$$D(\theta; y) = -2 \sum_{i=1}^{n} \log p_i(y_i | \theta_i)$$

$$p(y_{new} | \theta, y) = p(y_{new} | \theta)$$

$$= \prod_{i=1}^{n} p_i(y_{new,i} | \theta)$$

---

evidence před odhadem

$$p(y) = \int_{\Theta} p(y|\theta)\, p(\theta)\, d\theta$$

evidence po použití dat pro odhad

$$p(y_{new} | y) = \int_{\Theta} p(y_{new} | \theta)\, p(\theta | y)\, d\theta$$

$$\bar{D}_{pred} = E_{p(\theta|y)}(-2 \log p(y_{new} | \theta))$$

$$\overline{D}_{pred} = \sum_{i=1}^{n} E_{p(\theta|y)} D_i(\theta_i; y_{new,i})$$

$$\overline{D}^{CV}_{pred} = \sum_{i=1}^{n} \underbrace{E_{p(\theta|y_{(-i)})} D_i(\theta_i; y_i)}_{\overline{D}^{CV}_{pred,i}}$$

Snadno:

$$\overline{D} = \sum_{i=1}^{n} \underbrace{E_{p(\theta|y)} D_i(\theta_i; y_i)}_{\overline{D}_i}$$

$$p_{opt,i} = E\left(\overline{D}^{CV}_{pred,i} - \overline{D}_i \mid Y_{-i}\right)$$

$$PED = \sum_{i=1}^{n} \overline{D}_i + \sum_{i=1}^{n} p_{opt,i}$$

$$\approx \overline{D}^{CV}_{pred}$$

penalized expected deviance

→ souvisí s DIC

deviance information crit.

NĚKDY (např. exponenciální rodina):

$$p_{opt} = \sum_{i=1}^{n} p_{opt,i} \approx \overline{D} - D(\hat{\theta}(y); y)$$

$E(\theta|y)$

$\eta = t(\theta)$ $\quad E(\eta|y) \neq t(E(\theta|y))$

$PED = \bar{D} + p_{opt}$

$(AIC = -2\log \hat{w} + 2p_{\hat{w}}$

Další problém

Co je $p(y|\theta)$?

---

ad hierarchický model

např. LMM

$\theta = (\beta, \sigma^2, \alpha^2)$

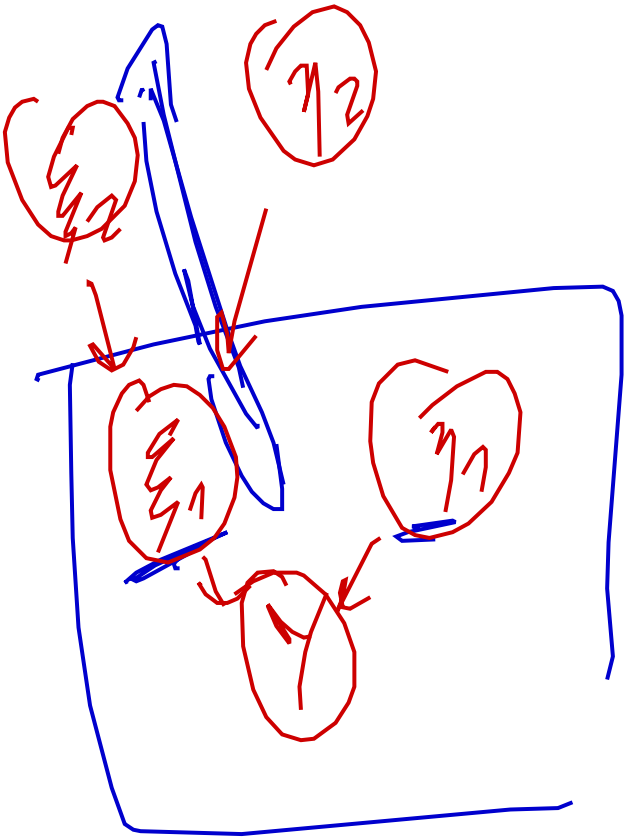
$p(y|\theta) = \int p(y|b,\theta)\, p(b|\theta)\, db$

pokud použito BDA, pak

"θ" $= (\theta, b)$

a "vedlejší":

