MLE og eksponentielle familier

$$L(\theta; y) = \exp(\varphi(\theta)\, t(y) - \tau(\theta)) \qquad (k=1)$$

likelihood ligning : $\varphi(\theta)\, t(y) - \tau(\theta) = 0$

diff. : $\varphi'(\theta)\, t(y) - \tau'(\theta) = 0 \quad \Rightarrow \quad t(y) = \dfrac{\tau'(\theta)}{\varphi'(\theta)}$,

* se formel (2.7) s. 46

hvoraf * $E[T(Y); \hat{\theta}] = t(y)$, når $\hat{\theta}$ eksisterer,

dvs. $\hat{\theta}$ er den værdi af θ, som sætter middelværdien af den sufficiente stikprøvefkt. lig den observerede værdi af samme

Løsn. for $\hat{\theta}$ når og kun når $t(y)$ er indre pkt. i det konvekse hylster af $t(Y)$'s støtte **

** kravet kun relevant for diskrete var.

fortsat diff. af $l(\theta)$:

$l''(\theta) = \varphi''(\theta)\, t(y) - \tau''(\theta)$, heraf

$$I(\theta) = E\left[-l''(\theta;Y);\theta\right] = -\varphi''(\theta)\underbrace{E\left[t(Y);\theta\right]}_{= t(y),\ \text{når } \theta = \hat{\theta}} + \tau''(\theta)$$

dvs. $I(\theta)\big|_{\theta=\hat{\theta}} = -l''(\hat{\theta};y) \Rightarrow l''(\hat{\theta};y) \leq 0$

løsn. af $E[T(Y);\hat{\theta}] = t(y)$ mht. $\hat{\theta}$ er altså entydig *

eks logistisk regression

fra eks. 2.4.4 s. 40-41 har vi, at

$(\sum y_i, \sum x_i y_i)$ er suff. stikpr. fkt. for (α,β).

desuden $E\left[(\sum Y_i, \sum x_i Y_i)\right] = (\sum \pi_i, \sum x_i \pi_i)$, hvor $\pi_i = \dfrac{\exp(\alpha+\beta x_i)}{1+\exp(\alpha+\beta x_i)}$

likelihood ligninger

$$\sum_i y_i = \sum_i \frac{\exp(\alpha+\beta x_i)}{1+\exp(\alpha+\beta x_i)}$$

$$\sum_i x_i y_i = \sum_i x_i \frac{\exp(\alpha+\beta x_i)}{1+\exp(\alpha+\beta x_i)}$$

har entydig løsn. for α og β

medmindre $t(y)$ ligger på randen

af det konvekse hylster af støtten

□

* entydigheden kan udstrækkes til regulære eksponentielle familier med $k>1$.

Betingningsprincippet

Betragt min. suff. stikpr. fkt. $s = (t, a)$ hvor fordelingen af A ikke afh. af θ.

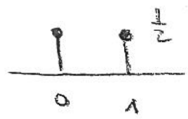
A kaldes en ancillar stikpr.-fkt., den bidrager ikke med information om θ.

Inferens bør baseres på fordelingen af T betinget med $A = a$.

Likelihood : $L_a(\theta) = c(t) f_{T|a}(t;\theta)$

Bemærk $L_a(\theta) \propto L(\theta)$, dvs. L_a bestemmer samme $\hat{\theta}$ som L, men ...

eks. møntkast :



A's ford.

efterfulgt af 10^{2a} obs. i $N(\theta,1)$

$(\bar{Y}, A)$'s ford. :

$$\frac{1}{2} \frac{1}{\sqrt{2\pi}\, 10^{-a}} \exp\left(-\frac{(\bar{y}-\theta)^2}{2 \cdot 10^{-2a}}\right)$$

$(\bar{y}, a)$ er min. suff. for θ

a er ancillar, $\hat{\theta} = \bar{y}$

$$\mathrm{Var}[\bar{Y}\,|\,a] = 10^{-2a} = \begin{cases} \frac{1}{100} & \text{for } a = 1 \\ 1 & \text{for } a = 0 \end{cases}$$

(a posteriori, den observerede varians)

$$\mathrm{Var}[\bar{Y}] = \frac{1}{2}\frac{1}{100} + \frac{1}{2}1 = \frac{101}{200}$$

(a priori, kan benyttes ved sammenl. m. andre eksperimentelle design)

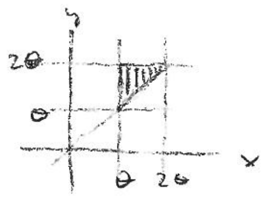
□

eks. $y = (y_1, \dots, y_n)$, $Y_i \sim U(\theta, 2\theta)$ uafh.

