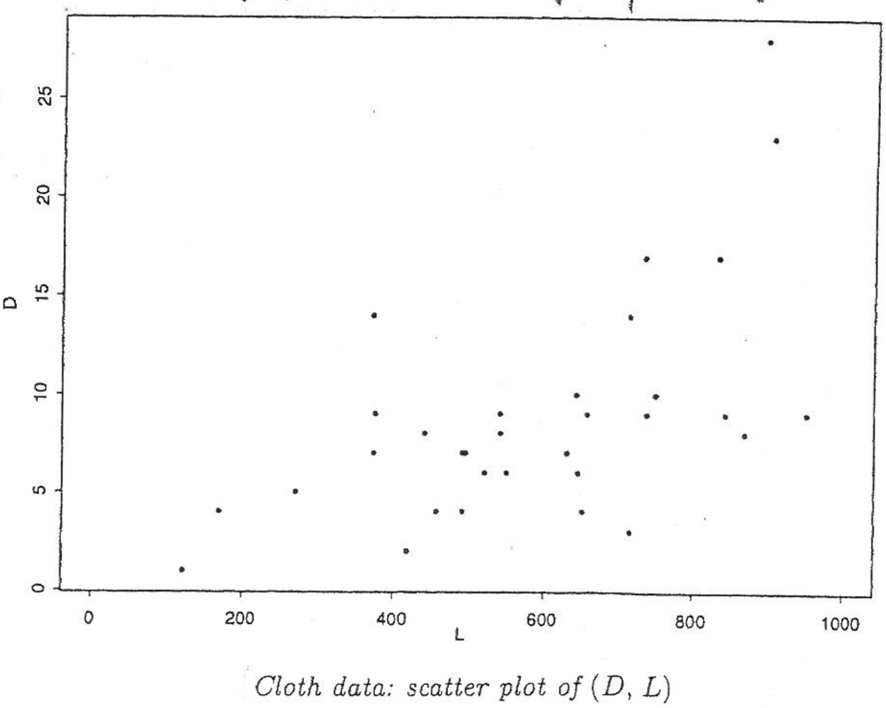
Generaliserede lineære modeller

Sædvanlige lineære modeller er ikke altid hensigtsmæssige, da

- evt. linearisering kan være for kunstig
- prediktion af fremtidige observationer giver ikke-tilladelige y-værdier
- variansen ikke konstant
- store afvigelser fra normalfordelingen (fx for diskrete variable)

De sædvanlige lineære modeller bliver nu specialtilfælde inden for en ny mere omfattende klasse af modeller, de generaliserede lineære modeller – GLM

Eks. antal fejl D i stof af længde L



Cloth data: scatter plot of (D, L)

Stat. 28.4.2010

Model: $D \sim p(\beta L)$, β parameter

Falder under GLM

$\square$

Klasse af fordelinger

Stokastiske variable Y med tætheds-funktion

$$f(y) = \exp\left(\frac{w}{\psi} (y\theta - b(\theta)) + c(y, \psi) \right),$$

hvor

θ og ψ er parametre,

w en kendt konstant (vægtfaktor)

$b(\cdot)$ og $c(\cdot,\cdot)$ kendte funktioner,

udgør for ethvert valg af dispertions-parameteren ψ en eksponentiel familie og noteres

$$Y \sim EF\left(b(\theta), \frac{\psi}{w} \right)$$

Udledning af et par formler:

$$0 = \frac{\partial}{\partial \theta} \int_y f(y) \, d\nu(y)$$

$$= \int_y f(y) \frac{w}{\psi} (y - b'(\theta)) \, d\nu(y)$$

$$= \frac{w}{\psi} (E[Y] - b'(\theta))$$

$$\Rightarrow \quad E[Y] = b'(\theta) = \mu$$

$$0 = \int_y f(y) \left(-\frac{w}{\psi} b''(\theta) + \left(\frac{w}{\psi} (y - b'(\theta)) \right)^2 \right) d\nu(y)$$

