesh is determined by counting the number of *NCs* found for each type. The product is considered defective if the numbers of *NCs* in a sample exceeds the desired maximums set for each *NC* type. These types of *NCs* are typically correlated and in the presence of such dependencies, separate control charts (*CCs*) for each *NC* type could be misleading; see Mason and Young (1998). For this reason, a multivariate control chart (*MCC*) is proposed to monitor multiple dependent *NCs*.

Work on *MCCs*, when *NCs* are distributed with a multivariate discrete distribution has rarely been visited in the literature. Patel (1973) proposed a *CC* for a multivariate Poisson (*MP*) and multivariate binomial random variable (*rv*). Patel (1973) assumed that the sample size is sufficiently large so that the Chi-square distribution approximation of the T^2 statistic for testing the means of binomial or Poisson *rvs* can be applied; see Montgomery (2001) for the Hotelling T^2 *CC*. Lu et al. (1998) proposed a multivariate *np* chart for monitoring the number of defective products, called a multivariate *np* (*MNP*) chart. The *MNP* chart uses a weighted sum of the count of the number of defective products as a test statistic. Shewhart 3-sigma limits are used for the *MNP*, which assumes that the test statistic is normally distributed. If the sample size is "small" or the test statistic is not distributed

multivariate normal, the performance of the *CCs* by Patel (1973) and Lu et al. (1998) could be problematic; see Ajmani (1997).

In this work, two forms of a *MCC* for correlated *NCs* is proposed; the first form is an adjusted chi-square *CC* (used when model parameters are known) and the second form is an adjusted T^2 *CC* (used when model parameters must be estimated from data). The proposed *MCC* takes into account higher order joint moments, which, unlike multivariate normal *rvs*, are not necessarily zero. The T^2 control chart (*CC*) using data for *NCs* is used to compare with the adjusted T^2 *CC*. Test results show that the adjusted T^2 *CC* is more robust to the distributional assumptions of the *NCs* than the T^2 *CC*. This allows one to construct an appropriate *CC* for monitoring *NCs* without complete certainty about the distribution of the *NCs*. In multivariate cases, it is not uncommon for the distribution to be uncertain or unknown. To construct the proposed *MCC* a *MP* with a multivariate lognormal distribution for the parameters (*MPLN*) is assumed for the *NCs*, i.e., the mean of the Poisson *NCs* are assume random. Measuring the deviation of the sample means of *NCs* from the targets is a natural approach for deciding if the process is in-control, i.e., the expected means of the *NCs* are chosen as targets. Then the technique called a Multivariate Edgeworth Expansion (*MEE*) is used to approximate the joint distribution of the sample means of the *NCs*. A test statistic for the *MCC* is developed that is a function of the sample means with an adjustment factor for higher order moments. This technique can be applied to the *NCs* or any correlated *rvs* that may have any multivariate distribution, discrete or

continuous. The single upper control limits of the proposed *MCC* were obtained empirically that does not rely on large sample properties.

The rest of the paper is organized as follows. The next section defines the notation and the distributional assumptions for the *NCs*. A numerical example shows how to fit the distribution to *NCs* using data from a wire mesh weaving process and the method for obtaining the parameter estimates is discussed. The section following shows how to apply the *MEE* technique to approximate the distribution of the sample means and obtain the test statistic. The next section explains how to determine the control limits of the proposed *MCC* and explains the details of the method used to obtain the control limits. A numerical example is presented using data from the wire mesh weaving process. The following section gives results showing the robustness of both the proposed adjusted chi-square *MCC* and the adjusted T^2 *MCC* when the *NCs* are not distributed *MPLN*. The last section provides a conclusion and discussion.

MODEL

To develop the *MCC* for multiple correlated *NCs*, a multivariate discrete distribution for the *NCs* is needed; see Johnson et al (1997) for the general definitions of discrete multivariate distributions. Because count data are reasonably modeled marginally by the Poisson distribution, the *MP* distribution seems a natural choice for the multivariate case. The *MP*, although, is restricted to the class of distributions where the correlation among the Poisson *rvs* is positive; see Johnson et al. (1997). To model the multivariate *NCs* with a general correlation structure while maintaining a

marginal Poisson distribution for the *NCs* a hierarchical modeling technique is commonly adopted.

Moreover in a batch process, i.e., weaving process consider here it is sensible to conceive that quality of product such as expected means of *NCs* could be different to some extent among different batches depending on machine set-up for each roll. Also the mean nonconformities even for the same nonconformity type could be varying among batches for some degree. The hierarchical modeling technique we adopted not only provides us as a general tool for modeling the correlated nonconformities with a general correlation structure but also accommodates the possible batch-quality dependent situation. Steyn (1976) considered a class of *MP* distributions where the parameters of the marginal Poisson *rvs* were themselves assumed to be *rvs* and distributed multivariate normal. Nelson (1985) considered a *MP* distribution where the parameters were distributed multivariate gamma. See Kotz et al (2000) for an excellent source of continuous multivariate distributions.

In this work, the multivariate lognormal is chosen for the distribution of the parameters. The multivariate lognormal distribution for the parameters of a multivariate Poisson (*MPLN*) was first considered by Aitchison and Ho (1989). Problems modeling very low counts may arise when assuming a normal distribution for the parameters, specifically, negative estimates could be obtained when estimating the parameters. Choosing multivariate lognormal for the parameters does not have this difficulty and produces a suitably wide range of correlation structures; see

Aitchison and Ho (1989). Closed forms of the joint moments can be derived and used to obtain estimates for the parameters.

Let the *rvs* in the column vector $\mathbf{x}_{jk} = \left(X_{1jk},...,X_{Ijk}\right)^T$ denote the count of each *NC* type in the j^{th} inspection unit of the k^{th} sample with *NC* types $i = 1,2,...I$. Each sample consists of an integer number of inspection units. Assume that $\mathbf{x}_{jk}$ are independently distributed *MPLN* $\forall\ j,k$ conditioning on the parameters $\boldsymbol{\lambda}_{jk} = \left(\lambda_{1jk},...,\lambda_{Ijk}\right)^T$. Let $\boldsymbol{\lambda}_{jk}\ \forall\ j,k$ be independently and identically distributed (*i.i.d.*) multivariate lognormal with mean $\boldsymbol{\mu} = \left(\mu_1,...,\mu_I\right)^T$ and variance-covariance matrix $\boldsymbol{\Sigma}$. A joint probability mass function for $\mathbf{x}_{jk}$ is expressed as

$$P\left(X_{1jk} = x_{1jk},...,X_{Ijk} = x_{Ijk}\right) = \int_{\{R^I \geq 0\}} \exp\left(-\sum_{i=1}^{I}\lambda_{ijk}\right)\prod_{i=1}^{I}\left(\lambda_{ijk}^{x_{ijk}}/x_{ijk}!\right)f(\boldsymbol{\lambda}_{jk})d\boldsymbol{\lambda}_{jk} \qquad (1)$$

where

$$f\left(\lambda_{1jk},...,\lambda_{Ijk}\right) = \left(\prod_{i=1}^{I}\lambda_{ijk}\right)^{-1}(2\pi)^{-I/2}\left|\boldsymbol{\Sigma}\right|^{-1/2}\exp\left\{-\frac{1}{2}\left(\log\boldsymbol{\lambda}_{jk}-\boldsymbol{\mu}\right)^T\boldsymbol{\Sigma}^{-1}\left(\log\boldsymbol{\lambda}_{jk}-\boldsymbol{\mu}\right)\right\} \qquad (2)$$

is a joint density function of the multivariate lognormal *rvs* $\boldsymbol{\lambda}_{jk}$. See Kotz et al (2000). The mean, and covariance matrix of $\mathbf{x}_{jk}$, denoted by $\boldsymbol{\tau} = \left(\tau_1,...,\tau_I\right)^T$ and $\mathbf{V}$ respectively, are

$$\tau_i = \exp\left[\mu_i + \Sigma_{ii}/2\right] \qquad i = 1,...,I \qquad (3)$$

$$\begin{aligned} \mathbf{V}_{ii'} &= \tau_i + \tau_i^2\left(\exp[\Sigma_{ii}]-1\right) \qquad i = i' \\ &= \tau_i\tau_{i'}\left(\exp[\Sigma_{ii'}]-1\right) \qquad i \neq i' \end{aligned} \tag{4}$$

We see that depending on the sign of $\Sigma_{ii'}$ in eqn. (4) the *MPLN* can represent both negative and positive correlation.

ESTIMATING $\boldsymbol{\mu}$ AND $\boldsymbol{\Sigma}$

Standard estimators such as the maximum likelihood estimator (*MLE*) and the method of moments estimator (*MOM*) can be used to estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$; see Casella and Berger (1990), Kendall et al (1991) for definitions of *MLE* and *MOM*. To estimate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ assume that data is available from a presumably in-control process that we are interested in statistically monitoring with the *MCC*. Say K samples are taken where the samples are of possibly different sizes. Let J_k denote the number of inspection units in the k^{th} sample. The *MLE* estimates of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ denoted as $\hat{\boldsymbol{\mu}}_{MLE}$ and $\hat{\boldsymbol{\Sigma}}_{MLE}$, respectively, are determined by finding a set of $\hat{\boldsymbol{\mu}}_{MLE}$ and $\hat{\boldsymbol{\Sigma}}_{MLE}$ that maximizes the total likelihood function value or equivalently the log-likelihood function value of the data such as

$$Maximize\ Argument\left[\prod_{k=1}^{K}\prod_{j=1}^{J_k} P\left(X_{1jk} = x_{1jk},\ldots, X_{ljk} = x_{ljk}\right)\right] \tag{5}$$

where $P\left(X_{1jk} = x_{1jk},..., X_{Ijk} = x_{Ijk}\right)$ is calculated from eqn.(1). Aitchison and Ho (1989) provide a good discussion on the mathematical issues regarding optimization of the log-likelihood function for *MLE* estimators.

