Quality improvement from the viewpoint of statistical method

de Mast, J.

Citation for published version (APA):
de Mast, J. (2002). Quality improvement from the viewpoint of statistical method
4 The multi-vari chart: a systematic approach

4.1 Introducing the multi-vari chart

The multi-vari chart is a useful graphical tool, which provides a graphical display of the behaviour of a CTQ in a running process. Variance components and suspicious patterns in the data are easily recognised.

Seder (1950) introduces the multi-vari chart and compares it with the \((\bar{X}, R)\)-control chart. The multi-vari chart presents an analysis of the variation in a process, hereby differentiating between three main sources:

1. *Intra-piece*, the variation within a piece, batch, lot et cetera;
2. *Inter-piece*, the additional variation between pieces;
3. *Temporal variation*, variation which is related to time.

Duncan (1986) presents the multi-vari chart as a graphical tool for testing joint homogeneity of means and variance in a nested structure. Bhote (1991) mentions the multi-vari chart as one of the Shainin DOE-tools. Its purpose is to provide clues on the factors that cause a substantial part of the process variation.

In view of the descriptions of these authors, I discern two applications for the technique:

1. As a tool to present the results of a homogeneity study, displaying variance components in the distribution of a CTQ. An example is provided in section 4.3.1, where a multi-vari chart present the results of a study after the homogeneity of the distribution of caffeine over a batch of coffee.
2. As a tool in exploratory studies, displaying patterns which could be related to assignable causes, and giving indications on important influence factors by showing in which class (intra-piece, inter-piece or temporal) they are likely to be found.

The literature on multi-vari charts has been descriptive rather than rigorous. Inferences based on the charts are made intuitively without sound statistical basis. It is the purpose of this chapter to make the analysis of multi-vari charts more rigid. The organization is as follows. Below, I present an introductory example and I describe how a multi-vari chart is composed. Thereupon, the models underlying the multi-vari chart analysis are given and I raise the basic questions to which the analysis should provide an answer. In section 4.2 formal test procedures are derived to facilitate the analysis of
the multi-vari chart. In section 4.3 the proposed analysis procedure is demonstrated from an example. A discussion concludes the chapter.


4.1.1 Example: weighing tea boxes

To illustrate the multi-vari chart I discuss a real life example which concerns a tea packing process. Tea is put into tea bags which in turn are put in a paper envelope. Finally, 20 of these are put in a box.

A characteristic which is considered important is the weight of the filled boxes. In order to evaluate the performance of the running process, data samples were collected. On each of 11 days 5 consecutive boxes were gathered from each of 4 packing machines. The weight of the $k^{th}$ box from the $j^{th}$ machine at the $i^{th}$ day is denoted by $Y_{ijk}$, $i = 1, \ldots, 11$, $j = 1, \ldots, 4$, $k = 1, \ldots, 5$.

Of interest is the structure of the variability in the process. To study this variability the multi-vari chart of figure 4.1 is constructed. On its $y$-axis it has the scale of the measurements, grammes in this case. On the $x$-axis are the indices of the samples. From each machine sample the minimum and the maximum value are indicated by a solid box. For each sample the two boxes are connected by a vertical line, thus representing the range of the machine sample. Within each day, sample means of the four machine samples are connected by a line. Finally, the crosses indicate the day means.

![Weight of tea boxes](image)

Figure 4.1 Multi-vari chart for weights of tea boxes.

Studying the chart in figure 4.1 we see that the sample taken at the seventh day at the second machine, and the eighth day’s samples of the first and second machines require a closer examination. Inspection of the corresponding individual weights learns that
there are four boxes that have a weight under 76 grammes or above 84 grammes. Since there ought to be 20 tea bags in a box and since the average weight of a tea bag is four grammes, it is likely that in these four boxes there is either one tea bag missing or one tea bag too much. I adjusted the weights of the four boxes by adding or subtracting the average weight of one tea bag. The multi-vari chart of the adjusted data is given in figure 4.2.

![Adjusted weight of tea boxes](image)

**Figure 4.2 Multi-vari chart for adjusted weights of tea boxes.**

From this chart inferences could be made concerning the process’s behaviour with respect to the weight of the tea boxes. The chart could answer questions such as: “Is the process stable or not?”, “Are there variance components other than the within-sample variation?”, “Can fixed differences among the various machines be found?”. From multi-vari charts such as they are used so far these questions can be answered only intuitively, since formal statistical tests for making inferences based on multi-vari charts are lacking.

