Statistische Modelle

- Lineare Modelle

$$Y_i \sim N(\mu_i, \sigma^2) \quad \text{und} \quad E[Y] = X\beta$$

Y stetig X stetig und/oder kategoriell

- Logistische Regression

$$Y_i \sim B(1, \pi_i) \quad \text{und} \quad E[Y] = \frac{1}{1 + e^{-X\beta}}$$

Y binär X stetig und/oder kategoriell

- Loglineare Modelle

$$Y_i \sim P(\lambda_i) \quad \text{und} \quad E[Y] = e^{X\beta}$$

Y ganzzahlig X kategoriell

Und in jedem Fall sind die $\{Y_i\}$ als unabhängig betrachtet.

Verallgemeinerte lineare Modelle

Gegeben $\{X_j = x_j\}$ seien die $\{Y\}$ Werte unabhängig und

$$E[Y|X] = \mu = h(X\beta)$$

$\eta = X\beta$ ist der lineare Prädiktor
$\eta = g(\mu)$ ist die Linkfunktion

Die (bedingte) Verteilung von Y soll zu einer einfachen Exponentialfamilie gehören:

$$f(y, |\ \theta, \varphi) = \exp\{w\frac{y\theta - b(\theta)}{\varphi} + c(y, \frac{\varphi}{w})\}$$

- $c(y, \frac{\varphi}{w}) \geq 0$
- θ der "kanonischer" (Lokations-) Parameter
- $\varphi > 0$ der Streuungs- oder Störparameter (unabhängig von θ), und w bekannte Gewichte

Momente

$b(\theta)$ ist im Innern des Parameterraums beliebig oft differenzierbar, und alle Momente von Y existieren.

$$\int f(y;\theta,\varphi)dy = 1$$

$$\frac{\partial \int f(y;\theta,\varphi)dy}{\partial\theta} = 0$$

Daraus folgt unter schwachen Regularitätsbedingungen

$$\int \frac{\partial f(y;\theta,\varphi)}{\partial\theta}dy = 0$$

$$\frac{\partial f(y;\theta,\varphi)}{\partial\theta} = (y - b'(\theta))f(y;\theta,\varphi)$$

und deshalb gilt

$$E_\theta[Y] = \mu(\theta) = \frac{\partial b(\theta)}{\partial\theta}$$

Mit analogen Argumenten können wir zeigen, dass

$$VAR_\theta[Y] = \Sigma(\theta) = \frac{\varphi}{w}\frac{\partial^2 b(\theta)}{\partial\theta\partial\theta'} = \frac{\varphi}{w}V[\mu]$$

Bsp: **Normalverteilung:** $Y_i \sim N(\mu_i, \sigma^2)$.

$$\begin{aligned}
&f(y_i, \mu_i, \sigma^2) \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \exp -\{\frac{1}{2\sigma^2}(y_i - \mu_i)^2\} \\
&= \exp\{\frac{1}{\sigma^2}(y_i\mu_i - \frac{1}{2}\mu_i^2 - \frac{1}{2}y_i^2) - \frac{1}{2}\ln(2\pi\sigma^2)\} \\
&= \exp\{\frac{1}{\sigma^2}(y_i\mu_i - \frac{1}{2}\mu_i^2) - \frac{1}{2}(\frac{y_i^2}{\sigma^2} + \ln(2\pi\sigma^2))\}
\end{aligned}$$

also,

$$\begin{aligned}
&f(y_i \mid \theta_i, \varphi) \\
&= \exp\{\frac{1}{\varphi}(y_i\theta_i - b(\theta_i)) - \frac{1}{2}(\frac{y_i^2}{\varphi} + \ln(2\pi\varphi))\}
\end{aligned}$$

mit

$$\begin{aligned}
\theta_i &= \mu_i \\
b(\theta_i) &= \theta_i^2/2 = \mu_i^2/2 \\
\varphi &= \sigma^2 \\
w_i &= 1 \\
c(y_i, \frac{\varphi}{w_i}) &= -\frac{1}{2}(\frac{y_i^2}{\varphi} + \ln(2\pi\varphi))
\end{aligned}$$

Poisson: $Y_i \sim P(\lambda_i)$

$$
\begin{aligned}
f(y_i, \lambda_i) &= \exp(-\lambda_i)\frac{\lambda_i^{y_i}}{y_i!} \\
&= \exp(-\lambda_i + y_i \ln \lambda_i - \ln(y_i!)) \\
&= \exp(y_i \ln \lambda_i - \lambda_i - \ln(y_i!))
\end{aligned}
$$

also,

$$
\begin{aligned}
\theta_i &= \ln \lambda_i \\
b(\theta_i) &= \exp(\theta_i) = \lambda_i \\
\varphi &= 1 \\
w_i &= 1 \\
c(y_i, 1) &= \ln(y_i!)
\end{aligned}
$$

Das Poisson Modell gilt nicht nur für loglineare Modelle, sondern auch für Poisson Regressionsmodelle, z.B. Y als Anzahl Unfälle an verschiedenen Stellen und X die erklärenden Variablen.