The *MOM* estimates denoted $\hat{\boldsymbol{\mu}}_{MOM}$ and $\hat{\boldsymbol{\Sigma}}_{MOM}$ are obtained by first computing from the data the overall sample means for each *NC* type denoted by $\bar{\bar{\mathbf{x}}} = \left(\bar{\bar{x}}_1,..., \bar{\bar{x}}_I\right)^T$ where $\bar{\bar{x}}_i = \sum_{k=1}^{K}\sum_{j=1}^{J_k} x_{ijk} \Big/ \sum_{k=1}^{K} J_k$ is a grand average of *NC* of type i^{th} over samples and inspection units. Next, represent the data in a matrix by concatenating $\mathbf{x}_{jk}$ side-by-side into a matrix, $\mathbf{X} = \left[\mathbf{x}_{11}\cdots\mathbf{x}_{J_11}\cdots\mathbf{x}_{1k}\cdots\mathbf{x}_{J_kk}\right]$ of size I by $\sum_{k=1}^{K} J_k$ and compute the sample variance-covariance matrix, $\mathbf{S} = \left[\mathbf{X} - \bar{\bar{\mathbf{x}}}\mathbf{1}\right]\left[\mathbf{X} - \bar{\bar{\mathbf{x}}}\mathbf{1}\right]^T \Big/ \left(\sum_{k=1}^{K} J_k - 1\right)$ where $\mathbf{1} = [1,\ldots,1]$ is a row vector of 1's of length $\sum_{k=1}^{K} J_k$; see for example Johnson and Wichern (1988) for details of these computations. To obtain the *MOM* estimates, equate the expected values of the means, variances, and covariances of $\mathbf{x}_{jk}$ from eqns. (3) and (4) to the sample means, variances, and covariances and solve for $\hat{\boldsymbol{\mu}}_{MOM}$ and $\hat{\boldsymbol{\Sigma}}_{MOM}$. The estimates are

$$\hat{\Sigma}_{ii} = \log\left(\left(\mathbf{S}_{ii} - \bar{\bar{x}}_i\right)\Big/\bar{\bar{x}}_i^2 + 1\right) \tag{6}$$

$$\hat{\Sigma}_{ii'} = \log\left(\mathbf{S}_{ii'} \Big/ \left(\bar{\bar{x}}_i \bar{\bar{x}}_{i'}\right) + 1\right) \tag{7}$$

$$\hat{\mu}_i = \log \bar{\bar{x}}_i - \frac{1}{2}\log\left(\left(\mathbf{S}_{ii} - \bar{\bar{x}}_i\right) \Big/ \bar{\bar{x}}_i^2 + 1\right). \tag{8}$$

There are advantages and disadvantages for each estimation method. For *MLE* an advantage is that a valid solution always exists and can be found theoretically. Another advantage of an *MLE* is the nice asymptotic properties, e.g., the *MLE* would result in an asymptotically unbiased estimator. The one major drawback of using a *MLE* is that it is considerably more difficult to compute, especially when the number of *NC* types increases.

For *MOM*, the primary benefit is in its computational ease, which is quite significant relative to the *MLE*. Unfortunately, depending on the covariance structure of the data, *MOM* estimates are sometimes unobtainable or invalid. Observe that for a solution to exist the quantity on which the log is taken in eqns. (6), (7), and (8) must be positive. For example, to satisfy eqn. (6) one of the following must hold; $\mathbf{S}_{ii} > \bar{\bar{x}}_i(1 - \bar{\bar{x}}_i)$. Likewise, to satisfy eqn. (7) $\mathbf{S}_{ii'} \Big/ \left(\bar{\bar{x}}_i \bar{\bar{x}}_{i'}\right) + 1$ must be positive and from eqn. (8) $\hat{\mu}_i$ must be positive definite for all i. Another drawback is that if a solution is obtained, it may be invalid. For example $\hat{\boldsymbol{\Sigma}}_{MOM}$ may not be positive definite. Although *MOM* estimates do not posses the asymptotic properties of the *MLE*, it can be argued that when the data set is large, the *MOM* estimates will, with high confidence, be very close to the true parameter values.

Example 1 – Estimating parameters: $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$.

Data from a wire mesh weaving process is used for this example. A single sample of 36 observations of *NC* counts for two types of *NCs* was collected from the process when it was considered to be in-control. The data for two *NC* types are shown in Table 1. The sample mean vector and variance-covariance matrix for the *NC* counts were computed as $\bar{\mathbf{x}} = (5.52, 2.25)^T$ and $\mathbf{S} = \begin{bmatrix} 28.82 & -4.02 \\ -4.02 & 8.36 \end{bmatrix}$.

Notice the sample variances are much larger than the sample means suggesting a lack of fit to the Poisson distribution marginally. A chi-square goodness of fit test was performed for each *NC* type against the Poisson distribution. The results show that there is statistically strong evidence that the *NC* types are not Poisson. Knowledge of the weaving process and the characteristics of the wire mesh suggest that these two types of *NCs* are dependent. Due to uncertainty about the distribution of the *NCs*, to avoid complications with distributional dependence the nonparametric Spearman rank test was performed to test for dependence between the *NC* types. The highly significant Spearman rank correlation suggests that the *NCs* are indeed correlated.

Both *MLE* and *MOM* estimates were obtained. To find the *MLE* estimates, eqn. (5) was optimized using the data in Table 1. The optimization was performed on a Linux operating system with dual processor AMD Athlon MP2000 CPUs using software Matlab version "6.5". The results for the *MLE* estimates are

$\hat{\boldsymbol{\mu}}_{MLE} = (1.47, 0.42)^T$ and $\hat{\boldsymbol{\Sigma}}_{MLE} = \begin{bmatrix} 0.43 & -0.24 \\ -0.24 & 0.67 \end{bmatrix}$. To find the *MOM* estimates, eqns. (6), (7), and (8) were solved. The *MOM* estimate results are

$\hat{\boldsymbol{\mu}}_{MOM} = (1.43, 0.41)^T$, $\hat{\boldsymbol{\Sigma}}_{MOM} = \begin{bmatrix} 0.57 & -0.39 \\ -0.39 & 0.79 \end{bmatrix}$. Both $\hat{\boldsymbol{\mu}}_{MLE}$ and $\hat{\boldsymbol{\mu}}_{MOM}$ are very similar.

The log-likelihood value can be used as a measure of how well the estimates fit the data. The log-likelihood measures for the *MLE* and *MOM* estimates are -167.86 and -187.54, respectively, and since we are interested in maximizing the log-likelihood measure, the results suggest that the *MLE* estimates are better.

Two alternative distributions for $\boldsymbol{\lambda}_{jk}$ were considered – a multivariate exponential and a multivariate gamma; see Gumbel (1960); and Gupta and Wong (1984). *MLE* estimates were obtained for each alternative distribution. Table 2 shows the estimated mean and variance $\boldsymbol{\tau}$ and $\mathbf{V}$ of $\mathbf{x}_{jk}$ calculated from the *MLE* of these two distributions which are very similar to the *MPLN* for the given data set.

Table 4: Count of *NCs* in Wire Mesh.

Observation	*NC* Type 1	*NC* Type 2	Observation	*NC* Type 1	*NC* Type 2
1	6	1	19	2	2
2	0	2	20	8	2
3	8	3	21	5	7
4	1	4	22	1	4
5	9	0	23	9	0
6	0	3	24	2	3
7	5	1	25	8	0
8	6	1	26	5	0
9	5	0	27	6	1
10	11	3	28	3	1
11	30	1	29	3	3
12	16	0	30	5	2
13	4	8	31	4	1
14	3	14	32	4	0
15	6	0	33	5	2
16	3	1	34	0	6
17	3	5	35	8	0
18	1	0	36	4	0

Table 5: The Estimates of Mean, Covariane and Correlation of $\mathbf{x}_{jk}$.

MLE estimates	Estimate of $\boldsymbol{\tau}$	Estimate of $\mathbf{V}$	Estimate Correlation
Multivariate Lognormal	$\begin{pmatrix} 5.41 \\ 2.12 \end{pmatrix}$	$\begin{bmatrix} 21.32 & -2.44 \\ -2.44 & 6.45 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.21 \\ -0.21 & 1 \end{bmatrix}$
Multivariate Exponential	$\begin{pmatrix} 5.33 \\ 2.19 \end{pmatrix}$	$\begin{bmatrix} 33.93 & -2.64 \\ -2.64 & 6.93 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.17 \\ -0.17 & 1 \end{bmatrix}$
Multivariate Gamma	$\begin{pmatrix} 4.92 \\ 2.01 \end{pmatrix}$	$\begin{bmatrix} 9.62 & -0.97 \\ -0.97 & 4.24 \end{bmatrix}$	$\begin{bmatrix} 1 & -0.15 \\ -0.15 & 1 \end{bmatrix}$

DEVELOPMENT OF THE MCC TEST STATISTIC

Assume we would like to statistically monitor a process for the change in one or more of the mean *NCs* $\boldsymbol{\tau}$ away from an expected target value. At equal time intervals, a sample is taken from the process and the number of each type of *NC* found on the product is counted, then a test is performed to determine if one or more of the mean *NCs* has changed. This can be restated as the hypothesis test

$$H_0 : E\left(X_{ijk}\right) = \tau_i \quad \forall i \in \{1,...,I\} \quad \textit{versus} \quad H_1 : E\left(X_{ijk}\right) \neq \tau_i \quad \textit{for some } i\,. \tag{9}$$

Suppose the in-control $\mathbf{V}$ and $\boldsymbol{\tau}$ are known. The traditional chi-square (χ^2) *CC* based on the χ^2 statistic

$$\chi^2 = J\left(\bar{\mathbf{x}}_k - \boldsymbol{\tau}\right)^T \mathbf{V}^{-1}\left(\bar{\mathbf{x}}_k - \boldsymbol{\tau}\right) \tag{10}$$

could be used to test the hypothesis in eqn. (9) where $\bar{\mathbf{x}}_k = \sum_{j=1}^{J} \mathbf{x}_{jk} \Big/ J$ are the sample means of the *NCs*. This test statistic is theoretically for $\bar{\mathbf{x}}_k$ distributed multivariate normal (*MVN*). See Alt (1985), Jackson and Morris (1957), Jackson (1985) and Lowry and Montgomery (1995) for the *MCC* for *MVN*. If $\bar{\mathbf{x}}_k$ is not approximately *MVN*, which is a typical assumption in practice, a different test statistic than the χ^2 in eqn. (10) should be used. To develop such a test statistic, one approach would be considering the differences $\mathbf{D} = \left(D_1,...,D_I\right)^T$ of $\bar{\mathbf{x}}_k$ from the targets $\boldsymbol{\tau}$, i.e.,

$\mathbf{D} = \bar{\mathbf{x}}_k - \boldsymbol{\tau}$. The *MCC* should signal an out of control condition if the magnitude of one or more $D_i : i = 1,...,I$ are larger than a corresponding critical value, i.e., if

$$|D_i| > C_i \quad for\ some\ C_i > 0. \qquad (11)$$

The critical values $\mathbf{C} = (C_1,...,C_I)^T$ are chosen so that the *MCC* has a desired false alarm probability α. Unfortunately, values for $\mathbf{C}$ that satisfy eqn. (11) are not unique.

