Advances and Applications in Statistics
Volume 19, Number 2, 2010, Pages 97-111
Published Online: March 17, 2011
This paper is available online at http://pphmj.com/journals/adas.htm

A GRAPHICAL QUALITY CONTROL PROCEDURE USING DATA DEPTH

HAJLAOUI MEKKI

Faculté des Sciences Economiques et de Gestion de Mahdia
Sidi Messoud 5111 Mahdia, Tunisia
e-mail: mekki.hajlaoui@fsegma.rnu.tn

Abstract

Quality is a multivariate concept by nature. So since the Mahalanobis work, the statisticians tried to develop devices that help in monitoring the quality process. Their efforts come to breed multiple procedures but all of them are dependent to multiple conditions to be applied in industry. Using the centre-outward ranking induced by the notion of data depth, we use a graphical method named *DD*-diagram to see the outlyingness of a given multivariate quality data. This one-dimensional curve in the plane is used to monitor a multivariate quality process. Hence, shifts in mean and/or in scale are analyzed through graphical representation of an observed sample. To see the benefits of this graphical method test, a real example from industry is presented. In addition, we suggest to rank data depths before sketching out the *DD*-diagram.

1. Introduction

Most statistical phenomena are multivariate by nature and the classical multivariate analysis relies heavily on the hypothesis of normality which is difficult to justify in reality.

This paper focuses on presenting multivariate quality data on the plane. This graphical method provides a visualization of a change in position and/or in scale

2010 Mathematics Subject Classification: 62H10, 90B25, 60G35, 62G30, 62G05.

Keywords and phrases: multivariate data, quality control, data depth, *DD*-diagram.

Received November 13, 2010

between an empirical sample with respect to a reference sample. Furthermore, this approach paves the way for an easier interpretation and a rapid adjusting action in the multivariate quality process.

This depth-based methodology may be viewed in part as a multivariate generalization of standard univariate rank methods. There is however a difference in the two ways of ranking: the ranking in the univariate case is linear; from the smallest to the largest, while the multivariate one is a centre-outward ranking.

The paper is organized as follows. In Section 2, we present some background material and basic definitions. In Section 3, data depth plots are presented. In Section 4, an application for the *DD*-diagram is given with interpretation of the results.

2. Data Depth and Background Material

According to [1] and to [5], a data depth is a way to measure the "depth" or the "outlyingness" of a given point with respect to a multivariate data cloud or its underlying distribution.

Let *F* be a probability distribution in $\Re^p$, $p \geq 1$. Throughout the following, unless stated otherwise, we assume that *F* is absolutely continuous and also that $(X_1, X_2, \ldots, X_n)$ is a random sample from *F*. Each sample point X_i is viewed as a $p \times 1$ column vector. Therefore, if the quality of the *i*th observed unit is noted by ω, then $X_i'(\omega) = (X_{i_1}(\omega), X_{i_2}(\omega), \ldots, X_{i_p}(\omega))$.

A data depth is a way of measuring how deep or central a given point $x(\omega) \in \Re^p$ is with respect to *F* or with respect to a given data cloud $(X_1(\omega), X_2(\omega), \ldots, X_n(\omega))$. Some useful examples of data depth are given in the literature, but here we will adopt only the Mahalanobis depth, see [4].

The Mahalanobis depth at $x(\omega)$ with respect to *F* is defined to be

$$D(F, x(\omega)) = \frac{1}{1 + (x(\omega) - \mu_F)\Sigma_F^{-1}(x(\omega) - \mu_F)'}, \tag{1}$$

where μ_F and Σ_F are the mean vector and dispersion matrix of *F*, respectively. The sample version of (1) is obtained by replacing μ_F and Σ_F with their respective sample estimates,

$$\hat{\mu} = \overline{X}_{p\times 1}(\omega) = \frac{1}{n}\sum_{i=1}^{n} X_i(\omega) \tag{2}$$

and

$$\hat{\Omega}(\omega) = S_{p\times p}(\omega) = \frac{1}{n-1}\sum_{i=1}^{n}[X_i(\omega) - \overline{X}(\omega)][X_i(\omega) - \overline{X}(\omega)]', \tag{3}$$

then

$$D_n(F,\ x(\omega)) = \frac{1}{1 + (x(\omega) - \hat{\mu}_F)S_F^{-1}(x(\omega) - \hat{\mu}_F)'}. \tag{4}$$

Henceforth, D or D_n will be used to indicate the datum depth notion and a larger value of $D_{\bullet}(F,\ x(\omega))$ always implies a deeper (or more central) $x(\omega)$ with respect to F.

