

Domáca úloha 7 - Moderná aplikovaná regresia

Prosím odovzdať do najbližších cvičení

1 (40b)

Majme dáta `crawl` z knižnice `faraway` obsahujúce štúdiu veku, kedy sa bábätká učia plaziť v závislosti od okolitej teploty. Dataset obsahuje aj ďalšie vhodné informácie, počet babätiek, ktoré sme v danom mesiaci pozorovali a ich varianciu. Fitnite do dát lineárny model s `crawling` ako odozvou, poriadne si naštudujte výskumnú otázku.

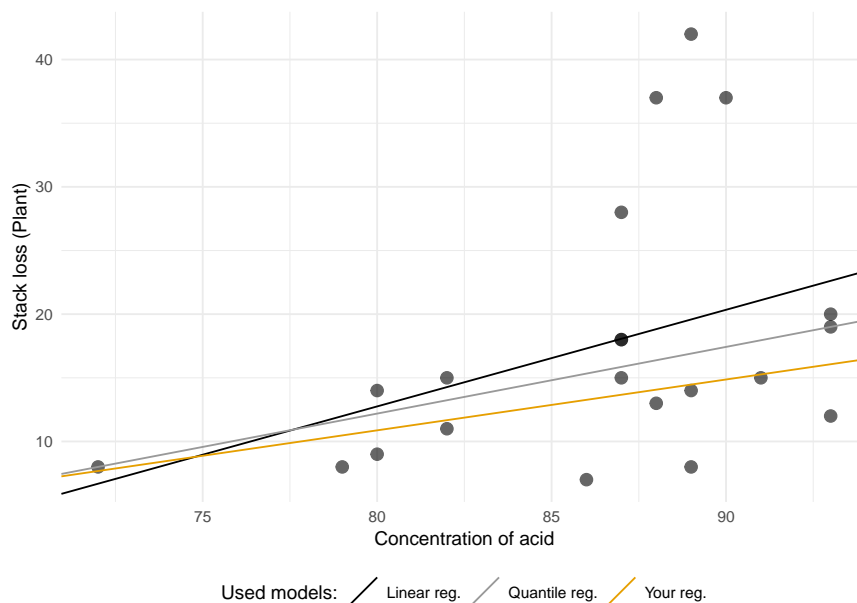
1. Vizualizujte a okomentujte varianciu reziduálov modelu.
2. Overte nezvyčajné pozorovania.
3. Navrhните kroky na vylepšenie modelu a realizujte ich.
4. Svoje zmeny overte celkovou diagnostikou modelu.

2 (60b)

Vezmite dáta `stackloss` z cvičení, uvažujme `stack.loss` ako odozvu a k nej prediktor `Acid.Conc...`. Dáta odmodelujte za pomoci:

1. Lineárny model (bez úprav)
2. Robustnú regresiu (ja som použil mediánovú, vy nemusíte)
3. Váš navrhovaný lineárny model (vylepšený, vaše kroky poriadne zdôvodnite)

Potom výsledky zakreslite do obrázka za pomoci knižnice `ggplot2` v ktorom nesmie chýbať legenda.



3 (Bonus)

Podobne ako v cvičení 1 ukážte, že odhad parametrov v lineárnom modeli so závislými normálne rozdelenými chybami a kovariančnou maticou Σ je $\hat{\beta}_{ML}^h = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}y$.