Stikprøve $y = (y_1, y_2, \ldots, y_n)$

y_i er realiseret værdi af en stok. var.

Parametrisk (statistisk) model

$$\mathcal{F} = \{ F(\cdot\,;\theta) \mid \theta \in \Theta \} \quad , \quad \Theta \subseteq R^k ,$$

men da i alm. kont. eller diskret ford., så i stedet

$$\mathcal{F} = \{ f(\cdot\,;\theta) \mid \theta \in \Theta \} \quad , \quad \Theta \subseteq R^k$$

θ : k-dim. parameter

Θ : parameterrummet

Det antages, at der findes en sand værdi, θ_* ; dog i alm. ukendt.

Ikke-parametriske modeller behandles ikke!

Udfaldsrum

$$\mathcal{Y} = \bigcup_{\theta \in \Theta} \mathcal{Y}_\theta \quad , \quad \mathcal{Y}_\theta \text{ støtten for } f(\cdot\,;\theta)$$

eks. $Y_1, Y_2 \sim N(\theta, 1)$ uafh., $y = (y_1, y_2)$

$$f(y;\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_1-\theta)^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_2-\theta)^2}$$

$$\mathcal{Y} = R \times R, \quad \Theta = R \qquad \square$$

Sted- og skalaparameter

Lad Y have tæth. $g(\cdot)$, $(\theta_1, \theta_2) \in R \times R_+$

Sæt $U = \theta_1 + \theta_2 Y$, $u = \theta_1 + \theta_2 y \Leftrightarrow y = \frac{u-\theta_1}{\theta_2}$

U's tæth.: $g\left(\frac{u-\theta_1}{\theta_2}\right)\left|\frac{1-0}{\theta_2}\right| = \frac{1}{\theta_2} g\left(\frac{u-\theta_1}{\theta_2}\right)$

θ_1 kaldes stedparameter

θ_2 kaldes skalaparameter

eks.

$$g(y;\omega,\lambda) = \begin{cases} \lambda e^{-\lambda(y-\omega)} & \text{for } y > \omega \\ 0 & \text{ellers} \end{cases}$$

$\theta = (\omega, \lambda) \in R \times R_+$

ω er stedparameter

$\frac{1}{\lambda}$ er skalaparameter (λ kaldes ofte intensiteten

bemærk $\mathcal{Y}_\theta = \{y \mid y > \omega\} \Rightarrow \mathcal{Y} = R \qquad \square$

Reparametrisering

$h: \Theta \to \Psi$, h entydig

$\mathcal{F} = \{f(\cdot;\psi) \mid \psi = h(\theta), \theta \in \Theta\}$

$= \{f(\cdot;\psi) \mid \psi \in \Psi\}$, $\Psi = \{\psi \mid \psi = h(\theta), \theta \in \Theta\}$

inferens uafh. af parametrisering
kan ikke altid opnås

eks $Y \sim e(\lambda)$, $f(y) = \lambda e^{-\lambda y}$, $y>0$; $\lambda>0$

$$f(y) = \frac{1}{(\frac{1}{\lambda})} e^{-\frac{1}{(\frac{1}{\lambda})} y} = \frac{1}{\tau} e^{-\frac{y}{\tau}} \quad , y>0 ; \tau>0$$

Når Y er tiden mellem ankomster i en Poissonproces, så er

- λ det gennemsnitlige antal ankomster pr. tidsenhed
- τ den gennemsnitlige tid mellem ankomster.

Identificerbarhed

En model er identificerbar, når der

$$\exists A \subseteq \mathcal{Y} : \quad \theta_1 \neq \theta_2 \Rightarrow P(A;\theta_1) \neq P(A;\theta_2)$$

(I modsat fald kan der ikke skelnes mellem θ_1 og θ_2, og inferens vil være flertydig.)