Given a notion of data depth, we can compute the depths of all quality measures $(X_1(\omega),\ X_2(\omega),\ ...,\ X_n(\omega))$ and order them according to decreasing depth values. This gives a ranking of the sample point associated with the ith highest depth value. We view $(X_{[1]}(\omega),\ X_{[2]}(\omega),\ ...,\ X_{[n]}(\omega))$ as the order statistics, with $X_{[1]}(\omega)$ being the deepest or the most central point or simply the centre, and $X_{[n]}(\omega)$ the most outlying point. The implication is that a larger rank is associated with a more outlying position with respect to the data cloud. These order statistics induced by a data depth are different from the usual order statistics on the real line, since the latter are ordered from the smallest sample point to the largest, while the former start from the middle sample point and move outwards in all directions. In this paper, only the depth-induced order statistics are studied. They will be referred to as depth order statistics denoted by *DO*-statistics, and their ordering or ranking as depth ranking or depth ordering.

When ties occur in the ordering, the corresponding sample points are viewed as depth equivalent, and the set of these points is termed a depth-equivalence class. In the particular case where there is more than one sample point with the highest depth value, we refer to their average as the deepest point, for convenience. If points $x_{i_1}(\omega),\ x_{i_2}(\omega),\ ...,\ x_{i_k}(\omega)$ all belong to the same depth-equivalence class where

$i_1 < i_2 < \cdots < i_k$ and there are exactly j sample points with higher depth values, then we assign $x_{i_1}(\omega), x_{i_2}(\omega), \ldots, x_{i_k}(\omega)$ to be $x_{[j+1]}(\omega), x_{[j+2]}(\omega), \ldots, x_{[j+k]}(\omega)$ in that order.

3. Data Depth Plots

In order to illustrate depth ordering and its ramifications we use Mahalanobis depth to order quality sample measures. But, before doing so we pick out a parameter that can be defined in an analogous manner to descriptive statistics.

Given the notion of data depth, (1) or (4), the sample becomes $X_{[1]}(\omega), X_{[2]}(\omega), \ldots,$ there is a natural choice of location parameter for the observed distribution. Specifically and according to [5], the centre is the most central point so

$$\nu(\omega) = X_{[1]}(\omega). \tag{5}$$

When the depth-equivalence class contains more than one point measure, $(X_{[11]}(\omega), X_{[12]}(\omega), \ldots, X_{[1k]}(\omega))$, the median or the centre is the average of the deepest points, so in this case

$$\nu(\omega) = \frac{1}{K}\sum_{k=1}^{K}(X_{[1k]}(\omega)). \tag{6}$$

On this basis and using data depth, equations (5) and (6) fix out a centre or a multivariate median. However, if the observed distribution is asymmetric the use of another definition of data depth leads to a different centre, see [1]. Moreover, [5] states that if Mahalanobis depth is used, the central point defined in equations (5) and (6) turns out to be the mean of the observed data. This suggests concepts of location which are intermediate between the mean and the median.

A data depth plots a graphical comparison between two multivariate distributions based on data depth. Let $(X_1(\omega), X_2(\omega), \ldots, X_n(\omega))$ and $(Y_1(\omega), Y_2(\omega), \ldots, Y_n(\omega))$ be two samples from both distributions that characterise a multivariate quality process during two different periods. The $\mathbf{X}(\omega)$ sample is a reference sample taken when the production process is in control and the $\mathbf{Y}(\omega)$ sample results from an inspection routine drawn during a later period. The

DD-diagram results from plotting data depth values obtained from their samples. These values are arranged from the most central point to the most extreme one.

