

4 Mnohonásobný lineární regresní model

Příklad 1. V souboru `mlrm-fat.txt` máme k dispozici antropometrická data mladých zdravých dospělých žen (převážně studentek vysokých škol z Brna). Chceme modelovat závislost tělesné hmotnosti (proměnná `body.W`) na tělesné výšce (proměnná `body.H`), tloušťce kožní řasy ve výši 10. žebra (proměnná `rib.F`), tloušťce kožní řasy na břicho (proměnná `abdo.F`), tloušťce kožní řasy na boku (proměnná `hip.F`) a tloušťce kožní řasy nad čtyřhlavým svalem stehenním (proměnná `quad.H`). Hmotnost byla měřena v kilogramech, ostatní veličiny v milimetrech.

- Zobrazte si závislost tělesné hmotnosti na jednotlivých vysvětlujících proměnných pomocí bodových diagramů.
- Sestavte lineární regresní model závislosti tělesné hmotnosti na ostatních proměnných.
- Ověřte splnění předpokladů tohoto modelu (graficky i pomocí vhodných testů). V případě, že jsou v diagnostických grafech vidět pozorování, která splnění předpokladů brání, odstraňte je z datového souboru a sestavte model znovu. Pro něj opět zkontrolujte splnění předpokladů.
- Pomocí výběrových korelačních koeficientů vysvětlujících proměnných a pomocí hodnot koeficientů VIF zhodnoťte multikolinearitu v modelu.
- Najděte odhady parametrů modelu a index determinace a interpretujte je.
- Na hladině významnosti 0.05 proveďte celkový F -test.
- Na hladině významnosti 0.05 proveďte dílčí t -testy pro jednotlivé koeficienty modelu.
- Sestrojte 95% intervaly spolehlivosti pro jednotlivé koeficienty.
- Vypusťte z modelu proměnné, které podle dílčích testů vycházejí nevýznamné (může se jednat i o absolutní člen). Porovnejte tento model s předchozím modelem. Pokud se zdá nový model lepší, ověřte u něj splnění předpokladů.
- Proveďte regresi metodou STEPWISE, zkuste verze backward, forward i both. Porovnejte výsledné modely s modelem, který jste sestavili na základě dílčích testů.