22.10.2015 S16, S17, S18

1

a) $y \sim Bin(n, \pi)$

$$f(y|n,\pi) = \binom{n}{y} \pi^{y} (1-\pi)^{n-y} =$$

$$= \binom{n}{y} \exp\left(y \log \pi + (n-y)\log(1-\pi)\right) =$$

$$= \binom{n}{y} \exp\Big(y \underbrace{\log\left(\frac{\pi}{1-\pi}\right)}_{\theta} + n \underbrace{\log(1-\pi)}_{a(\theta)} \Big) =$$

$$= \Big[c(y) = \binom{n}{y};\ \theta = \log\frac{\pi}{1-\pi},\ a(\theta) = \log(1-\pi) =$$

$$= n \log(1 + e^{\theta}) \Big] = c(y) \exp\left(\frac{y\theta - a(\theta)}{\varphi} \right) =$$

$$= [\varphi = 1] = c(y) \exp(y\theta - a(\theta))$$

$$\dot a(\theta) = \frac{n e^{\theta}}{1+e^{\theta}} = \pi n$$

$$\ddot a(\theta) = \frac{n e^{\theta}(1+e^{\theta}) + n e^{\theta} e^{\theta}}{(1+e^{\theta})^2} = \frac{n e^{\theta}}{(1+e^{\theta})^2} =$$

$$= \boxed{n\pi(1-\pi)}$$

1 $f(y) = f(y|n,\pi) \overset{d}{=} Binom(n,\pi) \in$ exponential family

2 $g(\theta_i) = \eta_i$ $\quad \boxed{g^{-1}(\eta_i) = \theta_i}$

$$\boxed{\theta_i = \log\left(\frac{\pi_i}{1-\pi_i}\right) = \eta_i}$$ — canonical

3 $\eta_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni}$

$g(\pi_i) = \varphi^{-1}(\pi_i) = \eta_i$ – Probit link.

$\pi_i = \varphi(\eta_i) \in (0,1)$

$g(\pi_i) = \log(-\log(1-\pi_i))$.

$\pi_i = g^{-1}(\eta_i) \in (0;1)$

c). $\eta_i = \alpha + \beta x_i$, $y_i \overset{d}{=} Binom(1, \pi_i)$

$H_0: \beta = 0$

$$l(\theta_i) = \log L(\theta_i) = \sum \log(c(y_i,\varphi)) + \frac{y_i\theta_i - a(\theta_i)}{\varphi}$$

$$\frac{\partial l(\theta_i)}{\partial \theta_i} = \frac{1}{\varphi}(y_i - \dot a(\theta_i)) = 0$$

$$\Longleftrightarrow y_i = \dot a(\theta_i) = \tilde\mu_i \qquad \theta_i = \dot a^{-1}(\tilde\mu_i)$$

$$\Delta = 2(\tilde l - \hat l)$$

$H_0: \beta = 0 \Rightarrow LR$ test.

$\eta_i = \alpha + \beta x_i$ – full

$\eta_i^* = \alpha$ – restricted model

$$\Delta_{full} = 2(\tilde l - \hat l_{full});\ \Delta_{res} = 2(\tilde l - \hat l_{restr})$$

$$LR = 2(\hat l_{full} - \hat l_{rest}) = \Delta_{restr} - \Delta_{full} \sim$$

$$\sim \chi^2_{p_{full} - p_{ress}} \overset{d}{=} \chi^2_1$$

$p_{full} = 2$, $p_{ress} = 1$

$$P(\chi_1^2 > LR)$$

d)

$$l(\alpha,\beta) = \sum \log\left(\frac{1}{y_i}\right) + y_i \log\left(\frac{\pi_i}{1-\pi_i}\right) +$$

$$+ 1\log(1-\pi_i) = \Big[\log\frac{\pi_i}{1-\pi_i} = \eta_i = \alpha + \beta x_i \Big] =$$

$$\sum \log\left(\frac{1}{y_i}\right) + y_i(\alpha + \beta x_i) - \log(1 + e^{\alpha+\beta x_i})$$