Let *F* and *G* be two distributions on $\Re^p$ and $D(\cdot, \cdot)$ be a data depth computed using (1). We define the *DD*-diagram as

$$DD(F, G) = \{(D(F, x(\omega)), D(F, y(\omega))), \text{ for any } x(\omega), y(\omega) \in \Re^p\}. \tag{7}$$

Since $DD(\cdot, \cdot)$ is a subset of $\Re^2$, the resulting graph is one-dimensional curve in the plane. If the two distributions are identical, the *DD*-diagram in (7) turns out to be the diagonal line from the point $(0, 0)$ to $(1, 1)$. Different patterns of deviations from the diagonal line in the *DD*-diagram are indications of differences in specific characteristics of *F* and *G*.

In general, the distributions are rarely known so instead we use an empirical version of the *DD*-diagram. If *F* and *G* are unknown distributions for the samples $(X_1(\omega), X_2(\omega), ..., X_n(\omega))$ and $(Y_1(\omega), Y_2(\omega), ..., Y_n(\omega))$, then the *DD*-diagram results by plotting

$$DD(F_n, G_n) = \{(D_n(F, x(\omega)), D_n(F, y(\omega))), \text{ for any } x(\omega), y(\omega) \in \Re^p\}, \tag{8}$$

if equation (4) is used to compute the data depth. This diagram exhibits all the differences between the distributions *F* and *G* for any departure from the diagonal line linking $(0, 0)$ to $(1, 1)$.

If $p \geq 2$ and if *F* and *G* are both absolutely continuous, then *DD*-diagram corresponds to a region with non zero area. The area of this region can serve as a measure of the discrepancy between *F* and *G*, see [5]. If the two distributions are identical, the data cloud of the *DD*-plot should be concentrated along the diagonal line. Other patterns are indications of differences in specific characteristics of *F* and *G*, i.e., in position, in scale, in skewness, ...

In most cases, the departure from the diagonal line usually takes the form of pulling towards the axis of the reference sample, beginning from the lower left corner and spreading out the points around the diagonal line to the upper right corner. In order to bring out scale differences, the centre of the samples should be equalized first by subtracting from the data the identified centre in (5) or in (6). Suppose that *G* is more spread out than the reference sample *F*, then the points in *DD*-diagram tend to arch toward the *F* sample in the shape of an early half moon.

In analogous manner to the multivariate Shewhart chart, the decision of an out of control signal may be made using control limits. These limits must be defined on the basis of data depth of the reference sample. In fact to detect visually the shifts in location and/or in dispersion, control limits are defined with respect to the whole set of values. So knowing that $0 < D(F, x(\omega)) \leq 1$ for any $x(\omega) \in \Re^p$ and using equation (1) or (4), these limits are given by

$$L_{\max} = \sqrt{D_{\bullet}(F, x(\omega))[2 - D_{\bullet}(F, x(\omega))]}, \tag{9}$$

$$L_{\min} = 1 - \sqrt{1 - D_{\bullet}^2(F, x(\omega))}. \tag{10}$$

The role of these limits is to allow the operator or the supervisor to decide graphically if the production process is in control or not. The region under control is located between (9) and (10). Therefore, if a point $\{d_i\}_{i=1, 2, \ldots, n}$ computed using (7) or (8) is located out of this region, the observed process is out of control and a decision must be made in order to investigate which characteristic (of the p-characteristics considered) is to be adjusted, see [3].

4. Application of the *DD*-diagram

In what follows, we present an application of the *DD*-diagram in a real case study from the industry and adopted by [2] in his thesis. The basic idea is to collect individual observations data from a production process during which the process is considered in control. These observations form a reference sample for data collected later as an empirical sample. Using data depth to plot *DD*-diagram enables evaluating the performance of the production process.

4.1. Data collection

Mega is a leading manufacturer and supplier of dressing products made of absorbent cotton gauzes and non woven materials. The company is operating in medical field ensuring that all products intended for human medical use, will be safe and effective in compliance with European and International norms. Since 1999, Mega firm is certified CE according to the requirement of the 93/42/EEC directive of medical devises, corresponding to the Good Manufacturing Practices and following the European Norm EN 46002.

The firm produces several types of 100% absorbent cotton gauze swabs with folded edges. Although, there exist different sizes the application will focus only on one single use swabs that measures 10×10 cm with 12 plies. For the reference sample and empirical one, the quality characteristics are determined through the sealing forces in the 4 edges $(p = 4)$. Their standard norm is to be between 2 and 5 Newtons for each edge of a packet.