$$= -\frac{w}{\psi} b''(\theta) + \left(\frac{w}{\psi}\right)^2 Var[Y]$$

$$\Rightarrow \quad Var[Y] = \frac{\psi}{w} b''(\theta)$$

Sæt $V(\mu) = b'' \circ b'^{-1}(\mu)$ variansfunktionen

$$\left(\mathrm{Var}[Y] = \frac{\varphi}{w} V(\mu) \right)$$

eks. $Y \sim P(\mu)$

$$P(Y=y) = \frac{e^{-\mu}\mu^{y}}{y!}$$

$$= \exp(y \ln\mu - \mu - \ln y!)$$

$$= \exp(y\theta - e^{\theta} - \ln y!)$$

dvs. $b(\theta) = e^{\theta} \Rightarrow b'(\theta) = e^{\theta} = \mu$

$$\Rightarrow b''(\theta) = e^{\theta} = \mu$$

$$w = \varphi = 1, \quad \mu = e^{\theta}, \quad V(\mu) = \mu$$

$$c(y;\varphi) = -\ln y!$$

$$\mathrm{Var}\,Y = \frac{1}{1}\mu = \mu$$ □

* se eks. side 9

Lineære normale modeller

Sæt $\eta = X\beta$, dvs. $\eta_i = x_i^T\beta$ $\left(= \sum_j x_{ij}\beta_j\right)$,

$X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{bmatrix}$, dvs. x_i indgår som den i'te rækkevektor i X

η_i benævnes lineær prædiktor

$$Y_i \sim N(\mu_i, \sigma^2) \qquad \underbrace{\mu_i = \eta_i}_{\text{middelværdien er lig den lineære prædiktor}} \qquad \eta_i = x_i^T\beta$$

Ny klasse af modeller

$Y_i \sim EF(b(\theta_i), \frac{\psi}{w_i})$ $\quad g(\mu_i) = \eta_i \quad$ $\eta_i = x_i^T\beta$

stokastisk struktur $\quad$ linkfunktion $\quad$ lineær prædiktor

mere detaljeret:

- $y_1, \dots, y_n$ er obs. af uafh. stok. var. $Y_1, \dots, Y_n$
- $Y_i \sim EF(b(\theta_i), \frac{\psi}{w_i})$, dvs. $E[Y_i] = \mu_i = b'(\theta_i)$
- der vælges en linkfunktion $g(\cdot)$, så $g(\mu_i) = \eta_i$, $\eta_i = x_i^T\beta$ lineær prædiktor, β er parametervektor
- funktionen $b(\cdot)$ og parameteren ψ er fælles for alle Y_i (hvorimod w_i gerne må variere)
- $b(\cdot)$, $c(\cdot,\cdot)$ og $g(\cdot)$ er kendte funktioner

når $g(\cdot)$ er den identiske afbildning, og der gælder $Y_i = \mu_i + U_i$, $U_i \sim N(0, \sigma^2)$, så fremkommer den lineære normale model.

En vilkårlig GLM har i almindelighed ikke en sådan struktur.

eks. $Y_i \sim p(\mu_i)$

Poisson regression: $\ln \mu_i = \eta_i$

$Var[Y_i] = \mu_i = \exp \eta_i$, dvs. $Y_i \neq \mu_i + U_i$ □

eks. logistisk regression, $Y_i \sim b(1,\mu_i)$

$$P(Y_i = y_i) = \frac{e^{\theta_i y_i}}{1+e^{\theta_i}}, \quad y_i = 0,1, \quad \mu_i = \frac{e^{\theta_i}}{1+e^{\theta_i}}$$

$$= \exp(\theta_i y_i - \ln(1+e^{\theta_i})),$$

$$\theta_i = \operatorname{logit} \mu_i = \ln \frac{\mu_i}{1-\mu_i}$$

$$\theta_i = \eta_i = \alpha + \beta x_i$$

$$b(\theta) = \ln(1+e^{\theta}) \quad \text{(indeks } i \text{ udeladt)}$$

$$b'(\theta) = \frac{1}{1+e^{\theta}} e^{\theta} = \mu$$

$$b''(\theta) = \frac{(1+e^{\theta})e^{\theta} - e^{\theta} \cdot e^{\theta}}{(1+e^{\theta})^2} = \frac{e^{\theta}}{(1+e^{\theta})^2} = \frac{\mu}{1+e^{\theta}}$$

$$= \mu(1-\mu)$$

$$V(\mu) = \mu(1-\mu) \quad (= \operatorname{Var}[Y], \text{ da } w = \psi = 1)$$

udvidelse af modellen:

- $\theta_i = x_i^T \beta$, dvs. flere forkl. var., kan umiddelbart indføres
- gentagne observationer for enhver kombination af de forkl. var.

