University of Wisconsin — Madison

Department of Statistics

On Empirical Likelihood for a Semiparametric Mixture Model

F. Zou

J. P. Fine

B. S. Yandell

Department of Statistics

Technical Report # 1030

December 2000

Department of Statistics
University of Wisconsin–Madison
1210 West Dayton Street
Madison WI 53706-1685

Phone: 608/262-2598
Fax: 608/262-0032
Internet: yandell@stat.wisc.edu

Summary

Plant and animal studies of quantitative trait loci provide data which arise from mixtures of distributions with known mixing proportions. Previous approaches to estimation involve modelling the distributions parametrically. We propose a semiparametric alternative which assumes that the log ratio of the component densities satisfies a linear model, with the baseline density unspecified. It is demonstrated that a constrained empirical likelihood has an irregularity under the null hypothesis that the two densities are equal. A factorization of the likelihood suggests a partial empirical likelihood which permits unconstrained estimation of the parameters. The partial likelihood is shown to give consistent and asymptotically normal estimators, regardless of the null. The asymptotic null distribution of the log-partial likelihood ratio is chi-square. Theoretical calculations show that the procedure may be as efficient as the full empirical likelihood in the regular set-up. The usefulness of the robust methodology is illustrated with a rat study of breast cancer resistance genes.

Some key words: Boundary condition; Breeding experiment; Exponential tilt; Lagrange multiplier; Molecular marker; Profile likelihood; Weak convergence.

1. INTRODUCTION

Our motivation is the identification of genetic loci influencing quantitative traits. This use of molecular marker data in breeding experiments has traditional applications in plant and animal studies, such as improving grain yield in rice and increasing milk production in cows. Recently, animal models have proved useful for complex human diseases. For example, controlled crosses of inbred rat strains (Lan et al, 2000) characterized several genomic regions conferring breast cancer resistance or susceptibility.

The standard method for quantitative trait loci is interval mapping (Lander & Botstein 1989). Since markers are observed at known locations, the genotypes between the locations are missing. In backcross studies, this leads to a two sample mixture model at putative loci. The component densities, f and g, are associated with the possible genotypes. The mixing probabilities are determined by the recombination fractions between a locus and the flanking markers (Knapp et al, 1990). The set-up differs from those in which the focus is inference for unknown mixing proportions when some data is from f and g (Titterington, Smith & Makov, 1985). Murray & Titterington (1978) and Hall (1981) discuss nonparametric approaches. With quantitative traits, the proportions are known, vary among observations, and direct information on the distributions may be unavailable. The emphasis is testing that a locus has no genetic influence, that is, $H_0 : f = g$.

Following early work on mixture models (Hosmer, 1973), most mapping methods employ a likelihood analysis with f and g specified parametrically (Doerge, Zeng, & Weir, 1997). Kruglyak & Lander (1995) proposed an ad hoc nonparametric test for H_0. A formal procedure for robust estimation of the distributions does not exist. In the rat study, the traits are tumor counts. A challenge is relaxing the usual parametric assumptions. We adopt a semiparametric model subsuming discrete and continuous outcomes. The densities are related by an exponential tilt but are otherwise unspecified (Anderson, 1979). That is,

$$g(x) = \exp(\beta_0 + \beta_1 x) f(x), \tag{1}$$

where $(\beta_0, \beta_1) \in \mathcal{H}$, a compact subset of $\mathcal{R}^2$. Normal variates with common variance follow (1), as do exponential densities. The binomial and poisson distributions are other specialisations. Including $x^2, x^3, \ldots$ in the loglinear model for g/f enhances its flexibility.

The exponential tilt model resembles the Cox (1972) regression model in which the ratio of two hazard functions is linear in covariates. A partial likelihood not involving the baseline hazard gives efficient estimators for the coefficients in the proportional hazards model (Cox, 1975). An analagous partial likelihood has yet to be developed for model (1). Qin (1999) used a profile empirical likelihood (Owen, 1988, 1990) to construct confidence intervals for the mixture proportions and for $F = \int f$ and $G = \int g$. However, estimation of (β_0, β_1) enforces constraints on F and G and is computationally involved. Furthermore, in §2, we show that the constraints induce a boundary condition and Theorems 1-4 (Qin, 1999) do not hold under H_0. That is, the profile likelihood has an irregularity when $f = g$.

To derive a valid test of the null hypothesis, the profile empirical likelihood is factored into two pieces. One part involves the constraints while the other does not: the partial profile empirical likelihood. The partial likelihood gives consistent and asymptotically normal estimates of (β_0, β_1) regardless of $f = g$. The log partial likelihood ratio for testing $\beta_0 = \beta_1 = 0$ has a chi-square distribution. Maximising the partial likelihood is straightforward, avoiding constrained optimisation of the full likelihood. Theoretical calculations show when $f \neq g$, the estimators may be as efficient as those from the full likelihood. New estimators for F and G are proved to be uniformly consistent and to converge to Gaussian processes.

In §3, simulations show that the partial profile empirical likelihood works well with realistic sample sizes. The semiparametric methods are illustrated on the breast cancer data in §4 and some remarks conclude in §5.

2. ESTIMATION AND INFERENCE

2.1 Data and Profile Empirical Likelihood

The data are independent observations from K mixtures with known proportions and component densities f and g satisfying model (1). Let X_{kj} be the jth observations from the kth mixture with density $\lambda_k f(x) + (1-\lambda_k)g(x), j = 1, 2, ..., n_k, k = 1, 2, ..., K$. Assume $0 \leq \lambda_k \leq 1$, $\lambda_1 \neq \ldots \neq \lambda_K$, and $f(x)$ is nondegenerate. If $K = 1$, then the model is nonidentifiable. That is, there are multiple (β_0, β_1, f) giving the same distribution for the data. With $d(x) = \{\lambda_1 + (1-\lambda_1)\exp(\beta_0^* + x\beta_1^*)\}f^*(x)$, $(\beta_0^*, \beta_1^*, f^*)$ and $(0, 0, d)$ yield equivalent models. In the sequel, $K \geq 2$.

Define

$$\omega_k(x, \beta) = \lambda_k + (1-\lambda_k)\exp(\beta_0 + x\beta_1)$$

The likelihood is

$$\begin{aligned} L(\beta, F) &= \prod_{k=1}^{K}\prod_{j=1}^{n_k} dF(x_{kj}) \prod_{k=1}^{K}\prod_{j=1}^{n_k} \omega_k(x_{kj}, \beta) = \prod_{i=1}^{n} dF(z_i) \prod_{k=1}^{K}\prod_{j=1}^{n_k} \omega_k(x_{kj}, \beta) \\ &= \prod_{i=1}^{n} p_i \prod_{k=1}^{K}\prod_{j=1}^{n_k} \omega_k(x_{kj}, \beta) \end{aligned} \qquad (2)$$

where $n = \sum_{k=1}^{K} n_k$, $z = (z_1, z_2, ..., z_n) = (x_{11}, x_{12}, ..., x_{Kn_K})$ and $p_i = dF(z_i)$.

Unconstrained maximisation of $L(\beta, F)$ does not provide a valid estimator for β. To see this, note that the likelihood increases monotonically in $p_i, i = 1, \ldots, n$ and β_0. For a given β, it is natural to constrain p to the set

$$C_\beta \doteq \left[p \mid \sum_{i=1}^{n} p_i = 1, p_i \geq 0, \sum_{i=1}^{n} p_i\{\exp(\beta_0 + z_i\beta_1) - 1\} = 0 \right].$$

This ensures that the estimators for F and G are cumulative distribution functions. To compute the maximum likelihood estimator of β, say $\tilde{\beta}$, one first maximises $L(\beta, F)$ over $p \in C_\beta$. This yields a profile likelihood in β which is then maximised to obtain $\tilde{\beta} = (\tilde{\beta}_0, \tilde{\beta}_1)$ (Qin, 1999). The estimators $\tilde{F}(x) = \sum_i^n \tilde{p}_i I(z_i \leq x)$ and $\tilde{G}(x) = \sum_i \exp(\tilde{\beta}_0 + z_i\tilde{\beta}_1)\tilde{p}_i I(z_i \leq x)$ are evaluated at $\tilde{p} = (\tilde{p}_1, \ldots, \tilde{p}_n)$, where $\tilde{p}$ maximises $L(\tilde{\beta}, F)$ over $p \in C_{\tilde{\beta}}$.

