

Pooling Designs for Microarrays

Christina Kendzierski, Ph.D.
University of Wisconsin

Abstract:

In a microarray experiment, messenger RNA samples are oftentimes pooled across subjects out of necessity, or in an effort to reduce the effect of biological variation. Two basic problems in such experiments are to estimate the nominal expression levels of a large number of genes and to identify genes that are differentially expressed. Pooling samples will affect these tasks, but the exact effects are not yet known as the approach has not been systematically studied in this context. We consider how mRNA pooling affects expression estimates by assessing the finite sample performance of different estimators for designs with and without pooling. Conditions under which it is advantageous to pool mRNA are defined; and general properties of estimates from both pooled and non-pooled designs are derived under these conditions. A formula is given for the total number of subjects and arrays required in a pooled experiment to obtain gene expression estimates and confidence intervals comparable to those obtained from the no pooling case. The assumptions that facilitate derivation of this formula are considered using data from a cDNA microarray experiment. The implications of pooling on the assessment of differential expression will also be discussed.