Generalized linear model

Peter Smits

Review of Linear Regression

Definition of linear regression

> *. . . a method that summarizes how the average values of a numerical outcome variable vary over subpopulations defined by linear functions of predictors. [. . .] Regression can be used to predict an outcome given a lienar function of these predictors, and regression coefficients can be thought of as comparisons across predicted values or as comparisons among averages in the data.*

Gelman and Hill, 2007, p.31

Compacy written form

$$y \sim \mathcal{N}(\mu, \sigma)$$
$$\mu = X\beta$$

$y \in \mathbb{R}$, $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$, $\beta_k \in \mathbb{R}$for $k = 1, ..., K$.

y is a length N vector of observations. X is a $N \times K$ matrix of covariates (and a column of 1s). β is a length K vector of regression coefficients (including intercept).

Interpreting regression parameters

Coeffcient β is the expected difference in y between two observations that differ by 1 in a single predictor.

Fitting and inspecting a regression model