Similar to Qin & Lawless (1994), for any fixed β such that C_β is not empty, maximising

$L(\beta, F)$ over C_β gives

$$p_i = \frac{1}{n} \frac{1}{r(z_i, \beta)\{1 + \alpha h(z_i, \beta)\}} \tag{3}$$

where α is the Lagrange multiplier determined by

$$\frac{1}{n} \sum_{i=1}^{n} \frac{h(z_i, \beta)}{1 + \alpha h(z_i, \beta)} = 0, \tag{4}$$

with $h(x, \beta) = \{\exp(\beta_0 + x\beta_1) - 1\} r(x, \beta)^{-1}$, $r(x, \beta) = 1 + \xi\{\exp(\beta_0 + x\beta_1) - 1\}$, and $\xi = \sum_{k=1}^{K} n^{-1} n_k (1 - \lambda_k)$. Plugging (3) into (2) gives the profile log-likelihood $l\{\beta, \tilde{\alpha}(\beta)\} = l_1\{\beta, \tilde{\alpha}(\beta)\} + l_2(\beta) - n \log n$, where

$$\begin{aligned} l_1\{\beta, \tilde{\alpha}(\beta)\} &= -\sum_{i=1}^{n} \log\{1 + \tilde{\alpha} h(z_i, \beta)\}, \\ l_2(\beta) &= -\sum_{i=1}^{n} \log\{r(z_i, \beta)\} + \sum_{k=1}^{K} \sum_{j=1}^{n_k} \log\{\omega_k(x_{kj}, \beta)\} \end{aligned}$$

and $\tilde{\alpha}(\beta)$ solves equation (4). Maximising $l(\beta, \tilde{\alpha})$ in $(\beta, \tilde{\alpha})$ may be unreliable because the function may have many saddlepoints and the maximiser must satisfy a simplex condition (Qin & Lawless, 1994). Another method evaluates $\tilde{\alpha}(\beta)$ explicitly for each β, which is computationally intensive. This contrasts with certain models (Qin, 1998) for which the parameter of interest and the Lagrange multiplier may be treated separately.

2.2 Irregularity of Profile Empirical Likelihood

The issue is that C_β may be empty for some β and the maximiser of $L(\beta, F)$ may not exist. The problem occurs when the true value of β, $\beta_T = (\beta_{0T}, \beta_{1T})$, is 0. The irregularity seems to have been overlooked in Theorems 1-4 in Qin (1999). This is precisely stated in the following result; see appendix for proof.

Theorem 1.
(i): C_β is not empty $\iff$ $\beta = (\beta_0, \beta_1) \in J_n(z) \stackrel{\text{def}}{=} \{(\beta_0, \beta_1) \mid \min_{i=1}^n(\beta_0 + z_i\beta_1) \leq 0 \leq \max_i^n(\beta_0 + z_i\beta_1)\}$.
(ii): If $\beta_T \neq 0$, then there exists a neighbourhood of β_T, $N(\beta_T)$, such that for every $\beta \in N(\beta_T)$, $\beta \in J_n(z)$ as $n \to \infty$.
(iii): If $\beta_T = 0$, then there exists no such $N(\beta_T)$.

If $\beta_T \neq 0$, then for n large enough, there exists a neighbourhood of β_T such that for every $\beta \in N(\beta_T)$, C_β is not empty. However, there is no neighborhood of 0 in which every $\beta \in J_n(z)$. This happens because $\beta = (\beta_0, 0)$ is not in $J_n(z)$ whenever $\beta_0 \neq 0$. In essence, the constraints produce a boundary condition at the origin in which all finite α satisfy (4).

As in Lemmas 1 and 2 of Qin (1993), we can show that when $\beta_T \neq 0$, the constraint has an implicit solution $\tilde{\alpha}(\beta)$ in a $O(n^{-1/3})$ neighbourhood of β_T and $\tilde{\alpha}(\beta)$ is uniformly $O(n^{-1/3})$. Furthermore, it is easy to prove that $\tilde{\beta}$ is consistent and asymptotically normal. When $\beta_T = 0$, there is no guarantee the implicit solution of (4) is $O(n^{-1/3})$ in a $O(n^{-1/3})$ neighbourhood of β_T. This means the techniques used to derive the limiting behaviour of $\tilde{\beta}$ when $\beta_T \neq 0$ do not apply under H_0.

2.3 Partial Profile Empirical Likelihood

The Lagrange multiplier is a nuisance parameter. The irregularity of $l\{\beta, \tilde{\alpha}(\beta)\}$ occurs because α has known value 0 but is estimated to ensure that $\tilde{F}$ and $\tilde{G}$ are distribution functions in finite samples. The partial profile empirical likelihood, $l_2(\beta)$, does not depend on the constraints, while $l_1\{\beta, \tilde{\alpha}(\beta)\}$ does. Hence, the boundary condition is due to l_1.

A reasonable estimator for β is $\hat{\beta} = \text{argmax}_\beta\{l_2(\beta)\}$. Since $l_1 = 0$ when $\tilde{\alpha}(\beta) = 0$, $\hat{\beta}$ is the unconstrained maximiser of the full profile empirical likelihood. The asymptotic properties of the partial likelihood procedure are given below; see appendix for proof.

Theorem 2. Assume $\|h\|^3$ and $\|\frac{\partial h}{\partial \beta}\|$ are bounded by integrable functions in $N(\beta_T)$.
(i): For large enough n, with probability 1, $\frac{\partial l_2}{\partial \beta} = 0$ has a solution $\hat{\beta}$ in the interior of the interval $|\beta - \beta_T| \leq n^{-1/3}$. That is, $\hat{\beta}$ is $n^{1/3}$-consistent for β_T. Further, $\sqrt{n}(\hat{\beta} - \beta_T) \xrightarrow{\mathcal{L}} N(0, B)$, where $B = S^{-1}VS^{-1}$, $S = E\{n^{-1}\partial^2 l_2(\beta_T)(\partial\beta\partial\beta^T)^{-1}\}$ and $V = \frac{1}{n}\text{var}\{\partial l_2(\beta_T)(\partial\beta)^{-1}\}$.
(ii): $2l_2(\hat{\beta}) \xrightarrow{\mathcal{L}} \chi^2_1$ under H_0.

The estimator $\hat{\beta}$ is consistent and asymptotically normal and the partial likelihood ratio test has a chi-square distribution under H_0. However, B may not equal $-S^{-1}$, as in classical likelihood theory. Inferences for β must be based on the sandwich variance estimator $\hat{B} = \hat{S}^{-1}\hat{V}^{-1}\hat{S}^{-1}$, where $\hat{S} = n^{-1}\partial^2 l_2(\hat{\beta})(\partial\beta\partial\beta^T)^{-1}$,

$$\hat{V} = n^{-1}\sum_{k=1}^{K}\sum_{j=1}^{n_k}\left\{\frac{\partial r(x_{kj}, \hat{\beta})(\partial\beta)^{-1}}{r(x_{kj}, \hat{\beta})} - \frac{\partial w_k(x_{kj}, \hat{\beta})(\partial\beta)^{-1}}{w_k(x_{kj}, \hat{\beta})}\right\}^{\otimes 2},$$

and for a vector v, $v^{\otimes 2} = vv^T$.