Tidl.: $(y_{(1)}, y_{(n)})$ er min. suff. for θ

Simultan tæthed for $(Y_{(1)}, Y_{(n)})$ jf. note fra sandsynlighedsregning om ordensvariable

$$f(x,y) = n(n-1)\, f(x)\, f(y)\, (F(y) - F(x))^{n-2}$$
$$= n(n-1)\, \frac{1}{\theta}\frac{1}{\theta}\left(\frac{y}{\theta} - \frac{x}{\theta}\right)^{n-2}$$
$$= n(n-1)\, \frac{1}{\theta^n}(y-x)^{n-2}, \quad \theta < x < y < 2\theta$$



(kontr.: $\int_\theta^{2\theta}\int_x^{2\theta} n(1-n)\frac{1}{\theta^n}(y-x)^{n-2}\,dy\,dx = 1$)

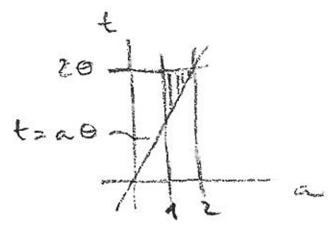
Variabelskift: sæt $T = Y_{(n)}$ og $A = \frac{Y_{(n)}}{Y_{(1)}}$

$x > \theta \Rightarrow t > a\theta$

$y < 2\theta \Rightarrow t < 2\theta$

$x < y \Rightarrow a > 1$

$a < \frac{2\theta}{\theta} = 2$



$$\left.\begin{array}{l} a = \frac{y}{x} \\ t = y \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} x = \frac{t}{a} \\ y = t \end{array}\right. \qquad J(a,t) = \begin{vmatrix} -\frac{t}{a^2} & \frac{1}{a} \\ 0 & 1 \end{vmatrix} = -\frac{t}{a^2}$$

$$g(a,t) = \frac{n(n-1)}{\theta^n}\left(t - \frac{t}{a}\right)^{n-2}\left|-\frac{t}{a^2}\right|$$

$$= \frac{n(n-1)}{\theta^n}\,\frac{(t(a-1))^{n-2}}{a^{n-2}}\,\frac{t}{a^2}$$

$$= \frac{n(n-1)}{\theta^n}\,\frac{t^{n-1}(a-1)^{n-2}}{a^n}, \quad 1<a<2, \quad a\theta<t<2\theta$$

$$g_A(a) = \int_{a\theta}^{2\theta}\frac{n(n-1)}{\theta^n}\,\frac{(a-1)^{n-2}}{a^n}\,t^{n-1}\,dt$$

$$= \frac{n(n-1)}{\theta^n}\,\frac{(a-1)^{n-2}}{a^n}\left[\frac{t^n}{n}\right]_{a\theta}^{2\theta}$$

$$= \frac{(n-1)(a-1)^{n-2}(2^n-a^n)}{a^n}, \quad 1<a<2$$

$$g_{T|a}(t) = \frac{n(n-1)t^{n-1}(a-1)^{n-2}}{\theta^n a^n}\,\frac{a^n}{(n-1)(a-1)^{n-2}(2^n-a^n)}$$

$$= \frac{n\,t^{n-1}}{\theta^n(2^n-a^n)}, \quad a\theta<t<2\theta, \quad a \text{ fast}$$

$$L_a(\theta) \propto \frac{1}{\theta^n}, \quad \frac{t}{2}<\theta<\frac{t}{a}$$

$$\hat\theta = \frac{t}{2} = \frac{y_{(n)}}{2} \quad (\text{som tidl.})$$

$$\mathrm{Var}[\hat\theta] = \frac{1}{4}\mathrm{Var}[Y_{(n)}] = \frac{n\theta^2}{4(n+1)^2(n+2)} \quad *$$

