

ROC 分析関連の推定コマンド

ROC というのは receiver operating characteristics (受信者動作特性) の略です。ROC 分析では診断テストにおける感度 (sensitivity) と特異度 (specificity)、偽陽性 (false positives) と検知されなかった真陽性 (true positives) の数を問題にします。“ROC” という用語はレーダが開発された当初、操作盤上にあったノブの名前に由来するものです。ノブを一方に回すと受信機はより高感度となるわけですが、その場合、飛行中の航空機を検出できる可能性が高まる代わりに、実際には航空機が存在しないにも関わらず応答が返ってくる可能性 (false positives) も高まることとなります。一方、反対向きにノブを回すと多くの false positives を防ぐことができるようになる代わりに、実在する航空機からの微弱な応答が検出できなくなります (undetected positives)。今日の診断テストにおいてこのようなノブを見かけることはまずないでしょうが、positive とみなすか negative とみなすかの境界値の選択はこのノブの調整操作に該当するわけです。

ROC 分析はプロビットやロジスティック回帰の対象となるような二値アウトカムデータに対して適用されます。あるモデルをフィットさせると positive なアウトカムに対する予測確率を計算できるようになります。その上である境界値を設定してやれば、予測確率に基づき positive か negative かの判定を行うことができるわけです。この閾値の選択によって感度と特異度のトレードオフがどう変化するかをモデル化するのが ROC 分析であると言えます。

Stata における ROC 分析関連のコマンドとしては `roctab`, `roccomp`, `rocfit`, `rocgold`, `rocreg`, `rocregplot` の 6 種類が用意されています。

1. `roctab` ([R] `roctab` (*mwp-050*) 参照) は ROC 曲線のノンパラメトリック推定を行い、曲線下面積に関する Bamber/Hanley 信頼区間値を算出します。
2. `roccomp` ([R] `roccomp` (*mwp-050*) 参照) は ROC 曲線下面積の等値性に関する検定を行います。
3. `rocfit` ([R] `rocfit` 参照) は潜在変数の 2 変量正規性を仮定する形で、最尤法に基づくパラメトリックな ROC モデルのフィットを行います。
4. `rocgold` ([R] `roccomp` (*mwp-050*) 参照) は至適基準 (gold standard) と目される ROC 曲線と対比する形で、ROC 面積の等値性に関する検定を行います。その際、多重比較検定に伴う補正を選択することもできます。
5. `rocreg` ([R] `rocreg` 参照) は ROC 回帰の機能を提供します。それは年齢や性別といった予測因子 (prognostic factors) に応じた感度/特異度の調整を可能とします。
6. `rocregplot` ([R] `rocregplot` 参照) は `rocreg` によってモデル化された ROC 曲線のプロットを行います。