Batch Process: Dry matter content in buttercream

Civit, S., Vilardell, M., Hess, A., Matthew, Z., Ge, Y., Caballe, A.

February 14, 2011

1 Motivation and Background

[1] report the results of an experiment that examined the influence of dry matter content in buttercream .

A typical problem in the application of statistical process control is that of the several variance components of the variable. We can encounter it, for example, in the control of the dry matter content in buttercream. This production is a batch process since the buttercream is mixed in a cutter. The variation of the dry matter content comes from the deviation of the compound, of the sample units and of the measurements of the same sample unit. The variance of the process is estimated from these components.

2 Data Modelling

The "own regulation" for dry matter content in buttercream is a minimum of 45%. This value is given on the production sheet, which was approved by the competent food authority. In the manufacturing process, the dry matter content has to be set so that despite the variations it does not drop under 45%. Prior to introducing the process control, measurements were carried out in order to examine the significance of the variance components and to estimate them. For this prior investigation out of r=8 compounds, which can be considered as batches, m=5 samples were taken, on which n=3 parallel measurements were carried out.

The nested model:
$X_{ijk} = \mu + A_i + B_{j(i)} + e_{ijk}$

1. X_{ijk} represents the kth measurement of the gth gene expression of jth chemical compound of the ith group evaluated in the $k-th$ cell line.
2. μ the mean of expression
3. A_i the effect of the ith group
4. $B_{j(i)}$ the effect of the jth compound within the ith group
5. e_{ijk} the measurement error term.

It is assumed that A_i, $B_{j(i)}$ and e_{ijk}t's are normal distributed with 0 expected value with standard deviations; σ_A σ_B and σ_e respectively, and that these two sequences of random variables are independent of each other. The data were evaluated by analysis of variance (ANOVA) based on this nested model.

3 Usage

3.1 The data

The data for this example is supplied in the VARCOMPCI package;

```
> library(varcompci)
> data(batch_process)
> dim(batch_process)
```

```
[1] 120   5
```

```
> head(batch_process)
```

```
  X Comp Batch Rep Measurement
1 1    1     1   1       47.66
2 2    1     1   2       47.33
3 3    1     1   3       47.76
4 4    1     2   1       47.43
5 5    1     2   2       47.95
6 6    1     2   3       47.90
```

3.1.1 A VARCOMPCI overview

To use VARCOMPCI package is necessary to specify the data set, the response, the factors that we would like to test and write a design matrix;

```
> dsn = "batch_process"
> response = "Measurement"
> totvar = c("Comp", "Batch")
> Matrix = cbind(c(1, 0), c(1, 1))
> Matrix
```

```
     [,1] [,2]
[1,]    1    1
[2,]    0    1
```

Applying the varcompci function to get an ANOVA for the nested model;

```
> x <- varcompci(dsn = dsn, response = response, totvar = totvar,
+     Matrix = Matrix, sasnames = TRUE)
> x["ANOVA"]
```

```
            df       SS       MS        F Pval
Comp         7 94.78577 13.54082 30.81993    0
Batch(Comp) 32 14.05929  0.43935  8.96395    0
resid       80  3.92107  0.04901       NA   NA
```

The random effect estimates are obtained from;

```
> x["EMS"]
```

```
            EMS
Comp        "var(Resid) + 3var(Comp:Batch) + 15var(Comp)"
Batch(Comp) "var(Resid) + 3var(Comp:Batch)"
resid       "var(Resid)"
```

```
> x["CI"]

            Method      LB Estimate      UB
Comp         TBGJL 0.36444  0.87343 3.70901
Batch(Comp)  TBGJL 0.07791  0.13011 0.23981
resid        Exact 0.03677  0.04901 0.06861
```

As we can see from the ANOVA table, all the factors taking account in this model were significant for the model.

This situation is typical for many types of batch processes and in this context it is advisable to monitor the different variance components

Thus, estimates for the random effects σ_A σ_B and σ_e are provided using the CI method with the varcompci function. The estimates were 0.87, 0.13 and 0.05 respectively.

The global performance of the model can be obtained from;

```
> x["aic"]

AIC (smaller is better)
                12.00989

> x["bic"]

BIC (smaller is better)
                126.2970
```

The covariance parameter estimates could be obtained from;

```
> x["variance"]

            Covariance paramater
Comp                  0.87343145
Batch(Comp)           0.13011319
resid                 0.04901333
```

The Random and Fixed Mean Squares are provided from;

```
> x["Meansq"]

               Mean Sq
Comp        13.54082464
Batch(Comp)  0.43935292
resid        0.04901333
```

3.2 VARCOMPCI Conclusions for buttercream study

According to the results of the ANOVA, the variation of the dry matter content in buttercream is due to several significant factors and the large variation between the samples implies with this compound the buttercream remained all too inhomogeneous in the cutter, even though the average dry matter content shows no problems. It would be justified to take corrective action in the process and to make the buttercream more homogeneous before filling the cartons with it.

4 Note

This vignette is based on: VARCOMPCI: A Package for Computation of Confidence Intervals for Variance Components of Mixed Models in R"

References

[1] ”Statistical Process Control with Several Variance Components In the Dairy Industry.” *Food Control*, **12**, 119-125