Binomial: $\tilde{Y}_i \sim B(n_i, \pi_i)$ ersetzt durch $Y_i = \frac{\tilde{Y}_i}{n_i}$

$$\begin{aligned}
& f(n_i y_i, \pi_i) \\
& \quad = \binom{n_i}{n_i y_i} \pi_i^{n_i y_i} (1-\pi_i)^{n_i - n_i y_i} \\
& \quad = \exp[n_i y_i \ln \pi_i + (n_i - n_i y_i) \ln(1-\pi_i) + \ln \binom{n_i}{n_i y_i}] \\
& \quad = \exp\{n_i [y_i \ln \frac{\pi_i}{1-\pi_i} + \ln(1-\pi_i)] + \ln \binom{n_i}{n_i y_i}\}
\end{aligned}$$

also,

$$\begin{aligned}
\theta_i &= \ln(\frac{\pi_i}{1-\pi_i}) \\
b(\theta_i) &= \ln(1+e^{\theta_i}) = -\ln(1-\pi_i) \\
\varphi &= 1 \\
w_i &= n_i \\
c(y_i, 1/n_i) &= \ln \binom{n_i}{n_i y_i}
\end{aligned}$$

Andere mögliche (nicht-kanonische Linkfunktionen) sind:

Probit (Invers Normal) $\quad g(\pi) = \Phi^{-1}(\pi)$

"Complementary" log-log $\quad g(\pi) = log(-log(1-\pi))$

Alle sind Inverse bekannter Verteilungsfunktionen.

$V(\mu)$ für obige Beispiele

$$V(\mu) = \frac{\partial^2 b(\theta(\mu))}{\partial\theta\partial\theta'}$$

Normalverteilung: $b(\theta) = \frac{\theta^2}{2}$

$$\begin{aligned} V(\mu) &= I \\ \Rightarrow VAR_\theta[Y] &= \frac{\varphi}{w_i} V(\mu) = \sigma^2 I \end{aligned}$$

Poissonverteilung: $b(\theta_i) = e^{\theta_i}$

$$\begin{aligned} V(\lambda) &= \frac{\partial^2 \left(e^{\theta_1}, \ldots, e^{\theta_n}\right)}{\partial\theta\partial\theta'} \\ &= \mathrm{diag}(e^{\theta_i}) \\ &= \mathrm{diag}(\lambda_i) \\ \Rightarrow VAR_\theta[Y] &= \frac{\varphi}{w_i} V(\lambda) = \mathrm{diag}(\lambda_i) \end{aligned}$$

Binomialverteilung: $b(\theta_i) = \ln\left(1 + e^{\theta_i}\right)$

$$\begin{aligned}
V(p) &= \frac{\partial\left(\frac{1}{1+e^{\theta_1}}, \ldots, \frac{1}{1+e^{\theta_n}}\right)}{\partial\theta'} \\
&= \text{diag}(e^{-\theta_i}\left(1 + e^{-\theta_i}\right)^{-2}) \\
&= \text{diag}(\frac{1-\pi_i}{\pi_i} \cdot \left(\frac{1}{\pi_i}\right)^{-2}) \\
&= \text{diag}(\pi_i(1-\pi_i)) \\
\Rightarrow VAR_\theta[Y] &= \frac{\varphi}{w_i} V(\pi) = \text{diag}(\frac{\pi_i(1-\pi_i)}{n_i})
\end{aligned}$$

Weitere Verteilungen in der Exponential-Familie

Gamma Verteilung

$$f(y; \mu, \nu) = \frac{1}{\Gamma(\nu)} \left(\frac{\nu}{\mu}\right)^{\nu} y^{\nu-1} e^{-\frac{\nu y}{\mu}}$$

$$y, \mu, \nu > 0$$

Invers Gauß Verteilung

$$f(y; \mu, \alpha) = \sqrt{\frac{\alpha}{2\pi y^3}} e^{-\frac{\alpha(y-\mu)^2}{2\mu^2 y}}$$

$$y, \mu, \alpha > 0$$

Diese Verteilung modelliert die erste Übertrittszeit einer bestimmten Schwelle in einer Brownschen Bewegung (Zufallsprozess) oder die asymptotische Anzahl von Versuchen in einer sequenziellen Studie.