$\mathrm{Var}[\hat\theta|a]$:

$$E[T|a] = \frac{n}{\theta^n(2^n-a^n)}\int_{a\theta}^{\theta} t\cdot t^{n-1}\,dt = \frac{n}{\theta^n(2^n-a^n)}\,\frac{(2\theta)^{n+1}-(a\theta)^{n}}{n+1}$$

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

$$E[T^2|a] = \frac{n}{\theta^n(2^n-a^n)}\int_{a\theta}^{2\theta} t^2\cdot t^{n-1}\,dt = \frac{n}{\theta^n(2^n-a^n)}\,\frac{(2\theta)^{n+2}-(a\theta)^{n+2}}{n+2}$$

$$= \frac{n\theta^2(2^{n+2}-a^{n+2})}{(n+2)(2^n-a^n)}$$

* se udregninger s. 10

$$\mathrm{Var}[T|a] = \frac{n\theta^2(2^{n+2}-a^{n+2})}{(n+2)(2^n-a^n)} - \frac{n^2\theta^2(2^{n+1}-a^{n+1})^2}{(n+1)^2(2^n-a^n)^2}$$

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

$$\mathrm{Var}[\hat\theta|a] = \mathrm{Var}[\hat\theta] \frac{(2^{n+1}-a^{n+1})^2-(n+1)^2(2-a)^2(2a)^n}{(2^n-a^n)^2} \quad **$$

Bemærk, at $\mathrm{Var}[\hat\theta|a] \to 0$ for $a \to 2$ (L'Hospital) ***

Benyttelse af $L_a(\theta)$ giver bedre vurdering af estimatorens 'kvalitet'. □

Observeret information

Under passende regularitetsbetingelser gælder

$\mathrm{Var}[\hat\theta] \to (I(\theta))^{-1}$ for $n \to \infty$, dvs.

$(I(\hat\theta))^{-1}$ kan benyttes som 'kvalitetsmål'.

Alternativt benyttes den observerede information

$$J(\hat\theta) = -l''(\hat\theta) = -\frac{d^2}{d\theta^2} l(\theta)\Big|_{\theta=\hat\theta}$$

Bemærk ved gentagne uafh. obs.

$$-\frac{1}{n}\frac{d^2}{d\theta^2} l(\theta) = \frac{1}{n}\sum_{i=1}^{n}\left(-\frac{d^2}{d\theta^2} l_i(\theta)\right) \to i(\theta) \text{ a.s.}$$

(store tals stærke lov)

Er $I(\hat\theta)$ eller $J(\hat\theta)$ bedst?

- asymptotisk er de ens
- altid ens i reg. eksp. familier

Der kan findes begrundelser for at vælge $J(\hat\theta)$.

** se udregninger side 10 *** se udregninger side 11

eks. $y = (y_1, \dots, y_n)$, $Y_i \sim e(\theta)$ uafh.

censorering ved $y = C$

$$z_i = \begin{cases} 1 & \text{når ing. censorering} \\ 0 & \text{når censorering} \end{cases} \qquad i = 1, \dots, n$$

$$L(\theta; y) = \prod_{i=1}^{n} (\theta e^{-\theta y_i})^{z_i} \left(1 - (1 - e^{-\theta y_i})\right)^{1-z_i}$$

$$= \theta^{\sum_i z_i} e^{-\theta \sum y_i}$$

$$l(\theta; y) = \ln \theta \sum_i z_i - \theta \sum_i y_i$$

$$\left.\begin{aligned} l'(\theta; y) &= \frac{\sum_i z_i}{\theta} - \sum_i y_i = 0 \\ l''(\theta; y) &= -\frac{\sum_i z_i}{\theta^2} < 0 \text{ overalt} \end{aligned}\right\} \Rightarrow \hat{\theta} = \frac{\bar{z}}{\bar{y}}$$

$$J(\hat{\theta}) = \frac{\sum_i z_i}{\hat{\theta}^2} = \frac{n \bar{y}^2}{\bar{z}}$$

$$I(\hat{\theta}) = \left.\frac{E[\sum_i Z_i]}{\theta^2}\right|_{\theta = \hat{\theta}} = \left.\frac{n P(Y_i \le C)}{\theta^2}\right|_{\theta = \hat{\theta}}$$

$$= \frac{n(1 - \exp(-C\hat{\theta}))}{\hat{\theta}^2} = \frac{n \bar{y}^2 (1 - \exp(-\frac{C\bar{z}}{\bar{y}}))}{\bar{z}^2}$$