On one hand, the reference sample of 40 observations is taken on Jan. 8th, 2000, when the process is in control. This sample is given in Table 1. On the other hand, the empirical sample of 40 observations is drawn during one day of work. It is taken on Oct. 30th, 2001. The results are also given in Table 1 issued during production process.

Table 1. Sealing force samples in the edges of swabs 10×10 cm with 12 plies

Obs.	Reference sample Edge sealing force				Empirical sample Edge sealing force				Data depth of a sample	
	Top	Bottom	Right	Left	Top	Bottom	Right	Left	from F	from G
1	3.28	4.16	2.12	3.23	5.54	4.36	3.13	3.48	0.143	0.060
2	4.16	4.75	4.07	3.38	3.48	3.97	5.15	4.61	0.260	0.149
3	4.07	3.76	4.26	4.16	2.69	3.18	4.75	6.22	0.519	0.052
4	3.18	4.26	4	4.26	2.69	3.16	4.75	5.23	0.247	0.076
5	3.38	4.07	4.16	4.75	2.4	4.75	3.09	3.13	0.228	0.082
6	4	3.23	4.61	3.13	2.64	3.48	4	3.72	0.179	0.119
7	4.7	4.75	4	4.16	2.35	5.15	1.32	2.55	0.140	0.043
8	3.38	3.13	3.23	4	3.09	2.97	1.47	3.33	0.363	0.093
9	3.01	4.38	3.38	4.15	4.26	4.26	2.22	3.38	0.193	0.113
10	4.26	3.76	3.09	3.13	2.45	3.18	4.21	5.18	0.362	0.075
11	3.48	2.12	2.45	3.38	2.15	3.77	4.75	5.35	0.142	0.053
12	4.16	3.13	4.16	3.18	3.09	1.21	4.85	4.06	0.293	0.043
13	4.75	4.26	4.75	3.76	2.39	4.07	4.64	5.1	0.182	0.070

14	4.61	3.18	4.26	4	3.45	5.15	4.36	3.09	0.250	0.121
15	4.16	4.26	3.18	4.26	3.09	4.26	5.29	3.48	0.187	0.077
16	3.02	4.75	3.38	3.18	3.13	5.38	2.94	4.07	0.152	0.090
17	3.38	4.16	4.07	4.26	3.72	3.09	1.88	3.77	0.327	0.118
18	3.18	3.36	4.16	4.16	3.26	3.23	2.03	3.28	0.238	0.160
19	4.26	4.16	4.75	3.38	3.48	2.95	3.03	3.37	0.227	0.360
20	4.16	3.18	4.16	3.18	3.77	1.88	4	3.95	0.304	0.124
21	4.71	3.13	4.02	4.16	3.13	3.02	4.21	3.18	0.203	0.127
22	4.26	3.02	4.16	3.18	4	4.36	4.36	4.16	0.262	0.366
23	3.82	3.38	4.32	4.75	3.77	7.09	4.75	4.36	0.267	0.042
24	3.18	2.12	3.13	4.41	3.09	4.22	5.22	4	0.119	0.097
25	3.76	2.46	2.45	4.26	3.18	3.97	5	4.02	0.126	0.128
26	3.18	3.36	3.36	2.12	4.26	4.85	4.75	3.13	0.106	0.145
27	4.16	4.16	3.36	3.18	5.05	2.79	1.36	4.16	0.400	0.035
28	4.07	4	4.61	3.23	3.77	5.15	1.98	7.45	0.243	0.015
29	4.26	4.75	2.35	2.12	4.16	4.07	2.02	6.36	0.090	0.025
30	4.16	3.38	2.12	2.45	5	5.94	3.22	4.07	0.142	0.044
31	3.76	3.36	3.36	3.23	4.21	3.69	3.38	4.16	0.635	0.320
32	4.26	4.12	3.18	3.36	3.38	3.97	5.1	5.48	0.324	0.102
33	4.1	3.02	3.13	4	3.23	4.21	5.02	4.37	0.305	0.137
34	3.76	4.26	4.18	4.16	5.15	1.16	4.75	3.75	0.432	0.047
35	3.23	4.06	4.32	4.37	4.26	1.26	2.69	2.34	0.235	0.066
36	4.07	3.13	4.12	4.26	3.97	7.77	3.09	4.21	0.384	0.024
37	3.23	4.07	4.02	3.18	4.07	5.38	5.36	4.37	0.209	0.100
38	3.42	4.16	3.22	4.16	4.21	4	5.38	4.85	0.287	0.141
39	4.71	3.02	4.32	4.41	3.18	4.07	4.08	4.26	0.169	0.259
40	3.82	2.12	2.45	2.12	4.07	6.22	3.82	3.13	0.099	0.069

