QUALITY AND RELIABILITY ENGINEERING INTERNATIONAL
Qual. Reliab. Engng. Int. 2005; **21**:177–195
Published online 27 January 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/qre.614

Research

Ordered Samples Control Charts for Ordinal Variables

Fiorenzo Franceschini*,†, Maurizio Galetto and Marco Varetto
Dipartimento di Sistemi di Produzione ed Economia dell'Azienda, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy

The paper presents a new method for statistical process control when ordinal variables are involved. This is the case of a quality characteristic evaluated by an ordinal scale. The method allows a statistical analysis without exploiting an arbitrary numerical conversion of scale levels and without using the traditional sample synthesis operators (sample mean and variance). It consists of a different approach based on the use of a new sample scale obtained by ordering the original variable sample space according to some specific 'dominance criteria' fixed on the basis of the monitored process characteristics. Samples are directly reported on the chart and no distributional shape is assumed for the population (universe) of evaluations. Finally, a practical application of the method in the health sector is provided. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: ordered samples control charts; ordinal variables; linguistic variables; ordinal scales; quality monitoring; service quality; dominance criteria

1. INTRODUCTION

Many quality characteristics are evaluated on linguistic or ordinal scales. This is the case when performing visual controls on manufactured products or when evaluating some characteristics of the quality of a service. The levels of these scales are terms such as 'good', 'bad', 'medium', etc., which can be ordered according to the specific meaning of the quality characteristic at hand. Ordered linguistic scales mainly differ from numerical or cardinal scales because the concept of distance is not defined. The ordering is the main property associated to such scales[1,2].

The problem of on-line monitoring of an ordinal quality characteristic requires the development of techniques able to deal with ordinal data. The assignment of weights, demerits and so on, to reflect the degree of severity of product non-conformity, has been adopted in many circumstances[3,4]. Different numbers of demerits are assigned to each class and the total number of demerits is monitored by some control chart for defectives. This is a *subjective* approach that requires the ability to uniquely classify each state into one of several mutually exclusive classes, with well-defined boundaries among them. Although the numerical conversion of verbal information simplifies the subsequent analysis, it also gives rise to two basic problems. The first is concerned with the validity of encoding a discrete verbal scale into a numerical form. The numerical codification implies fixing the distances among scale levels, thus converting the ordinal scale into a cardinal one. The second is related to the absence of consistent criteria for the selection of the type of numerical conversion. It is obvious that changing the numerical

*Correspondence to: Fiorenzo Franceschini, Dipartimento di Sistemi di Produzione ed Economia dell'Azienda, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy.
†E-mail: fiorenzo.franceschini@polito.it

Received 14 June 2002
Revised 3 September 2003

Table I. Results of the visual control of a sample of 30 corks

Reject	Poor quality	Medium quality	Good quality	Excellent quality
2 corks	5 corks	9 corks	7 corks	7 corks

encoding may determine a change in the obtained results. In this way the problem analyst directly influences the acceptance of results. Therefore, any conclusions drawn from the analysis on 'equivalent' numerical data could be partially or wholly distorted.

Consider, for example, the case concerning a production line of fine liqueurs, reported in Franceschini and Romano[5]. The visual control of the corking and closing process is carried out on the basis of the following assessments (see Table I):

- 'reject' if the cork does not work;
- 'poor quality' if the cork must not be rejected but has some defects;
- 'medium quality' if the cork has relevant aesthetic flaws but no other defects;
- 'good quality' if the cork only has small aesthetic flaws;
- 'excellent quality' if the cork is perfect.

Suppose we decide to introduce the following codification:

- 'reject' = 1;
- 'poor quality' = 2;
- 'medium quality' = 3;
- 'good quality' = 4;
- 'excellent quality' = 5.

Referring to the example in Table I, the resulting arithmetic mean is $\bar{x} = 3.4$. Hence, the sample mean seems to be between 'medium quality' and 'good quality' (nearer to the former than to the latter).

The adopted numerical conversion is based on the implicit assumption that all scale levels are equispaced. However, we are not sure that the evaluator perceives the subsequent levels of the scale as equispaced, nor even if s/he has been preliminarily trained. For example, the evaluator might perceive the upper levels as more distinguished from the others. The suitable codification of the levels of the scale for this inspector might be the following[1]:

- 'reject' = 1;
- 'poor quality' = 3;
- 'medium quality' = 9;
- 'good quality' = 27;
- 'excellent quality' = 81.

In this case the arithmetic mean is $\bar{x} = 28.5$, that is to say that the sample mean is between 'good quality' and 'excellent quality' (nearer to 'good quality').

We cannot say which is the right value of the sample mean at hand because an 'exact' codification does not exist.

Let us consider another example. Four masses (a, b, c and d) are compared by means of a two-plate balance. Their values are unknown and only the following relationships have been proven:

$$a > b > c > d \tag{1}$$

and

$$a > b + c + d \tag{2}$$

If we try to introduce a linear codification on the basis of Equation (1), for example,

$$a = 4, \quad b = 3, \quad c = 2, \quad d = 1$$

it is easy to demonstrate that Equation (2) is never verified. In fact, if Δx is the scale unit, we can write

$$c = d + \Delta x$$
$$b = c + \Delta x = d + 2 \cdot \Delta x$$
$$a = b + \Delta x = d + 3 \cdot \Delta x$$

Substituting into Equation (2), we obtain

$$a > b + c + d$$
$$\Downarrow$$
$$d + 3 \cdot \Delta x > d + 2 \cdot \Delta x + d + \Delta x + d$$
$$\Downarrow$$
$$d + 3 \cdot \Delta x > 3 \cdot d + 3 \cdot \Delta x$$

which generates an incongruence.

These two examples point out that a simple codification of scale levels could result in a misrepresentation of the original gathered information. A correct approach should be based on the usage of the only properties of ordinal scales themselves.

The main aim of the present paper is to propose a new method for on-line process control of a quality characteristic evaluated on an ordinal scale, without exploiting an artificial conversion of scale levels. Other approaches, based on the use of the so-called linguistic control charts, have already been presented[5–7]. They emulate traditional Shewart control charts, making use of two-sample synthesis operators: one for a measure of central tendency and the other for a measure of variability. The new proposal does not consider these synthesis operators. It allows on-line monitoring based on a new *process sample scale* obtained by ordering the original variable sample space according to some specific ‘dominance criteria’. Samples are directly reported on the chart and no distributional shape is assumed for the population (universe) of evaluations.

First, the paper describes the new methodology; next, it compares the method with other approaches. Finally, a practical application of the method in the health sector is provided.

2. PRELIMINARY CONSIDERATIONS

The sample space of a generic ordinal quality characteristic is not ordered in nature. However, samples can be ordered according to some specific dominance criteria. A dominance criterion allows attributing a position in the ordered sample space to each sample. If sample B dominates sample A, then sample A has a lower position in the ordering.

For each pair of samples a dominance criterion states a dominance or an equivalence relationship. If the *resolution* of the dominance criterion is high, the dimension of *equivalence classes* is very small (i.e. few samples have the same position in the ordered sample space). The most resolving criterion is the one assigning a different position to each ordered sample. This is the same as saying that every equivalence class has only one element.

Let us consider the following example. An operator evaluates the quality of a surface finishing of a mechanical part by a visual control. Every hour a sample of four parts is analyzed. Evaluations are given on a three-level scale: ‘High’, ‘Medium’ and ‘Low’. Table II shows the results of ten subsequent samples.