We propose an approach for choosing an acceptance region (*AR*) for the *MCC* from the set of $\bar{\mathbf{x}}_k$ that satisfies a given condition, i.e., we choose an *AR* for $\bar{\mathbf{x}}_k$ such that the density function value, denoted by $f_{\bar{\mathbf{x}}_k}(\bar{\mathbf{x}}_k)$, is greater than some limit, say *f*, i.e., $AR = \{\bar{\mathbf{x}}_k : f_{\bar{\mathbf{x}}_k}(\bar{\mathbf{x}}_k) > f\}$. If $\bar{\mathbf{x}}_k$ is *MVN*, then the χ^2 *CC* can be used to test the hypothesis in eqn. (9) and the *AR* in two dimensions is an ellipsoid or a region where the density function value of a *MVN* distribution is greater than some limit. Moreover as the number of inspection units *J* in a sample becomes large, the joint distribution of $\bar{\mathbf{x}}_k$ converges to the *MVN* distribution. If we choose $AR = \{\bar{\mathbf{x}}_k : f_{\bar{\mathbf{x}}_k}(\bar{\mathbf{x}}_k) > f\}$, when $f_{\bar{\mathbf{x}}_k}(\bar{\mathbf{x}}_k)$ converges to the *MVN* the *AR* converges to that of the χ^2 statistic as expected. It is worth noting that this method can be used for any multivariate distribution discrete or continuous.

It is apparent that the density function of $\bar{\mathbf{x}}_k$ must be known. With small to moderate sample sizes with discrete *rvs* such as counts of *NCs*, asymptotic properties

like the convergence to a known distribution do not usually exist. Therefore, the work presented here uses a distribution approximating technique called a Multivariate Edgeworth Expansion (*MEE*) to approximate the density of $\bar{\mathbf{x}}_k$; see McCullagh (1987) for the details and derivations of an *MEE*. Theoretically the more the higher order joint moments adjusted or used in the *MEE* function the more accuracy of the approximation. The *MVN* distribution approximation is simply the *MEE* function with only the first two joint moments. Let the approximate density for $\bar{\mathbf{x}}_k$ obtained from an *MEE* be denoted $\hat{f}_{\bar{\mathbf{x}}_k}(\bar{\mathbf{x}}_k)$. The computation of the *MEE* with joint moment of order four or more becomes exponential intensive. Thus the *MEE* approximation of $\hat{f}_{\bar{\mathbf{x}}_k}(\bar{\mathbf{x}}_k)$ uses only the first three moments in the approximation yet theoretically produces at least or more accuracy than the *MVN* approximation. When $\mathbf{x}_{jk}$ is distributed *MPLN*, with log normal mean and variance $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ and *MPLN* mean and variance $\boldsymbol{\tau}$ and $\mathbf{V}$ respectively, $\hat{f}_{\bar{\mathbf{x}}_k}(\bar{\mathbf{x}}_k)$ can be written as

$$\hat{f}_{\bar{\mathbf{x}}_k}(\bar{\mathbf{x}}_k) = \phi(\bar{\mathbf{x}}_k, \boldsymbol{\tau}, \mathbf{V}/J)\left[1 + \sum_{a=1}^{I}\sum_{b=1}^{I}\sum_{c=1}^{I} h_{abc}(\bar{\mathbf{x}}_k) K_{abc}(\bar{\mathbf{x}}_k)\right] \tag{12}$$

where

$$\phi(\bar{\mathbf{x}}_k, \boldsymbol{\tau}, \mathbf{V}/J) = (2\pi)^{-I/2} |\mathbf{V}/J|^{-1/2} \exp\left\{-\frac{J}{2}(\bar{\mathbf{x}}_k - \boldsymbol{\tau})^T \mathbf{V}^{-1}(\bar{\mathbf{x}}_k - \boldsymbol{\tau})\right\} \tag{13}$$

$$h_{abc}(\bar{\mathbf{x}}_k) = h_a(\bar{\mathbf{x}}_k) h_b(\bar{\mathbf{x}}_k) h_c(\bar{\mathbf{x}}_k) - J\left\{h_a(\bar{\mathbf{x}}_k)\left[\mathbf{V}^{-1}\right]_{b,c} + h_b(\bar{\mathbf{x}}_k)\left[\mathbf{V}^{-1}\right]_{a,c} + h_c(\bar{\mathbf{x}}_k)\left[\mathbf{V}^{-1}\right]_{a,b}\right\} \tag{14}$$

$$h_a(\bar{\mathbf{x}}_k) = J\left[\mathbf{V}^{-1}\right]_{a,\bullet}(\bar{\mathbf{x}}_k - \boldsymbol{\tau}) \tag{15}$$

$$K_{abc}(\bar{\mathbf{x}}_k) = \left\{\begin{matrix} E(X_{ajk}X_{bjk}X_{cjk}) - \tau_a E(X_{bjk}X_{cjk}) - \tau_b E(X_{ajk}X_{cjk}) \\ -\tau_c E(X_{ajk}X_{bjk}) + 2\tau_a\tau_b\tau_c \end{matrix}\right\} / J^2 \tag{16}$$

$$\begin{aligned} E(X_{aj}X_{bj}X_{cj}) &= \tau_a^3 \exp(3\Sigma_{aa}) + 3\tau_a^2 \exp(\Sigma_{aa}) + \tau_a \qquad for\ a = b = c \\ &= \tau_a\tau_c \exp(\Sigma_{ac}) + \tau_a^2\tau_c \exp(\Sigma_{aa} + 2\Sigma_{ac}) \quad for\ (a = b) \neq c \\ &= \tau_a\tau_b\tau_c \exp(\Sigma_{ab} + \Sigma_{ac} + \Sigma_{bc}) \qquad for\ a \neq b \neq c \end{aligned} \tag{17}$$

$$\begin{aligned} E(X_{aj}X_{bj}) &= \tau_a + \tau_a^2 \exp(\Sigma_{aa}) \quad for\ a = b \\ &= \tau_a\tau_b \exp(\Sigma_{ab}) \qquad for\ a \neq b \end{aligned} \tag{18}$$

The quantity $\left[\mathbf{V}^{-1}\right]_{a,b}$ is the $(a,b)^{th}$ element of $\mathbf{V}^{-1}$ and $\left[\mathbf{V}^{-1}\right]_{a,\bullet}$ is the a^{th} row of $\mathbf{V}^{-1}$. Theoretically, when modeling *NCs*, the *MEE* density function of $\bar{\mathbf{x}}_k$ is more accurate than the *MVN* density function; see example 2. Under the assumption that the process is in-control, i.e., $E(X_{ijk}) = \tau_i$, the condition $\hat{f}_{\bar{\mathbf{x}}_k}(\bar{\mathbf{x}}_k) > f$ can be simplified to

$$J(\bar{\mathbf{x}}_k - \boldsymbol{\tau})^T \mathbf{V}^{-1}(\bar{\mathbf{x}}_k - \boldsymbol{\tau}) - 2\log\left[1 + \sum_{a=1}^{I}\sum_{b=1}^{I}\sum_{c=1}^{I} h_{abc}(\bar{\mathbf{x}}_k) K_{abc}(\bar{\mathbf{x}}_k)\right] \leq L \quad for\ some\ L \tag{19}$$

where the functions h_{abc} and K_{abc} are defined as in eqns. (14), (15), and (16). So, the *MCC* for *NCs* will signal an out-of-control condition when the value of the expression

on the left side of eqn. (19) is larger than some limit L. Therefore, the test statistic for the *MCC* for *NCs* and the hypothesis test in eqn. (9) is

$$\chi^2_{adj}(\bar{\mathbf{x}}_k) = J(\bar{\mathbf{x}}_k - \boldsymbol{\tau})^T \mathbf{V}^{-1}(\bar{\mathbf{x}}_k - \boldsymbol{\tau}) - 2\log\left[1 + \sum_{a=1}^{I}\sum_{b=1}^{I}\sum_{c=1}^{I} h_{abc}(\bar{\mathbf{x}}_k) K_{abc}(\bar{\mathbf{x}}_k)\right]. \quad (20)$$

The proposed *MCC* statistic will be called an adjusted chi-square (χ^2_{adj}) *CC* statistic. Notice when $\bar{\mathbf{x}}_k$ is multivariate normal, the second term in eqn. (20) is zero and the test statistic reduces to eqn. (10) and the *MCC* to the traditional chi-square (χ^2) *CC*.

When $\boldsymbol{\Sigma}$ and $\boldsymbol{\mu}$ are unknown the estimates $\hat{\boldsymbol{\Sigma}}$ and $\hat{\boldsymbol{\mu}}$ and consequently the estimates $\hat{\boldsymbol{\tau}}$ and $\hat{\mathbf{V}}$ as in example 1 can be obtained from data. Using these estimates, the statistics χ^2 and $\chi^2_{adj}(\bar{\mathbf{x}}_k)$ become T^2 and $T^2_{adj}(\bar{\mathbf{x}}_k)$ respectively, i.e.,

$$T^2 = J(\bar{\mathbf{x}}_k - \hat{\boldsymbol{\tau}})^T \hat{\mathbf{V}}^{-1}(\bar{\mathbf{x}}_k - \hat{\boldsymbol{\tau}}) \quad (21)$$

$$T^2_{adj}(\bar{\mathbf{x}}_k) = J(\bar{\mathbf{x}}_k - \hat{\boldsymbol{\tau}})^T \hat{\mathbf{V}}^{-1}(\bar{\mathbf{x}}_k - \hat{\boldsymbol{\tau}}) - 2\log\left[1 + \sum_{a=1}^{I}\sum_{b=1}^{I}\sum_{c=1}^{I} h_{abc}(\bar{\mathbf{x}}_k) K_{abc}(\bar{\mathbf{x}}_k)\right] \quad (22)$$

where the functions h_{abc} and K_{abc} are computed as before now using $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$, $\hat{\boldsymbol{\tau}}$ and $\hat{\mathbf{V}}$. The *CC* based on the T^2 statistic in eqn. (21) is similar to a Hotelling T^2 *CC* except $\hat{\boldsymbol{\tau}}$ and $\hat{\mathbf{V}}$ are computed using eqn. (3) and (4).

Example 2 – Comparing the *MPLN* joint density with the *MEE* and the *MVN* joint densities.