$\square$

eks. Halds nittedata

klassedeling af obs. nr. z_i som klassemidtpkt., $i = 1, \dots, k$

intuitivt: $\tilde{\mu} = \dfrac{\sum_i f_i z_i}{n} = 13{,}4264$

$$\tilde{\sigma}^2 = \frac{\sum_i f_i (z_i - \tilde{\mu})^2}{n} = 0{,}013149$$

(f_i er hyppigheden i klassen)

Alternativt :

$$L(\theta) = \prod_i (p_i(\theta))^{f_i} \qquad \text{(multinomialford.)}$$

med $p_i(\theta) = \Phi\left(\frac{z_i + \frac{h}{2} - \mu}{\sigma}\right) - \Phi\left(\frac{z_i - \frac{h}{2} - \mu}{\sigma}\right)$

$i = 1, \ldots, k$

numerisk maksimering af $L(\theta)$ giver

$$\hat{\mu} = 13{,}4264$$

$$\hat{\sigma}^2 = 0{,}012941 \approx \tilde{\sigma}^2 - \frac{h^2}{12}$$

(Sheppards korr.) $\square$

eks. antal albinoer blandt børn i 60 familier med 5 børn, heraf mindst én albino

antal albinoer	1	2	3	4	5
hyppighed	25	23	10	1	1

$Y_i \sim$ trunkeret binomial $(y_i > 0)$

$$P(Y_i = k) = \frac{\binom{m}{k}\pi^k (1-\pi)^{m-k}}{1-(1-\pi)^m}, \quad k = 1, \ldots, m$$

(her $m = 5$)

$$\ell(\pi) = c + \sum_{k=1}^{m} f_k \left(k \ln \pi + (m-k)\ln(1-\pi) - \ln(1-(1-\pi)^m)\right)$$

$$= c + n\left(\bar{y} \ln \pi + (m - \bar{y}) \ln(1-\pi) - \ln(1-(1-\pi)^m)\right)$$

$$\left(n = \sum_i f_i = 60, \quad \bar{y} = \frac{1}{n}\sum_i f_i k = 1{,}8335\right)$$

$$\ell'(\pi) = n\left(\frac{\bar{y}}{\pi} - \frac{m-\bar{y}}{1-\pi} - \frac{m(1-\pi)^{m-1}}{1-(1-\pi)^m}\right) = 0$$

$$\frac{\bar{y}(1-\pi+\pi)}{\pi(1-\pi)} = \frac{m}{1-\pi} + \frac{m(1-\pi)^{m-1}}{1-(1-\pi)^m}$$

$$\frac{\bar{y}}{m} = \frac{\pi\left(1-(1-\pi)^m+(1-\pi)^m\right)}{1-(1-\pi)^m}$$

$$\hat{\pi} = \frac{\bar{y}}{m}\left(1-(1-\hat{\pi})^m\right), \text{ som}$$

kan løses ved gentagne substitutioner,

startværdi $\pi = \frac{\bar{y}}{m}$, idet

$$E[Y_i] = \frac{m\pi}{1-(1-\pi)^m} \simeq m\pi,$$

efter 10 iterationer : $\hat{\pi} = 0{,}3048$

information

$$l'(\pi) = n\left(\frac{\bar{y}}{\pi} + \frac{\bar{y}}{1-\pi} + c\right)$$

$$\begin{aligned} I(\pi) &= \mathrm{Var}[l'(\pi)] \\ &= n^2\,\mathrm{Var}[\bar{Y}]\left(\frac{1}{\pi}+\frac{1}{1-\pi}\right)^2 \\ &= n\,\mathrm{Var}[Y_i]\left(\frac{1}{\pi(1-\pi)}\right)^2, \end{aligned}$$

$$\mathrm{Var}[Y_i] = \frac{m\pi(1-\pi)+(m\pi)^2}{1-(1-\pi)^m} - \left(\frac{m\pi}{1-(1-\pi)^m}\right)^2$$

$$I(\hat{\pi}) = 972{,}59$$

estimat for $\mathrm{Var}[\hat{\pi}] = \frac{1}{I(\hat{\pi})} = 0{,}0010304$ □

eks.