2.4 Theoretical Comparison of $\tilde{\beta}$ and $\hat{\beta}$

Since $\hat{\beta}$ is easier to compute than $\tilde{\beta}$ and is valid regardless of β_T, the relative efficiency of the estimators when $\beta_T \neq 0$ is of interest. One might expect that l_1 and the constraint (4) have extra information about β. We show formally that $\tilde{\beta}$ has variance bounded by that of $\hat{\beta}$. The result is stated precisely below; see appendix for details.

Theorem 3. Under the regularity conditions in Theorem 2 when $\beta_T \neq 0$:
(i): The estimator $\tilde{\theta} = (\tilde{\beta}, \tilde{\alpha})^T$ from $l\{\beta, \tilde{\alpha}(\beta)\} \xrightarrow{p} (\beta_T, 0)^T$ and

$$\sqrt{n}\left\{\tilde{\theta} - (\beta_T, 0)^T\right\} \xrightarrow{\mathcal{L}} N(0, \tilde{B}) \quad \textit{where } \tilde{B} = \tilde{S}^{-1}\tilde{V}\tilde{S}^{-1}, \tag{5}$$

$$\tilde{S} = \begin{pmatrix} S & S_{12} \\ S_{21} & s_{22} \end{pmatrix}, \quad \textit{and} \quad \tilde{V} = \begin{pmatrix} -S - \delta S_{12}S_{21} & -\delta S_{12}s_{22} \\ -\delta S_{21}s_{22} & s_{22} - \delta s_{22}^2, \end{pmatrix},$$

where S_{12}, S_{21}, s_{22}, and δ are defined in the appendix. Thus, $\sqrt{n}(\tilde{\beta} - \beta_T) \rightarrow N(0, \tilde{B}_{11})$ where $\tilde{B}_{11} = -S^{-1} - \frac{1}{s_{22} - S_{21}S^{-1}S_{12}} S^{-1}S_{12}S_{21}S^{-1}$.
(ii): $\tilde{B}_{11} - B = (\delta - \frac{1}{s_{22} - S_{21}S^{-1}S_{12}})S^{-1}S_{12}S_{21}S^{-1} \leq 0$.

For regular β_T and two or more mixtures, $\tilde{\beta}$ has limiting covariance which equals that for $\hat{\beta}$ plus the negative semi-definite matrix in (ii).

The efficiency loss can be quantified in various settings using the formulas for $\tilde{B}_{11}$ and B in the appendix. In all settings, $\text{var}(\tilde{\beta}_1)\{\text{var}(\hat{\beta}_1)\}^{-1} = 1$ after roundoff, but not so for β_0. In Table 1, $\text{var}(\tilde{\beta}_0)\{\text{var}(\hat{\beta}_0)\}^{-1}$ is given for normal, exponential and poisson mixtures. The mixture proportions are $\lambda = (\lambda_1, \ldots, \lambda_K)$. The probability of an observation with proportion λ_i is $\tilde{\rho}_i$, where $\sum_i \tilde{\rho}_i = 1$ and $\tilde{\rho} = (\tilde{\rho}_1, \ldots, \tilde{\rho}_K)$. The relative efficiency is ≈ 1 when all data is directly from f and g and > 0.95 in most other cases, even when $K = 2$, $\lambda_1 = 0.7$, and $\lambda_2 = 0.5$. The smaller $|\lambda_1 - \lambda_2|$ is, the closer the true model is to $K = 1$.

An anomalous result occurs with normal densities when $f(x) = g(-x)$ and $0 < \lambda_1, \ldots, \lambda_K < 1$. In these set-ups, the variance ratios may be less than 0.50. An explanation is $\beta_0 = 0$ but $\beta_1 \neq 0$. This is confirmed by calculations under a variety of distributions meeting the condition. The peculiarity is absent when $f \approx g$ and both coefficients are roughly zero.

2.5 Estimating F and G

To make inference about F and G, one may first test H_0 using $l_2(\beta)$. If H_0 is not rejected, then both F and G may be estimated with the empirical distribution from the pooled data. Otherwise, $l\{\beta, \tilde{\alpha}(\beta)\}$ may be used to obtain the estimates (Qin, 1999). Difficulties are that the inferential properties of this two-step procedure are unclear and estimation of F and G after rejecting $F = G$ requires constrained optimization. We propose a simple alternative. Setting $\alpha = 0$ and $\beta = \hat{\beta}$ in (3) gives $\hat{p}_i = \{nr(z_i, \hat{\beta})\}^{-1}$. Estimators for $F(x)$ and $G(x)$ are

$$F_n(x) \doteq \sum_{i=1}^{n} \hat{p}_i I(z_i \leq x) \text{ and } G_n(x) \doteq \sum_{i=1}^{n} \hat{p}_i \exp(\hat{\beta}_0 + z_i\hat{\beta}_1) I(z_i \leq x).$$

By inspection, the estimators are monotone increasing step functions in x, with jumps at the observed values $(z_i, i = 1, \ldots, n)$. Because estimation is unconstrained, in small samples, F_n and G_n may exceed 1 in the tail. The adjusted estimators $F_n^*(x) = F_n(x)/F_n(\infty)$ and $G_n^*(x) = G_n(x)/G_n(\infty)$ are always distribution functions.

Recall $\hat{\beta} \xrightarrow{p} \beta_T$ and note that p_i and $\exp(\beta_0 + \beta_1 z_i)$ have bounded derivatives in β for bounded z_i and $\beta \in \mathcal{H}$. Thus, it is straightforward to establish $\sup_{x\in[\tau_l,\tau_u]} |F_n(x) - \sum_i p_i I(z_i \leq x)|$ and $\sup_{x\in[\tau_l,\tau_u]} |G_n(x) - \sum_i p_i \exp(\beta_0 + \beta_1 z_i) I(z_i \leq x)|$ vanish in probability, where $\text{pr}(z_i < \tau_l) > 0$ and $\text{pr}(z_i > \tau_u) > 0$. A uniform law of large numbers gives that $\sup_{x\in[\tau_l,\tau_u]} |F(x) - \sum_i p_i I(z_i \leq x)| \xrightarrow{p} 0$ and $\sup_{x\in[\tau_l,\tau_u]} |G(x) - \sum_i p_i \exp(\beta_0 + \beta_1 z_i) I(z_i \leq x)| \xrightarrow{p} 0$. As a result, F_n and G_n are uniformly consistent.

The next theorem is helpful in constructing confidence intervals for the distributions; see appendix for proof.

Theorem 4. Under the regularity conditions of Theorem 2,

$$\sqrt{n}\{F_n(x) - F(x)\} \xrightarrow{W} K_F(x) \text{ and } \sqrt{n}\{G_n(x) - G(x)\} \rightarrow K_G(x),$$

where $K_F(x)$ and $K_G(x)$ are mean zero Gaussian processes with continuous sample paths for $x \in [\tau_l, \tau_u]$ and covariance functions $\Sigma_F(x, y)$ and $\Sigma_G(x, y)$ given in the appendix.

Estimators for the covariance functions, $\hat{\Sigma}_F$ and $\hat{\Sigma}_G$, are computed with empirical estimates in place of theoretical quantities in Σ_F and Σ_G. The resulting plug-in formulas are tedious and are omitted. A 0.95 confidence interval for $F(x)$ is $F_n(x) \pm n^{-1/2}1.96\hat{\Sigma}_F(x,x)$ and similarly for $G(x)$.