Table II. Results of a visual control of the quality of a surface finishing of 10 subsequent samples of a mechanical part. Every hour a sample of four parts is analyzed. Evaluations are given on a three-level scale: 'High' (H), 'Medium' (M) and 'Low' (L)

Sample number	First part	Second part	Third part	Fourth part
1	*H*	*H*	*M*	*H*
2	*H*	*M*	*H*	*H*
3	*H*	*M*	*M*	*H*
4	*H*	*H*	*M*	*L*
5	*H*	*M*	*M*	*H*
6	*M*	*H*	*M*	*M*
7	*H*	*M*	*H*	*M*
8	*L*	*L*	*M*	*L*
9	*M*	*H*	*H*	*H*
10	*H*	*H*	*H*	*H*

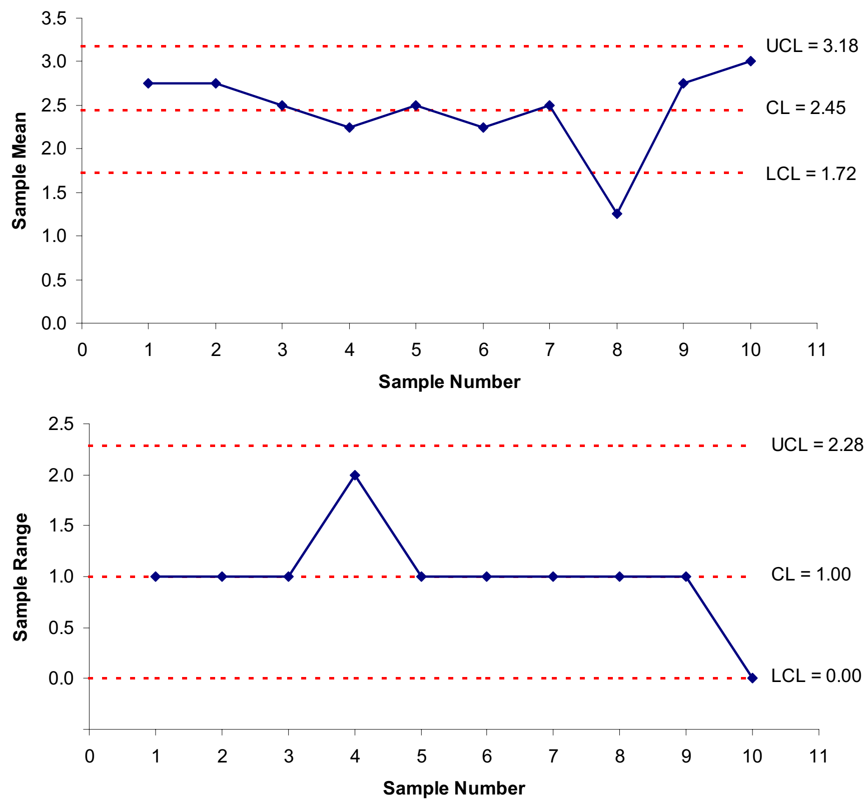

Figure 1. $\bar{X}$–R control charts for the example of visual control of the quality of a surface finishing of a mechanical part. The levels of the ordinal evaluation scale are numerically coded ('Low' = 1; 'Medium' = 2; 'High' = 3)

Can we use the results of Table II to build a process control chart? The first *classical* answer to this question is the assignment of a specific numerical value to each level of the evaluation scale. A possible codification could be the following:

$$\text{'Low'} = 1; \quad \text{'Medium'} = 2; \quad \text{'High'} = 3$$

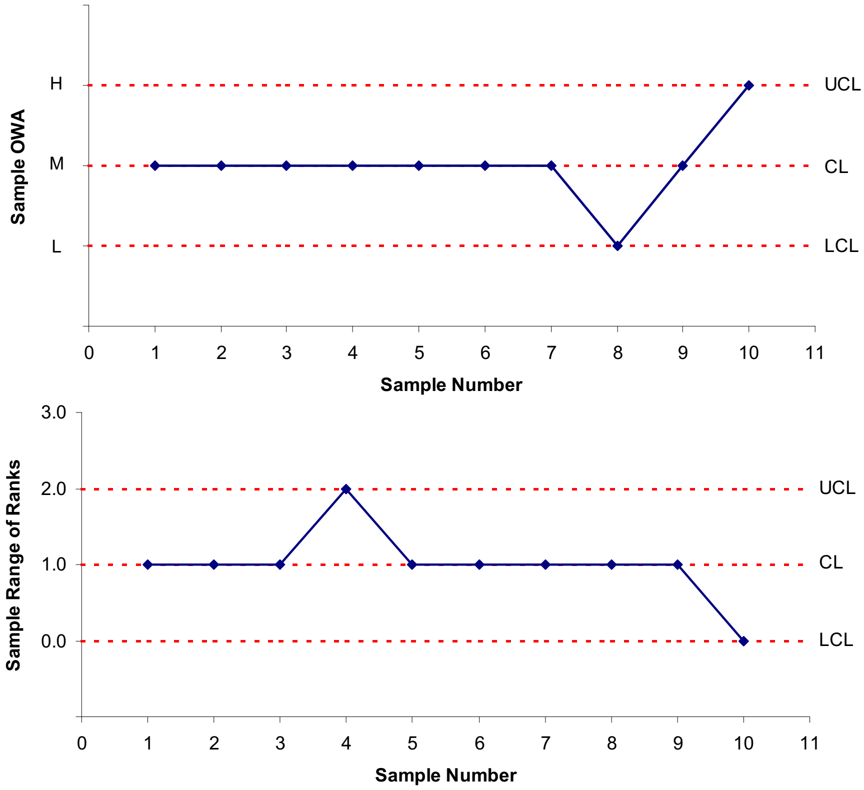

Figure 2. Linguistic control charts for the OWA emulator of arithmetic mean and the range of ranks of data reported in Table II[5]

The codification allows building *traditional* $\bar{X}$–R control charts. However, as anticipated, this procedure has three main contraindications. First, each conversion is arbitrary and different codifications can lead to different results. Second, codification introduces the concept of distance among scale levels, which is not originally defined. Third, since the original distribution of evaluations is discrete with a very small number of levels, the central limit theorem hardly applies to this context[8]. Figure 1 reports $\bar{X}$–R control charts for the example at hand. A classical 3σ couple of control limits is provided.

A second analysis of data in Table II can be executed by the method suggested by Franceschini and Romano[5]. This methodology is based on the use of operators that do not require the numerical codification of ordinal scale levels. The adopted location measure is the ordered weighted average (OWA) emulator of arithmetic mean (see Appendix A), firstly introduced by Yager and Filev[2,9]. The OWA operator can take values only in the set of levels of the original ordinal scale. The related control chart is built following a methodology very similar to the traditional chart for mean values. The adopted dispersion measure is the range of ranks r_S, defined as the total number of levels contained between the maximum and the minimum value of a sample (the rank $r(q)$ is the sequential integer number of a generic level q on a linguistic scale):

$$r_S = [r(q)_{\max} - r(q)_{\min}]$$

For the range of ranks too, the related control chart is constructed using the traditional approach. Figure 2 shows the control charts for the OWA and the range of ranks of data reported in Table II.

Although this methodology does not exploit the device of codification, the dynamics of the charts are poor and little information can be extracted about the process. Moreover, the method is not free from distributional assumptions. The dispersion measure assumes that the scale ranks do not depend on the position of levels of the ordinal variable.

In this paper we propose a third way of analyzing data reported in Table II. It exploits the only properties of ordinal scales, avoiding the synthesis of information contained in the sample. No distributional assumptions are required about the population (universe) of evaluations.

As traditional control charts, this new methodology is based on the use of two different charts: one for *ordered sample values*, and the other for *ordered sample ranges*. These charts provide different performance analysis of the ordinal quality characteristic at hand. As a consequence, they can be built and used separately. However, for an exhaustive analysis, a conjoint approach is highly recommended.

3. ORDERED SAMPLES CONTROL CHARTS

The new proposal is based on the ordering of the sample space of an ordinal quality characteristic. We introduce this concept by a simple example.

Let us consider the following ordered samples, defined on a three-level ordinal scale ('High' (H), 'Medium' (M) and 'Low' (L)):

- sample A: $\{H, M, M\}$;
- sample B: $\{H, H, L\}$;
- sample C: $\{M, M, M\}$.