$$\frac{\partial l(\alpha,\beta)}{\partial \alpha} = \sum y_i - \frac{e^{\alpha+\beta x_i}}{1+e^{\alpha+\beta x_i}} = [\eta_i = \alpha + \beta x_i]$$

$$= \sum y_i - \frac{e^{\eta_i}}{1+e^{\eta_i}}$$

$$\frac{\partial l(\alpha,\beta)}{\partial \beta} = \sum y_i x_i - \frac{x_i e^{\alpha+\beta x_i}}{1+e^{\alpha+\beta x_i}} =$$

$$= \sum x_i\left(y_i - \frac{e^{\alpha+\beta x_i}}{1+e^{\alpha+\beta x_i}}\right) = \sum x_i\left(y_i - \frac{e^{\eta_i}}{1+e^{\eta_i}}\right)$$

$$I = -E\left\{ \frac{\partial^2 l(\alpha,\beta)}{\partial(\alpha,\beta)\,\partial(\alpha,\beta)} \right\}$$

$$\boxed{\beta^{(k+1)} = \beta^{(k)} + I^{-1} S.}$$ — Fisher algo.

5.17 $f(y|\theta) = c(y)\exp(y\theta - a(\theta)) =$ 2

$= \exp(y\theta - a(\theta) + d(y))$.

$p(y) = \frac{\lambda^y e^{-\lambda}}{y!} = \frac{1}{y!}\exp(y\underbrace{\log\lambda}_{\theta} - \underbrace{\lambda}_{a(\theta)}) =$

$= [\lambda = e^\theta,\ a(\theta) = e^\theta \Rightarrow \dot a(\theta) = e^\theta = \lambda$

$\ddot a(\theta) = e^\theta = \lambda]$

b) $f(y) = \lambda e^{-\lambda y} = \exp(\underbrace{-\lambda}_{\theta} y + \underbrace{\log\lambda}_{a(\theta)}) =$

$= [\lambda = -\theta \Rightarrow a(\theta) = -\log(\lambda) = -\log(-\theta)]$

c) $\dot a(\theta) = -\frac{1}{\theta} = \frac{1}{\lambda} = E(y_i)$

$\ddot a(\theta) = \frac{1}{\theta^2} = \frac{1}{\lambda^2} = Var(y_i)$

$Var(y_i) = (E(y_i))^2,\ y_i \sim Expo(\lambda)$

$F(y|\lambda) = 1 - \left(\frac{1}{y}\right)^\lambda,\ y > 1$.

$V = \log(Y) \sim expo(\mu = \frac{1}{\lambda})$

$f(y|\lambda) = \frac{\partial F(y|\lambda)}{\partial y} = \lambda y^{-(\lambda+1)}$

$V = \log(Y)$

$Y = e^V$ $g(V) = f(y(V))\left|\frac{\partial y(V)}{\partial V}\right| =$

$= \lambda (e^V)^{-(\lambda+1)} e^V = \lambda e^{-V\lambda - V} e^V = \lambda e^{-V\lambda} =$

$\stackrel{d}{=} Expo(\lambda)$. $E\{V\} = \frac{1}{\lambda}$ $Var(V) = \frac{1}{\lambda^2}$

e) $Y_i \sim Pareto(\lambda)$

$\lambda_i = \exp(\alpha + \beta X_i) = \exp(\eta_i)$.

$g(\lambda_i) = \log(\lambda_i) = \eta_i$

1). $V_i = \log Y_i \stackrel{d}{=} Expo(\lambda_i)$

2). $E(V_i) = \frac{1}{\lambda_i}$. $\boxed{\eta_i = \log(\lambda_i)} \neq g(E(y_i))$.

3). $\eta_i = \alpha + \beta X_i$ - linear in parameters.

$(=)$

$\lambda_i = \exp(\eta_i) = \exp(\alpha + \beta X_i) \Rightarrow$

apply Fisher scoring to estimate $\hat\alpha$ & $\hat\beta$

f). $g(y,\theta) = \exp(\theta\, a(y) - c(\theta) + d(y))$.

