Practical 1: Simple Hierarchical Structures

Dankmar Böhning

Southampton Statistical Sciences Research Institute
University of Southampton, UK

1 Exercise

Use the contamination data of the Lecture 1 (`purity10.dta`) to answer the following questions:

1. determine the values of the variance components
2. is there a significant batch effect?
3. what is the overall mean of the amount of purity?
4. How does the mean or the confidence interval change when there is no random (batch) effect assumed?

2 Exercise

Consider the test data on the 12 beef carcasses (used in Lecture 1) where two tests determined the number of colonies (on log-scale).

Answer the following questions:

1. Try to explain why $\sigma_B^2/(\sigma_B^2+\sigma^2)$ is a reasonable measure for the reliability of the test!
2. Is there a significant beef carcass random effect?
3. Which test has the higher reliability, the new or the standard test?
4. How can the significance of any difference between be tested?

Solution for Exercise 2

Answer the following questions:

1. Try to explain why $\sigma_B^2/(\sigma_B^2+\sigma^2)$ is a reasonable measure for the reliability of the test!
Suppose the test produces identical measurements for each of the 2 determination on the 12 carcasses. Then the estimate for σ^2 would be zero and the estimate of reliability 1 which is what we observe.

2. Is there a significant beef carcass random effect?
For the standard test the LRT is 5.34 with P-value = 0.0104 and for the new test the LRT is 5.63 with P-value = 0.0088. Hence for both tests the random carcass effect is necessary. In the lecture the following table was provided:

test	σ_B^2	σ^2	*reliability* $\sigma_B^2/(\sigma_B^2+\sigma^2)$
standard	*0.0172*	*0.0112*	*0.6055*
new test	*0.0291*	*0.0180*	*0.6183*

3. *indicating that the new test has the higher reliability.*

4. How can the significance of any difference between be tested?
This can be approached by testing the following hypothesis:

$$H_0 : \sigma_{B,0}^2/(\sigma_{B,0}^2+\sigma_0^2) = \sigma_{B,1}^2/(\sigma_{B,1}^2+\sigma_1^2)$$

or the slightly more restrictive hypothesis

$$H_0 : \sigma_{B,0}^2 = \sigma_{B,1}^2 \text{ and } \sigma_0^2 = \sigma_1^2,$$

where the second index refers to the test. Under the null-hypothesis we have the four parameters μ_0, μ_1 (means for the two tests) and σ_B^2, σ^2 whereas under the alternative we have the six parameters μ_0, μ_1 (means for the two tests) and $\sigma_{B,0}^2, \sigma_{B,1}^2, \sigma_0^2, \sigma_1^2$.

The separate log-likelihoods for the two tests are 9.7371 and 4.0597, jointly 9.7371+4.0597=13.7968. For the null-hypothesis the log-likelihood is 11.4095. Hence the LRT is 4.7746 with a P-value 0.0918 which is not significant.