

MR5014809 62G05 62F03 62G20 62H30

Balabdaoui, Fadoua (CH-ETHZ-S); **Besdziek, Harald** (CH-ETHZ-S)

Simple estimators of the mixing proportion in a semi-parametric mixture with known component. (English. English summary)

Sankhya A **88** (2026), no. 1, 80–135.

This article studies the problem of estimating a mixture of background and signal components. The mixture model examined in this article is

$$(1) \quad F_0 = (1 - \alpha_0) F_b + \alpha_0 F_s,$$

where the background distribution F_b is completely known, and the mixing proportion $\alpha_0 \in (0, 1)$ and the signal distribution F_s are unknown and thus must be estimated. Throughout, it is assumed that the support of the signal distribution $\mathcal{S}_s$ is a strict subset of that of the background distribution $\mathcal{S}_b$, i.e., $\mathcal{S}_s \subset \mathcal{S}_b$. This assumption ensures identification of the mixture model (1).

Estimation of the mixture model (1) can be done under the following additional assumption: There is some $a_0 \in \mathbb{R}$ so that $(-\infty, a_0) \subset \mathcal{S}_s^c$, and $(-\infty, a_0)$ is the largest set with this property. Then, given n i.i.d. observations $\{X_i\}_{i=1}^n$, the mixing proportion α_0 can be estimated by

$$\hat{\alpha}_n = 1 - \inf_{x \geq \kappa_0} \frac{\mathbb{F}_n(x)}{F_b(x)},$$

where $\mathbb{F}_n$ is the empirical cdf, and κ_0 is some prespecified constant that satisfies $\kappa_0 \leq a_0$ and $F_b(\kappa_0) > 0$. The mixing proportion estimator $\hat{\alpha}_n$ is shown to be consistent for α_0 with a parametric rate of convergence, and the limiting distribution is also demonstrated.

An estimator of the threshold a_0 and a testing procedure for the null of $\kappa_0 \leq a_0$ are also proposed. The entire estimation procedure can be extended to the case where there is some $a_0 \in \mathbb{R}$ so that $(a_0, \infty) \subset \mathcal{S}_s^c$, and (a_0, ∞) is the largest set with this property. This is done by mirroring the data (i.e., multiplying the data by -1). Finally, nonparametric estimation of the signal density when it admits monotonicity, monotonicity and convexity, or log-concavity is extensively investigated.

Masayuki Hirukawa