A BAYESIAN FEATURE ALLOCATION MODEL FOR TUMOR HETEROGENEITY

Peter Mueller,[1]   Juhee Lee,[2]   Yuan Ji [3]   Kamalakar Gulukota[3]

[1] *UT Austin.*
[2]*UCSC.*
[3]*Northshore Hospital.*

Resumen

We characterize tumor variability by hypothetical latent cell types that are defined by the presence of some subset of recorded SNV's. (single nucleotide variants, that is, point mutations). Assuming that each sample is composed of some sample-specific proportions of these cell types we can then fit the observed proportions of SNV's for each sample. In other words, by fitting the observed proportions of SNV's in each sample we impute latent underlying cell types, essentially by a deconvolution of the observed proportions as a weighted average of binary indicators that define cell types by the presence or absence of different SNV's. Taking a Bayesian perspective, we proceed with a prior probability model for all relevant unknown quantities, including in particular a prior probability model on the binary indicators that characterize the latent cell types by selecting (or not) the recorded SNV's. Such prior models are known as feature allocation models. We define a simplified version of the Indian buffet process, one of the most traditional feature allocation models.