### 4.1.2 Multi-vari chart: basics and models

A multi-vari chart aims to assess the extent to which the variability in a quality characteristic can be attributed to the effect of two or more specific factors. Typically in multi-vari studies, one of these factors is related to time — hours, days, weeks, etcetera — and one or more are related to streams of products — e.g., batches, lots, machines, operators. In what follows, I limit the number of factors to two, although in practice more factors can be used. One of the factors is related to time and it is considered random. The factor related to streams of products can be considered either a random or a fixed effect, depending on the situation. This leads to two different models to describe the process:

**Random effects model** If an inquirer used for instance *batches* as the second factor, he would consider a random effects model in which the batches factor is nested
The multi-vari chart within the time factor:

\[ Y_{ijk} = \mu + \beta_i + \gamma_{j(i)} + \epsilon_{k(ij)}, \quad i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K. \quad (4.1) \]

In this model \( \mu \) is the over-all mean. The effect of the \( I \) time instants is modelled in the \( \beta_i \), which are independent and identically distributed (i.i.d.) following a \( \mathcal{N}(0, \sigma^2_\beta) \) distribution. The effects of the \( J \) batches that are collected within each time instant are represented by \( \gamma_{j(i)} \), which are i.i.d. \( \mathcal{N}(0, \sigma^2_\gamma) \). The errors \( \epsilon_{k(ij)} \) in the \( j^{th} \) batch collected at time instant \( i \) are i.i.d. \( \mathcal{N}(0, \sigma^2_{\epsilon}) \). All random variables are independent.

**Mixed effects model** If an inquirer considers machines as the second factor (and he considers all machines used), a mixed effects model in which the machines factor is crossed with the time factor seems more appropriate:

\[ Y_{ijk} = \mu + \beta_i + \gamma_j + \epsilon_{k(ij)}, \quad i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K. \quad (4.2) \]

where the \( \gamma_j \) are now fixed effects. The assumptions on the \( \beta_i \) and \( \epsilon_{k(ij)} \) remain similar.

The choice between the random and the mixed effects model is made mostly on physical grounds. If the levels of a factor are limited in number one would choose a mixed effects model. If the effects are drawn from a (theoretically) infinite number, one would use a random model.

The variation that stems from the \( \epsilon_{k(ij)} \) is referred to as **within-batch variation**. The additional variation induced by the second factor is called **between-batches variation** or **batch-to-batch variation**. The additional variation caused by the time factor is called **temporal variation**.

The aim is to use the models to decide from a multi-vari chart whether

1. The within-batch variation is constant over the various batches;
2. The additional batch-to-batch variation, either random or fixed, is significantly non-zero;
3. The additional temporal variation is significantly non-zero.

When a multi-vari chart is used in a homogeneity study, the questions above give operational meaning to the term 'homogeneity' and the answers constitute the study's result. In an exploratory study, the answers to the abovementioned questions indicate in which class the important influence factors are to be found (thus getting a zooming-in strategy off the ground; see section 2.2.4). However, the multi-vari chart provides more clues in this situation. The data, plotted in their order, make it possible to identify assignable causes of variation when the inquirer sees nonrandom patterns.

### 4.2 Test procedures

In this section I present statistical tests which are of help in drawing conclusions on the questions that were listed in the preceding section. Rather than using ratios of sums-of-squares as test statistics I base these tests on control limits comparable to the control limits in control charts methodology. Hence, tests are not over-all tests, such as the sums-of-squares based F-test, but instead batches are tested individually. The advantage is that the resulting tests form an easy to use graphical tool.
4.2.1 Tests under the random model

The statistics that are indicated graphically in the multi-vari chart are:

- $R_{j(i)} = \text{Range}_k \{Y_{ijk}\}$ for $j = 1, \ldots, J$ within $i = 1, \ldots, I$ (the ranges of the observations in each batch).
- $\bar{Y}_{ij}$ for $j = 1, \ldots, J$ within $i = 1, \ldots, I$ (the batch means).
- $\bar{Y}_i$ for $i = 1, \ldots, I$ (the means at the time instants).