To compare and order these samples we introduce a rule called 'dominance criterion', defined, case by case, on the basis of the characteristics of the monitored process. In accordance with this rule, if sample A dominates sample B, then sample A is preferred to sample B. As a result we can define a new ordinal scale whose levels are the positions of the samples in the ordered sample space. If there is no dominance relationship between sample A and sample B, they belong to the same 'equivalence class'.

The choice of the dominance criterion influences the resolution of the scale (i.e. the number of levels of the ordered sample space) and also the order of levels. For each process one or more dominance criterion may be established on the basis of the specific application.

In the following a series of three intuitive dominance criteria will be analyzed.

We begin analyzing the *Pareto-dominance criterion*. We state that sample X Pareto-dominates sample Y if all elements in Y do not exceed the corresponding elements in X, and at least one element in X exceeds the corresponding one in Y. This situation is formally denoted by $\mathrm{X} \triangleright \mathrm{Y}$.

In case samples X and Y belong to the same equivalence class, i.e. no dominance relationship can be defined between them, we use the following notation: $\mathrm{X} \approx \mathrm{Y}$. Referring to our example, we have

$$\mathrm{A} \triangleright \mathrm{C}; \quad \mathrm{A} \approx \mathrm{B}; \quad \mathrm{B} \approx \mathrm{C}$$

Sample A dominates sample C, while samples A and B belong to the same equivalence class, as well as samples B and C. Figure 3 represents these results; an arrow denotes a Pareto-dominance relationship and each circle defines an equivalence class.

As we can see from Figure 3, it is not possible to assign a well-defined position to samples A, B and C, because their intersection is not empty. The problem can be solved by introducing the concept of '*semi-equivalence class*'. A semi-equivalence class is composed of equivalence classes whose intersections are not empty. Samples in Figure 3 belong to the same semi-equivalence class.

In general, the Pareto-dominance criterion gives a 'poor' ordering for the sample space of an ordinal quality characteristic. A more discerning criterion is the '*rank dominance criterion*'. Its introduction requires the definition of the concept of 'optimal sample'. A sample is said to be optimal if all elements assume the highest level of an ordinal scale. In our example the optimal sample is *HHH*.

For each sample we define a *rank index* which quantifies its *positioning* with regard to the optimal sample. The index is built in by adding up the numbers of scale levels contained between each sample value and the corresponding value of the optimal sample. For example, if we consider *HML*, its rank index

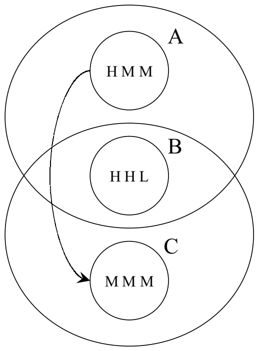

Figure 3. An application of the Pareto-dominance criterion. The arrow represents a Pareto-dominace relationship. Each circle defines an equivalence class

Table III. Ordered samples for a sample space defined by a three-level scale (L = 'Low'; M = 'Medium'; H = 'High') and a sample size $n = 4$. Column 3 reports the positions of each sample after the application of the rank dominance criterion. Column 4 shows the position of each sample after the sequential application of the rank and the dispersion dominance criteria

Sample space	Rank index	Position in the ordered sample space (equivalence class) [*rank dominance criterion*]	Position in the ordered sample space (equivalence class) [*rank and dispersion dominance criterion*]
LLLL	8	1st	1st
MLLL	7	2nd	2nd
MMLL	6	3rd	4th
MMML	5	4th	6th
MMMM	4	5th	9th
HLLL	6	3rd	3rd
HMLL	5	4th	5th
HMML	4	5th	8th
HMMM	3	6th	11th
HHLL	4	5th	7th
HHML	3	6th	10th
HHMM	2	7th	13th
HHHL	2	7th	12th
HHHM	1	8th	14th
HHHH	0	9th	15th

(with regard to *HHH*) is 3, obtained by adding up 0, 1 and 2 (i.e. zero levels between H and H, one level between H and M, and two levels between H and L).

A high value of the rank index corresponds to a 'bad' sample. All samples that are characterized by the same index belong to the same equivalence class. Therefore their positioning with respect to the optimal sample can be equivalently identified by the corresponding equivalence class. The number of elements of the new ordinal sample scale depends on the sample size and on the number of levels of the evaluation scale.

Denoting by t the number of levels of the evaluation scale and by n the sample size, the rank dominance criterion gives a number of equivalence classes equal to $n(t - 1) + 1$. This is also the number of levels of the resulting ordinal scale of sample positions. Table III (first column) reports all possible ordered samples of size $n = 4$, on an evaluation scale with $t = 3$ levels. For each sample, the corresponding position on the resulting scale is reported (third column). The greater the position number, the higher the sample evaluation.

A greater resolution, i.e. a larger number of levels, on the ordinal sample scale can be obtained by integrating the rank dominance criterion with the '*dispersion dominance criterion*'. This criterion allows distinguishing among samples belonging to the same equivalence class by analyzing sample dispersion. A lower position is

Table IV. Number of equivalence classes (i.e. number of scale levels of the ordered sample space) for different sample sizes (n) and different numbers of scale levels (t). Results are obtained after the sequential application of the rank and the dispersion dominance criteria

n	t	Number of equivalence classes
2	3	6
	5	15
	7	28
	9	45
3	3	10
	5	35
	7	84
	9	165
4	3	15
	5	70
	7	210
	9	495

associated with a sample with a greater dispersion. The fourth column of Table III reports the position of each sample in the new ordered sample space after the sequential application of the rank and the dispersion dominance criteria. As we can see, each ordered sample is associated with a different position; this is the greatest possible resolution.

Table IV shows the number of equivalence classes, for different sample sizes (n), and different numbers of scale levels (t). The three dominance criteria introduced are consistent. A richer dominance criterion splits the equivalence or semi-equivalence classes given by the poorer criteria, refining the order of the sample space.

Figure 4 reports the results of the subsequent application of the three dominance criteria to a three-level evaluation scale with samples of three elements. A vertical arrow represents a dominance relationship. A continuous ellipse represents an equivalence class, while a dashed ellipse represents a semi-equivalence class.

Figures 4(a)–(c) report respectively the results of the application of the Pareto-dominance, the rank-dominance and the rank plus dispersion dominance criteria. The transverse arrows describe how each equivalence or semi-equivalence class is split by the application of a more discerning dominance criterion. The resolution of the ordered sample space varies with the considered dominance criterion.

In accordance with a specific dominance criterion, sample charts report the positions of samples in the ordered sample space on the vertical axis. Given the particular meaning of sample charts, only the lower control limit (UCL) is defined. The central line (CL) represents the median of sample distribution. A set of initial samples is considered to determine the sample empirical frequency distribution. This empirical distribution is then used to calculate the lower control limit for a given type I error[4]. Control limits are determined by empirical estimates of probabilities based on observed frequencies in a set of initial samples. Therefore, because the probabilities are estimated, the estimates contain errors, which could become significant for very small probabilities. A large initial set of samples or an alternative approach based on *bootstrap* techniques[10] are needed to estimate the limits with a more reasonable accuracy.

Figure 5 represents the sample control charts and the corresponding sample frequency distributions of data reported in Table II, after the sequential application of the Pareto (a), the rank (b) and the rank plus dispersion (c) dominance criteria.

Ordered sample charts are rich in insights from different points of view. The resolution of charts increases while shifting from case (a) to cases (b) and (c). The information of depicted charts of Figure 5 are strictly connected; this can be observed on analyzing peaks and valleys in the same positions.

