Reprinted from IFPress (2019) 22(2) 3,4

Action Standards for Machines and Humans in Quality Assurance

Daniel M. Ennis and Benoît Rousseau

Background: Descriptive analysis of consumer products is extensively used to make quality decisions often involving a standard against which test products are compared. Current analysis of this type of data typically involves single attribute tests or even qualitative decisions based on spider plots of the attribute profiles.

In 1993, a U.S. patent was issued that described a machine process quite similar to human processing[1]. The patent concerned a process for automatic high speed image inspection of finished product labels in manufacturing. A feature of this patent is that the machine inspection method was based on theory concerning how humans represent percepts in memory and use them in decision making. A machine was presented with multiple typical variants of a package and stored the results in computer memory. Then a new item was presented and the machine's task was to decide if this item is likely to have been drawn from the same population as those in the training set. Each item in the training set was a photograph containing about 60,000 pixels and these pixels were converted to 64 package segment means with their associated variance-covariance matrix to record the variances of the 64 segments and their dependencies. These data were then used to calculate a statistic leading to accepting or rejecting the item depending on whether it was typical or not of current production.

In this technical report we will demonstrate how these ideas can be used to evaluate products based on sensory variables from a descriptive panel.

Scenario: Your expert descriptive panel conducts routine sensory tests of current production on texture attributes of your company's cookies daily. The panel considers 15 sensory quality attributes[2] shown in Table 1. Your method for comparing a particular day's production to the historical process mean involves single attribute tests using targeted upper and lower control values. You would like to take into account the fact that many of the descriptive attributes are correlated and have different variances as was discussed in the image inspection patent previously mentioned. You suspect that you have been missing detectable differences and also at times rejecting product that otherwise would have passed in a more comprehensive analysis. In other words, you need a multivariate control chart with an appropriate action standard that properly represents the current data.

Roughness	Loose Particles	Cohesiveness
Dryness	Fracturability	Tooth Pack
Hardness	Particle Size	Particles
Denseness	Uniformity of Chew	Oily
Grittiness	Moisture Absorption	Chalky

Table 1. Descriptive panel variables for texture (scored by an expert panel).

Limitations of Univariate Tests: A possible approach when evaluating multiple attributes is to consider them individually. In this case, a set of upper and lower control limits for each attribute is based on the mean and variance of each variable for the control set. In the cookie example, the control set is current production. Suppose, for example, that there are two independent attributes and that the data on these attributes have been standardized to normal z-scores. Hence these attributes have zero mean and unit variance. If 95% control limits are used on each attribute separately, then an acceptable product falls between −1.96 and +1.96 on each attribute. This leads to a square acceptance region for the two attributes as shown in Figure 1. The distribution of two independent normal variables with equal variance, such as the variables described above is circular, not square.

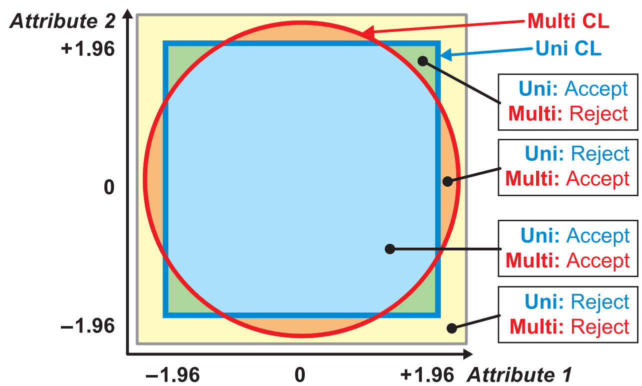

Figure 1. Acceptance/rejection regions for two attributes one at a time (square = Univariate Control Limit) and together (circle = Multivariate Control Limit).

This means that points of equal likelihood are arranged on a circle around the mean. If a square acceptance region is used to assess values from a circular distribution, the conclusions will be subject to Type I errors (rejecting an acceptable product) and Type II errors (accepting an unacceptable product) at a higher level than targeted. Figure 1 illustrates the various cases that can arise with univariate 95% tests compared to a bivariate circular acceptance region with a radius of 2.447 (the 95% circular boundary for a standard bivariate normal distribution.) The discrepancy between the multivariate tests and the single attribute tests may become even greater if attributes are correlated. In the general multivariate case, univariate tests involve a hypercubic acceptance region while the multivariate tests involve a hyperspheric acceptance region of which the square and circle of Figure 1 are special cases.

Limitations of Spider Plots: The spider plot has proven to be a useful tool to display sensory data. However, as a control chart, it has some limitations. Figure 2 shows the fifteen texture attributes for a cookie displayed in a spider plot with 95% confidence limits set on each attribute (the red lines). A test product is displayed in this figure (the blue line) that falls within the upper and lower boundaries on each attribute. Nevertheless, this product is defective because it falls outside the 95% multivariate boundary for these attributes.