3. NUMERICAL STUDIES

Simulations were run to investigate the small sample behavior of $\hat{\beta}$, $\hat{B}$, and $2l_2(\hat{\beta})$ in a genetic experiment. Two homozygous lines (P1 and P2) are mated, yielding heterozygous (F1) children. P1 individuals have genotype a/a at all loci, P2 individuals are A/A at all loci, and F1 individuals are a/A at all loci. F1 is bred to P1, yielding backcross progeny (BC) which are either a/a or a/A at a given locus. These breedings are designed to study a quantitative trait locus at 30 cM on a hypothetical chromosome. The BC generation is genotyped at markers at 20 cM and 40 cM.

The distribution of the trait is $f(x)$ for individuals a/a at 30 cM and $g(x)$ for individuals a/A. There are four possible genotypes at the flanking markers: aa/aa, aa/aA, aa/Aa, and aa/AA. The recombinant genotypes, aa/aA and aa/Aa, each occur with probability 0.082. Conditional on these genotypes, the probability of a/a at the trait locus is 0.5. These values are based on recombination fractions from the Haldane (1919) map function. In a like manner, the probabilities of aa/aa and aa/AA at the flanking markers are each 0.418, and the conditional probabilities of a/a at 30 cM are 0.99 and 0.01. This gives $\lambda = (0.99, 0.5, 0.01)$ and $\tilde{\rho} = (0.418, 0.164, 0.418)$.

Normal, poisson, and exponential mixtures were investigated. Five hundred samples were simulated for each mixture model with $n = 100$ or 250. In each sample, $\hat{\beta}$, $\hat{B}$, and $2l_2(\hat{\beta})$ were computed. The average values of $\hat{\beta}$ and $\hat{B}$ are in Table 2. The empirical rejection rate for a nominal 0.05 level test using $2l_2(\hat{\beta})$ and the empirical variance of $\hat{\beta}$ are also provided. The bias is small and the empirical and model-based variances agree. The performance improves as n increases. The test statistic rejects at the nominal level under H_0 and has good power when β_0 and $\beta_1 \neq 0$.

4. MAMMARY CARCINOMA DATA

Female rats from the Wistar-Kyoto (WKy) strain resistant to mammary carcinogenesis were crossed with male rats from the Wistar-Furth (WF) strain susceptible to cancer (Lan et al., 2000). Each strain was purebred, hence WF/WF or WKy/WKy at all loci. The progreny were mated to WF animals, producing 383 female rats which were either WF/WF or WKy/WF at each locus. These backcross rats were scored for number of mammary carcinomas and were genotyped at 58 markers on Chromosome 5. Using several interval mapping strategies, Lan et al. (2000) found that marker *D5Rat22* on Chromosone 5 was strongly associated with low tumor counts. That is, female rats with a copy of the WKy allele at *DFRat22* had fewer carcinomas than rats with no WKy alleles.

The data are reanalyzed with our semiparametric method. At a putative locus, let $f(x)$ be the distribution of tumor counts for a WF/WF animal and let $g(x)$ be the distribution for a WKy/WF animal. The mixture is $\lambda f(x) + (1-\lambda)g(x)$, where λ is the probability of WF/WF at the locus conditional on flanking marker genotypes. In Fig. 1, the partial likelihood statistic is shown as a function of location on Chromosome 5. The LOD score, $\log\{l_2(\hat{\beta}\}\{2\log(10)\}^{-1}$, the conventional measure of genetic linkage, is also given. For comparison, the profile from a normal mixture using MapMaker/QTL is displayed.

A practical issue is that the analysis requires testing H_0 at all loci on the chromosome. The simultaneous type I error probability is inflated from the pointwise level. Lander and Botstein (1989) presented critical values for the normal mixture which preserve a genome-wide error rate. The limiting distribution of the test statistic across the genome was approximated by an Ornstein-Uhlenbeck diffusion. The extreme value properties of the process were used to derive the thresholds. Interestingly, we can show that the asymptotic equivalent for $2l_2(\hat{\beta})$ is exactly identical to that in Lander and Botstein (1989). This means the same guidelines apply to the semiparametric model.

The curves are quite similar and their peaks are very near D5Rat22 and are well above the usual thresholds. The estimated distribution functions for Wky/WF and WF/WF genotypes were computed at the locus giving the maximum LOD score under the semiparametric and normal mixtures. These are displayed in Fig. 2 along with 0.95 pointwise confidence intervals using model (1). The plots exhibit that WF/WF rats have higher tumor counts. The estimated means for carcinomas in WKy/WF and WF/WF rats are $\int x d\hat{G}(x) = 2.69$ and $\int x d\hat{F}(x) = 5.45$, respectively. The estimated distributions from the normal mixture are rather different from the semiparametric estimates and may lie outside the confidence intervals. Other estimates (not shown) from a negative binomial model (Drinkwater & Klotz, 1981) fall entirely within the 0.95 limits.

To assess the goodness-of-fit of the exponential tilt assumption at the peak locus, the rats were divided into four groups according to flanking marker genotypes. Recombination was infrequent and $> 90\%$ of rats were either WFWF/WFWF or WKyWKy/WFWF. The empirical distribution functions were calculated for these groups. The distributions were also computed using the fitted semiparametric model. In Fig. 3, the model-based and nonparametric estimates match closely, indicating the model fits well.

5. REMARKS

The profile empirical likelihood for a semiparametric mixture model arising in quantitative genetics was shown to have an irregularity under the null hypothesis of no linkage. After factoring the likelihood, a partial likelihood was identified and was shown to give valid inferences. The estimator $\hat{\beta}_1$ had the same variance as $\tilde{\beta}_1$ and $\hat{\beta}_0$ had good efficiency relative to the full likelihood when data is observed directly from f and g. This is realistic in backcross studies with dense marker maps. Recombination with flanking markers is unlikely and most observations have mixture proportions close to 0 or 1.

The methodology can be adapted to more complicated breeding experiments. For example, in an intercross (F2) mating of heterozygous animals, there are three distributions in the mixture. In theory, the model can accomodate an arbitrary number of components. Another important extension is to incorporate higher powers of x in (1). This is easily accomplished with our approach.

Empirical likelihood may pose computational difficulties (Owen, 1988, 1990). The partial profile empirical likelihood for the exponential tilt model enables unconstrained estimation of the parameters of interest. It would be worthwhile to investigate whether empirical likelihood has useful factorizations in other scenarios.

REFERENCES

ANDERSON, J. A. (1979). Multivariate logistic compounds. *Biometrika* **66**, 17-26.

COX, D. R. (1972). Regression models and life tables. *J. R. Statist. Soc. Ser. B* **34**, 187-220.

COX, D. R. (1975). Partial likelihood. *Biometrika* **62**, 269-76.

DOERGE, R. W., ZENG, Z.-B., & WEIR B. S. (1997). Statistical issues in the search for genes affecting quantitative traits in experimental populations. *Statist. Sci.* **12**, 195-219.

DRINKWATER, N. R. & KLOTZ, J. H. (1981). Statistical methods for the analysis of tumor multiplicity data. *Cancer Res.* **41**, 113-9.

DURRETT, R. (1991). *Probability: Theory and Examples.* New York: Wadsworth.

HALDANE, J. B. S. (1919). The combination of linkage values and the calculation of distances between the loci of linked factors. *J. Genet.* **8**, 299-309.

HALL, P. (1981). On the non-parametric estimation of mixture proportions. *J. R. Statist. Soc. Ser. B* **43**, 147-56.

KNAPP, S. J., BRIDGES, W. C., & BIRKES, D. (1990). Mapping quantitative trait loci using molecular marker linkage maps. *Theor. Appl. Genet.* **79**, 583-92.

KRUGLYAK, L. & LANDER, E. S. (1995). A nonparametric approach for mapping quantitative trait loci. *Genet.* **139**, 1421-28.

LAN, H., KENDZIORSKI, C. M., HAAG, J. D., SHEPEL, L. A., NEWTON, M. A., & GOULD, M. N. (2000). Genetic loci controlling breast cancer susceptibility in the Wistar-Kyoto rat. *Genet.*, in press.