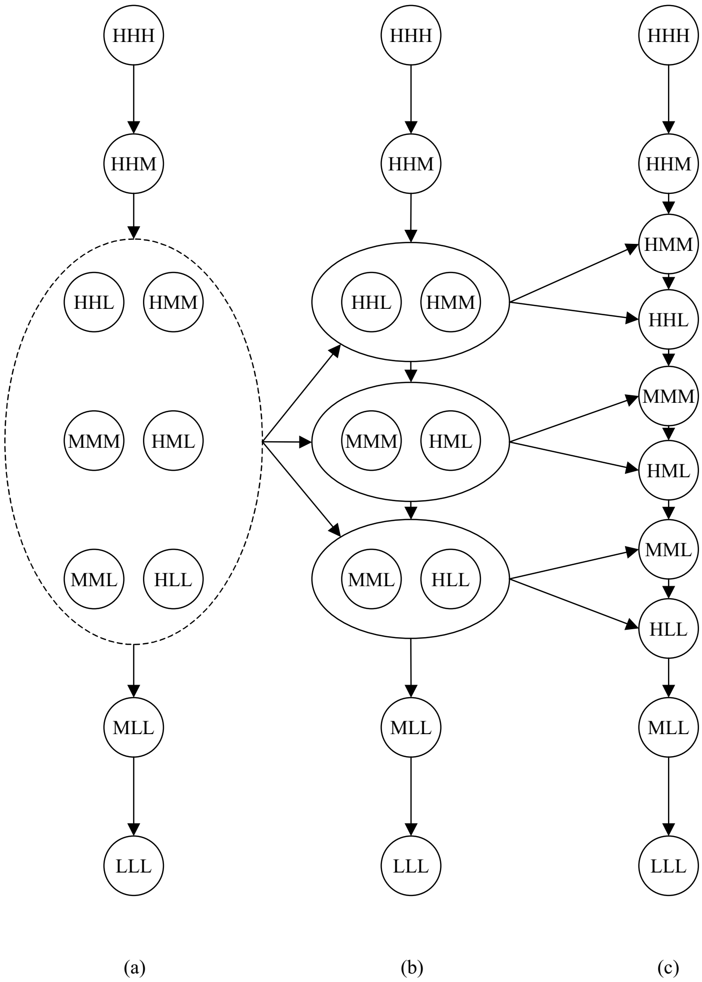

Figure 4. Orderings of the sample space for a three-level evaluation scale with samples of three elements. The applied dominance criteria are the Pareto (a), the rank (b) and the rank plus dispersion (c) criteria. A vertical arrow represents a dominance relation, while an ellipse represents an equivalence or a semi-equivalence class. For the sake of simplicity, (a) does not represent dominance relationships and equivalence classes involving samples within the semi-equivalence class

Furthermore, comparing these results with those obtained in Figure 1, some differences appear. An example is the 'out-of-control' points occurring in the corresponding eighth sample of the $\bar{X}$ chart in Figure 1.

It must be noted that the two approaches give, in the example, very similar results because the adopted ordering criteria (applied to ordinal scales) have the same 'monotonic' properties as the mean operator (applied to interval scales)[11]. This is ever more evident because in the example of Figure 1 we (arbitrarily) adopted a linear codification of levels. With different codifications and criteria the difference between the proposed approach and the traditional one would be more marked.

For example, referring to data in Table II, suppose that, due to the particular kind of manufacturing, any sample including a 'Low' rating is considered worse than any sample that does not include it. So, for example, *HHHL* is worse than *HMMM*.

Figure 6 represents the sample control charts and the corresponding sample frequency distributions of data reported in Table II, after the sequential application of the rank plus dispersion dominance criteria and the 'no-Low' criterion. We see that the chart behavior is very different from those reported in the previous figures (Figures 1 and 5).

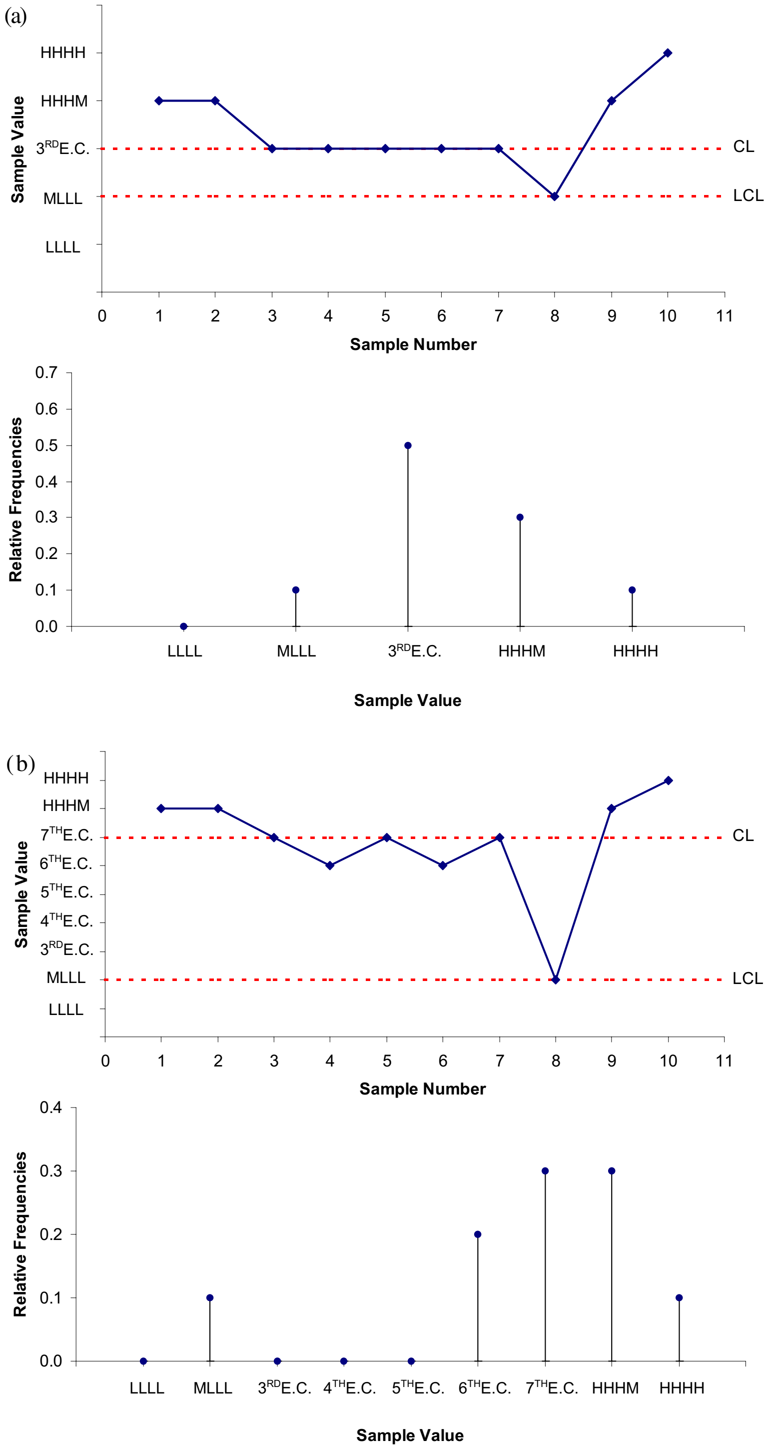

Figure 5. Ordered samples control chart and sample value frequency distribution of data reported in Table II, after the application of (a) the Pareto dominance criterion, (b) the rank dominance criterion and (c) the rank and dispersion dominance criteria. The 3rd equivalence class (E.C.) contains all the remaining samples in the ordered sample space in (a) some equivalence classes in the ordered sample space in (b) and no equivalence classes in the ordered sample space in (c) (see Table III). The type I error is fixed at 0.05