4.2. Data processing

In order to get data clouds, the reference sample is processed and analyzed so that the process parameters are estimated during the in-control state. The computing equations (2) and (3) give,

- for the mean $\overline{X}_{4\times 1} = \frac{1}{40}\sum_{j=1}^{40} X_j = (3.8625,\ 3.64725,\ 3.6605,\ 3.655)'$,

- for the dispersion matrix

$$S_{p\times p} = \frac{1}{39}\sum_{j=1}^{40}(X_j - \overline{X})(X_j - \overline{X})'$$

$$= \begin{bmatrix} 0.2699 & -0.0051 & 0.0944 & -0.0316 \\ * & 0.5328 & 0.1168 & 0.0120 \\ * & * & 0.5582 & 0.2094 \\ * & * & * & 0.4950 \end{bmatrix}.$$

Given that these quality process parameters are estimated, the data depth of all observations is calculated.

Processing the reference sample gives a centre in the form of a depth-equivalence class of order one containing a unique swab with rank 31. According to the corresponding column in Table 1 of data depth and the equation (4), the observation $X_{31} = (3.76,\ 3.36,\ 3.36,\ 3.23)'$ is specified by the highest value of the statistic data depth, $D_{40}(F,\ X_{31}) = 0.63529$.

Note that in the reference sample, the point of order 31 is characterized by the maximum data depth in either case before and after centering with respect to the computed centre vector, see Table 1 and Table 2.

During the second phase, the quality control is conducted by observing the outlyingness of data depth of a sample extracted from the production process. Hence, for this case the quality control is based on observing the sealing forces in a sample of cotton gauze swabs with respect to the reference sample. That means represent graphically the quality performance using data depth as indicated in equation (8). Therefore, the *DD*-diagram consists of plotting data depth of the reference sample taken during in-control state versus the data depth of any sample drawn later, named empirical sample. For this application the resulting schemes are

Figure 1 and Figure 2. They correspond respectively to the case of an empirical sample with a change in position and in scale simultaneously in Table 1, and to the case of an empirical sample with a change in scale only in Table 2. The change in position is eliminated by centering the observations of the empirical sample G with respect to the most central point found in the reference sample F.

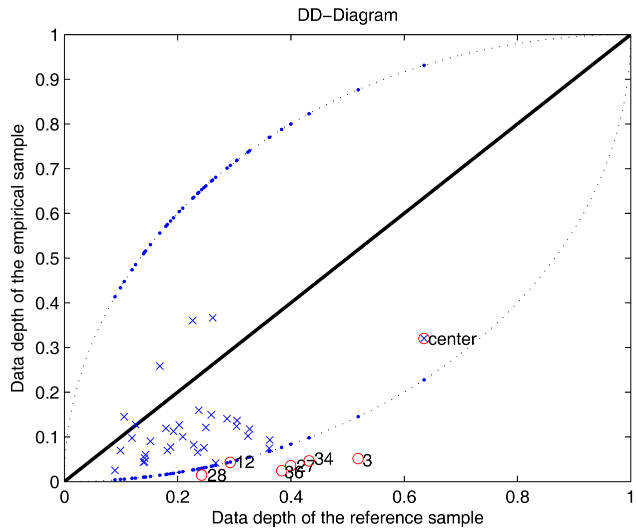

Figure 1. Quality data on edge sealing forces of swabs 10×10.