blodtype (fænotype)	A		B		AB	0
genotype	AA	AO	BB	BO	AB	OO
sandsynligh.	p^2	$2pr$	q^2	$2qr$	$2pq$	r^2
antal obs.	n_A		n_B		n_{AB}	n_O

$$p+q+r=1 \Rightarrow r=1-p-q$$

$$p+q<1$$

$$l(p,q) = n_A \ln(p^2+2pr) + n_B \ln(q^2+2qr) + n_{AB}\ln(2pq) + 2n_0 \ln r$$

niveaukurver for $l(p,q)$, se bog s. 99

Newton-Raphson beregning i bog s. 98-100 (jf. s. 63)

Udregninger fra side 5

* $$f_{Y_{(n)}}(y) = \int_\theta^y n(n-1)\frac{1}{\theta^n}(y-x)^{n-2}\,dx = n(n-1)\frac{1}{\theta^n}\left[-\frac{(y-x)^{n-1}}{n-1}\right]_\theta^y$$

$$= n\frac{1}{\theta}\left(\frac{y-\theta}{\theta}\right)^{n-1}, \quad \theta<y<2\theta$$

$$1-U = \frac{1}{\theta}(Y_{(n)}-\theta), \quad 1-u = \frac{y-\theta}{\theta} \Leftrightarrow y = \theta(1-u)+\theta, \quad \frac{dy}{du} = -\theta$$

$$f_U(u) = n\frac{1}{\theta}(1-u)^{n-1}|-\theta| = n(1-u)^{n-1}$$

$$= \frac{\Gamma(1+n)}{\Gamma(1)\Gamma(n)} u^{1-1}(1-u)^{n-1}, \quad 0<u<1$$

dvs. $U \sim B(1,n)$

$$\mathrm{Var}\,U = \frac{1\cdot n}{(1+n)^2(1+n+1)} = \frac{n}{(n+1)^2(n+2)}$$

$$\mathrm{Var}\,Y_{(n)} = (-\theta)^2\,\mathrm{Var}\,U = \frac{n\theta^2}{(n+1)^2(n+2)}$$

Udregninger fra side 6

** $$(n+1)^2(2^n-a^n)(2^{n+2}-a^{n+2}) - n(n+2)(2^{n+1}-a^{n+1})^2$$

$$= (n+1)^2(2^n-a^n)(2^{n+2}-a^{n+2}) - (n(n+2)+1)(2^{n+1}-a^{n+1})^2 + (2^{n+1}-a^{n+1})^2$$

$$= (2^{n+1}-a^{n+1})^2 - (n+1)^2\left((2a)^n a^2 + (2a)^n 4 - 2(2a)^{n+1}\right)$$

$$= (2^{n+1}-a^{n+1})^2 - (n+1)^2\left((2a)^n a^2 + (2a)^n 4 - (2a)^n 4a\right)$$

$$= (2^{n+1}-a^{n+1})^2 - (n+1)^2(2-a)^2(2a)^n$$

fortsættes

Stat. 5.3. 2010

***

$$\lim_{a \to 2} \frac{(2^{n+1} - a^{n+1})^2 - (n+1)^2 (2-a)^2 (2a)^2}{(2^n - a^n)^2}$$

$$= \lim_{a \to 2} \frac{2(2^{n+1} - a^{n+1})(-(n+1)a^n) + 2(n+1)^2(2-a)(2a)^n - (n+1)^2(2-a)^2 2^n n a^{n-1}}{2(2^n - a^n)(-n a^{n-1})}$$

$$= \lim_{a \to 2} \frac{-(n+1)(2^{n+1} - a^{n+1})a + 2^n (n+1)^2 (2-a) a - 2^{n-1} n (n+1)^2 (2-a)^2}{-n(2^n - a^n)}$$

$$= \lim_{a \to 2} \frac{(n+1)^2 a^n - (n+1)(2^{n+1} - a^{n+1}) + 2^{n-1}(n+1)^2(-a+2-a) + 2^n n (n+1)^2 (2-a)}{n^2 a^{n-1}}$$

$$= \frac{(n+1)^2 2^n - 0 - 2^n (n+1)^2 + 0}{n^2 2^{n-1}} = \frac{2(n+1)^2 - 2(n+1)^2}{n^2} = 0$$