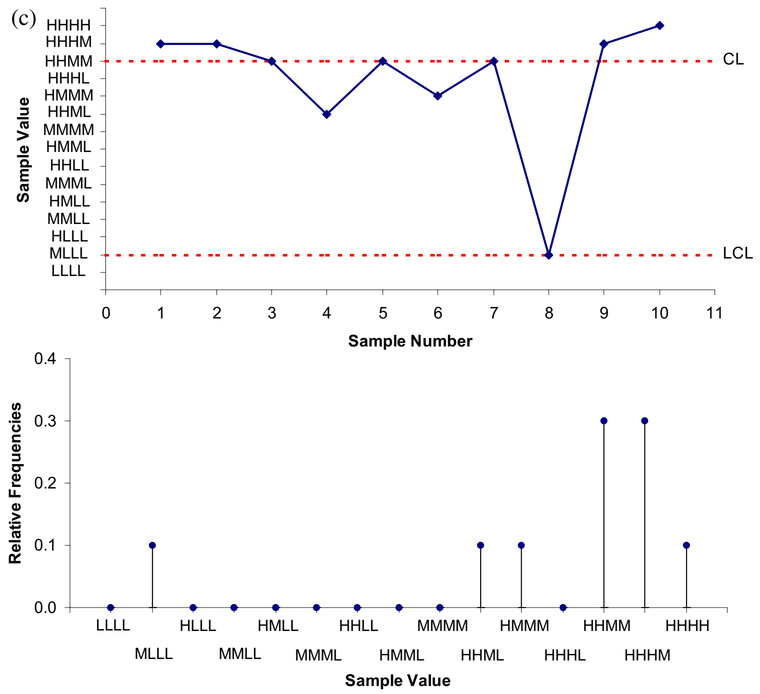

Figure 5. *Continued*

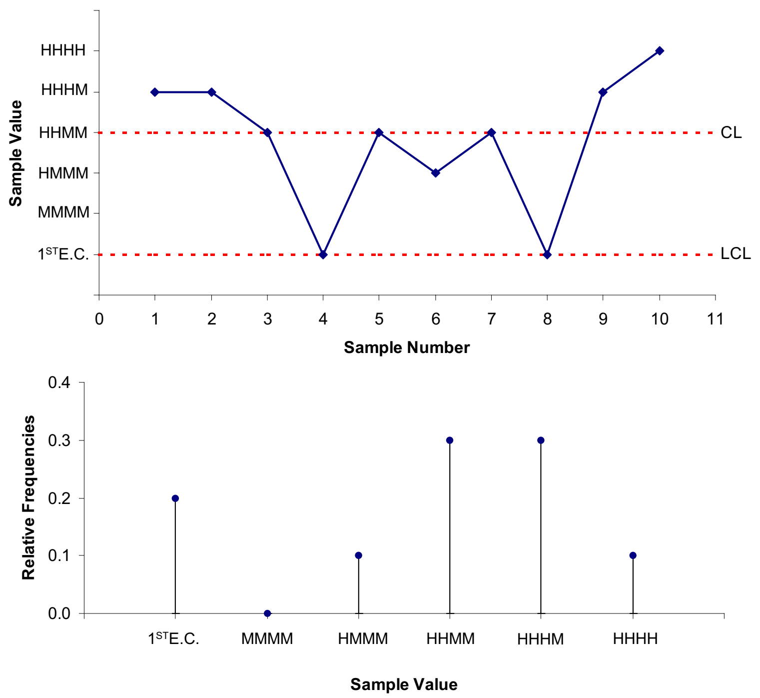

Figure 6. Ordered sample control chart and sample value frequency distribution of data reported in Table II, after the sequential application of the rank and dispersion dominance criteria and the 'no-Low' criterion. The type I error is fixed at 0.05. The figure shows one equivalence class including all the samples with at least one 'Low' included in them. The 1st E.C. includes the following sample configurations: *HHHL*, *HHML*, *HMML*, *HHLL*, *MMML*, *HMLL*, *MMLL*, *HLLL*, *MLLL* and *LLLL*

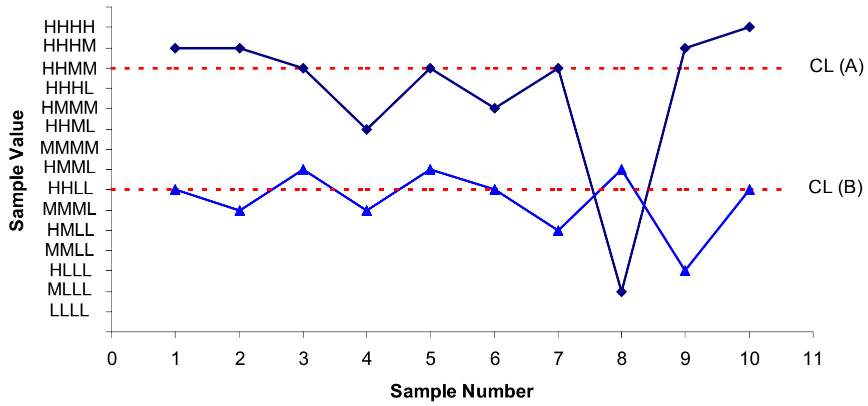

Figure 7. Example of two different ordered samples control charts for the same process. Chart (A) is better than chart (B). It is centered on higher values. CL (A) and CL (B) are respectively the central limits of the two charts

This last example clearly shows the effect of the subjective choice of the ordering criterion. According to the specific process characteristics, different ordering criteria can determine different process control charts.

In the end, it must be highlighted that the interpretation of the proposed chart is a bit different from the traditional Shewhart charts. The search for out-of-control points, trends and special patterns is integrated with a process positioning analysis. A good quality process will present a concentration of samples at the highest positions of the ordered sample space scale (see Figure 7).

4. *ORDINAL RANGE CONTROL CHARTS*

Location and dispersion measures in an ordinal environment should be defined without the introduction of the concept of distance among scale levels. In a recent paper, we already introduced a new dispersion measure, the so-called '*ordinal range*'[8]. This dispersion measure is proposed to develop control charts for sample ranges.

Let us briefly recall the concept of ordinal range. Given a t-point evaluation scale, the ordinal range is a new ordinal scale with $t(t+1)/2$ levels. These levels are obtained according to a sequential application of a dispersion and a '*dangerousness*' criteria. Dispersion is given by the number of scale levels between the minimum and the maximum sample elements. This number being equal, dispersion is more 'dangerous' for a sample centered at a lower value of the original evaluation scale. For example, *HMM* and *MLL* have the same number of scale levels between their maximum and minimum elements (i.e. 1), nevertheless *HMM* is less 'dangerous' than *MLL* because it is centered at a lower value of the original evaluation scale.

Table V reports the ordinal ranges for a sample space defined by a three-level scale and a sample size $n = 3$.

The concept of dangerousness can be seen as a dominance criterion to order a sample space, according to the sample dispersion. Different dominance criteria could be defined and used by practitioners for specific applications.

The control chart for sample ranges is built in the same way as the control chart for ordered samples. Given the particular meaning of ordinal range charts, only the UCL is defined (we are interested in detecting upwards shifts of ordinal range charts). The CL represents the median of the ordered sample range positions.

A set of initial samples is considered to estimate the ordinal range distribution, which is used to calculate the UCL with a given type I error.

Figure 8 represents the ordinal range control chart and the corresponding ordinal range frequency distribution of data reported in Table II. Samples are ordered by the dangerousness dominance criterion. The upper control limit is calculated with a type I error equal to 0.05.