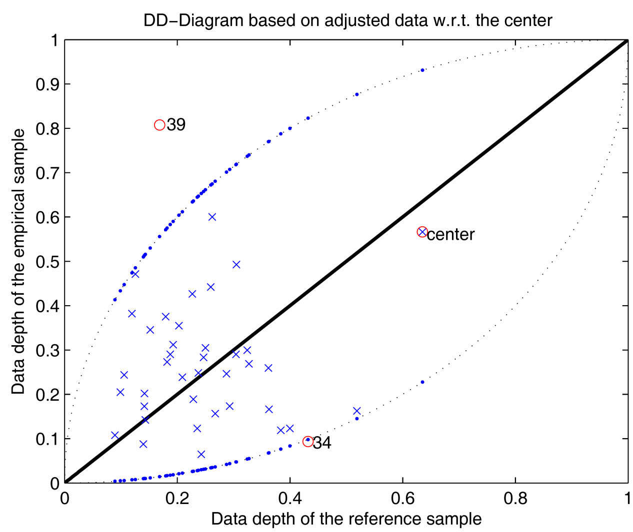

Figure 2. Centered quality data on edge sealing forces of swabs 10×10.

Table 2. Centered sealing force samples in the edges of swabs 10×10 cm with 12 plies

Obs.	Centered reference sample				Centered empirical sample				Data depth of a sample	
	Edge sealing force				Edge sealing force					
	Top	Bottom	Right	Left	Top	Bottom	Right	Left	from F	from G
1	–0.48	0.8	–1.24	0	1.78	1	–0.23	0.25	0.143	0.142
2	0.4	1.39	0.71	0.15	–0.28	0.61	1.79	1.38	0.260	0.442
3	0.31	0.4	0.9	0.93	–1.07	–0.18	1.39	2.99	0.519	0.163
4	–0.58	0.9	0.64	1.03	–1.07	–0.2	1.39	2	0.247	0.283
5	–0.38	0.71	0.8	1.52	–1.36	1.39	–0.27	–0.1	0.228	0.189
6	0.24	–0.13	1.25	–0.1	–1.12	0.12	0.64	0.49	0.179	0.375
7	0.94	1.39	0.64	0.93	–1.41	1.79	–2.04	–0.68	0.140	0.088
8	–0.38	–0.23	–0.13	0.77	–0.67	–0.39	–1.89	0.1	0.363	0.166
9	–0.75	1.02	0.02	0.92	0.5	0.9	–1.14	0.15	0.193	0.312
10	0.5	0.4	–0.27	–0.1	–1.31	–0.18	0.85	1.95	0.362	0.259
11	–0.28	–1.24	–0.91	0.15	–1.61	0.41	1.39	2.12	0.142	0.202
12	0.4	–0.23	0.8	–0.05	–0.67	–2.15	1.49	0.83	0.293	0.173
13	0.99	0.9	1.39	0.53	–1.37	0.71	1.28	1.87	0.182	0.273
14	0.85	–0.18	0.9	0.77	–0.31	1.79	1	–0.14	0.250	0.305
15	0.4	0.9	–0.18	1.03	–0.67	0.9	1.93	0.25	0.187	0.290
16	–0.74	1.39	0.02	–0.05	–0.63	2.02	–0.42	0.84	0.152	0.345
17	–0.38	0.8	0.71	1.03	–0.04	–0.27	–1.48	0.54	0.327	0.269
18	–0.58	0	0.8	0.93	–0.5	–0.13	–1.33	0.05	0.238	0.248
19	0.5	0.8	1.39	0.15	–0.28	–0.41	–0.33	0.14	0.227	0.426
20	0.4	–0.18	0.8	–0.05	0.01	–1.48	0.64	0.72	0.304	0.290
21	0.95	–0.23	0.66	0.93	–0.63	–0.34	0.85	–0.05	0.203	0.355