Likelihood

Med observationen y givet afhænger $f(y;\theta)$ kun af θ

To forskellige θ-værdier kan vurderes ved at betragte kvotienten

$$\frac{f(y;\theta_1)}{f(y;\theta_2)} = \frac{c\, f(y;\theta_1)}{c\, f(y;\theta_2)} \quad , c = c(y) \neq 0$$

Likelihood funktionen

$$L : \Theta \to R_0 \quad , L(\theta;y) = c(y)\, f(y;\theta)$$

egentlig en ækvivalensklasse af funktioner

Log-likelihood funktionen

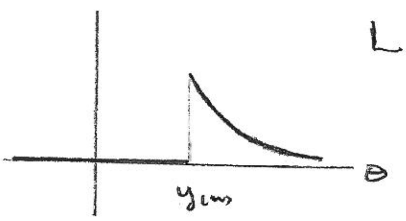
$$\ell(\theta;y) = \ln L(\theta;y) = c_1(y) + \ln f(y;\theta)$$

$$(\ \ell(\theta;y) = -\infty \quad \text{for} \quad L(\theta;y) = 0\ )$$

eks. $y = (y_1, \dots, y_n)$, $Y_i \sim U[0;\theta]$ uafh.

$f(y_i) = \frac{1}{\theta} I_{(0;\theta)}(y_i)$ ($I.$ er indikator fkt.)

bemærk $I_{(0;\theta)}(y_i) = I_{(0;1)}\left(\frac{y_i}{\theta}\right) = I_{(1;\infty)}\left(\frac{\theta}{y_i}\right)$



$$L(\theta;y) = \prod_{i=1}^{n} \frac{1}{\theta} I_{(0;\theta)}(y_i) = \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(1;\infty)}\left(\frac{\theta}{y_i}\right)$$

$$= \frac{1}{\theta^n} I_{(1;\infty)}\left(\frac{\theta}{y_{(n)}}\right) = \frac{1}{\theta^n} I_{(y_{(n)};\infty)}(\theta)$$

$$\ell(\theta;y) = \begin{cases} -\infty & \text{for } \theta < y_{(n)} \\ -n\ln\theta & \text{for } \theta > y_{(n)} \end{cases} \qquad \square$$

eks. Markovkæde med $Y_1 \sim b(1,\pi)$ og med

overgangsmatrix $\begin{bmatrix} 1-\pi_0 & \pi_0 \\ 1-\pi_1 & \pi_1 \end{bmatrix}$

$y = (y_1, \dots, y_n)$, $y_i \in \{0,1\}$

$$L(\pi, \pi_0, \pi_1; y)$$

$$= \pi^{y_1}(1-\pi)^{1-y_1} \prod_{j=2}^{n} (\pi_{y_{j-1}})^{y_j} (1-\pi_{y_{j-1}})^{1-y_j}$$

Likelihood princippet

stat. model : $\{ f(\cdot;\theta) \mid \theta \in \Theta \}$; $y, z \in \mathcal{Y}$

Når $L(\theta;y) \propto L(\theta;z)$ skal begge

obs. føre til samme inferens om θ

Det stærke likelihood princip

$\{f(\cdot\,;\theta) \mid \theta \in \Theta\}$ stat. model for $y \in \mathcal{Y}$

$\{g(\cdot\,;\theta) \mid \theta \in \Theta\}$ stat. model for $z \in \mathcal{Z}$

Når $L_f(\theta;y) \propto L_g(\theta;z)$ skal begge obs. føre til samme inferens om θ

Principperne – især det stærke – kan ikke altid opfyldes.

eks. $Y \sim b(n,\theta)$, $f(y;\theta) = \binom{n}{y}\theta^y(1-\theta)^z$, $z = n-y$

$y = 0, 1, \ldots, n$

$Z \sim nb(y,\theta)$, $g(z;\theta) = \binom{y+z-1}{z}\theta^y(1-\theta)^z$,

$z = 0, 1, \ldots$

Bemærk $L_f(\theta;y) \propto L_g(\theta;z) \propto \theta^y(1-\theta)^z$ □

Stikprøvefunktion (– variabel)

$T: \mathcal{Y} \to R^r$, hvor $T(y)$ ikke afhænger af θ

eks: $T_1 = \Sigma y_i \in R$, $T_2 = (\Sigma y_i, \Sigma y_i^2) \in R^2$ □

bemærk $A_t = \{y \in \mathcal{Y} \mid T(y) = t\}$
klassedeling af udfaldsrummet

eks. $T(y) = \Sigma y_i$
A_t'erne er parallelle
hyperplaner $y_1 + y_2 + \ldots + y_n = t$ □