Table V. Ordinal ranges for a sample space defined by a three-level scale (L = Low; M = Medium; H = High) and a sample size $n = 3$

Sample space	Position in the ordered range space (equivalence class)
LLL	3rd
MLL	5th
MML	5th
MMM	2nd
HLL	6th
HML	6th
HMM	4th
HHL	6th
HHM	4th
HHH	1st

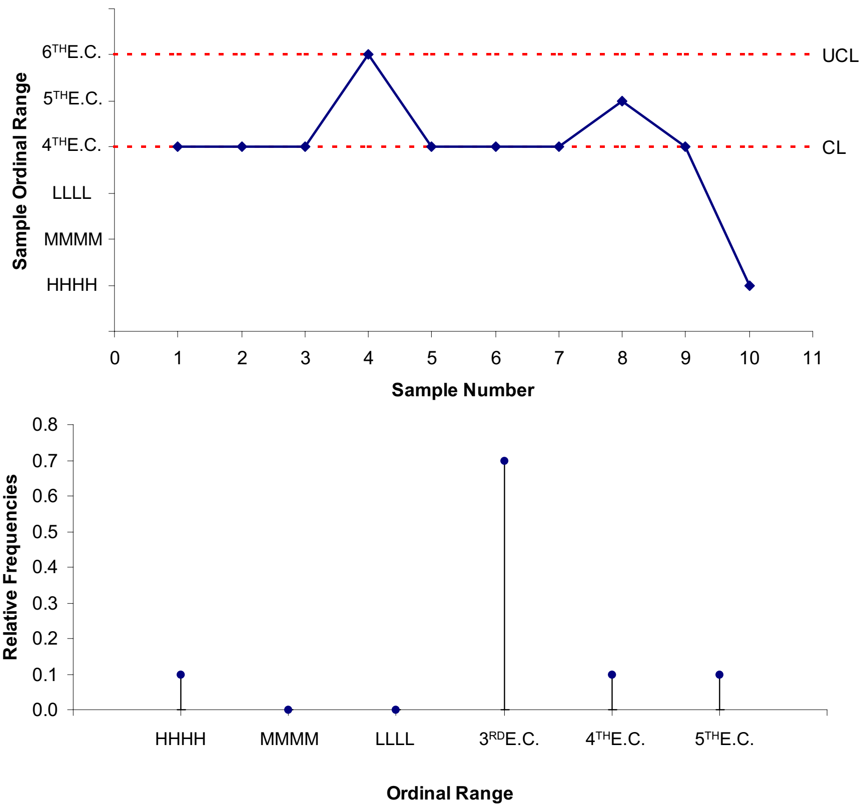

Figure 8. Ordinal range control chart and corresponding sample ordinal range frequency distribution of data reported in Table II. Samples are ordered by the dangerousness dominance criterion. The figure shows some equivalence classes in the ordered sample space. The 4th E.C. includes the following sample configurations: *HMMM*, *HHMM* and *HHHM*; the 5th E.C.: *MLLL*, *MMLL* and *MMML*; and the 6th E.C.: *HLLL*, *HMLL*, *HHLL*, *HMML*, *HHML* and *HHHL*

By the adopted criteria, the example presents some significant differences compared with the approach based on the numeric codification of levels. Using different criteria the difference between the proposed approach and the traditional one becomes more evident, such as for the ordinal sample charts.

Furthermore, ordinal range charts also allow a process positioning analysis. A good quality process will present a concentration of samples at the lowest positions of the ordinal range space scale.

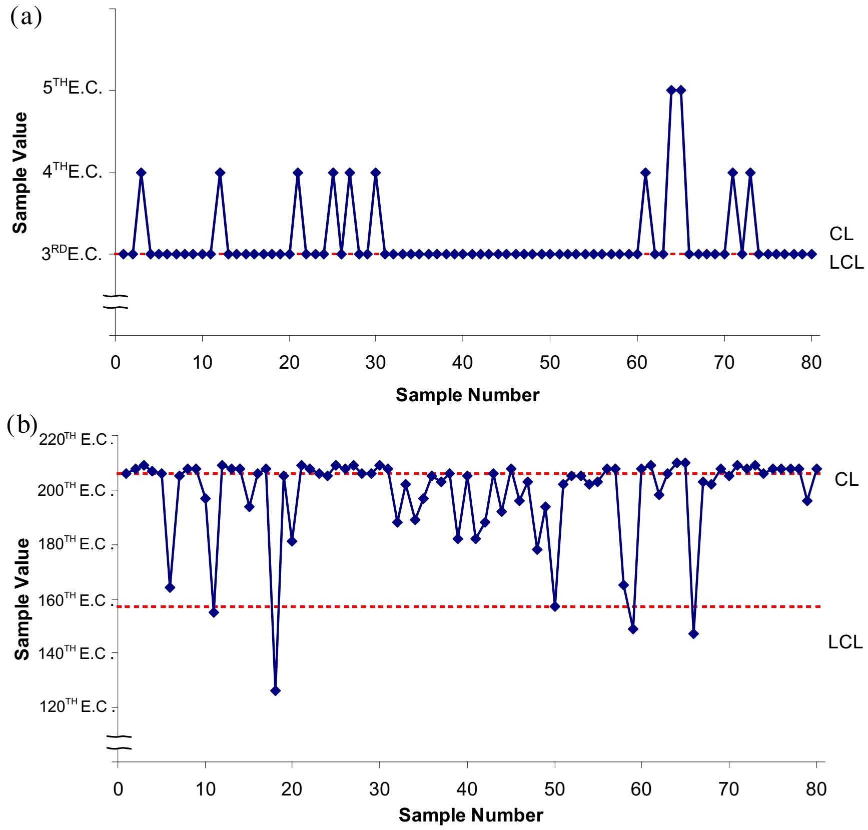

Figure 9. (a) Sample control chart for the perceived quality of the booking service at the Evangelical Waldensian Hospital in Turin (Italy)[12]. The Pareto-dominance criterion has been applied. Charts are based on a preliminary run of 80 samples of four elements. The figure shows some equivalence classes in the ordered sample space. In this example CL and LCL coincide. (b) Sample control chart obtained with the same data as (a) after the sequential application of the rank and dispersion dominance criteria. Charts are based on a preliminary run of 80 samples of four elements. In this case each equivalence class contains only one element

5. A CASE STUDY

The proposed methodology has been applied to the on-line monitor of the perceived quality of the booking service at the Evangelical Waldensian Hospital in Turin (Italy)[12].

Figure 9 reports the sample control charts (with $t = 7$, $n = 4$) obtained respectively by the application of the Pareto (a) and the rank plus dispersion (b) dominance criteria. Charts are built after a preliminary run of 80 samples of four elements. The lower control limit (LCL) is calculated for a type I error equal to 0.05. The CL represents the median of sample positions. The figure shows some out-of-control points.

Figure 10 represents the frequency distributions of sample positions for the two dominance criteria.

Figure 11 shows the ordinal range control chart. The UCL is calculated for a type I error equal to 0.05. The CL represents the median of the ordered sample ranges. The chart highlights four out-of-control points.

Figure 12 reports the frequency distribution of ordinal ranges, which is used to calculate the UCL.

Figure 13 reports the corresponding traditional $\bar{X}$–R control charts, if we perform the following (arbitrary) numerical codification:

$$\text{level } 1 = 1, \quad \text{level } 2 = 2, \quad \text{level } 3 = 3, \quad \text{level } 4 = 4,$$
$$\text{level } 5 = 5, \quad \text{level } 6 = 6, \quad \text{level } 7 = 7$$

