Boundary Crossing Probabilities in Linkage Analysis

Josée Dupuis
Genome Therapeutics Corporation, Waltham, MA

David Siegmund
Department of Statistics, Stanford University, Stanford, CA

Abstract

Two novel problems of boundary crossing probabilities that arise in genetic linkage analysis based on sib pairs are addressed by modifications of techniques developed to solve problems of sequential analysis.

1 Introduction.

Genome scans in linkage analysis lead to problems involving boundary crossing probabilities (cf. Feingold, Brown and Siegmund, 1993), which can be addressed using methods developed in sequential analysis during the 1970's. In this paper we discuss two problems where genetically natural conditions lead to novel variations.

The goal of linkage analysis is to identify regions of the genome harboring genes affecting particular traits. In humans these are often genes that increase susceptibility to particular diseases, and it is convenient to speak of "disease" genes, although other traits affected by an individual's genetic makeup can be studied similarly.

A convenient unit for the linkage analysis of human diseases is an affected sib pair. Given $N \geq 1$ unrelated sib pairs, we let $X_{i,t}^{(N)}$ denote the number of pairs that share i alleles identical by descent ($i = 2, 1, 0$) at locus t, and let $X_t^{(N)} = (X_{0,t}^{(N)}, X_{1,t}^{(N)}, X_{0,t}^{(N)})$. (An allele is shared identical by descent by two relatives if it is inherited from a common ancestor.) With probability $1/2$ a sibling pair can inherit zero or one allele identical by descent from their mother and similarly from their father. These events are independent, so the probability that two siblings share i alleles identical by descent at locus t is given by $EX_{2,t}^{(1)} = 1/4, EX_{1,t}^{(1)} = 1/2, EX_{0,t}^{(1)} = 1/4$.

For an *affected* sib pair, on a chromosome containing a disease locus at τ the (conditional on being affected) distribution of alleles shared identical