Sufficiens

I en stat. model er T sufficient for θ, hvis $\forall\ y, z \in \mathcal{Y}$:

$$T(y) = T(z) \Rightarrow L(\theta; y) \propto L(\theta; z) \text{ for alle } \theta \in \Theta$$

dvs. $L(\theta; y)$ afhænger kun af y gennem $T(y)$

eks. Y diskret var., $y = 0, 1, 2$; $\theta \in \{0, 1\}$

	$P(Y=0)$	$P(Y=1)$	$P(Y=2)$
$\theta = 0$	$\frac{8}{12}$	$\frac{1}{12}$	$\frac{3}{12}$
$\theta = 1$	$\frac{4}{12}$	$\frac{2}{12}$	$\frac{6}{12}$

bemærk $L(1; \theta) \propto L(2; \theta)$

$T(y) = I_{\{0\}}(y)$ er sufficient for θ,

idet $T(y) = \begin{cases} 1 & \text{for } y = 0 \\ 0 & \text{for } y = 1, 2 \end{cases}$ $\square$

$T(\cdot)$ sufficient for $\theta \Rightarrow \exists\, g : L(\theta; y) \propto g(T(y); \theta)$

men også $L(\theta; y) \propto f(y; \theta)$, derfor er

$$\frac{f(y; \theta)}{g(T(y); \theta)} \text{ uafhængig af } \theta$$

$\Rightarrow f(y; \theta) = h(y)\, g(T(y); \theta)$; bemærk også $\Leftarrow$

Neymanns faktorisering

T sufficient for θ

$\Leftrightarrow f(y; \theta) = h(y)\, g(T(y); \theta)$

Bemærk $f_T(t;\theta) = \int_{A_t} f(y;\theta)\, d\nu(y)$

$$= g(t;\theta) \int_{A_t} h(y)\, d\nu(y)$$

$$= g(t;\theta)\, h^*(t)$$

Heraf ses, at $L_T(\theta;t) \propto L(\theta;y)$

Sætning: T sufficient for θ

$\Leftrightarrow$ ford. af $Y \mid T(y)=t$ afh. ikke af θ

$\Rightarrow$: $f(y;\theta \mid T=t) = \dfrac{f(y;\theta)}{f_T(t;\theta)} = \dfrac{h(y)\,\cancel{g(T(y);\theta)}}{h^*(t)\,\cancel{g(t;\theta)}}$,

som ikke afhænger af θ.

$\Leftarrow$: $f(y;\theta) = f(y \mid T=t)\, f_T(t;\theta)$

$$= h_*(y,t)\, h^*(t)\, g(t;\theta)$$

som svarer til Neymans faktorisering

To-trins eksperiment

(1) vælg t, dvs. vælg $A_t = \{y \mid T(y)=t\}$ med ford. $g(\cdot;\theta)$, som afh. af θ

(2) vælg $y \in A_t$, ford afh. ikke af θ

Bemærk, at $U \circ T(\cdot)$, hvor U er entydig, også er sufficient for θ, idet $T(y)=t$ og $U \circ T(y) = U(t)$ bevirker samme klassedeling af $\mathcal{Y}$.

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

bemærk $I_{(\theta;2\theta)}(y_i) = \ldots = I_{(\frac{1}{2};1)}\left(\frac{\theta}{y_i}\right)$

$L(\theta) = \prod_{i=1}^{n} \frac{1}{\theta} I_{(\theta;2\theta)}(y_i) = \frac{1}{\theta^n} I_{(\frac{1}{2}y_{(n)};\, y_{(1)})}(\theta)$,

idet $\frac{1}{2} < \frac{\theta}{y_i}$ for alle i $\Leftrightarrow$ $\frac{1}{2} y_{(n)} < \theta$

og $\frac{\theta}{y_i} < 1$ for alle i $\Leftrightarrow$ $\theta < y_{(1)}$

$(y_{(1)}, y_{(n)})$ er altså sufficient for θ $\square$