Wysowl Pty Ltd
ACN 010 677 022

18 Miller Street
CHERMSIDE QLD 4032
Ph: Intl+ 61 7 3359 9918
Mob: Intl+ 61 407 149 966
www.wysowl.com.au
john@wysowl.com.au

HIERARCHICAL STUDIES

Separating Batch, Sampling and Test Error

In some instances, sampling and analytical error can be greater than process variation, but how can we now? A hierarchical study can separate process, sampling and analytical error so that manufacturers are able to identify the true source of variation.

Consider an example where product is made in batches. The total variation can be expressed as:

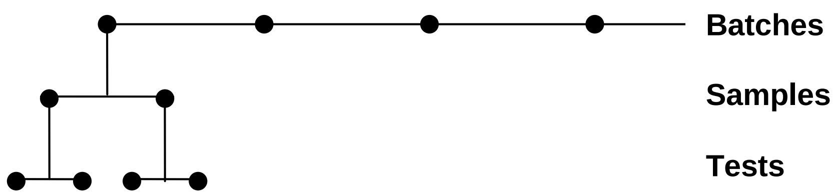
Total variation = batch to batch variation + sample to sample variation + test to test variation

Or as: $\sigma^2 = \sigma^2_B + \sigma^2_S + \sigma^2_T$

Where:
σ^2 = Total variance
σ^2_B = Batch variance
σ^2_S = Sample variance
σ^2_T = Test variance

Experiment Design
The experimental design for this case study was for two samples to be taken from each of fifteen batches and for each sample to be tested twice, as per Figure 1.

FIGURE 1



In this case study, there were 15 batches. Two samples were taken from each batch, and each sample was tested twice.

Experiment design is a critical element of this process. As will be seen, the maths is straightforward. It is wise to use the services of a competent statistician when conducting the design. As the great R. A. Fisher once observed:

> 'Sometimes the only thing you can do with a poorly designed experiment is to try to find out what it died of.'

However, provided a long enough series of batches is used (say, 15 or more) S and T can be as small as 2 in most cases. If you suspect high sampling error, increase S to 3. If you suspect high analytical error, increase T to 3.

Figure 2 shows the data laid out by batch, sample and test. It shows also the calculations, explained later.

We take B batches, S samples from each batch and conduct T tests on each sample. In this case:
$B = 15 \quad S = 2 \quad T = 2$

Table 2

A	B	C	D	E	F	G	H	I	J	K
Batch	Sample	Test	Moisture	Variance for samples	VT	Sample Means	Variance of sample means	VS	batch ave	VB
1	1	1	40	0.5000	0.900	39.5	45.125	29.05	34.75	21.72
		2	39							
	2	1	30	0.0000		30				
		2	30							
2	1	1	26	2.0000		27	1.125		26.25	
		2	28							
	2	1	25	0.5000		25.5				
		2	26							
3	1	1	29	0.5000		28.5	98		21.5	
		2	28							
	2	1	14	0.5000		14.5				
		2	15							
4	1	1	30	0.5000		30.5	21.125		27.25	
		2	31							
	2	1	24	0.0000		24				
		2	24							
5	1	1	19	0.5000		19.5	3.125		18.25	
		2	20							
	2	1	17	0.0000		17				
		2	17							
6	1	1	33	0.5000		32.5	28.125		28.75	
		2	32							
	2	1	26	2.0000		25				
		2	24							
7	1	1	23	0.5000		23.5	40.5		28	
		2	24							
	2	1	32	0.5000		32.5				
		2	33							
8	1	1	34	0.0000		34	12.5		31.5	
		2	34							
	2	1	29	0.0000		29				
		2	29							
9	1	1	27	0.0000		27	8		29	
		2	27							
	2	1	31	0.0000		31				
		2	31							
10	1	1	13	4.5000		14.5	60.5		20	
		2	16							
	2	1	27	4.5000		25.5				
		2	24							
11	1	1	25	2.0000		24	2		25	
		2	23							
	2	1	25	2.0000		26				
		2	27							
12	1	1	29	0.0000		29	3.125		30.25	
		2	29							
	2	1	31	0.5000		31.5				
		2	32							
13	1	1	19	0.5000		19.5	50		24.5	
		2	20							
	2	1	29	0.5000		29.5				
		2	30							
14	1	1	23	0.0000		23	2		24	
		2	23							
	2	1	25	0.0000		25				
		2	25							
15	1	1	39	2.0000		38	60.5		32.5	
		2	37							
	2	1	26	2.0000		27				
		2	28							

Procedure

Calculate VT

1. Calculate the variance for the test results for each sample, as noted in column E at Figure 2.
2. Calculate the average of these variances; sum them and divide by BS, as seen in column F of Figure 2.

This provides VT, an estimate of σ^2_T, the estimated test error.

Calculate VS

1. For each sample, calculate the average of the test results, as noted at column G of Figure 2.
2. Now calculate the variance of these sample means, seen at column H of Figure 2.
3. Calculate the average of these variances; sum them and divide by B, see column I of Figure 2.

This gives VS, an estimate of $\sigma^2_S + \sigma^2_T/T$, or the sampling error plus the test error.

Calculate VB

1. Calculate the average of the test results for each batch, see column J at Figure 2.
2. Calculate the variance of these averages, as noted in column K of Figure 2.

This provides VB, an estimate of $\sigma^2_B + \sigma^2_S/S + \sigma^2_T/ST$, or the batch error plus the sampling error plus the test error.

$$VT = \sigma^2_T = 0.9$$

$$\begin{aligned}
VS &= \sigma^2_S + \sigma^2_T/T \\
\sigma^2_S &= VS - \sigma^2_T/T \\
&= 29.05 - (0.9/2) \\
&= 29.05 - 0.45 \\
&= 28.6
\end{aligned}$$

$$\begin{aligned}
VB &= \sigma^2_B + \sigma^2_S/S + \sigma^2_T/ST \\
\sigma^2_B &= VB - \sigma^2_S/S - \sigma^2_T/ST \\
&= 21.72 - (28.6/2) - (0.9/4) \\
&= 21.72 - 14.3 - 0.225 \\
&= 7.195
\end{aligned}$$

Therefore:

Test σ^2	= 0.9	and Test σ	= 0.95
Sample σ^2	= 28.6	and Sample σ	= 5.35
Batch σ^2	= 7.19	and Batch σ	= 2.68
Total σ^2	= 36.69	and Total σ	= 6.06

In this case the largest contributor to total variability was sampling error.

The data for this example was taken from, *Statistics for Experimenters*, Box, Hunter and Hunter, Wiley, Second Edition, 2005