Consider two types of *NCs* distributed *MPLN*. To see how well the *MEE* and the *MVN* densities approximate a *MPLN* for the $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ computed from the weaving process in example 1, consider a *MPLN* distribution using $\boldsymbol{\mu} = \hat{\boldsymbol{\mu}}_{MLE} = (1.47, 0.42)^T$ and $\boldsymbol{\Sigma} = \hat{\boldsymbol{\Sigma}}_{MLE} = \begin{bmatrix} 0.43 & -0.24 \\ -0.24 & 0.67 \end{bmatrix}$. For each vector $\mathbf{x}_{jk} = (X_{1jk}, X_{2jk})$ with $0 \le X_{1jk} \le 10$ and $0 \le X_{2jk} \le 15$ compute the probability $P(X_{1jk} = x_{1jk}, X_{2jk} = x_{2jk})$ using eqns. (1) and (2). These *MPLN* probabilities are plotted in Figure 1. Next compute the sample mean vector and sample variance-covariance matrix of $\mathbf{x}_{jk}$, i.e., $\boldsymbol{\tau} = (5.41, 2.12)^T$ and $\mathbf{V} = \begin{bmatrix} 21.32 & -2.43 \\ -2.43 & 6.44 \end{bmatrix}$, respectively. For each $\mathbf{x}_{jk}$, compute the *MVN* probability from eqn. (13). The *MVN* probabilities are plotted against the *MPLN* in Figure 2. Similarly, for each $\mathbf{x}_{jk}$ an *MEE* probability is computed using the same values for $\boldsymbol{\tau}$ and $\mathbf{V}$ in eqns. (12) – (18). The *MEE* probabilities are plotted against the *MPLN* in Figure 3. Figure 2 shows that the *MEE* for this set of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ is a better approximation to the density of $\mathbf{x}_{jk}$ than the *MVN* density.

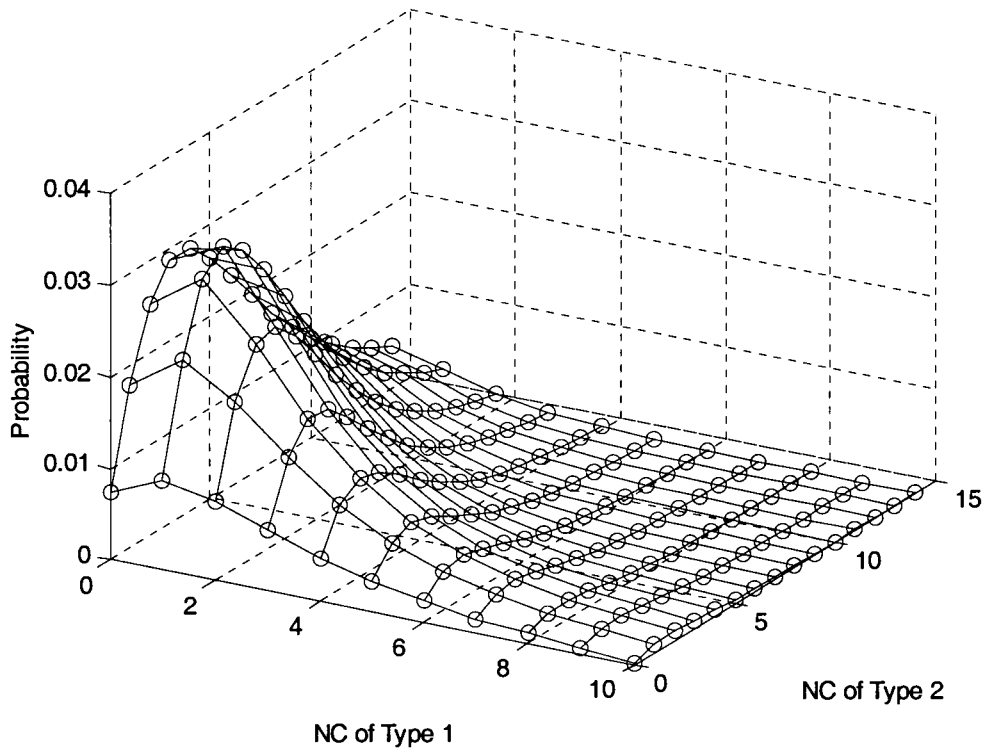

Figure 7: Plot of The *MPLN* Density for Given Values for μ and Σ.

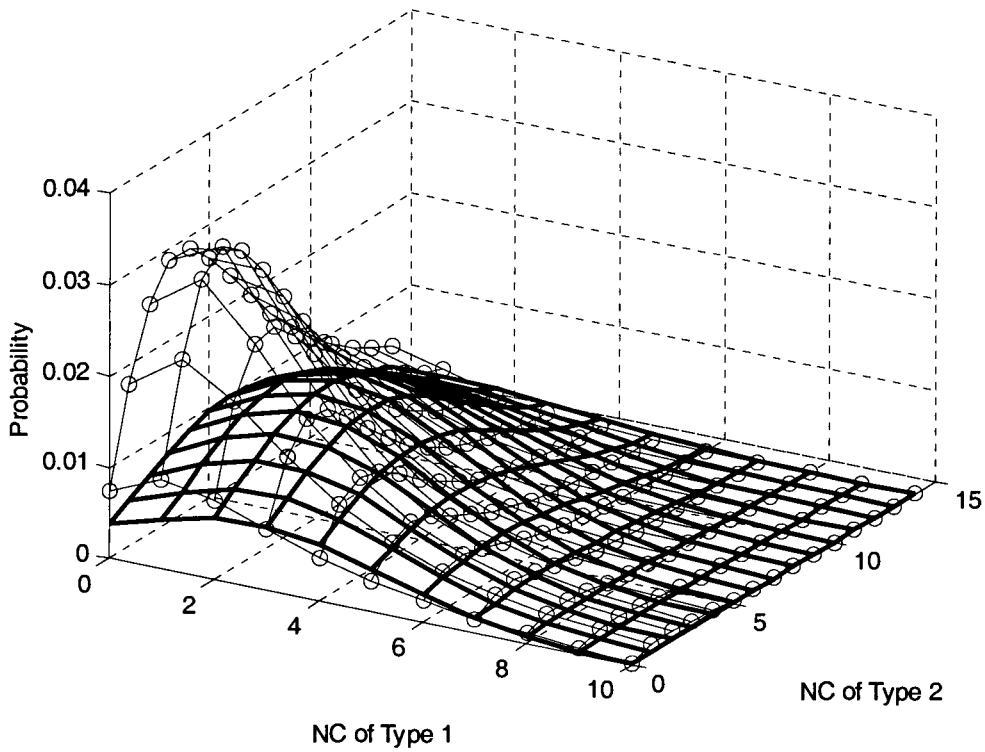

Figure 8: Plot of The *MPLN* and The *MVN* Densities.

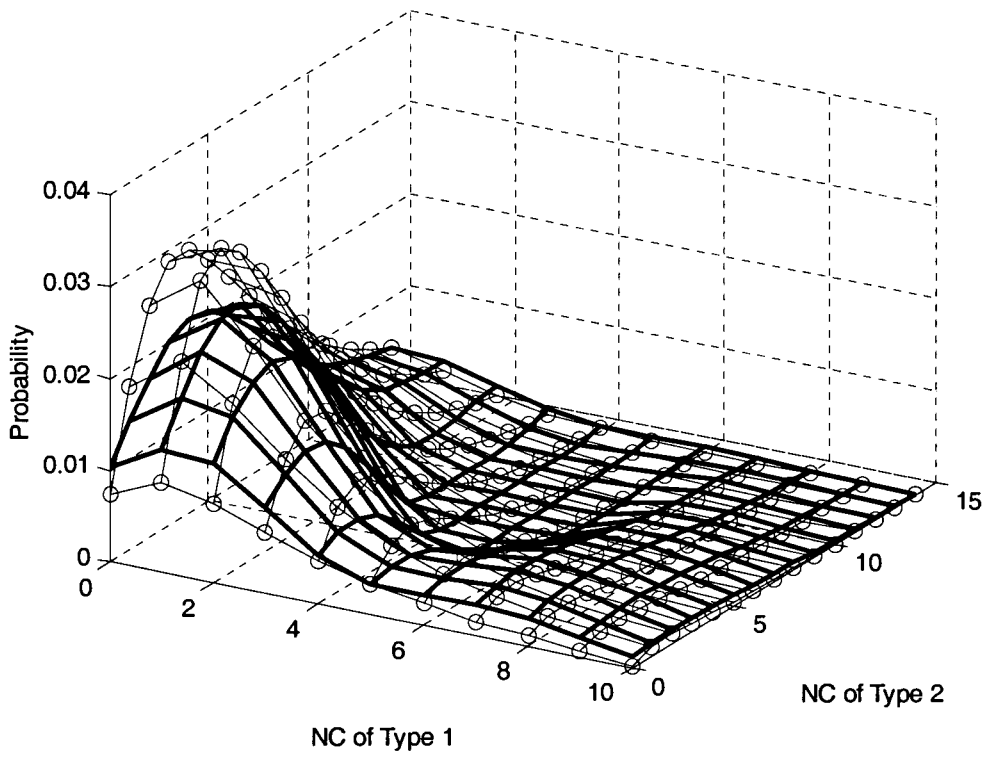

Figure 9: Plot of The *MPLN* and The *MEE* Densities.

DETERMINING AN UPPER CONTROL LIMIT FOR THE MCC

The exact distributions for $\chi^2_{adj}(\bar{\mathbf{x}}_k)$ or $T^2_{adj}(\bar{\mathbf{x}}_k)$ are unknown and cannot be derived explicitly. Therefore, the upper control limit (*UCL*) for the *MCC* is derived empirically. When $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, $\boldsymbol{\tau}$, and $\mathbf{V}$ known the upper control limit for the *MCC* is $UCL(\chi^2_{adj})$ and when $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, $\boldsymbol{\tau}$, and $\mathbf{V}$ are unknown, the *UCL* is $UCL(T^2_{adj})$. To find the empirical *UCL* for the *MCC* the empirical distribution of $\chi^2_{adj}(\bar{\mathbf{x}}_k)$ or $T^2_{adj}(\bar{\mathbf{x}}_k)$ is used.

