Satistical modelling of clinical trials

Sept. 21st, 2012

Patients' recruitment in a multicentric clinical trial :
Recruit N_f patients via C centres. Typically, $N_f \sim 300$ and $C \sim 50$.

- Trial can be very long (several years) and very expensive
- Previous models based on linear or exponential interpolation

$\Longrightarrow$Need for a probabilistic model that takes into account large variability of recruitment rates in different centres

Centre i opens at timeu_i.
Γ-model proposed by V. Anisimov (2; 1) :

- Recruitment processes of centres are independent Poisson processes
- Recruitment rate of centre i : λ_i (unknown)
- Γ-Poisson model : $(\lambda_1, \ldots, \lambda_C)$ iid with Gamma distribution of parameters α and β.

$$p_{\alpha,\beta}(x) = e^{-\beta x} x^{\alpha-1} \beta^{\alpha} \Gamma(\alpha)^{-1}$$

On-going study at time t_1. Observed data is the number of patients recruited by each of C centres : $(k_1, \dots, k_C)$. Independent but not identically distributed since $\lambda_i \overset{\mathcal{L}}{\sim} \Gamma(\alpha, \beta)$ and $k_i \overset{\mathcal{L}}{\sim} \mathcal{P}(\lambda_i \tau_i)$.
Let $\tau_i = t_1 - u_i$ (0 if negative). Estimation of (α, β) maximum likelihood technique :

$$\mathbb{P}\left[N_{t_1}^i = k_i; 1 \leq i \leq C\right] = \mathbb{E}\left[\prod_{i=1}^{C} \frac{(\lambda_i \tau_i)^{k_i}}{k_i!} e^{-\tau_i \lambda_i}\right] = \prod_{i=1}^{C} \frac{\Gamma(\alpha + k_i)}{k_i! \Gamma(\alpha)} \frac{\beta^\alpha \tau_i^{k_i}}{(\beta + \tau_i)^{\alpha + k_i}}$$

log-likelihood (up to some constant) :

$$\mathcal{L}_C(\alpha, \beta) = \alpha \ln \beta - \ln \Gamma(\alpha) + \frac{1}{C} \sum_{i=1}^{C} \ln \Gamma(\alpha + k_i) - (\alpha + k_i) \ln(\beta + \tau_i)$$

$$(\hat{\alpha}_C, \hat{\beta}_C) = \arg\max_{(\alpha,\beta) \in \Theta} \mathcal{L}_C(\alpha, \beta)$$

As $C \to +\infty$, the MLE is consistent and asymptotically normal.
Approximated Fischer information matrix : if
$f(\theta, k_i, \tau_i) = \alpha \ln \beta - \ln \Gamma(\alpha) + \ln \Gamma(\alpha + k_i) - (\alpha + k_i) \ln(\beta + \tau_i)$, then
$(\theta = (\alpha, \beta))$:

$$I(\theta_0) \simeq \frac{1}{C} \sum_{i=1}^{C} I_i(\theta_0) = -\frac{1}{C} \sum_{i=1}^{C} \mathbb{E}_{\theta_0} \left[\nabla^2 f(\theta_0, k_i, \tau_i) \right].$$

and

$$\sqrt{C}(\hat{\theta}_C - \theta_0) \overset{\mathcal{L}}{\simeq} \mathcal{N}(0, I(\theta_0)^{-1})$$

$$(I_i)_{11} = -\mathbb{E}_{\alpha_0, \beta_0} \left[\psi^{(1)}(\alpha_0 + k_i) \right] + \psi^{(1)}(\alpha_0),$$

$$(I_i)_{12} = (I_i)_{21} = -\frac{1}{\beta_0^2} \left(1 - \frac{1}{1 + \tau_i / \beta_0} \right),$$

$$(I_i)_{22} = \frac{\alpha_0}{\beta_0^2} \left(1 - \frac{1}{1 + \tau_i / \beta_0} \right),$$