**Homogeneity of within-batch spread**

The first of the three questions that was raised in the preceding section concerns homogeneity of the within-batch variation. It can be formalized as a hypothesis. Given a mean value $\sigma^2 = \frac{1}{IJ} \sum \sigma_{ij}^2$ for the within-batch variance we test for each batch $j(i)$

$$
H_0 : \sigma_{ij}^2 = \sigma^2, \text{ vs.} \\
H_A : \sigma_{ij}^2 \neq \sigma^2.
$$

Visually this test is established by comparing the lengths of the vertical lines which indicate the batch ranges. The more formal test is based on control limits. I calculate an upper and a lower limit in between which the range of the measurements within each batch is supposed to be. If the range is beyond one of these limits, this is considered evidence that the within-batch variation in the particular batch differs significantly from the within-batch variation in other batches.

**Figure 4.3 Control limits for within-batch spread for an artificial data set.**

To compute the control limits I follow the approach of Does and Schriever (1992). The Upper and Lower Control Limit ($UCL$ and $LCL$ respectively) are

$$
UCL = L + c_1 S, \\
LCL = L + c_2 S,
$$
where $L$ is an estimate of the location of the ranges $R_{j(i)}$ and $S$ is an estimate of their spread. The constants $c_1$ and $c_2$ are chosen such that the probabilities under the null-hypothesis of an $R_{j(i)}$ being below the upper control limit or below the lower control limit are $1 - \alpha/2$ and $\alpha/2$ respectively. Thus, the probability of wrongly rejecting the null-hypothesis is $\alpha$ for each batch. In control charting methodology it is customary to take $\alpha = 0.0027$, which corresponds to the well-known ‘three sigma’-limits.

Let us take $L = \bar{R}$ and $S = \frac{d_2}{d_3} \bar{R}$. Here, $\bar{R} = \frac{1}{IJ} \sum_{i,j} R_{j(i)}$ is the average range, $d_3$ is chosen such that $\text{Var}(R_{j(i)}/d_3) = \sigma^2$ and $d_2$ such that $E(R/d_2) = \sigma$ (see Duncan, 1986).

Writing $b_m = 1 + \frac{d_2}{d_3} c_m$, $m = 1, 2$, we have $L + c_m S = b_m \bar{R}$. To establish $b_1$:

$$1 - \frac{\alpha}{2} = P_{H_0}(R_{j(i)} \leq UCL) = P_{H_0} \left( \frac{\bar{R}}{R_{j(i)}} \geq \frac{1}{b_1} \right) =
$$

$$P_{H_0} \left( \frac{R_{j(i)}}{\frac{1}{IJ-1} \sum_{l(k) \neq j(i)} R_{l(k)}} \leq \frac{b_1 (IJ - 1)}{IJ - b_1} \right). \quad (4.3)$$

This probability can be computed using an approximation due to Patnaik (1950). In this paper the distribution of a squared average of ranges is approximated by a chi-square distribution:

$$\nu \left( \frac{1}{IJ-1} \sum_{l(k) \neq j(i)} R_{l(k)} \right)^2 \sim \chi^2 \nu \text{ approximately.}$$

The $a$ and $\nu$ are chosen such that the first two moments of both the approximated distribution and the chi-square distribution are equal. These constants can be found in tables, for instance in table D3 in Duncan (1986). The value $a$ is Duncan’s constant $d_2^2$. It can be obtained, together with $\nu$, by taking the number of samples $g$ equal to $IJ - 1$ and the size of the samples $m$ equal to $K$. Approximately, in (4.3) we have a random variable consisting of a range divided by the square root of a statistic that follows a chi-square distribution and that is independent of the range in the numerator. Such a statistic has a studentized range (Tukey) distribution. We have:

$$1 - \frac{\alpha}{2} \approx Q_{K,\nu} \left( \frac{b_1 (IJ - 1)}{IJ - b_1} \right),$$

where $Q_{n,\nu}$ is the studentized range distribution function for a range based on $n$ observations and having $\nu$ degrees of freedom in its denominator. From this result an expression for the $UCL$ can be derived:

$$UCL = \frac{IJ Q_{K,\nu}^{-1} \left( 1 - \frac{\alpha}{2} \right)}{a(IJ - 1) + Q_{K,\nu}^{-1} \left( 1 - \frac{\alpha}{2} \right)} \bar{R}.$$