If $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, $\boldsymbol{\tau}$, and $\mathbf{V}$ are known, the empirical distribution of $\chi^2_{adj}(\bar{\mathbf{x}}_k)$ is computed as follows. First, generate J *i.i.d.* $\boldsymbol{\lambda}_{jk}$ mean vectors one for each of the J inspection units in the sample. For each $\boldsymbol{\lambda}_{jk}$ in the sample, randomly generate an independent Poisson random vector $\mathbf{x}_{jk}$. Next compute the sample averages of the *NCs* of each type $\bar{\mathbf{x}}_k = (\bar{X}_1, ..., \bar{X}_I)^T$ and calculate the $\chi^2_{adj}(\bar{\mathbf{x}}_k)$ statistic for the sample using eqn. (20) using the parameters $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, $\boldsymbol{\tau}$ and $\mathbf{V}$. Repeat this procedure many times; for this study, 250,000 samples were generated for each sample size J in Table 3. The $UCL(\chi^2_{adj})$ is determined by selecting a probability of false alarm, α and then letting $UCL(\chi^2_{adj}) = [(1-\alpha)*100]^{th}$ percentile of the empirical distribution of $\chi^2_{adj}(\bar{\mathbf{x}}_k)$.

When $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, $\boldsymbol{\tau}$ and $\mathbf{V}$ are unknown, the empirical distribution of $T^2_{adj}(\bar{\mathbf{x}}_k)$ is generated to compute $UCL(T^2_{adj})$. The empirical distribution and subsequently the corresponding *UCL* is obtained using the same procedure as for $UCL(\chi^2_{adj})$ except that the estimates $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$, $\hat{\boldsymbol{\tau}}$ and $\hat{\mathbf{V}}$ are used in eqn. (22). For comparison, the *UCL* of T^2 statistic, $UCL(T^2)$, is also empirically computed using the estimates and eqn. (21).

Example 3: Computing $UCL(T^2)$ and $UCL(T^2_{adj})$.

Recall the example 1 in which two types of *NCs* are considered. We assumed the data in Table 1 was collected from an in-control condition. The estimates for the mean and variances of *NCs* of type 1 and 2 were calculated to be $\hat{\boldsymbol{\tau}} = (5.41, 2.12)^T$, $\hat{\mathbf{V}} = \begin{bmatrix} 21.32 & -2.43 \\ -2.43 & 6.44 \end{bmatrix}$ from the estimates $\hat{\boldsymbol{\mu}}_{MLE}$, $\hat{\boldsymbol{\Sigma}}_{MLE}$. Based on a sample of size, say three, i.e., three independent vectors of *MPLN* $\mathbf{x}_{jk}$ with parameters $\hat{\boldsymbol{\mu}}_{MLE}$, $\hat{\boldsymbol{\Sigma}}_{MLE}$ were generated for a single sample. The sample means $\bar{\mathbf{x}}_k = (\bar{X}_1, \bar{X}_2)^T$ were calculated and the T^2 and $T^2_{adj}(\bar{\mathbf{x}}_k)$ statistics were computed for the sample using eqn. (21) and (22) with the estimates $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$, $\hat{\boldsymbol{\tau}}$ and $\hat{\mathbf{V}}$. A total of 250,000 samples were generated with a total of 250,000 computed values of T^2 and $T^2_{adj}(\bar{\mathbf{x}}_k)$. Suppose α is chosen as 0.05 which is the 95^{th} percentile of the T^2 and $T^2_{adj}(\bar{\mathbf{x}}_k)$ representing the $UCL(T^2)$ and $UCL(T^2_{adj})$ were computed to be 6.52 and 5.79 respectively.

To use the *CC* on-line plot values of $T^2_{adj}(\bar{\mathbf{x}}_k)$ for each sample and if $T^2_{adj}(\bar{\mathbf{x}}_k)$ exceeds the control limit of 5.79 the T^2_{adj} *CC* signals indicating that one mean *NC* or a combination of mean *NCs* are different from the in-control target(s) $\hat{\boldsymbol{\tau}} = (5.41, 2.12)^T$. The process is stopped and investigated for an assignable cause; Once found it is repaired and the process continues. For example the data in Table 1 was grouped into 12 samples of size 3 and then the sample means and the T^2_{adj} statistics were calculated for each sample and plotted on the control chart with upper limit of 5.79 in figure 4. Figure 4 shows that sample number four is above the control limit and considered out-of-control.

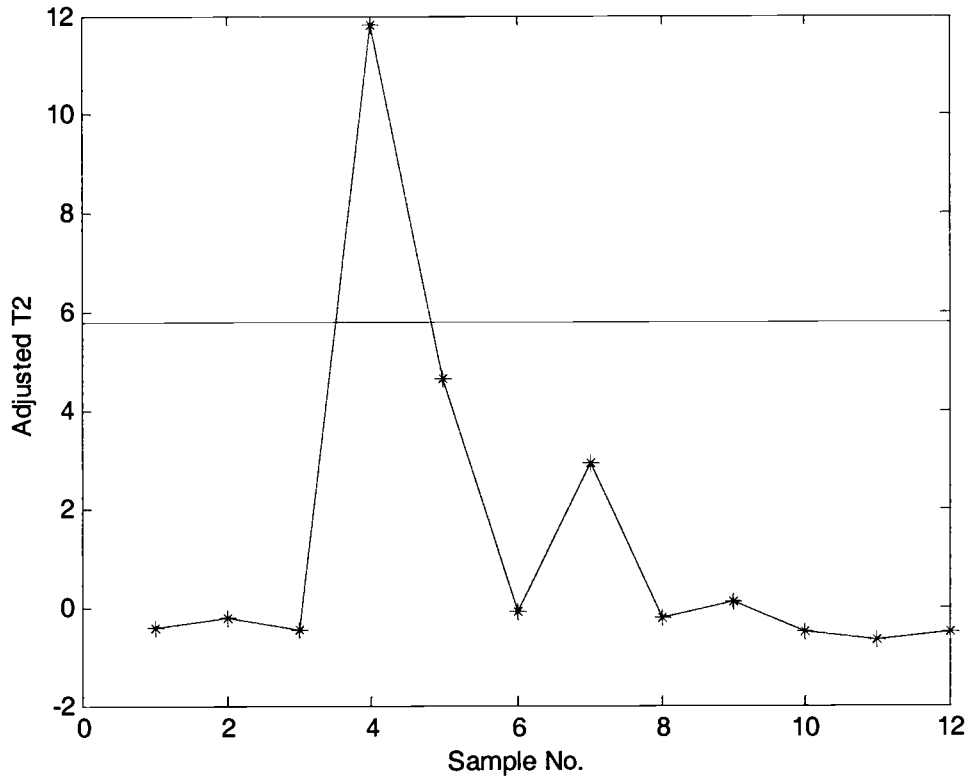

Figure 10: Control Chart for *NCs* – T^2_{adj}.

For α=0.05, Table 3 shows the values for the $UCL(T^2)$ and $UCL(T^2_{adj})$, along with the actual $\alpha's$, for various sample sizes. The actual values of α were computed as the proportion of 150,000 samples whose computed values, $T^2(\bar{\mathbf{x}}_k)$ and $T^2_{adj}(\bar{\mathbf{x}}_k)$, exceeded their corresponding control limits. Note that $T^2_{adj}(\bar{\mathbf{x}}_k)$ converges to the Hotelling T^2 statistic for large sample sizes. The simulation study shows that the $UCL(T^2)$ and $UCL(T^2_{adj})$ are similar for very large sample sizes, i.e., $J \geq 150$.

Table 6: The $UCL(T^2)$ and $UCL(T^2_{adj})$ for $\alpha = 0.05$ and Various Sample Sizes.

Sample Size	T^2 CC		T^2_{adj} CC	
J	$UCL(T^2)$	Actual α	$UCL(T^2_{adj})$	Actual α
1	6.49	0.0472	8.81	0.0603
2	6.68	0.0500	5.26	0.0502
3	6.52	0.0507	5.79	0.0508
4	6.39	0.0496	5.52	0.0502
5	6.34	0.0502	5.28	0.0510
10	6.14	0.0505	5.42	0.0501
15	6.10	0.0494	5.62	0.0493
20	6.03	0.0507	5.64	0.0507
25	6.08	0.0491	5.75	0.0493

ROBUSTNESS OF THE T^2_{adj} CONTROL CHART WHEN NCs ARE NOT DISTRIBUTED MPLN

Typically the distribution of $\boldsymbol{\lambda}_{jk}$ and $\mathbf{x}_{jk}$ may not be known exactly. It is possible that the distributional assumption of $\boldsymbol{\lambda}_{jk}$ and hence $\mathbf{x}_{jk}$ could be mis-specified or the distribution of $\boldsymbol{\lambda}_{jk}$ could change even though the target $\boldsymbol{\tau}$ could remain the same. We would like the *CC* to perform as desired whenever the means of the *NCs* are still on target. If *CC* less affected by the underlying assumption of $\boldsymbol{\lambda}_{jk}$ but is still able to perform as expected the *CC* is considered better or robust in some sense. In example 3, $UCL(T^2)$ and $UCL(T^2_{adj})$ were computed under the assumption that the

NCs were distributed *MPLN*. The question is, how well would the *MCC* for *NCs* with those limits perform if the *NCs* were not really distributed *MPLN*. The following sections show the results of such a robustness study for the T^2_{adj} *CC*.

Robustness of T^2_{adj} when $\boldsymbol{\lambda}$ are Not All Lognormal

Three cases are considered where the distribution of $\boldsymbol{\lambda}_{jk}$ are not multivariate lognormal but the means of $\mathbf{x}_{jk}$ remain as before, i.e., $\hat{\boldsymbol{\tau}} = (5.41, 2.12)^T$. The three cases are shown in Table 4. The first case considers the mean of *NC* type 1 as a constant and the mean of *NC* type 2 distributed as lognormal. The second case assumes the first mean is distributed lognormal but the second mean is a constant and the third case lets both means be constant. As a result, $\boldsymbol{\lambda}_{jk}$ are no longer distributed as *multivariate* lognormal. The third case assumes the *NCs*, between inspection units are independent Poisson *rvs* and Cases 1 and 2 make the *NC* types independent.

The reasoning behind choosing these types of changes for the parameters is that for a process the values of the parameters $\hat{\boldsymbol{\mu}}_{MLE}$ and $\hat{\boldsymbol{\Sigma}}_{MLE}$ are likely to change but if these changes do not translate into a change in the mean number of *NCs*, the *MCC* should not show an out-of-control condition. Table 5 shows the computed values for $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, $\boldsymbol{\tau}$, $\mathbf{V}$, and correlation matrix $\boldsymbol{\rho}$ of $\mathbf{x}_{jk}$.

Table 7: The Three Cases for λ_{jk}.