22	0.5	–0.34	0.8	–0.05	0.24	1	1	0.93	0.262	0.600
23	0.06	0.02	0.96	1.52	0.01	3.73	1.39	1.13	0.267	0.157
24	–0.58	–1.24	–0.23	1.18	–0.67	0.86	1.86	0.77	0.119	0.382
25	0	–0.9	–0.91	1.03	–0.58	0.61	1.64	0.79	0.126	0.471
26	–0.58	0	0	–1.11	0.5	1.49	1.39	–0.1	0.106	0.244
27	0.4	0.8	0	–0.05	1.29	–0.57	–2	0.93	0.400	0.123
28	0.31	0.64	1.25	0	0.01	1.79	–1.38	4.22	0.243	0.065
29	0.5	1.39	–1.01	–1.11	0.4	0.71	–1.34	3.13	0.090	0.108
30	0.4	0.02	–1.24	–0.78	1.24	2.58	–0.14	0.84	0.142	0.173
31	0	0	0	0	0.45	0.33	0.02	0.93	0.635	0.566
32	0.5	0.76	–0.18	0.13	–0.38	0.61	1.74	2.25	0.324	0.300
33	0.34	–0.34	–0.23	0.77	–0.53	0.85	1.66	1.14	0.305	0.493
34	0	0.9	0.82	0.93	1.39	–2.2	1.39	0.52	0.432	0.094
35	–0.53	0.7	0.96	1.14	0.5	–2.1	–0.67	–0.89	0.235	0.123
36	0.31	–0.23	0.76	1.03	0.21	4.41	–0.27	0.98	0.384	0.119
37	–0.53	0.71	0.66	–0.05	0.31	2.02	2	1.14	0.209	0.239
38	–0.34	0.8	–0.14	0.93	0.45	0.64	2.02	1.62	0.287	0.246
39	0.95	–0.34	0.96	1.18	–0.58	0.71	0.72	1.03	0.169	0.808
40	0.06	–1.24	–0.91	–1.11	0.31	2.86	0.46	–0.1	0.099	0.205

5. Interpretations and Results

The *DD*-diagram is a graphical comparison that exhibits:

• Location shifts and/or scale increase when moving from the distribution F of the reference sample to the distribution G of the empirical one.

• Variations in dispersion between both distributions F and G.

When constructing *DD*-diagram, there is no restriction made to the data distribution generated by the production process. This approach is completely a nonparametric method. Both graphs Figure 1 and Figure 2 sketch out the performance of the multivariate quality process characterized by the sealing forces in the edges of gauze swabs of size 10×10 cm with 12 plies.

In fact,

• Both figures show the desired values of the quality of a product, these are represented by the diagonal line.

• Figure 1 gives a sequence of points that represents the quality distribution of the empirical sample with respect to the reference one, where the data are subject to changes in location and/or in scale.

• Figure 2 displays a sequence of other points, plotted differently, that represents the quality distribution of the empirical sample with respect to the reference one, where the data are centered with respect to the deepest point of *F* in order to uncover any change in the scale.

In general, consider the test of a null hypothesis asserting stability of a production process versus an alternative one that concerns the existence of shifts in location and/or in scale then the empirical sample has higher dispersion than that of the reference one. This is deduced from the fact that the resulting clouds, of centered depths or not, are located between both control limits, $L_{\max}$ and $L_{\min}$ arching around the diagonal line. The data depth points are spread out in the shape of an early half moon toward the reference sample axis.

6. Concluding Remarks

According to (8), the suggested *DD*-diagram outlines points d_i such that

$$d_i = (D_n(F, x(\omega)), D_n(F, y(\omega))),$$

for any $i = 1, 2, ..., n.$ So, the coordinates of each point are considered according to their chronological order. As a result, the first component of a point d_i may correspond to a data depth of a central point with respect to the sample issued from *F*, and the second component of the same point d_i may be a data depth of an outlying point with respect to the sample issued from *G*. Therefore, it is the best to compare data depths of points of the same kind in the drawn samples from F and G, respectively. Accordingly, the DD-diagram is redefined by

$$DD_0(F, G) = \{(D^0(F, x(\omega)), D^0(F, y(\omega))), \text{ for any } x(\omega), y(\omega) \in \Re^p\}, \quad (11)$$

