

## 2値変量が混在する場合のマハラノビス距離 による異常検出

慶應義塾大学理工学部 飯田 孝久

目白大学人文学部 福島 崇博

慶應義塾大学理工学部 篠崎 信雄

**要旨** 2値変量と連続変量が混在する場合の異常検出を、2値変量をダミー変数化しマハラノビス平方距離を用いて行うことがある。2値変量の値により連続変量の平均が変化するというロケーションモデルを仮定するとき、2値変量の一つの場合、マハラノビス平方距離は位置をずらした2つの $\chi^2$ 分布の混合分布にしたがうことを示した。2値変量が複数ある場合は、連続変量の平均が2値変量により加法的に変化するならば、マハラノビス平方距離は位置をずらした複数の $\chi^2$ 分布の混合分布にしたがうことを示した。

この結果を基に、異常検出を検定問題として定式化し、平均ならびに分散共分散行列が既知のときに、正確な有意水準を持つ修正マハラノビス距離法を構成した。この方法を用いたときの、正常状態における2値変量の値を与えたときの条件付棄却確率および棄却限界値の性質を明らかにした。

この方法を、2値変量の値を与えたという条件の下で平均ベクトルの検定を行う条件付異常検出法と、2値変量が1つの場合について比較した。2値変量の各値における非心度を固定したとき、両方法で検出力が等しくなる境界確率値が定まり、2値変量の分布により優劣が入れ替わるが、パラメータの広い範囲で修正マハラノビス距離法の方が高いまたは同程度の検出力を維持していることが数値計算で確認された。特に、正常状態で確率が小さい場合の確率が異常状態では増大するとき、修正マハラノビス距離法の方が有効であることが示された。