Cases	λ_{1jk}	λ_{2jk}
Case 1	Constant $\lambda_{1jk} = 5.41$	$\lambda_{2jk} \sim \log Normal\left(\hat{\mu}_2, \hat{\Sigma}_{22}\right)$
Case 2	$\lambda_{1jk} \sim \log Normal\left(\hat{\mu}_1, \hat{\Sigma}_{11}\right)$	Constant $\lambda_{2jk} = 2.12$
Case 3	Constant $\lambda_{1jk} = 5.41$	Constant $\lambda_{2jk} = 2.12$

Table 8: Parameters μ and Σ of λ_{jk} for Three Cases.

	MPLN	Case 1	Case 2	Case 3
μ	$\begin{pmatrix} 1.47 \\ 0.42 \end{pmatrix}$	$\begin{pmatrix} 1.69 \\ 0.42 \end{pmatrix}$	$\begin{pmatrix} 1.47 \\ 0.75 \end{pmatrix}$	$\begin{pmatrix} 1.69 \\ 0.75 \end{pmatrix}$
Σ	$\begin{bmatrix} 0.43 & -0.24 \\ -0.24 & 0.67 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 0.67 \end{bmatrix}$	$\begin{bmatrix} 0.43 & 0 \\ 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
τ	$\begin{pmatrix} 5.41 \\ 2.12 \end{pmatrix}$	$\begin{pmatrix} 5.41 \\ 2.12 \end{pmatrix}$	$\begin{pmatrix} 5.41 \\ 2.12 \end{pmatrix}$	$\begin{pmatrix} 5.41 \\ 2.12 \end{pmatrix}$
V	$\begin{bmatrix} 21.32 & -2.44 \\ -2.44 & 6.45 \end{bmatrix}$	$\begin{bmatrix} 5.41 & 0 \\ 0 & 6.45 \end{bmatrix}$	$\begin{bmatrix} 21.32 & 0 \\ 0 & 2.12 \end{bmatrix}$	$\begin{bmatrix} 5.41 & 0 \\ 0 & 2.12 \end{bmatrix}$
ρ	$\begin{bmatrix} 1 & -0.21 \\ -0.21 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

To assess the performance of the *MCCs*, the in-control *ARL* (ARL_0) is used, i.e., the average number of points plotted on the *MCC* between false alarms. The theoretical ARL_0 is computed as the reciprocal of the probability of a false alarm, i.e., $ARL_0 = 1/\alpha$. Theoretically for $\alpha = 0.05$ the T^2 and T^2_{adj} *MCCs* will have an $ARL_0 = 20$. This is called the *advertised* ARL_0. In practice, the ARL_0 may be different than the advertised. The ARL_0 one experiences during actual use of the *MCC* is called the *actual* ARL_0. To test the robustness of the T^2 and T^2_{adj} *MCCs* to changes in λ_{jk} while τ remains unchanged, the actual ARL_0 is compared to the advertised ARL_0 for both *MCCs*. The comparison is performed for the number of inspection units equal to $J = 1,..., 5, 10, 15, 20, 25$ and $\alpha = 0.05$ ($ARL_0 = 20$) for the three cases in Table 4. The control limits for the T^2 *MCC* from Table 3 are used when computing the ARL_0 values for both the T^2 and T^2_{adj} *MCCs*. A *MCC* is considered more robust to changes if its actual ARL_0 is closer to the advertised ARL_0. For example, to compute the actual ARL_0 for the T^2 *CC* with a sample of size five with $\alpha = 0.05$ for case 1, first randomly generate five independent vectors $\mathbf{x}_{1k}, \mathbf{x}_{2k}, \ldots, \mathbf{x}_{5k}$ from an *MPLN* using the parameters values from column two of Table 5. Each vector represents one inspection unit and the five vectors make up the single sample *k*. This is the same as, for a single sample, randomly generating five *i.i.d.* Poisson *rvs* with mean 5.41 and five independent Poisson *rvs* with random mean λ_{2jk} where λ_{2jk} is distributed *i.i.d.* lognormal with mean 0.42 and variance 0.67. The sample means

$\bar{\mathbf{x}}_k = \left(\bar{\bar{X}}_1, \bar{\bar{X}}_2\right)^T$ are calculated and the T^2 statistic is computed for the sample using eqn. (21) and the estimates $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$, $\hat{\boldsymbol{\tau}}$ and $\hat{\mathbf{V}}$ from example 3. A total of 150,000 samples are generated for a total of 150,000 values of T^2. The actual α is computed as the proportion of the 150,000 T^2 statistics that exceed the control limit 6.34; see Table 64 for $J = 5$. The process is repeated for cases two and three with their respective parameters.

Table 9: Case 1 – Comparison of Actual ARL_0 for T^2 and T^2_{adj} *MCCs.*

Sample Size	T^2 *CC*		T^2_{adj} *CC*	
J	Actual α	Actual ARL_0	Actual α	Actual ARL_0
1	0.0173	57.80	0.0557	**17.95***
2	0.0270	37.04	0.0412	**24.27***
3	0.0270	37.04	0.0376	**26.60***
4	0.0248	40.32	0.0281	**35.59***
5	0.0259	**38.61***	0.0234	42.74
10	0.0235	**42.55***	0.0206	48.54
15	0.0223	**44.84***	0.0201	49.75
20	0.0223	**44.84***	0.0202	49.50
25	0.0220	**45.45***	0.0202	49.50

Table 10: Case 2 – Comparison of Actual ARL_0 for T^2 and T^2_{adj} *MCCs.*

Sample Size	T^2 CC		T^2_{adj} CC	
J	Actual α	Actual ARL_0	Actual α	Actual ARL_0
1	0.0499	**20.04***	0.0610	16.39
2	0.0260	38.46	0.0442	**22.62***
3	0.0250	40.00	0.0451	**22.17***
4	0.0255	39.22	0.0394	**25.38***
5	0.0242	41.32	0.0324	**30.86***
10	0.0231	43.29	0.0261	**38.31***
15	0.0236	42.37	0.0246	**40.65***
20	0.0229	43.67	0.0248	**40.32***
25	0.0216	46.30	0.0230	**43.48***

Table 11: Case 3 – Comparison of Actual ARL_0 for T^2 and T^2_{adj} *MCCs.*

Sample Size	T^2 CC		T^2_{adj} CC	
J	Actual α	Actual ARL_0	Actual α	Actual ARL_0
1	0.0021	>400	0.0593	**16.86***
2	0.0004	>500	0.0240	**41.67***
3	0.0003	>500	0.0227	**44.05***
4	0.0002	>500	0.0106	**94.34***
5	0.0002	>500	0.0030	>300*
10	0.0002	>500	0.0003	>500
15	0.0001	>500	0.0001	>500
20	0.0001	>500	0.0002	>500
25	0.0001	>500	0.0001	>500

Table 6, Table 7, and Table 8 correspond to the three cases, respectively, and compare the actual ARL_0 for the T^2 and T^2_{adj} *CCs*. Values in bold and marked with an * indicate the *CC* whose actual ARL_0 closest to the advertised. From the results, it we see that the T^2_{adj} *CC* is nearer to the advertised $ARL_0 = 20$ and therefore quite robust for small sample sizes.

In Table 8 the ARL_0 for the T^2 *CC* is not robust for all sample sizes while the T^2_{adj} *CC* is very robust especially if it's applied to the individual observation. For the sample size of two and three both are not robust but the T^2_{adj} *MCC* has a significantly smaller difference between the actual and the advertised ARL_0. Considering the case of applying a *MCC* for an individual observation the T^2 *MCC* is only robust for case 2 but the propose T^2_{adj} *MCC* is very robust for all three cases.

Note that in all three cases the variance of $\mathbf{x}_{jk}$ are significantly smaller than the variance of $\mathbf{x}_{jk}$ in the *MPLN* case, i.e., in case 1 the variance of *NCs* of type 1 was significantly reduced from 21.32 to 5.41 which translates to an approximate decrease of 75%. In this situation the $\mathbf{x}_{jk}$ and accordingly the observed sample means $\bar{\mathbf{x}}_k$ will tend to cluster closer to the target mean $\hat{\boldsymbol{\tau}}$. When this occurs the T^2 and T^2_{adj} statistics tend toward zero. So we would expect any *MCC* to signal even less frequently than the advertised α. Moreover, this trend is intensified as the sample gets larger, i.e., the α delivered by the control chart decreases as J increases. Both the T^2 *MCC* and the

T^2_{adj} *MCC* have this property but mostly it is less severe with the T^2_{adj} *MCC* the T^2 *MCC.*

Robustness of T^2_{adj} when the variances of MPLN increases

Consider three cases where the distribution of $\boldsymbol{\lambda}_{jk}$ are multivariate lognormal or $\mathbf{x}_{jk}$ distribute *MPLN* but with higher variances while the means of $\mathbf{x}_{jk}$ remain as before. Again we want the *CC* to perform as desired whenever the means of the *NCs* are still on target. Consider the cases where both variances of *NCs* of type 1 and 2 were increased simultaneously by 50%, 75% and 100%. The three cases are shown in Table 9 along with the computed values for $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, $\boldsymbol{\tau}$, $\mathbf{V}$, and correlation matrix $\boldsymbol{\rho}$ of $\mathbf{x}_{jk}$.

Table 12: Parameters μ and Σ of λ_{jk} for Increasing Variances.

