Assessing DNA copy numbers in large scale studies using genomic arrays

Rob Scharpf and Ingo Ruczinski

Department of Biostatistics
Johns Hopkins Bloomberg School of Public Health

August 26, 2010

Genomic arrays

Copy number estimates are noisy

Plate effects

Confounding of plate and disease

Genotype estimates are more robust
Allele specific copy numbers

At locus $i$, for subject $j$ in plate $p$, we have for allele $k \in \{A, B\}$

$$I_{kijp} = \nu_{kijp} + \phi_{kijp}C_{kijp}$$

$$\hat{c}_{kijp} = \max \left\{ \frac{1}{\phi_{kijp}} (I_{kijp} - \nu_{kijp}), 0 \right\}$$