LANDER, E.S. & BOTSTEIN, D. (1989). Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genet.* **121**, 185-99.

LANDER, E.S. & KRUGLYAK, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11**, 241-7.

MURRAY, G. D. & TITTERINGTON, D. M. (1978). Estimation problems with data from a mixture. *J. Roy. Statist. Soc. Ser. C* **27** 325-34.

OWEN, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* **75**, 237-49.

OWEN, A. B. (1990). Empirical likelihood confidence regions. *Ann. Statist.* **18**, 90-120.

QIN, J. (1993). Empirical likelihood in biased sample problems. *Ann. Statist.* **21**, 1182-96.

QIN, J. (1998). Inferences for case-control and semiparametric two-sample density ratio models. *Biometrika* **85**, 619-30.

QIN, J. (1999). Empirical likelihood ratio based confidence intervals for mixture proportions. *Ann. Statist.* **27**, 1368-84.

QIN, J. & LAWLESS, J. F. (1994). Empirical likelihood and general estimating equations. *Ann. Statist.* **22**, 300-25.

TITTERINGTON, D. M., SMITH, A. F. M., & MAKOV, U. E. (1985). *Analysis of Finite Mixture Distributions.* New York: Wiley.

APPENDIX

Proofs of Theorems 1-4

Lemma 1 is needed for the proof of Theorem 1. The proof is trivial and is omitted.

Lemma 1. For any given $a = (a_1, a_2, ..., a_n)$, the set $\{p = (p_1, p_2, ..., p_n) \mid \sum_{i=1}^{n} p_i = 1, p_i \geq$

0 and $\sum_{i=1}^{n} p_i a_i = 1\}$ is non-empty $\iff$ $\min_{i=1}^{n}(a_i - 1) \leq 0 \leq \max_{i=1}^{n}(a_i - 1)$.

Proof of Theorem 1.

(i): For any given $\beta \in J_n(z)$, $\min_i(\beta_0+z_i\beta_1) \leq 0 \leq \max_i(\beta_0+z_i\beta_1) \Rightarrow \min_i\{\exp(\beta_0+z_i\beta_1)-1\} \leq 0 \leq \max_i\{\exp(\beta_0+z_i\beta_1)-1\}$. By Lemma 1, there exists $p = (p_1, p_2, ..., p_n) \in C_\beta$. On the other hand, if C_β is not empty, then there exists $p = (p_1, p_2, ..., p_n) \in C_\beta$ such that $\sum_{i=1}^{n} p_i = 1, p_i \geq 0$ and $\sum_{i=1}^{n} p_i\{\exp(\beta_0 + z_i\beta_1) - 1\} = 0$. Lemma 1 gives that $\min_i\{\exp(\beta_0 + z_i\beta_1) - 1\} \leq 0 \leq \max_i\{\exp(\beta_0 + z_i\beta_1) - 1\}$, or $\min_i(\beta_0 + z_i\beta_1) \leq 0 \leq \max_i(\beta_0 + z_i\beta_1)$.

(ii): We first show that $\beta_T \in J_n(z)$. If $\beta_T \notin J_n(z)$, then either all $\beta_{0T} + z_i\beta_{1T} > 0$ or all $\beta_{0T} + z_i\beta_{1T} < 0$. Without loss of generality, assume $\beta_{0T} + z_i\beta_{1T} > 0$, or $\exp(\beta_{0T} + z_i\beta_{1T}) > 1$ for $i = 1, 2, \ldots, n$, which indicates $\exp(\beta_{0T} + x\beta_{1T}) \geq 1$ for all x. Because $F(x)$ is nondegenerate, $1 = \int \exp(\beta_{0T} + x\beta_{1T})dF(x) > \int dF(x) = 1$. But this is a contradiction. Now, again without loss of generality, assume $\exp(\beta_{0T}+z_1\beta_{1T}) < 1$ and $\exp(\beta_{0T}+z_2\beta_{1T}) > 1$. Because $\exp(\beta_0 + z_1\beta_1)$ and $\exp(\beta_0 + z_2\beta_1)$ are continuous with respect to $\beta = (\beta_0, \beta_1)$, there exists a neighbourhood of β_T such that $\exp(\beta_0 + z_1\beta_1) < 1$ and $\exp(\beta_0 + z_2\beta_1) > 1$.

(iii): If $\beta_T = 0$, then for any $\beta_0 \neq 0$, $C_{(\beta_0,0)}$ is empty by (i). Thus, there does not exist an $N(0)$ in which C_β is empty for every β.

Proof of Theorem 2.

(i): Suppose $\beta_0 = \beta_{0T} + t_1 n^{-1/3}$ and $\beta_1 = \beta_{1T} + t_2 n^{-1/3}$ where $\sqrt{t_1^2 + t_2^2} = 1$. By Taylor expansion in β around β_T:

$$l_2(\beta) = l_2(\beta_T) + \sum_{k=1}^{K}\sum_{j=1}^{n_k}\left\{\frac{1-\lambda_k}{\omega_k(x_{kj},\beta_T)} - \frac{\xi}{r(x_{kj},\beta_T)}\right\}(t_1 + x_{kj}t_2)\exp(\beta_{0T} + \beta_{1T}x_{kj})n^{-1/3}$$

$$+\frac{1}{2}\sum_{k=1}^{K}\sum_{j=1}^{n_k}\left\{\frac{\lambda_k(1-\lambda_k)}{\omega_k^2(x_{kj},\beta_T)} - \frac{\xi(1-\xi)}{r^2(x_{kj},\beta_T)}\right\}(t_1 + t_2x_{kj})^2\exp(\beta_{0T} + \beta_{1T}x_{kj})n^{-2/3} + o(n^{1/3}).$$

Define $\rho_k = \lim_{n\to\infty} \frac{n_k}{n}, k = 1, 2, ..., K$, $\phi = \lim_{n\to\infty} \xi = \sum\limits_{k=1}^{K} \rho_k(1-\lambda_k)$, and

$$R(x,\beta) = \lim_{n\to\infty} r(x,\beta) = 1 + \phi\{\exp(\beta_0 + x\beta_1) - 1\} = \sum_{k=1}^{K} \rho_k \omega_k(x,\beta)$$

Note $\frac{1}{n}\sum_{k=1}^{K}\sum_{j=1}^{n_k}\left\{\frac{1-\lambda_k}{\omega_k(x_{kj},\beta_T)} - \frac{\xi}{r(x_{kj},\beta_T)}\right\}(t_1 + x_{kj}t_2)\exp(\beta_{0T} + \beta_{1T}x_{kj})$ approaches 0 as $n \to \infty$. By Theorem 9.6 in Durrett (Chap. 7, 1991) and the strong law of large numbers:

$$l_2(\beta) - l_2(\beta_T) = O(n^{1/6}(\log\log n)^{1/2}) +$$

$$\frac{1}{2}\left[\int\left\{\sum_{k=1}^{K}\frac{\rho_k\lambda_k(1-\lambda_k)}{\omega_k(x,\beta_T)} - \frac{\phi(1-\phi)}{R(x,\beta_T)}\right\}\left\{\exp(\beta_{0T} + x\beta_{1T})(t_1 + xt_2)^2\right\}dF(x)\right] n^{1/3} + o(n^{1/3}).$$

Next, we show

$$\triangle(x,\beta_T) \doteq \sum_{k=1}^{K}\frac{\rho_k\lambda_k(1-\lambda_k)}{\omega_k(x,\beta_T)} - \frac{\phi(1-\phi)}{R(x,\beta_T)} < 0 \text{ for all } x.$$