	MPLN	Case 4	Case 5	Case 6
μ	$\begin{pmatrix}1.47\\0.42\end{pmatrix}$	$\begin{pmatrix}1.36\\0.25\end{pmatrix}$	$\begin{pmatrix}1.32\\0.19\end{pmatrix}$	$\begin{pmatrix}1.28\\0.14\end{pmatrix}$
Σ	$\begin{bmatrix}0.43 & -0.24\\-0.24 & 0.67\end{bmatrix}$	$\begin{bmatrix}0.65 & -0.24\\-0.24 & 0.99\end{bmatrix}$	$\begin{bmatrix}0.74 & -0.24\\-0.24 & 1.11\end{bmatrix}$	$\begin{bmatrix}0.82 & -0.24\\-0.24 & 1.22\end{bmatrix}$
τ	$\begin{pmatrix}5.41\\2.12\end{pmatrix}$	$\begin{pmatrix}5.41\\2.12\end{pmatrix}$	$\begin{pmatrix}5.41\\2.12\end{pmatrix}$	$\begin{pmatrix}5.41\\2.12\end{pmatrix}$
V	$\begin{bmatrix}21.32 & -2.44\\-2.44 & 6.45\end{bmatrix}$	$\begin{bmatrix}31.99 & -2.44\\-2.44 & 9.67\end{bmatrix}$	$\begin{bmatrix}37.32 & -2.44\\-2.44 & 11.28\end{bmatrix}$	$\begin{bmatrix}42.64 & -2.44\\-2.44 & 12.89\end{bmatrix}$
ρ	$\begin{bmatrix}1 & -0.21\\-0.21 & 1\end{bmatrix}$	$\begin{bmatrix}1 & -0.14\\-0.14 & 1\end{bmatrix}$	$\begin{bmatrix}1 & -0.12\\-0.12 & 1\end{bmatrix}$	$\begin{bmatrix}1 & -0.10\\-0.10 & 1\end{bmatrix}$

To test the robustness of the T^2 and T^2_{adj} *MCCs* to changes in **V** while **τ** remains unchanged, the actual ARL_0 is compared to the advertised ARL_0 for both *MCCs* as in previous cases with the same control limits for the T^2 *MCC* from Table 64 and the same total number of samples 150,000. Table 10, Table 11, and Table 12, correspond to the three cases respectively and compare the actual ARL_0 for the T^2 and T^2_{adj} *CCs*. Values in bold and marked with an * indicate the *MCC* whose actual ARL_0 is closest to the advertised.

Table 13: Case 4 – Comparison of Actual ARL_0 for T^2 and T^2_{adj} *MCCs.*

Sample Size	T^2 *CC*		T^2_{adj} *CC*	
J	Actual α	Actual ARL_0	Actual α	Actual ARL_0
1	0.0702	14.25	0.0684	**14.62***
2	0.0812	12.32	0.0768	**13.02***
3	0.0860	11.63	0.0800	**12.50***
4	0.0898	11.14	0.0878	**11.39***
5	0.0918	**10.89***	0.0960	10.42
10	0.1040	**9.62***	0.1113	8.98
15	0.1102	**9.07***	0.1174	8.52
20	0.1148	**8.71***	0.1209	8.27
25	0.1163	**8.60***	0.1228	8.14

Table 14: Case 5 – Comparison of Actual ARL_0 for T^2 and T^2_{adj} *MCCs.*

Sample Size	T^2 *CC*		T^2_{adj} *CC*	
J	Actual α	Actual ARL_0	Actual α	Actual ARL_0
1	0.0765	13.07	0.0726	**13.77***
2	0.0928	10.78	0.0871	**11.48***
3	0.0972	10.29	0.0896	**11.16***
4	0.1029	9.72	0.1022	**9.78***
5	0.1089	**9.18***	0.1145	8.73
10	0.1268	**7.89***	0.1383	7.23
15	0.1367	**7.32***	0.1482	6.75
20	0.1459	**6.85***	0.1553	6.44
25	0.1493	**6.70 ***	0.1578	6.34

Table 15: Case 6 – Comparison of Actual ARL_0 for T^2 and T^2_{adj} *MCCs*.

Sample Size	T^2 *CC*		T^2_{adj} *CC*	
J	Actual α	Actual ARL_0	Actual α	Actual ARL_0
1	0.0824	12.14	0.0750	**13.33***
2	0.0994	10.06	0.0925	**10.81***
3	0.1088	9.19	0.0990	**10.10***
4	0.1154	8.67	0.1144	**8.74***
5	0.1225	**8.16***	0.1304	7.67
10	0.1469	**6.81***	0.1631	6.13
15	0.1612	**6.20***	0.1765	5.67
20	0.1759	**5.69***	0.1892	5.29
25	0.1783	**5.61** *	0.1909	5.24

Similarly in these three latter cases the variance of $\mathbf{x}_{jk}$ are larger than those of $\mathbf{x}_{jk}$ in the *MPLN* case. In this situation even though the $\mathbf{x}_{jk}$ is still distributed around the target mean $\hat{\boldsymbol{\tau}}$ the proportions of $\mathbf{x}_{jk}$ that are farther away from $\hat{\boldsymbol{\tau}}$ has increased along with the observed sample means $\bar{\mathbf{x}}_k$. This results in a considerably larger proportion of T^2 and the T^2_{adj} values exceeding the control limit. In this case, the *MCCs* will signal more frequently. Even though the changes in α for the T^2_{adj} *MCC* is less than that of the T^2 *MCC* for small sample sizes the performance does not seem significantly different.

Robustness of T^2_{adj} when $\boldsymbol{\lambda}_{jk}$ are Multivariate Exponential or Multivariate Gamma

To further test the robustness of the proposed *MMC* for *NCs*, suppose that instead of $\boldsymbol{\lambda}_{jk}$ following the multivariate lognormal distribution, $\boldsymbol{\lambda}_{jk}$ follow a multivariate exponential or multivariate gamma instead. We show that the proposed *MCC* is robust to such departures in the assumption of $\boldsymbol{\lambda}_{jk}$. Suppose the T^2 and the T^2_{adj} *MCCs* are tested with the control limits in Table 3 even though $\mathbf{x}_{jk}$ are distributed the multivariate exponential and multivariate gamma distribution for the parameters of a multivariate Poisson, *MPE* and *MPG* respectively. Since the values of $\hat{\boldsymbol{\tau}}$ computed based on *MPE* and *MPG* are very close to the target $\hat{\boldsymbol{\tau}}$ of *MPLN* in Table 3. We can assume that the actual unknown values of $\boldsymbol{\tau}$ based on *MPE,* and *MPG* are insignificantly different from $\hat{\boldsymbol{\tau}}$ of the *MPLN* in Table 3. Hence the $\mathbf{x}_{jk}$ we generate from the *MPE* and *MPG* are considered to be generated from an in-control condition.

The test procedure is, randomly generate 250,000 samples from both the *MPE* and *MPG* distributions. Compute T^2 and T^2_{adj} from eqns. (21) and (22) for each sample using $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$, $\hat{\boldsymbol{\tau}}$ and $\hat{\mathbf{V}}$. Find the actual α's as the proportion of T^2 and T^2_{adj} that exceed the corresponding control limit. Table 13 and Table 14 show the actual ARL_0 for both *MCCs*. Values in bold and marked with an * indicate the *MCC* whose actual ARL_0 is closest to the advertised. The results show that for small samples sizes, the proposed *MCC* is more robust to the two different distributional assumptions for $\boldsymbol{\lambda}_{jk}$.

Table 16: Comparison of Actual ARL_0 for T^2 and T^2_{adj} *MCCs* when $\mathbf{x}_{jk}$ are distributed *MPE.*

Sample Size	T^2 *CC*		T^2_{adj} *CC*	
J	Actual α	Actual ARL_0	Actual α	Actual ARL_0
1	0.0835	11.98	0.0742	**13.49***
2	0.0850	11.76	0.0815	**12.27***
3	0.0866	**11.55***	0.0877	11.40
4	0.0892	**11.21***	0.1016	9.84
5	0.0903	**11.08***	0.1089	9.18
10	0.0998	**10.02***	0.1141	8.76
15	0.1053	**9.50***	0.1133	8.83
20	0.1079	**9.27***	0.1144	8.74
25	0.1084	**9.23***	0.1129	8.85

Table 17: Comparison of Actual ARL_0 for T^2 and T^2_{adj} *MCCs* when $\mathbf{x}_{jk}$ are distributed *MPG.*

Sample Size	T^2 *CC*		T^2_{adj} *CC*	
J	Actual α	Actual ARL_0	Actual α	Actual ARL_0
1	0.0127	78.74	0.0499	**20.04***
2	0.0095	>100	0.0210	**47.62***
3	0.0075	>100	0.0199	**50.25***
4	0.0069	>100	0.0139	**71.94***
5	0.0066	>100	0.0086	**>100***
10	0.0052	>100	0.0068	**>100***

CONCLUSION AND DISCUSSION

The *MPLN* distribution was fittted to a real data set of bivariate *NCs* as well as the *MPE* and the *MPG*. The *MLE* was used to estimate the parameters of the *MPLN* which require the optimization procedure whereas the *MOM* estimate can be obtained readily and could be used because of its simplicity. To choose between the two estimates one may consider the total loglikelihood value as a statistical measure or subject to the practical conveniences. Nonetheless statistical property of the estimates chosen must be taken into account.

MCCs were developed for monitoring changes in mean *NCs* when the in-control means and variances of the *NCs* are unknown, the T^2_{adj} *CC*. The comparison study when *NCs* are not distribute *MPLN* shows that the T^2 *CC* has a false alarm rate not very close to the advertised for small sample size whereas the proposed *MCC* is. So this justifies the T^2_{adj} *CC* to be worth of the needed computation and hence recommended to use as the *MCC* for individual or small sample size.

The proposed statistic is invariant under affine transformation, i.e., $\mathbf{x}_{jk} \rightarrow A\mathbf{x}_{jk} + B$ for all invertible matrix A. The computed value of the T^2_{adj} statistic remains the same whether we works on $\mathbf{x}_{jk}$ or $A\mathbf{x}_{jk} + B$ and so is the control limit. The T^2_{adj} statistic can be deemed as independent of a particular scale in which the *rvs* are considered. So the T^2_{adj} statistics is, so to speak, a legitimate statistic.

The $T^2_{adj}(\bar{\mathbf{x}}_k)$ statistic can be both positive and negative while the χ^2 or T^2 statistics are nonnegative definite. It can be shown that the χ^2 or T^2 statistics is a measure of Euclidean distance $\mathbf{D}$ between $\bar{\mathbf{x}}_k$ and the in-control target $\boldsymbol{\tau}$, $\mathbf{D} = \bar{\mathbf{x}}_k - \boldsymbol{\tau}$. The further the difference or the distance the greater value of the χ^2 or T^2 statistics. The value of T^2_{adj} statistic decreases in some case even the difference $\mathbf{D}$ increases. In fact the T^2_{adj} statistic does not only measure the Euclidean distance as the T^2 statistics does. It measures both the Euclidean difference $\mathbf{D}$ and the deviation of the observation $\bar{\mathbf{x}}_k$ from the higher moment targets. Thus the T^2_{adj} statistic is not actually the Euclidean measure but rather a distributional measure, in some sense, that account for the deviation from the target distributional moments.

Finally this study shows that the T^2_{adj} *CC* is superior than the T^2 *CC* for the bivariate case and a specific set of parameters using $\hat{\boldsymbol{\mu}}$, $\hat{\boldsymbol{\Sigma}}$ and hence worth the extra computation. Even though generalization of the results to an arbitrarily case may or may not be appropriate we believe that the T^2_{adj} *CC* is superior to the T^2 *CC* for small sample sizes. Nonetheless further investigation on a wider range of parameter and more general type of *NC* is needed.