Likewise, we can calculate an expression for the $LCL$. The result is

$$LCL = \frac{IJ Q_{K,\nu}^{-1} \left( \frac{\alpha}{2} \right)}{a(IJ - 1) + Q_{K,\nu}^{-1} \left( \frac{\alpha}{2} \right)} \bar{R}.$$

In figure 4.3 these control limits are indicated in a multi-vari chart for an artificial data set. The data were collected from a simulation of model (4.1) with
o $\beta_i = 0$ for all $i$;

o $\gamma_{ij(i)} \sim \mathcal{N}(0, 0.25)$ and

o $\epsilon_{k(ij)} \sim \mathcal{N}(0, 1)$.

For each batch, the batch’s minimum is taken as zero point. The lower horizontal bar indicates the LCL, the upper one indicates the UCL. The null-hypothesis is rejected if the batch’s maximum is not in between these two bars.

Constancy of variance is necessary for the remaining tests. Although not rejecting the hypothesis of homogeneous variance does not imply accepting it, in practice the variance is considered constant if there is no strong evidence supporting heterogeneity of variance.

**Homogeneity of the batch means**

The next hypothesis that can be tested is that of no additional batch-to-batch variation. To test this hypothesis constancy of within-batch spread is assumed. I do not assume absence of temporal variation. The model that is implied by the null-hypothesis is

$$Y_{ijk} = \mu + \beta_i + \epsilon_{k(ij)},$$

that is, model (4.1) without batch means term. Under the null-hypothesis all observations collected at a single time instant share a common distribution. We test whether the mean of the $j^{th}$ batch of the $i^{th}$ time instant differs from the average batch mean more than is to be expected given the within-batch variation. The null- and alternative hypothesis in symbols are

$$H_0 : \text{Var}(\bar{Y}_{ij} - \bar{Y}_i) = \frac{I-1}{JK} \sigma^2, \text{ vs. }$$

$$H_A : \text{Var}(\bar{Y}_{ij} - \bar{Y}_i) > \frac{I-1}{JK} \sigma^2.$$

The formal test is based on control limits for the batch means. If a batch mean is beyond these limits we consider this evidence that there is additional batch-to-batch variation. The control limits are chosen such that for each batch mean the probability of wrongly rejecting the null-hypothesis equals $\alpha$. Since absence of temporal variation is not assumed, the location of the control limits is shifted according to the $\bar{Y}_i$...

Define

$$MS_{wb} = \frac{1}{IJ(K-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} (Y_{ijk} - \bar{Y}_{ij})^2.$$

Analogous to the control limits for the ranges, we have, taking $L = \bar{Y}_i$ and $S = \sqrt{MS_{wb}},$

$$UCL = \bar{Y}_i + c_1 \sqrt{MS_{wb}}, \text{ and }$$

$$LCL = \bar{Y}_i + c_2 \sqrt{MS_{wb}}.$$

Note that $(\bar{Y}_{ij} - \bar{Y}_i)/\sigma \sqrt{\frac{I-1}{JK}}$ follows a standard normal distribution. From calculating

$$1 - \frac{\alpha}{2} = P_{H_0} \left( \frac{\bar{Y}_{ij} - \bar{Y}_i}{\sqrt{\frac{I-1}{JK} \sqrt{MS_{wb}}}} \leq \frac{c_1}{\sqrt{\frac{I-1}{JK}}} \right) =$$

$$P_{H_0} \left( \frac{\bar{Y}_{ij} - \bar{Y}_i}{\sqrt{\frac{I-1}{JK} \sqrt{MS_{wb}}}} \leq \frac{c_1}{\sqrt{\frac{I-1}{JK}}} \right)$$

77
Figure 4.4 Control limits for batch means for an artificial data set.

we find the control limits to be

\[ UCL = \bar{Y}_i + \frac{\sqrt{J-1}}{JK} t_{I,(J-1)}^{-1} \left( 1 - \frac{\alpha}{2} \right) \sqrt{M_{wb}}, \]

\[ LCL = \bar{Y}_i - \frac{\sqrt{J-1}}{JK} t_{I,(J-1)}^{-1} \left( \frac{\alpha}{2} \right) \sqrt{M_{wb}}, \]

\( t_{\nu} \) being the \( t \)-distribution with \( \nu \) degrees of freedom. The chart of figure 4.4 demonstrates these control limits. Several batch means are beyond the control limits, which was to be expected in view of the way the data were generated.