→ Vi kan benytte $\tilde{Y}_i \sim b(m_i, \mu_i)$, men

dvs. m_i gentagelser

bedre med $Y_i = \frac{\tilde{Y}_i}{m_i}$ $(\Rightarrow E[Y_i] = \mu_i)$

$y = 0, \frac{1}{m}, \frac{2}{m}, \ldots, 1$

$$P(Y = y) = \binom{m}{my} \mu^{my} (1-\mu)^{m(1-y)} \quad \text{(indeks } i \text{ udeladt)}$$

$$= \binom{m}{my} \left(\frac{\mu}{1-\mu}\right)^{my} (1-\mu)^m$$

$\theta = \ln \frac{\mu}{1-\mu} = \operatorname{logit} \mu$, $w = m$

$$= \exp\left(w\left(y\theta - \ln(1+e^{\theta})\right) + \ln\binom{w}{wy}\right)$$

□

	Normal	Poisson	Binomial/m	Gamma
	$N(\mu, \sigma^2)$	$P(\mu)$	$Bin(m, \mu)/m$	$Gamma(\omega, \omega/\mu)$
Support	$(-\infty, \infty)$	$\{0, 1, 2, \ldots\}$	$\{0, \frac{1}{m}, \frac{2}{m}, \ldots, 1\}$	$(0, \infty)$
ψ	σ^2	1	1	ω^{-1}
w	1	1	m	1
$b(\theta)$	$\theta^2/2$	$\exp(\theta)$	$\log(1+e^\theta)$	$-\log(-\theta)$
$c(y; \psi)$	$-\frac{1}{2}\left(\frac{y^2}{\psi} + \log(2\pi\psi)\right)$	$-\log y!$	$\log\begin{pmatrix} m \\ my \end{pmatrix}$	$\omega\log(\omega y) - \log y$ $- \log\Gamma(\omega)$
$\mu(\theta)$	θ	$\exp(\theta)$	$e^\theta/(1+e^\theta)$	$-1/\theta$
Canonical link	identity	logarithm	logit	reciprocal
$V(\mu)$	1	μ	$\mu(1-\mu)$	μ^2

Characteristics of some distributions

Likelihood

X $n \times p$ (husk x_i indgår i matricen X som den i'te rækkevektor)

$\beta \in \mathbb{R}^p$

$y_1, \ldots, y_n$ obs. af uafh. stok. var. med middelværdier $\mu_1, \ldots, \mu_n$

$g(\mu_i) = \eta_i = x_i^T \beta$ linear prediktor

β og ψ parametre

ψ opfattes ofte som 'nuisance' parameter

log-likelihoodfunktionen

$$\ell(\beta; y_1, \ldots, y_n) = \sum_i \ell_i(\beta; y_i)$$

$$= \sum_i \left(\frac{w_i}{\psi} (y_i \theta_i - b(\theta_i)) + c_i(y_i, \psi) \right)$$

$$\frac{\partial l_i}{\partial \beta_j} = \frac{\partial l_i}{\partial \theta_i} \frac{\partial \theta_i}{\partial \mu_i} \frac{\partial \mu_i}{\partial \eta_i} \frac{\partial \eta_i}{\partial \beta_j}$$

$$\frac{\partial l_i}{\partial \theta_i} = \frac{w_i}{\phi}(y_i - b'(\theta_i)) = \frac{w_i}{\phi}(y_i - \mu_i)$$

$$\frac{\partial \mu_i}{\partial \theta_i} = b''(\theta_i) = \frac{w_i}{\phi} \mathrm{Var}[Y_i]$$

$$\frac{\partial \eta_i}{\partial \beta_j} = x_{ij}$$

$$\frac{\partial l_i}{\partial \beta_j} = \frac{w_i}{\phi}(y_i - \mu_i) \frac{\phi}{w_i} \frac{1}{\mathrm{Var}[Y_i]} \frac{\partial \mu_i}{\partial \eta_i} x_{ij}$$

$$= \frac{(y_i - \mu_i) x_{ij}}{\mathrm{Var}[Y_i]} \frac{\partial \mu_i}{\partial \eta_i}$$

likelihoodligninger for β er altså:

$$\sum_i \frac{(y_i - \mu_i) x_{ij}}{\mathrm{Var}[Y_i]} \frac{\partial \mu_i}{\partial \eta_i} = 0 \quad , \quad j = 1, \ldots, p$$

