Biostatistik II – Regressions Modelle

Arten der Regression:

|  |  |  |
| --- | --- | --- |
| **Regression** | **Y** | **Weitere Annahmen** |
| (Multiple) Lineare Regression | kontinuierlich | Residuen normalverteilt und unabhängig |
| Logistische Regression | binär | Residuen normalverteilt und unabhängig |
| Poisson Regression | diskret (Zähldaten) | Residuen normalverteilt und unabhängig, Achtung Overdispersion, Zero-Inflation! |
| Ordinale Regression | ordinal | Residuen normalverteilt und unabhängig |
| Multnomiale Regression | multinomial | Residuen normalverteilt und unabhängig |
| Cox PH Model | Zeit bis Event (Zensierung) | proportional Hazards, nicht-informative Zensierung |
| Gemischtes Modell | alles | „X“e können auch als zufällige Variablen modelliert werden, wenn diese abhängig sind |

**Wie baue ich ein Modell jetzt auf?**

Daten

Fragestellung + Hypothesen

Wahl des richtigen Modells (Art der Regression)

grobe Datenreinigung

Deskriptive Statistik

* Korrelationsplots (Auswahl einer der korrelierten Variablen)
* Table 1 (stratifiziert)
* Plots (Boxplots, Histogramme etc.)
* MDS Plot (Multidimensionale Skalierung)
* Untersuchung Outlier

Missings (evtl. Imputation)

Modellschätzten (evtl. Variablenselektion, Interaktion?)

Modeldiagnostik

* Residuenplot
* AIC
* (adj) R^2
* VIF
* (Kaplan-Meier-Kurve)

Ergebnis