Define $\theta = \exp(\beta_{0T} + x\beta_{1T}) - 1$. After tedious calculation:

$$\triangle(x,\beta_T)\left\{\prod_{k=1}^{K}\omega_k(x,\beta_T)R(x,\beta_T)\right\} = \prod_{l\neq k}\rho_k\lambda_k(1-\lambda_k)\omega_l(x,\beta_T)R(x,\beta_T) - \phi(1-\phi)\prod_{l=1}^{K}\omega_l(x,\beta_T)$$

$$= (\theta+1)\sum_{i\neq j}\{-\rho_i\rho_j(\lambda_i-\lambda_j)^2\prod_{l\neq i, l\neq j}(1+\lambda_l\theta)\} < 0$$

with unequal $\lambda_i, i = 1, \ldots, K$. So, for n large enough, $l_2(\beta) < l_2(\beta_T)$. It follows that $l_2(\beta)$ attains a local maximum at a point $\hat\beta$ in the interior of the interval $\mid \beta - \beta_T \mid \leq n^{-1/3}$. Solving

$$0 = \frac{1}{n}\frac{\partial l_2(\hat\beta)}{\partial\beta} = \frac{1}{n}\frac{\partial l_2(\beta_T)}{\partial\beta} + \frac{1}{n}\frac{\partial^2 l_2(\beta_T)}{\partial\beta\partial\beta^T}(\hat\beta - \beta_T) + o(n^{-1/2}),$$

$$\hat{\beta} - \beta_T = -S_n^{-1} Q_n + o(n^{-1/2}), \text{ where}$$

$$\begin{aligned} S_n &= \frac{1}{n}\frac{\partial^2 l_2(\beta_T)}{\partial\beta\partial\beta^T} = \frac{1}{n}\sum_{k=1}^{K}\sum_{j=1}^{n_k}\frac{\lambda_k(1-\lambda_k)(1,x_{kj})^T(1,x_{kj})\exp(\beta_{0T}+x_{kj}\beta_{1T})}{\omega_k(x_{kj},\beta_T)^2} \\ &\quad -\frac{1}{n}\sum_{i=1}^{n}\frac{\xi(1-\xi)(1,z_i)^T(1,z_i)\exp(\beta_{0T}+z_i\beta_{1T})}{r(z_i,\beta_T)^2}. \end{aligned}$$

The matrix S_n tends to

$$\begin{aligned} S &= \sum_{k=1}^{K}\rho_k\lambda_k(1-\lambda_k)\int\frac{\partial^2\exp(\beta_{0T}+x\beta_{1T})}{\partial\beta\partial\beta^T}\frac{1}{\omega_k(x,\beta_T)}dF(x) \\ &\quad -\phi(1-\phi)\int\frac{\partial^2\exp(\beta_{0T}+x\beta_{1T})}{\partial\beta\partial\beta^T}\frac{1}{R(x,\beta_T)}dF(x) \text{ as } n\to\infty. \end{aligned}$$

The matrix

$$Q_n = \frac{1}{n}\frac{\partial l_2(\beta_T)}{\partial\beta} = \frac{1}{n}\sum_{k=1}^{K}\sum_{j=1}^{n_k}\frac{(1-\lambda_k)}{\omega_k(x_{kj},\beta_T)}\frac{\partial\exp(\beta_{0T}+x_{kj}\beta_{1T})}{\partial\beta}$$

$$-\frac{1}{n}\sum_{i=1}^{n}\frac{\xi}{r(z_i,\beta_T)}\frac{\partial\exp(\beta_{0T}+z_i\beta_{1T})}{\partial\beta}$$

and $E(Q_n)\to 0$. By Lindeberg-Feller central limit theorem: $\sqrt{n}Q_n \to N(0,V)$, where

$$\begin{aligned} V &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{K}\sum_{j=1}^{n_k}\int\left[\left\{\frac{1-\lambda_k}{\omega_k(x,\beta_T)}-\frac{\xi}{r(x,\beta_T)}\right\}\frac{\partial\exp(\beta_{0T}+x\beta_{1T})}{\partial\beta}\right]^2\omega_k(x,\beta_T)dF(x) \\ &= -S-\delta\int\frac{\partial\exp(\beta_{0T}+x\beta_{1T})}{\partial\beta}\frac{1}{R(x,\beta_T)}dF(x)\left\{\int\frac{\partial\exp(\beta_{0T}+x\beta_{1T})}{\partial\beta}\frac{1}{R(x,\beta_T)}dF(x)\right\}^T \end{aligned}$$

and $\delta = \sum_{k=1}^{K}\rho_k(1-\lambda_k)^2 - \phi^2 > 0$. Thus, $\sqrt{n}(\hat{\beta}-\beta_T)\to N(0,B)$, where $B = S^{-1}VS^{-1}$.

(ii): A Taylor expansion of $\frac{1}{n}\frac{\partial l_2(\hat{\beta})}{\partial\beta}$ in $\hat{\beta}$ around $(0,0)$ gives

$$\begin{aligned} 0 &= \frac{1}{n}\frac{\partial l_2(\hat{\beta})}{\partial\beta} = \frac{1}{n}\frac{\partial l_2(0,0)}{\partial\beta} + \frac{1}{n}\frac{\partial^2 l_2(0,0)}{\partial\beta\partial\beta^T}\hat{\beta} + o(n^{-1/2}) \\ &= U - \delta X\hat{\beta} + o(n^{-1/2}), \end{aligned}$$

where $U = \{0, n^{-1}\partial l_2(0,0)(\partial\beta_1)^{-1}\}^T$ and $X = \int(1,x)^T(1,x)dF(x)$. It follows that $\hat{\beta} = \delta^{-1}X^{-1}U + o(n^{-1/2})$. Since $0 = 2l_2(0,0) = 2l_2(\hat{\beta}) + 2\frac{\partial l_2(\hat{\beta})}{\partial\beta}\hat{\beta} + \hat{\beta}^T\frac{\partial^2 l_2(\hat{\beta})}{\partial\beta\partial\beta^T}\hat{\beta} + o(1)$,

$$2l_2(\hat{\beta}) = -\hat{\beta}^T\frac{\partial^2 l_2(\hat{\beta})}{\partial\beta\partial\beta^T}\hat{\beta} + o(1) = n\delta\hat{\beta}^T X\hat{\beta} + o(1) = n\delta^{-1}U^TX^{-1}U + o(1)$$

$$= \frac{n}{\delta\sigma_F^2}\left\{\frac{1}{n}\frac{\partial l_2(0,0)}{\partial\beta_1}\right\}^2 + o(1) \xrightarrow{\mathcal{L}} \chi_1^2,$$

where $\sigma_F^2 = \int x^2 dF(x) - \left\{\int x dF(x)\right\}^2$. The convergence in distribution occurs because

$$\sqrt{n}\left\{\frac{1}{n}\frac{\partial l_2(0,0)}{\partial\beta_1}\right\} = n^{-1/2}\left\{\sum_{k=1}^{K}(1-\lambda_k)\sum_{j=1}^{n_k}x_{kj} - \xi\sum_{i=1}^{n}z_i\right\} \xrightarrow{\mathcal{L}} N(0, \delta\sigma_F^2).$$

Proof of Theorem 3.

(i): When $\beta_T \neq 0$, methods similar to those used in the proof of Theorem 2(i) give the consistency and asymptotic normality of $\tilde{\beta}$. The details are omitted.