$p(y|\theta, b)$



---

cenzor. zprava

$T_1, \dots, T_k$: žádá pozor. časy

$C_{k+1}, \dots, C_n$: cenzor. časy

Φ $T_j > C_j$, $j = k+1, \dots, n$

$P(T_j > C_j; \theta)$

věrohod. k

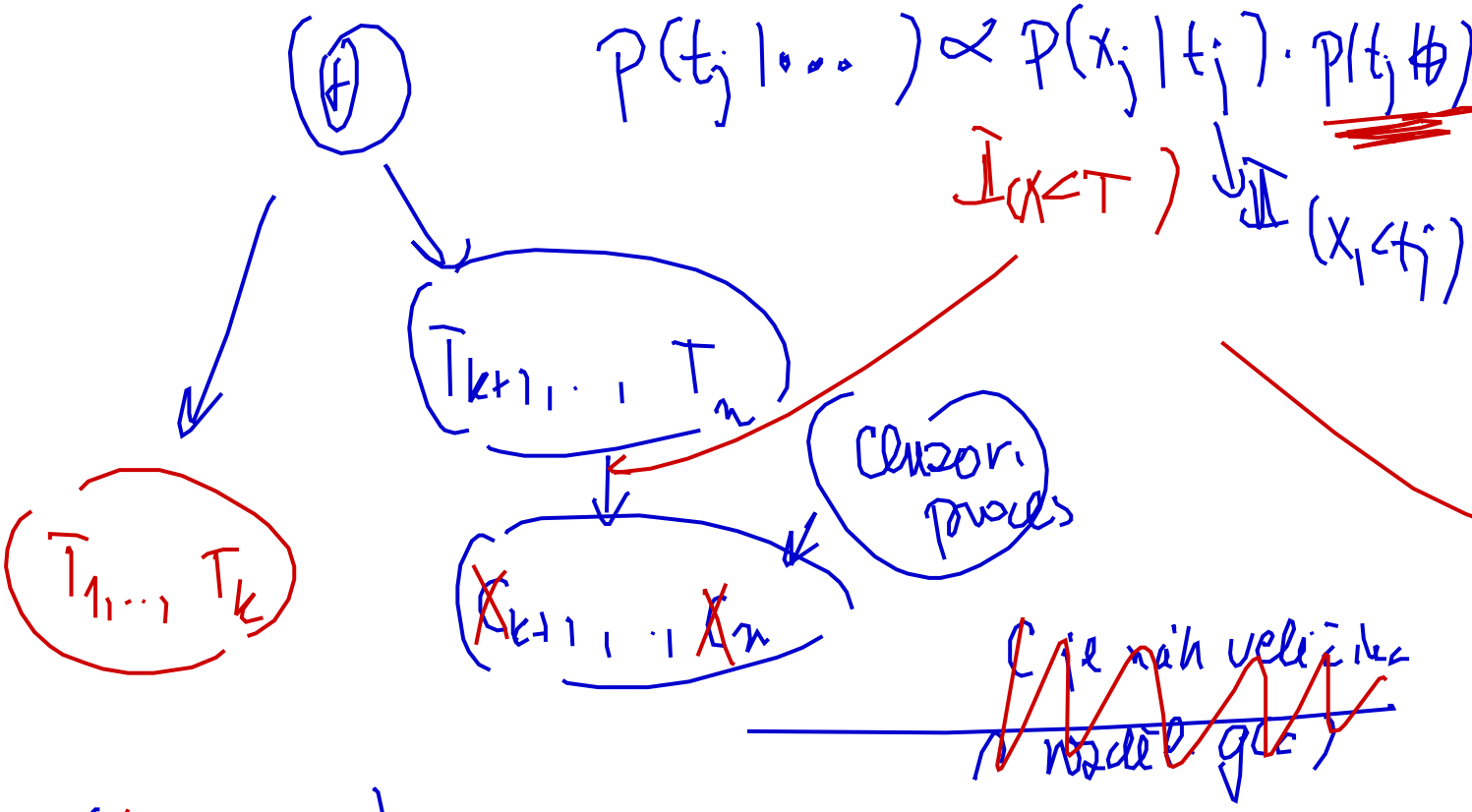
$$p(y|\theta) = \prod_{i=1}^{k} f(t_i;\theta) \prod_{i=k+1}^{n} (1 - F(c_i;\theta))$$