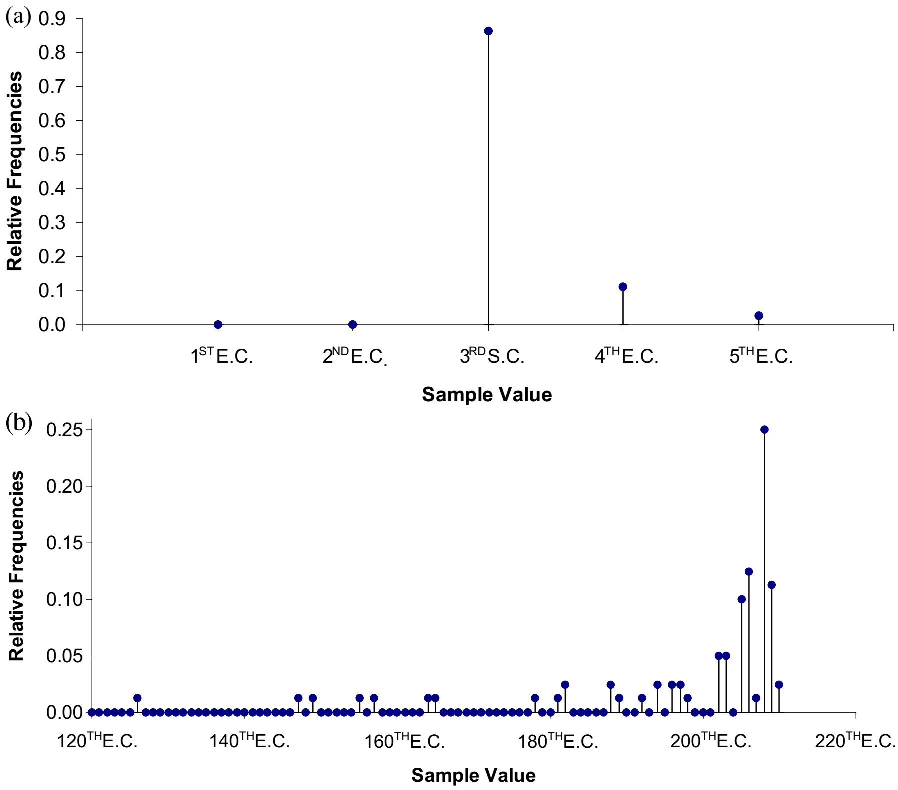

Figure 10. Frequency distribution of data reported on the control chart of (a) Figure 9(a) and (b) Figure 9(b)

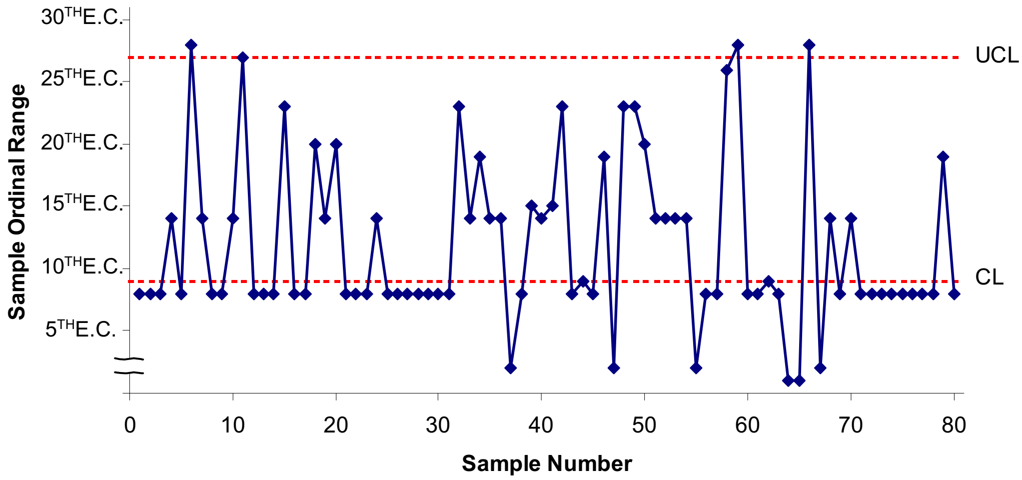

Figure 11. Ordinal range control chart for the perceived quality of the booking service at the Evangelical Waldensian Hospital in Turin[12]. The chart is based on a preliminary run of 80 samples of four elements

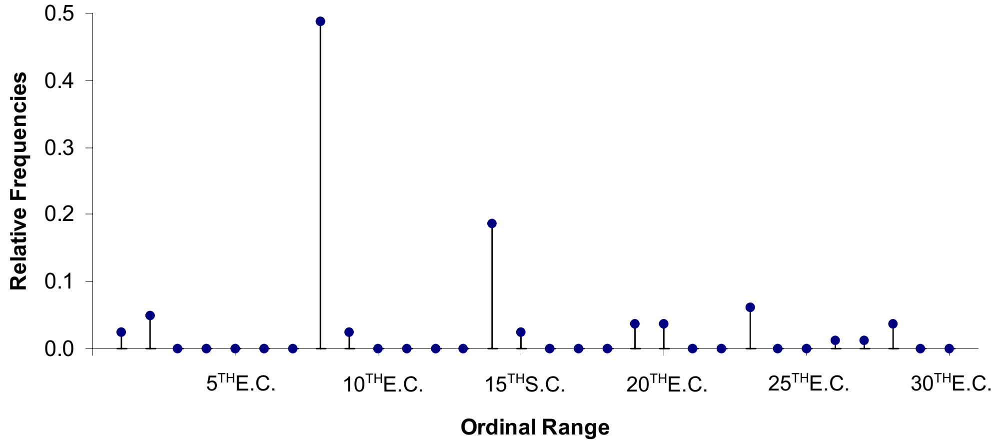

Figure 12. Frequency distribution of data reported on the control chart of Figure 11

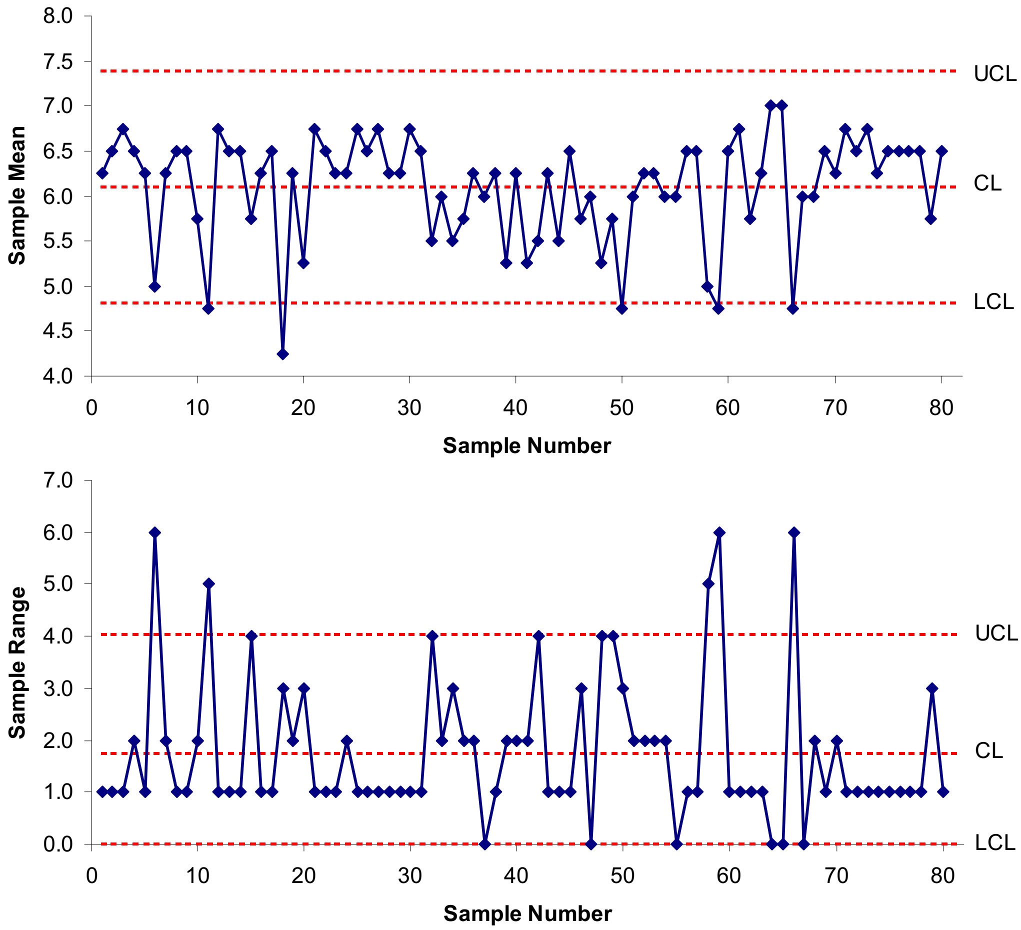

Figure 13. $\bar{X}$–R control charts for the perceived quality of the booking service at the Evangelical Waldensian Hospital in Turin[12]. The chart is based on a preliminary run of 80 samples of four elements

Many differences can be individuated between the results of $\bar{X}$–R and ordinal approaches (see Figures 9(b) and 11). In particular, despite the imposed type I error being higher in the ordinal approach, the traditional R chart shows more 'out-of-control' points than the corresponding ordinal range control chart. This is a direct consequence of the introduction of the specific criterion adopted for the construction of the ordinal range control chart.