(ii): When operating on matrices, > 0 and ≥ 0 denote positive and positive semi-definite, and < 0 and ≤ 0 denote negative and negative semi-definite. Define

$$S_{12} = S_{21}^T = \int\frac{\partial\exp(\beta_{0T} + x\beta_{1T})}{\partial\beta}\frac{1}{R(x,\beta_T)}dF(x), \quad \delta = \sum_{k=1}^{K}\rho_k(1-\lambda_k)^2 - \phi^2 > 0$$

$$\text{and} \quad s_{22} = \int\frac{\{1-\exp(\beta_{0T} + x\beta_{1T})\}^2}{R(x,\beta_T)}dF(x).$$

Note that $\tilde{V} = \begin{pmatrix} -S - \delta S_{12}S_{21} & -\delta S_{12}s_{22} \\ -\delta S_{21}s_{22} & s_{22} - \delta s_{22}^2 \end{pmatrix} > 0$ and $-S - \delta S_{12}S_{21} > 0 \Rightarrow S_{21}S^{-1}(-S - \delta S_{12}S_{21})S^{-1}S_{12} > 0$. This implies $-S_{21}S^{-1}S_{12} - \delta S_{12}S^{-1}S_{12}S_{21}S^{-1}S_{12} > 0$. Hence, $-S_{21}S^{-1}S_{12}(1 + \delta S_{21}S^{-1}S_{12}) > 0$, since $-S_{21}S^{-1}S_{12} > 0 \Rightarrow 1 + \delta S_{21}S^{-1}S_{12} > 0$.

Now, because $\tilde{V} > 0$, the last element of $\tilde{V}^{-1}$, say v_{33}^{-1}, is > 0 too. Calculating $\tilde{V}^{-1}$:

$$\begin{aligned} v_{33} &= (s_{22} - \delta s_{22}) - (-\delta S_{21} s_{22})(-S - \delta S_{12} S_{21})^{-1}(-\delta S_{12} s_{22}) \\ &= s_{22} - \delta s_{22}^2 + \delta^2 s_{22}^2 S_{21} S^{-1} S_{12} - \frac{\delta^3 s_{22}^2 S_{21} S^{-1} S_{12} S_{21} S^{-1} S_{12}}{1 + \delta S_{21} S^{-1} S_{12}} \\ &= \frac{s_{22} + \delta(S_{21} S^{-1} S_{12} - s_{22}) s_{22}}{1 + \delta S_{21} S^{-1} S_{12}} > 0 \end{aligned}$$

Using the first part of the proof and the fact that $s_{22} > 0$, $\delta - \frac{1}{s_{22} - S_{21} S^{-1} S_{12}} < 0$. Thus, $\tilde{B}_{11} - B = (\delta - \frac{1}{s_{22} - S_{21} S^{-1} S_{12}}) S^{-1} S_{12} S_{21} S^{-1} \leq 0$.

Proof of Theorem 4.

Note $F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{r(z_i, \hat{\beta})} I(z_i \leq x)$. A Taylor expansion of $r(z_i, \hat{\beta})$ at β_T gives

$$\begin{aligned} F_n(x) - F(x) &= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{r(z_i, \beta_T)} I(z_i \leq x) - F(x) \\ &\quad - \frac{1}{n} \sum_{i=1}^{n} \frac{1}{r^2(z_i, \beta_T)} \frac{\partial r(z_i, \beta_T)}{\partial \beta} I(z_i \leq x)(\hat{\beta} - \beta_T) + R_{1n}(x) \\ &= \frac{1}{n} \sum_{i=1}^{n} \frac{1}{r(z_i, \beta_T)} I(z_i \leq x) - F(x) + d_{1,F}(x) S^{-1} Q_n + R_{2n}(x) \end{aligned}$$

where $R_{in}(x), i = 1, 2$, satisfy $\sup_{\tau_l < x < \tau_u} |R_{in}(x)| = o(n^{-1/2})$ and

$$\begin{aligned} d_{1,F}(x) &= \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \frac{1}{r^2(z_i, \beta_T)} \frac{\partial r(z_i, \beta_T)}{\partial \beta} I(z_i \leq x) \\ &= \int_{-\infty}^{\infty} \frac{\phi}{R(u, \beta_T)} \frac{\partial \exp(\beta_{0T} + u \beta_{1T})}{\partial \beta} I(u \leq x) dF(u), \ a.s. \end{aligned}$$

Let $d_{2,F}(x) = d_{1,F}(x) S^{-1}$,

$$\epsilon_{F,k}(u, x) = \frac{I(u \leq x)}{r(u, \beta_T)} - \int_{-\infty}^{x} \frac{\omega_k(u, \beta_T) dF(u)}{R(u, \beta_T)}, \text{ and}$$

$$q_k(u) = -\xi \frac{\partial \exp(\beta_{0T} + u\beta_{1T})/\partial \beta}{r(u, \beta_T)} + (1-\lambda_k)\frac{\partial \exp(\beta_{0T} + u\beta_{1T})/\partial \beta}{\omega_k(u, \beta_T)}, k = 1, 2, ..., K.$$

Then

$$\sqrt{n}\{F_n(x) - F(x)\} = \frac{1}{\sqrt{n}} \sum_{k=1}^{K} \sum_{j=1}^{n_k} \{\epsilon_k(x_{kj}, x) + d_2(x) q_k(x_{kj})\} + o_p(1)$$

Using arguments from Qin (1999), $\sqrt{n}\{F_n(x) - F(x)\} \to K_F(x)$ in distribution, where $K_F(x)$ is a mean zero Gaussian process with continuous sample paths and covariance structure $\Sigma_F(x_1, x_2) = \sum_{k=1}^{K} \rho_k \text{cov}\{\epsilon_{F,k}(Y_k, x_1) + d_{2,F}(x_1) q_k(Y_k), \epsilon_{F,k}(Y_k, x_2) + d_{2,F}(x_2) q_k(Y_k)\}$, where $Y_k \sim \omega_k(y, \beta_T) f(y)$. Similarly, $\sqrt{n}\{G_n(x) - G(x)\} \to K_G(x)$ in distribution with $\Sigma_G(x_1, x_2) = \sum_{k=1}^{K} \rho_k \text{cov}\{\epsilon_{G,k}(Y_k, x_1) + d_{2,G}(x_1) q_k(Y_k), \epsilon_{G,k}(Y_k, x_2) + d_{2,G}(x_2) q_k(Y_k)\}$, where

$$d_{1,G}(x) = \int_{-\infty}^{\infty} \frac{\phi - 1}{R(u, \beta_T)} \frac{\partial \exp(\beta_{0T} + u\beta_{1T})}{\partial \beta} I(u \leq x) dF(u)$$

$d_{2,G}(x) = d_{1,G}(x) S^{-1}$, and

$$\epsilon_{G,k}(u, x) = \frac{I(u \leq x) \exp(\beta_{0T} + u\beta_{1T})}{r(u, \beta_T)} - \int_{-\infty}^{x} \frac{\exp(\beta_{0T} + u\beta_{1T}) \omega_k(u, \beta_T) dF(u)}{R(u, \beta_T)}.$$

Table 1. *Relative efficiency of $\tilde{\beta}_0$ to $\hat{\beta}_0$. n, e, and p denote normal, exponential, and poisson distributions.*

g(x)	f(x)	β_0	β_1	(a)	(b)	(c)	(d)
n(0,1)	n(2,1)	2	-2	0.998	0.973	0.918	0.924
n(2.01,1)	n(2,1)	-0.02	0.01	1.000	1.000	1.000	1.000
n(4,1)	n(2,1)	-6	2	1.000	0.997	0.990	0.972
p(1)	p(3)	2	-1.10	1.000	0.981	0.960	0.969
p(3.01)	p(3)	-0.01	0.01	1.000	1.000	1.000	1.000
p(6)	p(3)	-3	0.69	0.999	0.991	0.984	0.970
e(1)	e(3)	-1.10	2	0.993	0.980	0.966	0.958
e(3.01)	e(3)	0.01	-0.01	1.000	1.000	1.000	1.000
e(6)	e(3)	0.69	-3	1.000	0.986	0.987	0.989
n(-3,1)	n(2,1)	-2.5	-5	0.993	0.989	0.941	0.869
n(-2,1)	n(2,1)	0	-4	0.983	0.958	0.444	0.241
n(-1,1)	n(2,1)	1.5	-3	0.990	0.966	0.777	0.698