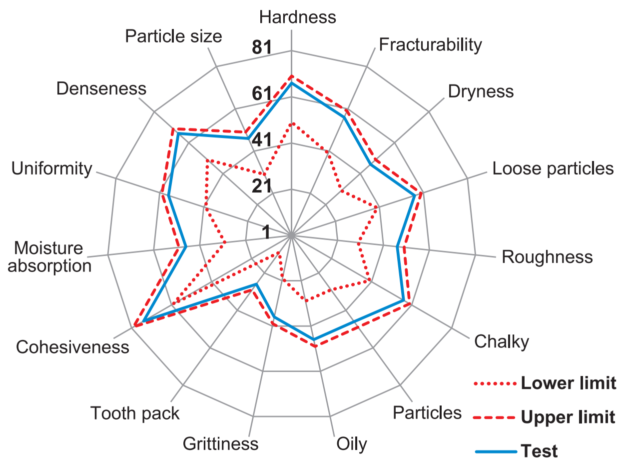

Figure 2. A spider plot may erroneously display an out-of-specification product as acceptable.

It can be seen that this product has cumulative evidence that it is *not* typical of the control set, although we cannot find fault with it based on individual attribute analyses. The failure of Figure 2 to diagnose that the test product is outside the multivariate specification illustrates a weakness in using spider plots to display multivariate sensory data. Spider plots are really univariate displays of multivariate data and are not ideally suited to set action standards for multivariate quality assessment of the type discussed.

Chi-square as an Action Standard: A product's sensory effects can be represented as a set of multivariate attributes. Assuming that this set is distributed multivariate normally, the attributes may be correlated or have different variances. We can transform the original attributes to a set of values that are distributed as a standard multivariate normal with mean zero, unit variance on all dimensions and zero correlation among dimensions[3]. This is achieved using a Cholesky factorization of the variance-covariance matrix of the control samples. The opportunity to apply the results of this analysis arises from the fact that sums of squared standard normal variables follow a chi-square distribution. If a control set is reduced to standard form, a set of attributes for a test product can be checked against the control set. In order to do this, the same transformation used for the control products is applied to the test product and the sum of squared elements of the resulting set are computed. This value is a χ^2 with n degrees of freedom corresponding to the number of attributes. This χ^2 can be compared to standard values to see if it is significantly larger than zero. If it is, the sample tested is not typical of current production. Geometrically, we have transformed a hyper-ellipsoid to a hyper-sphere. Squared distances to the mean of the sphere are χ^2 distributed with degrees of freedom equal to the number of attributes.

Sensory Testing of Current Production: From the data-base on current production you compute the means and variances for each attribute along with the correlation co-efficients among the attributes. The data are transformed to multivariate standard z-scores for which the mean is zero, variances are one, and attribute correlations are zero. This function is readily available in many computer programs including the R software environment[4]. Application of the new method over daily production for four weeks is shown in Figure 3. Figure 3 also shows which production samples would have been rejected or accepted using univariate tests. This multivariate control approach to sensory quality assurance will give you higher confidence in the accept/reject decisions you need to make on a daily basis.

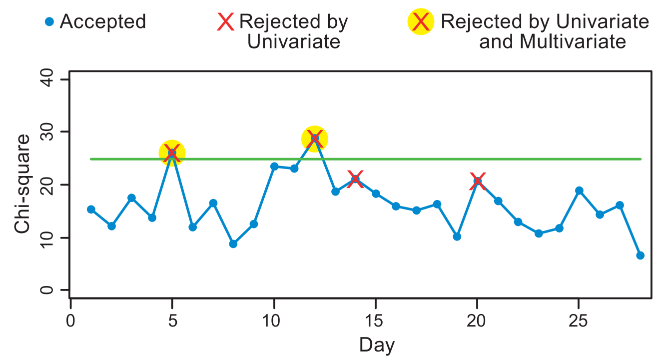

Figure 3. Each point in this figure represents a chi-square value for each of 28 days. Compared to the χ^2 action standard (shown as the green line), there are two occasions when the sample was rejected on days 5 and 12. Otherwise, the process was in control. Univariate inspection rejects two other samples on days 14 and 20.

Although both univariate and χ^2 tests agree on the rejection of samples on day 5 and 12, samples rejected on days 14 and 20 by the univariate method would *not* be rejected by the multivariate (χ^2) method.

Conclusion: Multiple, interdependent attributes are usually associated with consumer products. In order to measure and manage the quality of these products, multivariate techniques are important and useful. Individual attribute evaluations lead to two basic errors: Rejecting acceptable products and accepting products that should have been rejected. Multivariate quality control tools manage these errors by appropriately accounting for multivariate effects and providing a sound theoretical basis for decision-making. One more important aspect of this approach must be considered: The consumer-relevance of the variables and their levels. Some variables, although detectable, may not be important to consumers and others may be highly important. Including this aspect of quality would extend and improve not only multivariate inspection but other types of process control that are only based on statistical tests.

References

1. Ennis, D. M. (1993). Image inspection method and apparatus. *U.S. Patent* 5,208,870.

2. Meilgaard, M., Civille, G. V., and Carr, B. T. (1987). *Sensory Evaluation Techniques.* CRC Press: Boca Raton.

3. Ennis, D. M. and Bi, J. (2000). Multivariate quality control with applications to sensory data. *Journal of Food Quality*, **23**(6), 541-552.

4. R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/