

Bohumil Kába

katedra statistiky, PEF, Vysoká škola zemědělská
165 21 Praha 6 - Suchdol

Anotace:

Příspěvek pojednává o některých technikách regresní diagnostiky, umožňujících identifikaci vybočujících pozorování v modelu mnohonásobné lineární regrese. Popsáno bylo užití projekční matice a studentizovaných (jack-knife) reziduí a byly zmíněny nejčastěji užívané skalární míry pro detekci vlivných bodů.

Summary:

The outlying observations (i.e. observations which are well separated from the remainder of data) may often have undesirable effects on the fitted least squares regression function. The paper deals with selected methods for identification of outlying observations and appropriate applications of each method are discussed.

Klíčová slova:

regresní diagnostika, odlehlé pozorování, vzdálené pozorování, projekční matice, plně studentizované reziduum, vlivné pozorování, Cookova vzdálenost

Key words:

regression diagnostics, outlier, leverage point, hat matrix, studentized deleted residual, influential observation, Cook's distance measure

Lineární regresní modely patří mezi nejdůležitější a nejfrekventovanější metodologické postupy ekonomické analýzy. Kvalita jejich výsledků úzce souvisí s kvalitou a vlastnostmi analyzovaných dat. Často se stává, že ve zpracovávaném datovém souboru se objeví několik hodnot, jež nejsou konsistentní s ostatními pozorováními. Tyto hodnoty, výrazně se lišící od ostatních, sice nemusí zásadně ovlivnit regresní analýzu dat, avšak zpravidla jsou hlavním zdrojem problémů, spojených s užitím regresních modelů. V případě těchto modelů s jednou respektive dvěma vysvětlujícími proměnnými je analýza vybočujících pozorování relativně snadná a může být uskutečněna např. některými jednoduchými grafickými prostředky - korelační graf respektive graf reziduálních hodnot (v anglické terminologii scatterplot respektive residual plot - viz např. statistický