(a): $\tilde{\rho} = (0.5, 0.5), \lambda = (1, 0)$

(b): $\tilde{\rho} = (0.4, 0.2, 0.4), \lambda = (1, 0.5, 0)$

(c): $\tilde{\rho} = (0.33, 0.34, 0.33), \lambda = (0.7, 0.5, 0.3)$

(d): $\tilde{\rho} = (0.5, 0.5), \lambda = (0.7, 0.5)$

Table 2. *Simulation results. n, e, and p denote normal, exponential, and poisson distributions.*

						$\hat{\beta}_0$			$\hat{\beta}_1$		$2l_2(\hat{\beta})$
g(x)	f(x)	β_0	β_1	n	ave	var1	var2	ave	var1	var2	rr
n(2,1)	n(2,1)	0	0	100	0.02	0.237	0.219	-0.01	0.060	0.055	0.064
				250	0.00	0.089	0.083	0.00	0.022	0.021	0.046
n(0,1)	n(2,1)	2	-2	100	2.17	0.469	0.410	-2.17	0.366	0.331	1.000
				250	2.07	0.140	0.128	-2.06	0.113	0.100	1.000
p(3)	p(3)	0	0	100	0.05	0.178	0.159	-0.02	0.020	0.018	0.056
				250	0.01	0.063	0.062	0.00	0.007	0.007	0.050
p(1)	p(3)	2	-1.10	100	2.10	0.264	0.258	-1.16	0.089	0.085	1.000
				250	2.07	0.101	0.096	-1.14	0.033	0.031	1.000
e(3)	e(3)	0	0	100	0.00	0.063	0.057	0.01	0.616	0.540	0.056
				250	0.00	0.020	0.021	0.02	0.191	0.196	0.050
e(1)	e(3)	-1.10	2	100	-1.16	0.121	0.108	2.15	0.565	0.485	0.996
				250	-1.13	0.037	0.041	2.09	0.166	0.177	1.000

ave: average of $\hat{\beta}$

var1: empirical variance of $\hat{\beta}$

var2: average of $\hat{B}$

rr: empirical rejection rate

Fig. 1. Likelihood ratio statistic and LOD score as a function of location on chromosome 5. Solid line is the semiparametric mixture and dashed is the normal mixture.

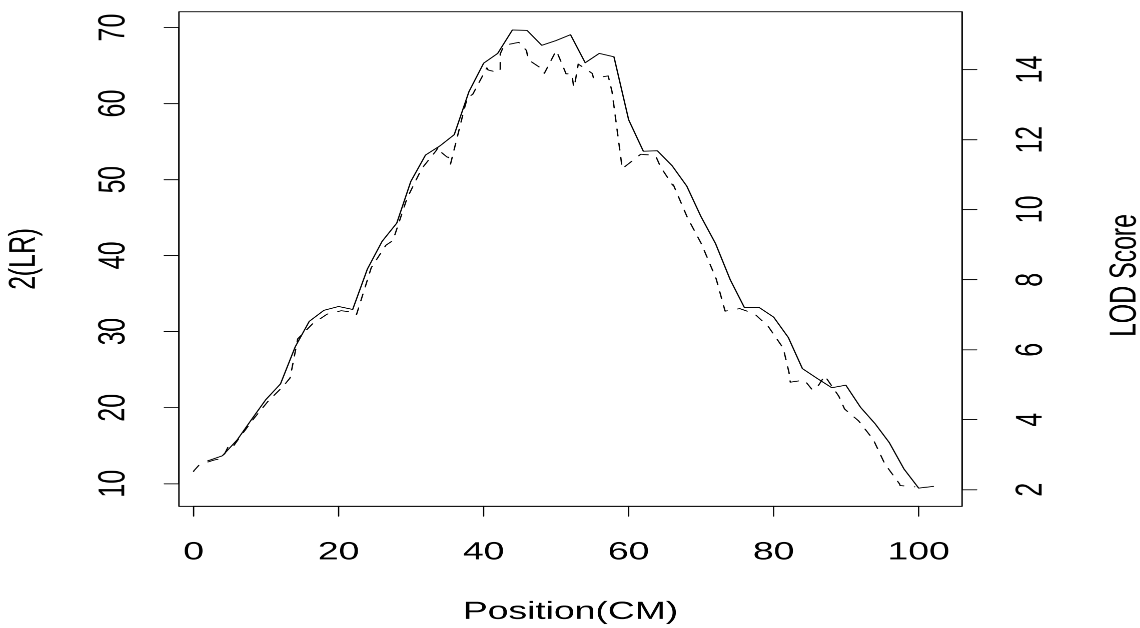

Fig. 2. Points estimates (+) and 0.95 pointwise confidence limits (0) for cumulative distributions at location of maximum partial likelihood ratio statistic. Dashed lines are point estimates from the normal mixture model. (a) WF/WF; (b) WKy/WF.

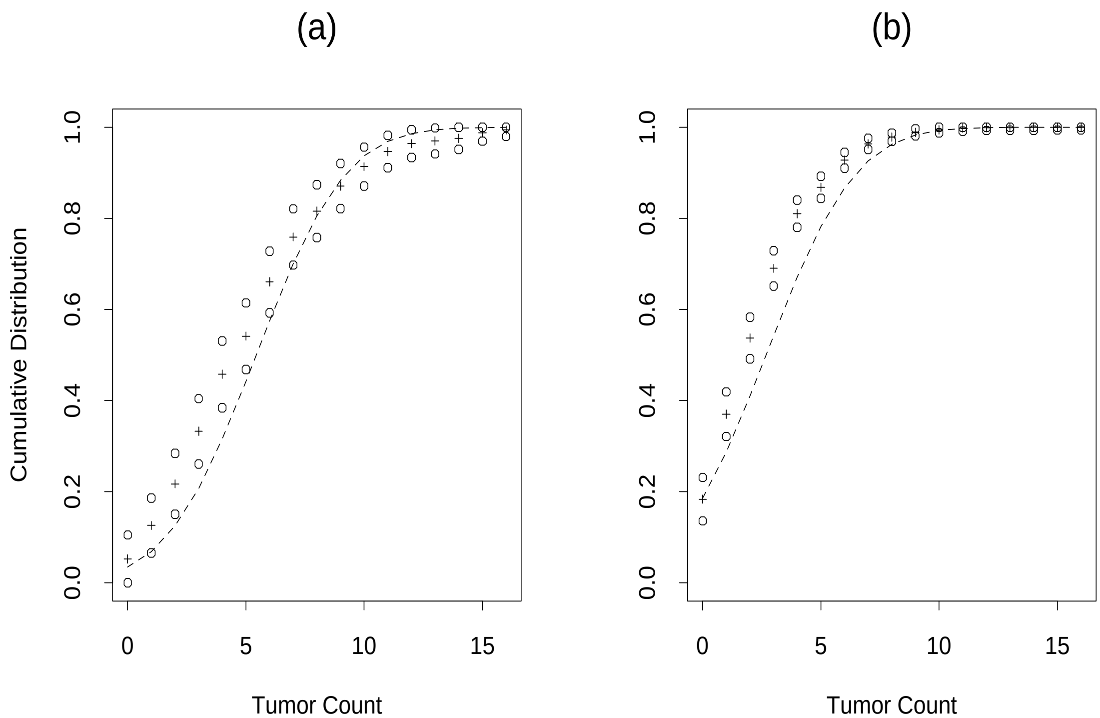

Fig. 3. Comparison of model-based (solid line) and nonparametric (dashed line) estimates of the cumulative distributions for flanking marker groups.

(a) WFWF/WFWF; (b) WKyWKy/WFWF.