

---

## ABSTAT17

IGC, April 10-13, 2017

### Small Project: Gibbs Sampling

Consider a sample of 947 portuguese individuals, for which we know their Blood Types:

A	B	AB	O
420	83	38	422

This is a typical problem, when we know the phenotype and want to estimate the allele frequencies and information is missing. If we knew the genotypes, the estimation would be simple. Since it is not known, one can use iterative procedures like Newton-Raphson, EM algorithm or Gibbs Sampling.

#### 1. The Problem

- Understand the problem.
- Why Gibbs?
- Identify the random vector and its distribution.
- Identify the distribution of the parameters vector.

#### 2. Estimate de allele frequencies using the Gibbs sampler. Use R.

#### 3. Prepare some slides including:

- Description of the problem;
- Description of the method (workflow);
- Eventually some parts of the script;
- The answer to the problem;
- Comparison of the obtained frequencies with the known frequencies for the portuguese population: O ( 65%) | A ( 29%) | B ( 6%)