REFERENCES

Aitchison, J., and Ho, C. H. (1989), "The multivariate Poisson-log normal distribution", Biometrika, 76, 643-653.

Ajmani, Vivek , Randles, Ronald , Vining, Geoff , and Woodall, William (1997), "Robustness of multivariate control charts", ASA Proceedings of the Section on Quality and Productivity, 152-157.

Alt, F. B (1985), "Multivariate Quality Control" in Encyclopedia of Statistical Sciences 6, edited by S. Kotz and N.L. Johnson, John Wiley&Sons Inc.

Casella, George, Berger, L. Roger (1990), "Statistical Inference", Duxbury Press.

Gumbel, E. J. (1960), "Bivariate exponential distributions", Journal of the American Statistical Association, 55, 698-707.

Gupta, A. K., and Wong, C. F. (1984), "On a Morgenstern-type bivariate gamma distribution", Metrika, 31, 327-332.

Jackson, J. E., Morris, R. H. (1957), "An application of Multivariate Quality Control to Photographic Processing", Journal of American Statistical Association, 52, 186-199.

Jackson, J. E (1985), "Multivariate Quality Control", Communications in Statistics, Theory and Method, 14, 2657-2688.

Johnson, R. A., and Wichern, D. W. (1988), "Applied multivariate statistical analysis", Prentice-Hall Inc (Englewood Cliffs, NJ).

Johnson, Norman Lloyd, Kotz, Samuel, and Balakrishnan, N. (1997), "Discrete multivariate distributions", John Wiley & Sons Inc.

Kendall, Maurice, Stuart, Alan, and Ord, J. Keith (1991), "Kendall's advanced theory of statistics Volume 2: Classical inference and relationship (Fifth edition)", Oxford University Press.

Kotz, Samuel, Balakrishnan, N., Johnson, N. L (2000), "Continuous multivariate distributions", John Wiley & Sons Inc.

Lowry, Cynthia A., and Montgomery, Douglas C. (1995), "A review of multivariate control charts", IIE Transactions, 27, 800-810.

Lu, X. S., Xie, M., Goh, T. N., and Lai, C. D. (1998), "Control Chart for Multivariate Attribute Processes", International Journal of Production, 36, 3477-3489.

Mason, Robert L., Young, John C. (1998), "Why Multivariate Statistical Process Control", Quality Progress, 31, 88-93.

McCullagh, P. (1987), "Tensor methods in statistics", Chapman & Hall Ltd.

Montgomery, Douglas C. (2001), "Introduction to statistical quality control", John Wiley & Sons Inc.

Nelson, James F. (1985), "Multivariate gamma-Poisson models", Journal of the American Statistical Association, 80, 828-834.

Patel, H. I. (1973), "Quality Control Methods for Multivariate Binomial and Poisson Distributions", Technometrics, 15, 103-112.

Steyn, H. S. (1976), "On the multivariate Poisson normal distribution", Journal of the American Statistical Association, 71, 233-236.

GENERAL CONCLUSION

We proposed multivariate control charts for a batch process in which the quality of the finished product is determined by how large the number and types of *NCs* are on the product. The weaving process for wire mesh is used as a motivating example in which a large number of *NC* types with different *NC* severities can determine the overall quality of the product.

The multivariate control chart proposed in the form of the demerit control chart is used when the types of *NC* are independent. Demerit control charts in the past have rarely received attention. If it is desired that a large number of *NC* types are to be statistically monitored with a minimal number of control charts then the demerit control chart is one of the right control charts to use. If separate control charts are used, the number of control charts becomes unmanageable for most practical situations and using such a large number of control charts will likely be ineffective and instead of being cost saving could be an expense. The demerit control chart, if designed properly, can really enhance control chart usage, especially in a shop floor level, because of its simplicity and ease of use. To generalize the demerit control chart to any number and types of *NC* two characteristics of the demerit control chart must be addressed; how to determine the control limits for general classification schemes and how to determine the weights for that scheme.

To generalize the demerit control chart we proposed an integrated methodology that both determines the weights and computes the control limits for an

arbitrary classification scheme. The asymptotic technique developed for determining the weights was able to achieve sensible and basic properties since it assigns more weight to the more severe types of *NCs* which is determined by drawing out the direction of the shifts that might occur in an out-of-control condition. That direction is then adjusted by the variances of the number of *NCs* since we normally want the type of *NCs* with larger variance to receive comparatively less weight. The asymptotic technique considers the direction of the shift rather than the magnitude of the shift.

The distribution approximation technique called an Edgeworth Expansion is proposed to construct the control limits of the demerit control chart that has an arbitrary weighting scheme. As expected, in general, the Edgeworth Expansion technique provides a more accurate estimation than the normal distribution approximation and gives more accurate results when determining the characteristics and properties of the control chart even for small sample sizes.

When the *NCs* are correlated we presented a methodology to first model the *NCs* and then construct a multivariate control chart, the adjusted T^2 control chart. In the batch process of the type considered here it is sensible to assume that the quality of the product may be different in some degree among batches. It is also reasonable to assume that the mean number of *NCs* even for the same *NC* type could be varying among batches to some extent yet still be acceptable product. The hierarchical modeling technique adopted here not only provides a general tool for modeling correlated *NCs* with a general correlation structure but also accommodates the batch-

quality dependent scenario. The hierarchical modeling is likely to result in an implicit joint marginal distribution for the *NCs* whose parameter estimation is usually implicit. When modeling correlated NCs the multivariate Poisson Lognormal distribution has the advantage that there is an explicit and straightforward way to estimate the parameters using method of moment estimators. If a large amount of in-control data is available for fitting the *NCs* to the multivariate Poisson Lognormal distribution then the method of moments works well and is comparatively computationally easier than maximum likelihood estimation.

A test statistic for the the adjusted T^2 control chart was developed using the differences between the sample means of the *NCs* and pre-defined targets for the *NCs*. The deviation of these differences is translated in terms of the density function value of the sample means. A Multivariate Edgeworth Expansion is used to approximate the joint distribution of the sample means and allows us to determine the test statistic. The test statistic is a function of the joint sample means with an adjustment factor for higher order joint moments. A single upper control limit for the adjusted T^2 control chart is obtained empirically. For small samples, especially for individual observations, simulation results showed that the proposed control chart is robust to the distributional assumption chosen for the correlated *NCs*. The proposed statistic is well approximated by a chi-square distribution approximation. The proposed control chart performs well and is recommended for the individual observation or for small samples.

BIBLIOGRAPHY

Aitchison, J., and Ho, C. H. (1989), "The multivariate Poisson-log normal distribution", Biometrika, 76, 643-653.

Ajmani, Vivek, Randles, Ronald, Vining, Geoff, and Woodall, William (1997), "Robustness of multivariate control charts", ASA Proceedings of the Section on Quality and Productivity, 152-157.

Alt, F. B (1985), "Multivariate Quality Control" in Encyclopedia of Statistical Sciences 6, edited by S. Kotz and N.L. Johnson, John Wiley&Sons Inc.

Casella, George, Berger, L. Roger (1990), "Statistical Inference", Duxbury Press.

Dodge, Harold F. and Torrey, M. N. (1977), "A check inspection and demerit rating plan (Reprint from Industrial Quality Control)", Journal of Quality Technology, 9, 146-153.

Gumbel, E. J. (1960), "Bivariate exponential distributions", Journal of the American Statistical Association, 55, 698-707.

Gupta, A. K., and Wong, C. F. (1984), "On a Morgenstern-type bivariate gamma distribution", Metrika, 31, 327-332.

Jackson, J. E., Morris, R. H. (1957), "An application of Multivariate Quality Control to Photographic Processing", Journal of American Statistical Association, 52, 186-199.

Jackson, J. E (1985), "Multivariate Quality Control", Communications in Statistics, Theory and Method, 14, 2657-2688.

Johnson, Norman Lloyd, Kotz, Samuel, and Balakrishnan, N. (1997), "Discrete multivariate distributions", John Wiley & Sons Inc.

Johnson, R. A., and Wichern, D. W. (1988), "Applied multivariate statistical analysis", Prentice-Hall Inc (Englewood Cliffs, NJ).

Jones, Allison L., Woodall, William H., and Conerly, Michael D. (1999), "Exact properties of demerit control charts", Journal of Quality Technology, 31, 207-216.

Kendall, Maurice, Stuart, Alan, and Ord, J. Keith (1991), "Kendall's advanced theory of statistics Volume 2: Classical inference and relationship (Fifth edition)", Oxford University Press.

Kolassa, John Edward (1997), "Series approximation methods in statistics", Springer-Verlag Inc., New York.

Kotz, Samuel, Balakrishnan, N., Johnson, N. L (2000), "Continuous multivariate distributions", John Wiley & Sons Inc.

Lehmann, E. L. (1997), Testing Statistical Hypotheses, 2nd edition, Springer.

Lowry, Cynthia A., and Montgomery, Douglas C. (1995), "A review of multivariate control charts", IIE Transactions, 27, 800-810.

Lu, X. S., Xie, M., Goh, T. N., and Lai, C. D. (1998), "Control Chart for Multivariate Attribute Processes", International Journal of Production, 36, 3477-3489.

Mason, Robert L., Young, John C. (1998), "Why Multivariate Statistical Process Control", Quality Progress, 31, 88-93.

McCullagh, P. (1987), "Tensor methods in statistics", Chapman & Hall Ltd.

Montgomery, Douglas. C. (1996), "Introduction to Statistical Quality Control", 3rd edition, John Wiley & Sons, New York.

Montgomery, Douglas. C. (2001), "Introduction to Statistical Quality Control", John Wiley & Sons, New York.

Nelson, James F. (1985), "Multivariate gamma-Poisson models", Journal of the American Statistical Association, 80, 828-834.

Patel, H. I. (1973), "Quality Control Methods for Multivariate Binomial and Poisson Distributions", Technometrics, 15, 103-112.

Steyn, H. S. (1976), "On the multivariate Poisson normal distribution", Journal of the American Statistical Association, 71, 233-236.

Stuart, Alan, and Ord, J. Keith (1994), "Kendall's Advanced Theory of Statistics. Volume 1. Distribution theory (Sixth edition)", Edward Arnold Publishers Ltd., London.

Vaart, van der A. W. (1998), "Asymptotic statistics", Cambridge University Press, Cambridge, UK.