$$\int_{c_i}^{\infty} f(s;\theta)\,ds$$

θ

cenzor. proces

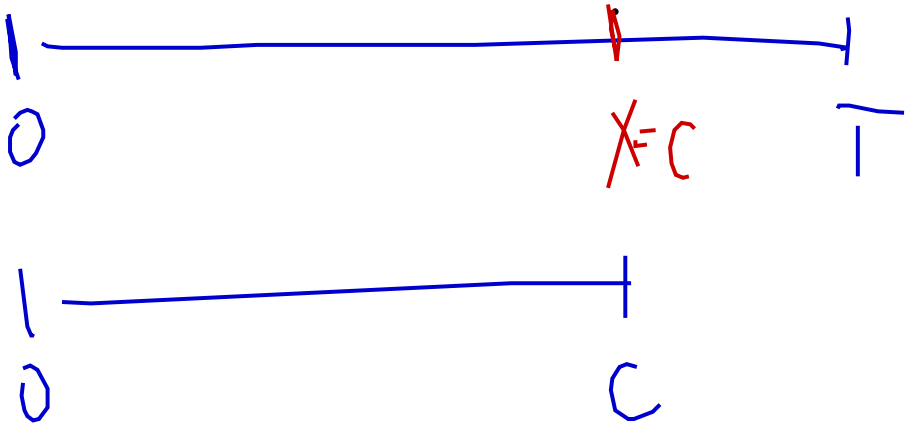
$T_1, \dots, T_k$ $C_{k+1}, \dots, C_n$



$$p(C_{k+1}, \dots, C_n \mid T_{k+1}, \dots, T_n, \text{cenzor proces})$$

$$= ? = \prod_{i=k+1}^{n} p(C_i \mid T_1, \text{cenzor proces})$$

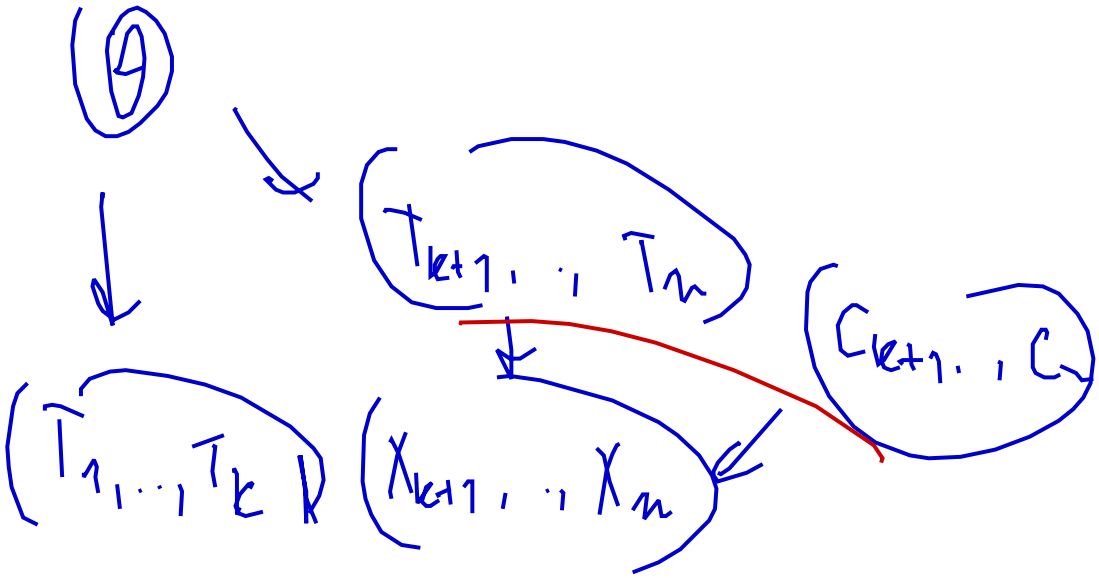
$$X_i = \min(T_i, C_i)$$



původní věrohodnost:

$$p(y|\theta) = \prod_i f(t_i;\theta) \prod_j (1 - F(x_j;\theta)) =$$

$$= \prod_i f(t_i;\theta) \prod_j \int_{x_j}^{\infty} f(s;\theta)\,ds$$



nová věrohodnost

$$p(y \mid \theta, T_{k+1}, \ldots, T_n)$$

---

pokračujeme:

$$p(X_i \mid T_i, C_i) = ?$$

X_i ... cenzorovaný čas

$$\equiv \prod (T_i > X_i)$$