**Homogeneity over time**

Control limits for testing homogeneity over time can be computed in the same way. Here we have two choices: the batch-to-batch variation can be assumed negligible or not. I choose for the latter option here. The testing problem is

\[ H_0 : \text{Var}(\bar{Y}_i) = \frac{\sigma_i^2}{J} + \frac{\sigma^2}{JK}, \text{ vs.} \]

\[ H_A : \text{Var}(\bar{Y}_i) > \frac{\sigma_i^2}{J} + \frac{\sigma^2}{JK}, \]

stating that the variation in the time instant means can (or cannot) be accounted for completely by the within-batch and the between-batches variation.

Notice that \( (\bar{Y}_i - \bar{Y}_.) / \sqrt{\sigma_i^2 + \frac{\sigma^2}{J} \sqrt{I_{ij}}} \) has a standard normal distribution under the null-hypothesis. We define

\[ \text{MS}_{wb} = \frac{K}{I(J-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} (\bar{Y}_{ij} - \bar{Y}_i) \]

Since \( I(J-1)\text{MS}_{wb}/(K\sigma_i^2 + \sigma^2) \) has a chi-square distribution with \( I(J-1) \) degrees of freedom, the \( t \)-distribution can be used to calculate the control limits. Taking \( L = \bar{Y}_i \);
4.2 Test procedures

\[ S = \sqrt{\frac{\text{MS}_{bb}}{K}}, \text{ we find} \]

\[
\begin{align*}
UCL &= \bar{Y}_- + \sqrt{\frac{I-1}{IJ} t_{I(J-1)}(1 - \frac{\alpha}{2})} \sqrt{\frac{\text{MS}_{bb}}{K}}, \\
LCL &= \bar{Y}_- - \sqrt{\frac{I-1}{IJ} t_{I(J-1)}(1 - \frac{\alpha}{2})} \sqrt{\frac{\text{MS}_{bb}}{K}}.
\end{align*}
\]

These control limits are drawn in the chart of figure 4.5. No points are beyond the control limits. Notice that together the three charts in figures 4.3–4.5 provide an easy to understand analysis of the variance components present in the given data set.

![Chart showing control limits](image)

**Figure 4.5 Control limits for day means for an artificial data set.**

### 4.2.2 Tests under the mixed model

In case the second factor has only a limited number of levels the inquirer should use the mixed effects model (4.2). An example is the machines factor in the tea packing example. The statistics indicated in the multi-vari chart are now:

- \( R_{ij} = \text{Range}_k \{Y_{ijk}\} \) (the range of the observations collected from machine \( j \) at time instant \( i \)).
- \( \bar{Y}_{ij} \) (the mean of the observations collected from machine \( j \) at time instant \( i \)).
- \( \bar{Y}_{i.} \) (the means at the time instants).

The hypothesis concerning homogeneity of within-machine samples variation and the related test are analogous to the hypothesis and test under the random effects model. The control limits for testing consistency over time can be calculated from straightforward adjustments to the test under the random effects model. Using \( \text{MS}_{w,b} = \)
The multi-vari chart

\[
\frac{1}{IJ(K-1)} \sum_i \sum_j \sum_k (Y_{ijk} - \bar{Y}_{ij})^2
\]
as an estimator for error variance we find

\[
UCL = \bar{Y} + \sqrt{\frac{I-1}{IJK} t_{IJ(K-1)}^{-1}(1 - \frac{\alpha}{2}) \sqrt{\text{MS}_{wb}}}
\]
\[
LCL = \bar{Y} - \sqrt{\frac{I-1}{IJK} t_{IJ(K-1)}^{-1}(\frac{\alpha}{2}) \sqrt{\text{MS}_{wb}}}
\]

A relevant hypothesis in the mixed effects model is the hypothesis concerning fixed differences among machines:

\[
H_0 : \gamma_j = 0 \text{ for } j = 1, \ldots, J, \text{ vs.} \\
H_A : \gamma_j \neq 0 \text{ for at least one } j.
\]

In a multi-vari chart fixed differences emerge as a pattern in the \(Y_{ij}\), that is repeated from time instant to time instant. To test the significance of observed differences, however, it is more natural to consider the machine means \(\bar{Y}_j\) rather than considering the \(Y_{ij}\). For this reason, the multi-vari chart, which does not indicate the \(\bar{Y}_j\), is a tool not as appropriate for testing this hypothesis as an other graphical technique: the Analysis of Means (ANOM). A thorough treatise can be found in Schilling (1973). The ANOM technique is illustrated using the data from the tea packing process.