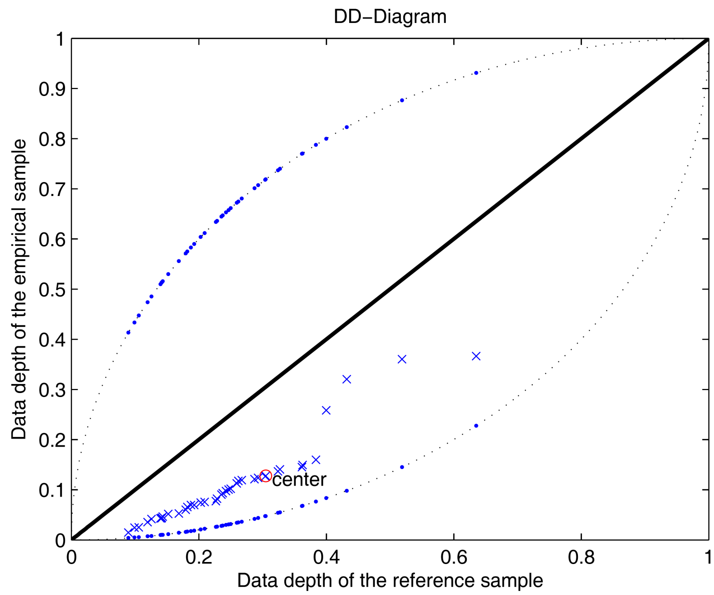

Figure 3. Ranked data depth of quality data on edge sealing forces of swabs.

where $D^0(\cdot, \cdot)$ indicates an arranged data depth computed using (1). The sample version of this new definition (11) of the DD-diagram is given by

$$DD_0(F_n, G_n) = \{(D_n^0(F, x(\omega)), D_n^0(F, y(\omega))), \text{ for any } x(\omega), y(\omega) \in \Re^p\}. \quad (12)$$

Using this definition (12), the resulting DD-diagrams of the data that we have, are drawn up in Figure 3 and in Figure 4.

From Figures 3 and 4, it is clear that the common points of the quality process are located in the lower-left corner of the DD-diagram. This means that the greater part of the observed measures shows shifts in position and/or in scale. As long as these changes in the multivariate quality data go weaker the sample points in the DD-diagram tend to the identified central point in the sample drawn from F. In Figure 4, the DD-diagram represents the observed measures with scale change only, i.e., shifts in position with respect to the identified centre are eliminated.

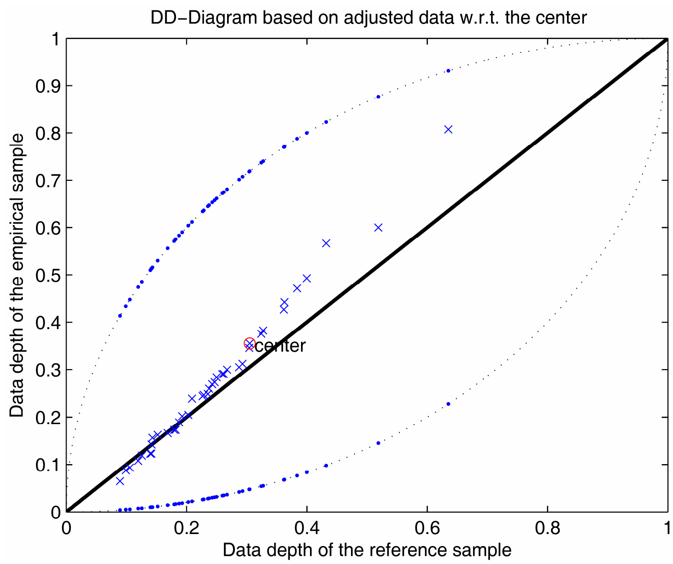

Figure 4. Ranked data depth of centered quality data on edge sealing forces of swabs.

In the light of what is given, the *DD*-diagram is a graphical device to visualise the behavior of multivariate quality process using data depth notion. The same tool may help in detecting changes of dispersion if data are centered with respect to the most central point of *F*. It is a successful work for SPC in monitoring multivariate quality measures simultaneously and to detect any dispersion change.

References

[1] D. L. Donoho and M. Gasko, Breakdown properties of location estimates based on half space depth and projected outlyingness, Ann. Statist. 20 (1992), 1803-1827.

[2] M. Hajlaoui, Maîtrise de la qualité univariée et multivariée: théories et applications, unpublished Doctoral thesis, University of Tunis, I.S.G. de Tunis, 2004.

[3] M. Hajlaoui and I. Nagati, A nonparametric quality control procedure using data depth, Proc. of the MSDM, Tunisia, 2 (2010), 34-40.

[4] R. Y. Liu, On a notion of data depth based on random simplices, Ann. Statist. 18 (1990), 405-414.

[5] R. Y. Liu, J. M. Parelius and K. Singh, Multivariate analysis by data depth: descriptive statistics, graphics and inference, Ann. Statist. 27 (1999), 783-858.