$= c(y)\exp(\theta\, a(y) - c(\theta))$.

$f(y) = \lambda y^{-(\lambda+1)} = \exp(\log\lambda - (1+\lambda)\log y) =$

$= \exp(-\lambda\log y + \log\lambda - \log y) =$

$= [\theta = -\lambda$ $a(y) = \log(y)$. $d(y) = -\log(y)$

$c(\theta) = \log\lambda]$

f). $g(y|\theta) = \exp(\theta\, a(y) - c(\theta) + d(y))$

$\xi = a(y)$. $y(\xi) = a^{-1}(\xi)$

$\gamma(\xi) = g(y(\xi))\left|\frac{\partial y(\xi)}{\partial \xi}\right| =$

$g(a^{-1}(\xi))\left|\frac{\partial a^{-1}(\xi)}{\partial \xi}\right| =$

$\exp(\theta\, a(a^{-1}(\xi)) - c(\theta) + d(a^{-1}(\xi)))$

$= \exp(\theta\xi - c(\theta) + d^*(\xi)) =$

$= [c(\theta) =: a(\theta),\ c^*(\xi) = \exp d^*(\xi)] =$

$= c^*(\xi)\, e^{\theta\xi - a(\theta)}$ exponential family.

GLM OK!

S 18 / a) Analysys of deviance.

3

$y \in (0, 1)$.

X_1 - factor of year of birth $\{1, \dots, 5\}$ $\beta_{11} \dots \beta_{14}$

X_2 - factor of gender $\in \{0, \underline{1}\}$ β_{21}

X_3 - weight $\in \mathbb{R}^+ \Rightarrow \beta_3$

	DF	Δ	LR.	DF	$p(\chi_i^2 > \Delta)$
Null (β_0)	$n-1$	1101	$\Delta_i - \Delta_{i+1}$	-	
vekt (β_0, β_3)	$n-2$	842	1101-842	1	⋮
factor(kohort) ($\beta_0, \beta_3, \beta_{11} \dots \beta_{14}$)	$n-6$	527		4	
kjon ($\beta_0, \beta_3, \beta_{11}, \dots, \beta_{14}, \beta_{21}$)	$n-7$	434	.	1	.
vekt : factor(kohort) (...)	$n-11$	428	⋮	4	
vekt : kjonn	$n-12$	427		1	
factor(kohort) : kjonn	$n-16$	413	.	4	
vekt : factor(kohort) : kjonn	$n-22$	408	413-408	4	

$LR = \Delta_{rest} - \Delta_{full}$

J - causes of death.

J+1 - being alive

$$y_{J+1} \overset{d}{=} \text{Multinomial}\left(n = J+1, p_1 \dots, p_J, p_{J+1} = 1 - \sum_{i=1}^{J} p_i\right)$$

$$f(y_1, \dots, y_{J+1}) = \frac{n!}{y_1! \dots y_{J+1}!} p_1^{y_1} \dots p_J^{y_J} p_{J+1}^{y_{J+1}}$$

$$n = \sum_{i=1}^{J+1} y_i \qquad \sum_{i=1}^{J+1} p_i = 1$$

$$E\{y_i\} = n p_i, \quad i \in \overline{1, J+1}$$

$$\mathrm{Var}(y_i) = n p_i (1 - p_i)$$

$$\mathrm{cov}(y_j, y_k) = -n p_j p_k \quad \forall j, k \in \overline{1, J+1}$$

$$\log\left(\frac{p_{ji}}{p_{0i}}\right) = \eta_{ji} = \beta_j^T x_i$$

$$p_{ji} = \frac{\exp(\eta_{ji})}{1 + \sum_{k=1}^{J} \exp(\eta_{ki})}$$

$$P(z_j = 0) = P\left(\sum_{i=1}^{J} y_{ij} = 1\right)$$

$$z_j = \begin{cases} 0, & \sum_{i=1}^{J} y_{ij} = 1 \\ 1, & y_{J+1\,j} = 1 \end{cases} \Rightarrow$$

$$z_j = \mathrm{Binom}(1, p_{J+1})$$