Bayesian re-estimation of the distribution of λ_i :
'forward distribution' of λ_i conditionnaly of the information at t_1 :
$\Gamma(\alpha + k_i, \beta + \tau_i)$
Overall recruitment rate : $\Lambda = \sum_{i=1}^{C} \lambda_i = \sum_{i=1}^{C} \Gamma(\alpha + k_i, \beta + \tau_i)$,
that can be approximated by a $\Gamma(A, B)$ distribution by matching moments.
In this framework, the (remaining) recruitment time $\hat{T} - t_1$ has density :

$$p_T(x) = \frac{\Gamma(A + N_1)}{\Gamma(A)\Gamma(N_1)} B^A \frac{x^{N_1 - 1}}{(x + B)^{N_1 + A}}$$

Then it is easy to get :

- $\mathbb{P}\left[\hat{T} \leq T_f\right]$
- $\mathbb{E}\left[\hat{T}\right]$

If $\mathbb{P}\left[\hat{T} \leq T_f\right]$ is too small (say lower than $p = 95\%$), we open M centres.

$$\Lambda = \sum_{i=1}^{C} \lambda_i + \sum_{i=1}^{M} \lambda_i'$$

- λ_i has the 'forward distribution' $\Gamma(\alpha_0 + k_i, \beta_0 + \tau_i)$
- λ_i' has a $\Gamma(\alpha_0, \beta_0)$ distribution

Here, centres are supposed to open instantaneously, but it is possible to assume they open later in some interval $[r_i', s_i']$.

If $\mathbb{P}\left[\hat{T} \leq T_f\right]$ is too high, one can close centres to save money.

Sensitivity parameters (see (3)) : $\mathbb{P}\left[\hat{T} \leq T_f\right]$ and $\mathbb{E}\left[\hat{T}\right]$ are calculated with $(\hat{\alpha}_C, \hat{\beta}_C)$ estimated at t_1, instead of real parameters (α_0, β_0). The subsequent error is evaluated thanks to sensitivity parameters (e.g. $\partial_\alpha \mathbb{P}\left[\hat{T} \leq T_f\right]$).
When the overall rate is approximated by a Gamma distribution :

$$m = \sum_{i=1}^{C} \frac{\alpha + k_i}{\beta + \tau_i} \qquad \text{et} \qquad v = \sum_{i=1}^{C} \frac{\alpha + k_i}{(\beta + \tau_i)^2}$$

althenors

$$\partial_\alpha \mathbb{E}\left[\hat{T}\right] \approx -N \frac{\partial_\alpha (m - \frac{v}{m})}{(m - \frac{v}{m})^2},$$
$$\partial_\beta \mathbb{E}\left[\hat{T}\right] \approx -N \frac{\partial_\beta (m - \frac{v}{m})}{(m - \frac{v}{m})^2}.$$

Extension of the model

- Pareto distribution instead of Gamma distribution :
  $p_{k_p,x_m}(x) = k_p x_m^{k_p} \mathbf{1}_{x \geq x_m} x^{-1-k_p}$
- opening times u_i of centres unkown, but we know the time of first recruitment v_i. Then we assume u_i uniformly distributed in $[0, v_i]$.

Real data (centres' opening times unkown) :
$N = 610$, $C = 77$, $T_f = 3$ years
Study actually finished in 2.31 years.

t_1	1	1.5	2
Pareto - parameters	(1.19, 1.39)	(1.23, 1.30)	(1.18, 1.22)
Pareto - $\mathbb{E}[\hat{T}]$	2.79	2.30	2.35
Gamma - parameters	(1.17, 0.25)	(1.08, 0.26)	(1.31, 0.33)
Gamma - $\mathbb{E}[\hat{T}]$	2.81	2.34	2.36

At $t_1 = 1$, in the Γ-Poisson model, we get $\mathbb{P}\left[\hat{T} \leq T_f\right] \geq 0.999$. Closing the smallest 10 centres steal leads to $\mathbb{P}\left[\hat{T} \leq T_f\right] \sim 0.999$.

