minsage (**min**imal **sa**mple size for **ge**notypes)

User documentation

minsage is designed to calculate the sample size of genotypes minimally required to ensure that all alleles with a specified frequency at one locus are detected with a given confidence.

The program is based on the following publications:

Gregorius, H.-G. (1980) The probability of losing an allele when diploid genotypes are sampled. Biometrics, 36, 643-652.

Gillet, E. M. (1999) Minimum sample sizes for sampling genetic marker distributions. In: Gillet, E. M. (ed.), Sampling strategies for marker analysis. Final compendium of the research project development, optimisation and validation of molecular tools for assessment of biodiversity in forest trees in the European Union DGXII Biotechnology FW IV rResearch programme molecular tools for biodiversity.

The required formulae are given in corollaries 2 and 3 in Gregorius (1980).

minsage is started by typing "minsage".

Within the program, you are prompted to specify the following parameters:

1. “allele frequency”:
   Minimum allele frequency α that is to be detected. The entered value must be less than 1.
2. “confidence”:
   Probability to detecting the specified allele. The entered value must be less than 1.
3. “uniformly distributed alleles or biallelic markers”:
   The specified allele can be set to be the less frequent allele of a biallelic marker. Alternatively, if neither the number of alleles nor the genotypic frequencies are known, alleles can set to be uniformly distributed. In that

case, the number of alleles at the marker is given by $1/\alpha$. By default, a biallelic marker is assumed.

The output renders the minimal sample size N of genotypes needed to detect alleles of frequency α with the specified confidence. The results are given both for the case that Hardy- Weinberg equilibrium can be or cannot be assumed.
The maximal sample size that can be calculated by this program is 100,000 genotypes.

Copyright: Prof. Dr. Andreas Ziegler
For questions or comments please contact:
Dr. Inke R. König
Institute of Medical Biometry and Statistics
University at Lübeck, Germany
email: Inke.Koenig@imbs.uni-luebeck.de