![ANOM for tea packing data](image)

**Figure 4.6** Analysis of means applied to tea packing data.

From the multi-vari chart with control limits for the variation within the machine samples and for the day means I found no evidence for inconsistencies. To test for fixed differences among the machines, I plotted the ANOM-chart in figure 4.6. Indicated in the chart are the \(\bar{Y}_j\) for \(j = 1, 2, 3\) and 4. The decision limits UDL and LDL were calculated from Schilling’s formulas:

\[
UDL = \bar{Y} + h\alpha(IJ(K-1), 4) \sqrt{\frac{J - 1}{J} \frac{\text{MS}_{wb}}{IK}}
\]
\[
LDL = \bar{Y} - h\alpha(IJ(K-1), 4) \sqrt{\frac{J - 1}{J} \frac{\text{MS}_{wb}}{IK}}
\]
with $h_\alpha$ the upper bound for the studentized maximum absolute deviation from population mean in normal samples as given in Schilling’s table. The probability under the hypothesis of no fixed differences that at least one $Y_{ij}$ is beyond the decision limits is smaller than $\alpha$. If one or more machine means are beyond the decision limits, this is considered evidence that there are fixed differences among the machines.

From the analysis of the tea packing data I would conclude that the process is stable in time and that the error variance is constant. There is no evidence of fixed differences among the machines. I also analysed the data set under the random effects model. This analysis as well did not provide evidence to reject the hypothesis of consistency. Appart from the apparent problem with the number of tea bags in a box no assignable causes were found.

4.3 Using a multi-vari chart in practice

4.3.1 Example: homogeneity of caffeine distribution

As an example of a fruitful combination of an analysis of variance using F-tests and a multi-vari study using the tests that I presented above, I consider a set of data collected from a caffeine extraction process. The purpose of the experiment was to give a characterization of the distribution of caffeine over a batch of coffee. From five consecutive batches samples were taken at two locations in the process, namely, at the extractor’s outlet and at a conveyor belt. Four samples were taken at each sampling location and from each batch. From each sample an operator determined the caffeine content in three measurements.

![Multi-vari chart for caffeine extraction data.](image)

Figure 4.7 Multi-vari chart for caffeine extraction data.

A good starting point for the analysis of the data is to draw a multi-vari chart of the
The multi-vari chart
data and study it to obtain a first impression of the data’s structure. The multi-vari
chart is given in figure 4.7. Next, the following model, describing the measured caffe-
feine content, was fit to the data.

\[ Y_{ijkl} = \mu + \beta_i + \gamma_j + \delta_{k(ij)} + \epsilon_{l(ijk)}, \]

\( i = 1, \ldots, 5; \ j = 1 \) (Extractor’s outlet), 2 (Conveyor); \( k = 1, \ldots, 4 \) and \( l = 1, 2, 3. \) The
batch effect is modelled in the \( \beta_j \), assumed to be \( \mathcal{N}(0, \sigma^2_b) \) distributed. The sampling
locations \( \gamma_j \) are modelled as a fixed effect, crossed with the batch effect. Nested in this
product is the sample effect \( \delta_{k(ij)} \), which is supposed to have a \( \mathcal{N}(0, \sigma^2_\delta) \) distribution.
The error is modelled in \( \epsilon_{l(ijk)} \), normally distributed with variance \( \sigma^2. \)
Apart from two outliers in the first batch no evidence was found that the normality
assumption for the error was inappropriate. The analysis of variance is presented in
table 4.1.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>EMS</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Batch</td>
<td>4</td>
<td>7.94x10^2</td>
<td>1.99x10^2</td>
<td>24\sigma_\delta^2 + 3\sigma_\beta^2 + \sigma^2</td>
<td>1075</td>
<td>0.00</td>
</tr>
<tr>
<td>Sampl. Locat.</td>
<td>1</td>
<td>3.39x10^-5</td>
<td>3.39x10^5</td>
<td>( 60 \sum \gamma_j^2 ) + 3\sigma_\gamma^2 + \sigma^2</td>
<td>1.84</td>
<td>0.18</td>
</tr>
<tr>
<td>Sample</td>
<td>34</td>
<td>6.28x10^-4</td>
<td>1.85x10^-5</td>
<td>( 3\sigma_\gamma^2 + \sigma^2 )</td>
<td>5.70</td>
<td>0.00</td>
</tr>
<tr>
<td>Error</td>
<td>80</td>
<td>2.59x10^-4</td>
<td>3.24x10^-6</td>
<td>\sigma^2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
<td>8.03x10^-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1 Analysis of variance for caffeine percentage.