Model validation : ν_i = number of centres having recruited exactly i patients

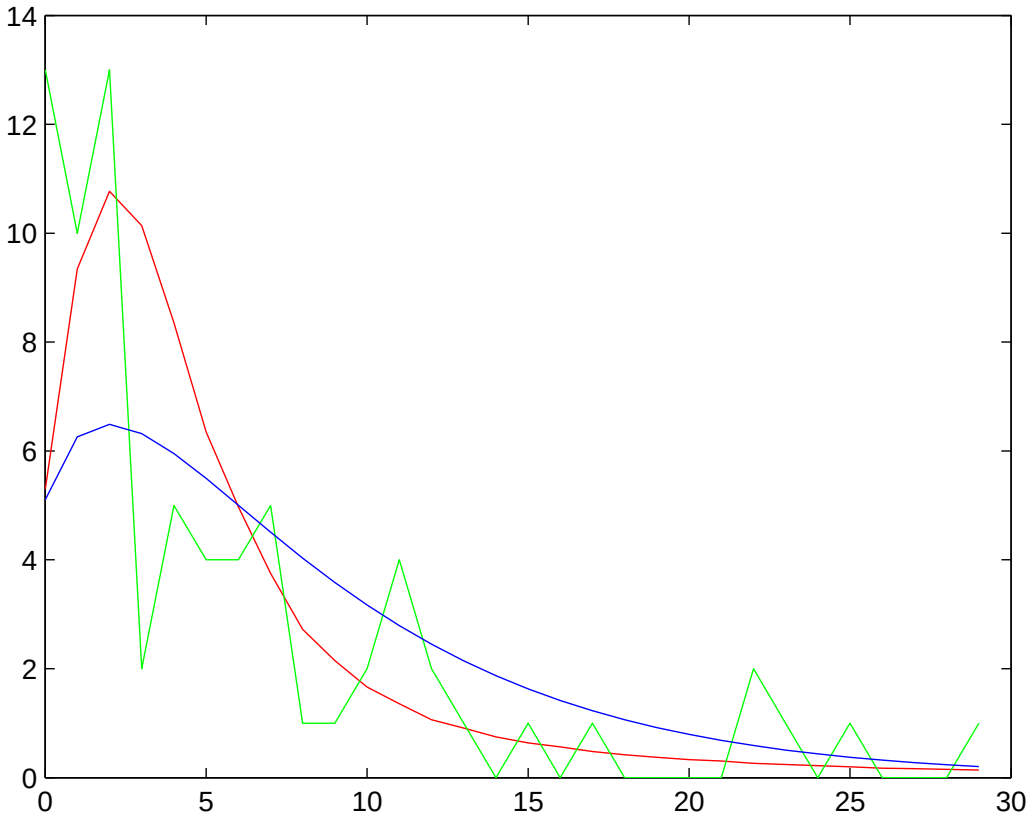

FIGURE: Green : real data ; Bleu : Gamma model ; Rouge : Pareto model

Further research :

- simultaneaous modelling of screened, randomized and lost patients
- cost modelling
- exogenous variables ?

[1] Vladimir V. Anisimov, *Using mixed poisson models in patient recruit in multicentre clinical trials*, Proceegings of th World Congress on Ingineering (London, United Kingdom), vol. II, 2008.

[2] Vladimir V. Anisimov and Valerii V. Fedorov, *Modelling, prediction and adaptive adjustment of recruitment in multicentre trials*, Stat. Med. **26** (2007), no. 27, 4958–4975. MR MR2405491

[3] Guillaume Mijoule, Stéphanie Savy, and Nicolas Savy, *Models for patients' recruitment in clinical trials and sensitivity analysis*, Statistics in Medicine **31** (2012), no. 16, 1655–1674.