Fisher information

Bidrag fra $l_i(\beta)$ til det j,k'te element i $I(\beta)$

$$-E\left[\frac{\partial^2 l_i}{\partial \beta_j \partial \beta_k}\right] = E\left[\frac{\partial l_i}{\partial \beta_j} \frac{\partial l_i}{\partial \beta_k}\right]$$

$$= E\left[\frac{(Y_i - \mu_i) x_{ij}}{\mathrm{Var}[Y_i]} \frac{\partial \mu_i}{\partial \eta_i} \frac{(Y_i - \mu_i) x_{ik}}{\mathrm{Var}[Y_i]} \frac{\partial \mu_i}{\partial \eta_i}\right]$$

$$= \frac{\mathrm{Var}[Y_i] x_{ij} x_{ik}}{(\mathrm{Var}[Y_i])^2} \left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2$$

$$= \frac{x_{ij} x_{ik}}{\mathrm{Var}[Y_i]} \left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2$$

j, k'te element i $I(\beta)$:

$$-\sum_i E\left[\frac{\partial^2 l_i}{\partial\beta_j\,\partial\beta_k}\right] = \sum_i \frac{x_{ij}\,x_{ik}}{Var[Y_i]}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2$$

Sæt

$$\tilde{W} = \begin{bmatrix} \tilde{w}_1 & & 0 \\ & \ddots & \\ 0 & & \tilde{w}_n \end{bmatrix}$$

med

$$\tilde{w}_i = \frac{1}{Var[Y_i]}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2 = \frac{w_i}{\phi\, V(\mu_i)}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2$$

$$I(\beta) = X^T\tilde{W}X$$

eks. $Y_i \sim e(\exp(\beta_1+\beta_2 x_i))$, $\sum x_i = 0$

$$= e(\exp \eta_i)$$

$$= e(\rho_i)$$

$f(y) = \rho \exp(-\rho y)$ (indeks i udeladt)

$$= \exp(-\rho y + \ln \rho)$$

$$= \exp(\theta y + \ln(-\theta))$$

$$\theta = -\rho = -e^{\eta}$$

$$b(\theta) = -\ln(-\theta)$$

$$b'(\theta) = -\frac{1}{-\theta}(-1) = -\frac{1}{\theta} = \mu = e^{-\eta}$$

$$b''(\theta) = \frac{1}{\theta^2} = V(\mu)$$

$$\frac{d\mu}{d\eta} = -e^{-\eta} = \frac{1}{\theta}$$

likelihoodlikningene

$$\sum_i \frac{(y_i + \frac{1}{\theta_i})\, x_{ij}}{\frac{1}{\theta_i^2}} \; \frac{1}{\theta_i} = 0$$

$$\Leftrightarrow \quad \sum_i \frac{(y_i - \frac{1}{\rho_i})\, x_{ij}}{\frac{1}{-\rho_i}} = 0$$

$$\Leftrightarrow \quad \sum_i \left(y_i - \frac{1}{\rho_i}\right) \rho_i\, x_{ij} = 0 \qquad \begin{cases} x_{i1} = 1 \\ x_{i2} = x_i \end{cases}$$

$$\Rightarrow \quad \begin{cases} \sum_i y_i \rho_i - n = 0 \\ \sum_i x_i y_i \rho_i = 0 \end{cases} \qquad \left(\sum_i x_i = 0\right)$$

$$\tilde{w}_i = \frac{1}{\frac{1}{\theta_i^2}} \left(\frac{1}{\theta_i}\right)^2 = 1$$

$I(\beta) = X^T X$, resultat som i tidl. eks. (note 4, side 2-3) □

---

* (fra side 3)

eks $\quad Y \sim N(\mu, \sigma^2)$

$$f(y) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$$

$$= (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}(y^2 - 2y\mu + \mu^2)\right)$$

$$= \exp\left(\frac{1}{\sigma^2}\left(y\mu - \frac{\mu^2}{2}\right) - \frac{1}{2}\left(\frac{y^2}{\sigma^2} + \ln(2\pi\sigma^2)\right)\right)$$

dvs. $\quad \theta = \mu, \quad b(\theta) = \frac{\theta^2}{2}, \quad b'(\theta) = \theta = \mu, \quad b''(\theta) = 1$

$w = 1, \quad \phi = \sigma^2, \quad \mu = \theta, \quad V(\mu) = 1$

$c(y; \sigma^2) = -\frac{1}{2}\left(\frac{y^2}{\sigma^2} + \ln(2\pi\sigma^2)\right)$

$\operatorname{Var} Y = \frac{\phi}{w} V(\mu) = \frac{\sigma^2}{1} 1 = \sigma^2$ □