Form der Verteilungen

Sowohl stetige Verteilungen (Normal, Gamma,...) als auch diskrete Verteilungen (Binomial, Poisson) sind Mitglieder der Exponential-Familie.

Kanonische Link-Funktionen

Für jede Verteilung gibt es eine besondere Linkfunktion, die kanonische Linkfunktion, wo $\theta = \eta = X\beta$. In diesem Fall gibt es suffiziente Statistiken $X'y$, wie aus der Maximum Likelihood Gleichungen zu sehen ist:

$$\ell(\theta, \varphi; Y) = \sum_{i=1}^{n} \left\{ \frac{w_i}{\varphi} (y_i\theta_i - b(\theta_i)) + c(y_i, \frac{\varphi}{w_i}) \right\}$$

Mit $w_i = w$ und $\frac{\partial \ell(\theta,\varphi;Y)}{\partial \beta} = 0$ gilt:

$$X'y = X'\widehat{\mu}$$

Verteilung	Kanonischer Link	
Normal	Identität	$\eta_i = \mu_i$
Binomial	Logit	$\eta_i = \log\left[\frac{\pi_i}{1-\pi_i}\right] = \log\left[\frac{\mu_i}{n_i-\mu_i}\right]$
Poisson	Log	$\eta_i = \log(\mu_i)$
Gamma	Invers	$\eta_i = \frac{1}{\mu_i}$
Invers Gauß	Invers^2	$\eta_i = \frac{1}{\mu_i^2}$

Schätzung im GLM ML-Ansatz:
Log-Likelihood eines GLM:

$$\ell(\theta, \varphi; Y) = \sum_{i=1}^{n} \left\{ \frac{w_i}{\varphi} (y_i \theta_i - b(\theta_i)) + c(y_i, \frac{\varphi}{w_i}) \right\}$$

Wir setzen $a_i(\varphi) = \frac{\varphi}{w_i}, \quad i = 1, \ldots, n$.
Für ML-Schätzer für β setzen wir die Ableitung der Log-likelihood nach gesuchten Parametern gleich 0, also

$$\frac{\partial \ell(\theta, \varphi; Y)}{\partial \beta} = 0$$

Mit den gegebenen Parametrisierungen gilt:

$$\frac{\partial \ell(\theta, \varphi; Y)}{\partial \beta} = \frac{\partial \ell(\theta, \varphi; Y)}{\partial \theta} \frac{\partial \theta}{\partial \mu} \frac{\partial \mu}{\partial \eta} \frac{\partial \eta}{\partial \beta}$$

Es gilt

$$\frac{\partial \eta}{\partial \beta} = X,$$

und die Ableitung

$$\frac{\partial \mu}{\partial \eta}$$

hängt von der Link-Funktion ab.

$$E_\theta[Y] = \mu(\theta) = \frac{\partial b(\theta)}{\partial \theta}$$

$$\frac{\partial^2 b(\theta)}{\partial \theta \partial \theta'} = \mathrm{diag}(\frac{1}{a_i(\varphi)}) \cdot VAR[Y]$$

Des weiteren gilt

$$\frac{\partial \mu}{\partial \theta} = \frac{\partial^2 b(\theta)}{\partial \theta \partial \theta'} = V(\mu)$$

Somit erhalten wir mit $v_i = (V(\mu))_{ii}$

$$\begin{aligned} \frac{\partial \ell(\theta_i, \varphi; Y)}{\partial \beta_j} &= \frac{w_i(y_i - \mu_i)}{\varphi} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} x_{ij} \\ &= \frac{w_i(y_i - \mu_i)}{v_i \varphi} \frac{\partial \mu_i}{\partial \eta_i} x_{ij} \end{aligned}$$

Setzen wir nun

$$\begin{aligned} W &= \mathsf{diag}(v_i \left(\frac{\partial \eta_i}{\partial \mu_i}\right)^2 \frac{\varphi}{w_i}) \\ &= \mathsf{diag}(W_i) \end{aligned}$$

und

$$z_i = \eta_i + (y_i - \mu_i)\frac{\partial \eta_i}{\partial \mu_i}$$

erhalten wir die Normalengleichungen eines gewichteten KQ-Problems

$$E[z] = \eta, \quad VAR[z] = W,$$

also

$$X'W^{-1}X\beta = X'W^{-1}z$$

jedoch mit iterativen Gewichten W_i^{-1}.
Lösung dieser Normalengleichungen liefert

$$\widehat{\beta} = \left(X'W^{-1}X\right)^{-1} X'W^{-1}z.$$

Die Schätzer für η und μ erhält man dann durch

$$\widehat{\eta} = X\widehat{\beta}, \qquad \widehat{\mu} = g^{-1}(X\widehat{\beta}).$$