6. CONCLUSIONS

The paper presents two new control charts for the process control of quality characteristics evaluated on an ordinal scale, without exploiting an artificial conversion of scale levels. The basic concept of the charts is the ordering of the sample space of the quality characteristic at hand.

The main novelties of the methodology are the following.

- Charts do not consider sample synthesis operators (sample mean and variance). They give an on-line monitoring based on a new *process sample scale* obtained by ordering the ordinal variable sample space.
- Sample space ordering is obtained by some specific dominance criteria. However, the method allows practitioners to formulate their own ordering criteria, according to quality characteristics at hand.
- Charts do not suffer from the poor *resolution* shown by other linguistic charts, where the original evaluation scale is used to evaluate samples[5,7].
- No distributional shape is assumed for the population (universe) of evaluations.
- Charts facilitate process positioning analysis: a good quality process will present a concentration of samples on the highest positions of the ordered sample space scale, and on the lowest positions of the ordinal range space scale.

The proposed methodology allows one to extract enough information from available data without artificially adding properties to the scale of evaluations. However, the initial effort for establishing the control limits can require an abundant amount of preliminary data.

For short run productions or services, where large initial sets of samples are not available, an alternative approach can be carried out using bootstrap techniques.

REFERENCES

1. Roberts FS. *Measurement Theory*. Addison-Wesley: Reading, MA, 1979.
2. Yager R, Filev DP. *Essentials of Fuzzy Modeling and Control*. Wiley: New York, 1994.
3. Hoadley B. The quality measurement plan (QMP). *The Bell System Technical Journal* 1981; **60**:215–273.
4. Montgomery DC. *Introduction to Statistical Quality Control* (4th edn). Wiley: New York, 2001.
5. Franceschini F, Romano D. Control chart for linguistic variables: A method based on the use of linguistic quantifiers. *International Journal of Production Research* 1999; **37**(16):3791–3801.
6. Franceschini F, Rossetto S. On-line service quality control: The 'Qualitometro' method. *Quality Engineering* 1998; **10**(4):633–643.
7. Franceschini F, Rossetto S. Service qualimetrics: The QUALITOMETRO II method. *Quality Engineering* 1999; **12**(1):13–20.
8. Franceschini F, Galetto M, Varetto M. Qualitative ordinal scales: The concept of ordinal range. *Quality Engineering* 2004; **16**(4):515–524.
9. Yager RR. Non-numeric multi-criteria multi-person decision making. *Group Decision and Negotiation* 1993; **2**(1):81–93.
10. Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. Chapman and Hall: New York, 1993.
11. Davey BA, Priestley HA. *Introduction to Lattices and Order*. Cambridge University Press: Cambridge, 1990.
12. Franceschini F, Stangalini M. Un'applicazione del metodo 'Qualitometro' per la valutazione della qualità di un servizio di prenotazione ospedaliero. *De Sanitate* 2000; **3**(14):43–53.

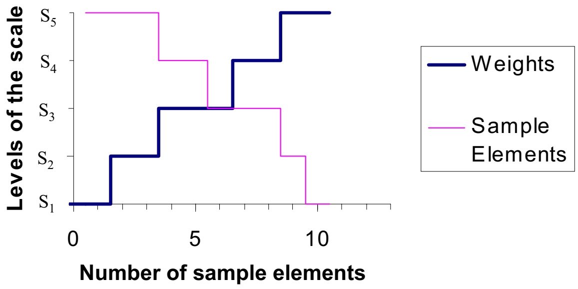

Figure A1. Graphical representation of the OWA calculation. The value of the OWA emulator of the arithmetic mean is given by the intersection of the 'ascending stair' (OWA weights) and the 'descending stair' (ordered sample elements)

APPENDIX A

The OWA emulator of the arithmetic mean was first introduced by Yager[2,9]. This operator is typically used with linguistic scales. It is defined as

$$\text{OWA} = \operatorname*{Max}_{k=1}^{n}[\text{Min}\{Q(k), b_k\}]$$

where:

- $Q(k) = S_{f(k)},\ \ k = 1, 2, \ldots, n$ is the average linguistic quantifier (the weights of the OWA operator), with

$$f(k) = \text{Int}\left\{1 + \left[k\frac{t-1}{n}\right]\right\}$$

  $S_{f(k)}$ is the $f(k)$th level of the linguistic scale (for example, $S_{f(k)} = S_1$ if $f(k) = 1$);
- $\text{Int}(a)$ is a function which gives the integer closest to a;
- t is the number of scale levels;
- n is the sample size;
- b_k is the kth element of the sample previously ordered in a decreasing order.

This OWA operator is said to be an emulator of the arithmetic mean since it operates, in an ordinal environment, in the same way as the arithmetic mean works in a cardinal one. It can take values only in the set of levels of the original ordinal scale, while a numerical codification of these levels could lead to some intermediate mean values.

As an example, let us assume a scale with $t = 5$ levels, namely S_1, S_2, S_3, S_4 and S_5, and a sample of size $n = 10$, whose elements, previously ordered in a decreasing order, are $\{S_5, S_5, S_5, S_4, S_4, S_3, S_3, S_3, S_2, S_1\}$.

The 'weights' of the OWA operator are:

- $Q(1) = S_1$;
- $Q(2) = Q(3) = S_2$;
- $Q(4) = Q(5) = Q(6) = S_3$;
- $Q(7) = Q(8) = S_4$;
- $Q(9) = Q(10) = S_5$.

Therefore, we have

$$\begin{aligned}\text{OWA} = \text{Max}[&\text{Min}\{S_1, S_5\}, \text{Min}\{S_2, S_5\}, \text{Min}\{S_2, S_5\}, \text{Min}\{S_3, S_4\}, \text{Min}\{S_3, S_4\}, \text{Min}\{S_3, S_3\},\\ &\text{Min}\{S_4, S_3\}, \text{Min}\{S_4, S_3\}, \text{Min}\{S_5, S_2\}, \text{Min}\{S_5, S_1\}] = S_3\end{aligned}$$

Figure A1 shows a graphical representation of the OWA calculation[7]. The value of the OWA emulator of the arithmetic mean is given by the intersection of the 'ascending stair' (OWA weights) and the 'descending stair' (ordered sample elements).

Authors' biographies

Fiorenzo Franceschini is Professor of Quality Management at the Polytechnic Institute of Turin (Italy)—Department of Manufacturing Systems and Economics. He is the author or co-author of three books and many published papers in scientific journals and international conference proceedings. His current research interests are in the areas of quality engineering, QFD and quality management. He is a member of the editorial board of *Quality Engineering*, and the *International Journal of Quality and Reliability Management*. He is a senior member of ASQ.

Maurizio Galetto is an Assistant Professor at the Department of Manufacturing Systems and Economics at the Polytechnic Institute of Turin. He holds a PhD in Metrology from the Polytechnic Institute of Turin. His current research interests are in the areas of quality management and statistical process control.

Marco Varetto graduated in Engineering Management at the Polytechnic Institute of Turin. His main scientific interests are in the area of statistical techniques in quality management.