A first interesting conclusion is that there is no evidence that the location where sam-
ple s are taken has influence on the measured values. A second conclusion is that there
is additional variation present between different samples from a single batch. This
implies that the caffeine is not homogeneously distributed over a batch of coffee.
We proceed the analysis using the multi-vari chart. The sampling location is omitted as
effect. In the chart of figure 4.8 we added control limits for the within-sample variation.
The two signals in the first batch are caused by the two outliers that were already
found. Due to the rounding off of the data, the three measurements in the last sample
of the last batch are identical, leading to a zero within-sample variation. This causes
the signal in this sample. We conclude that the error is homogeneous.
In the chart of figure 4.9 we added control limits for the sample means. We see that
in the batches 1, 2, 4 and 5, which have a lower average caffeine content, there is
no evidence of additional sample-to-sample variation. Thus, the evidence that was
found in the analysis of variance for a heterogeneous caffeine distribution is solely
accounted for by the third batch, which has a higher caffeine content. A conservative
conclusion is that the caffeine distribution is quite homogeneous in batches which
have a relatively low caffeine content.
The fact that the F-test found sample to be a significant effect whereas the multi-vari-
test did not, is a demonstration of one of the differences between the two tests. The
F-test is an over-all test whereas the multi-vari test, testing individual samples, is not.
4.3 Using a multi-vari chart in practice

Figure 4.8 Control limits for within-sample spread for caffeine extraction data.

4.3.2 Discussion

The similarities between multi-vari charts and control charts go beyond the use of control limits. In fact, one could regard the multi-vari chart as an extension of the control chart. To see this, remember that a regular \((\bar{X}, R)\)-control chart is intended to 1. establish that the within-samples variation is constant, and 2. study whether there is additional sample-to-sample variation. The multi-vari chart is an extension of this methodology in that it allows for more additional variance components to be part of the analysis. Hence, if the products follow different streams, the expected stream-to-stream differences can be integrated in the analysis.

The hypotheses that are studied by means of the multi-vari chart can as well be tested using analysis of variance techniques such as the F-test. There are some differences between the two approaches. An important difference is the fact that an F-test is a test on all batch-means or day-means simultaneously, whereas the approach with control limits leads to individual tests for each mean separately. As a disadvantage of the latter option I mention that the confidence level for the combination of the individual tests is not well defined. I chose for the individual tests anyway for the reason that in general an inquirer is interested not only in the fact that there are inhomogeneities, but also in the question to which sample, batch or machine they refer (a similar discussion is presented in section 5.3.1).

A second difference in approach is the graphical nature of the analysis in charting methodology versus the numerical approach in ANOVA. This difference leads to several advantages for the multi-vari chart. First of all, a graphical display of the results of an analysis is easier understood by non-statisticians than a result that is summarized in a number of statistics. Secondly, if the process is disturbed at some time instant and/or at a particular machine, the multi-vari chart gives immediate clues about the
The multi-vari chart

Figure 4.9 Control limits for sample means for caffeine extraction data.

time instant and/or the machine where the inquirer should search for trouble. The case of the tea packing process, where the number of tea bags in a box showed a few irregularities, illustrates this point.

Finally, the graphical display of the data can give much more indications on properties of the process than just the conclusions on hypotheses concerning homogeneity of means and variance. As Tukey (1977, p. vi) wrote it, the value of a good graphical presentation is that "it forces us to notice what we never expected to see" (emphasis is Tukey's). Especially if the technique is used in an exploratory study, a graphical display of the data in their original order is vital to identify nonrandom patterns hinting at assignable causes.