Population filter inference: A novel nonlinear mixed effects inference approach for snapshot time series data

David Augustin[1], Ben Lambert[2], Ken Wang[3], Antje-Christine Walz[3], Martin Robinson[1], David Gavaghan[1]

1 University of Oxford, 2 University of Exeter, 3 Hoffmann-La Roche

Take home:
Population filter inference makes intractable tractable – it enables inference of nonlinear mixed effects models from snapshot time series data.

What is nonlinear mixed effects (NLME) modelling?

NLME modelling is an approach to model the dynamics of heterogeneous populations. NLME models have 3 building blocks, see example below.

Example: Early cancer growth across patients

1. Time series model

$$\bar{y}(y_0, \lambda, t) = y_0 \mathrm{e}^{\lambda t}$$

2. Error model

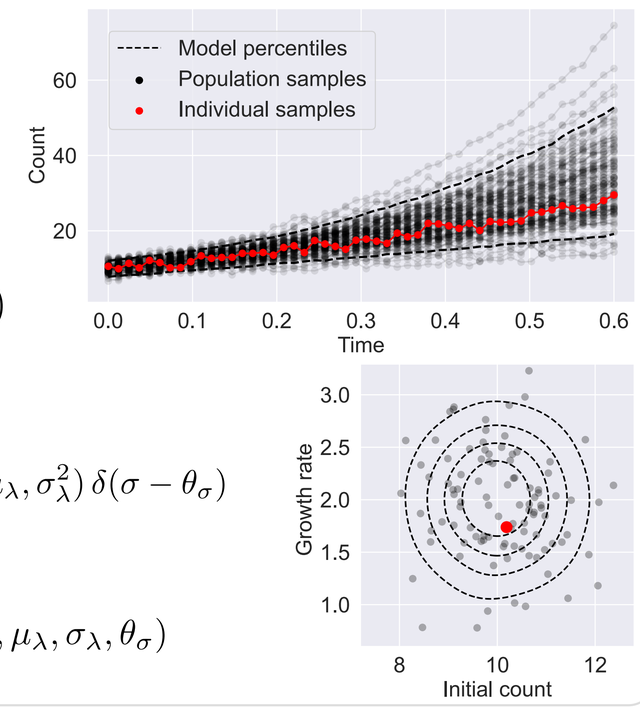
$$p(y|\psi, t) = \mathcal{N}\left(y|\bar{y}(y_0, \lambda, t), \sigma^2\right)$$

3. Population model

$$p(\psi|\theta) = \mathcal{N}(y_0|\mu_{y_0}, \sigma^2_{y_0})\,\mathcal{N}(\lambda|\mu_\lambda, \sigma^2_\lambda)\,\delta(\sigma - \theta_\sigma)$$

Model parameters:

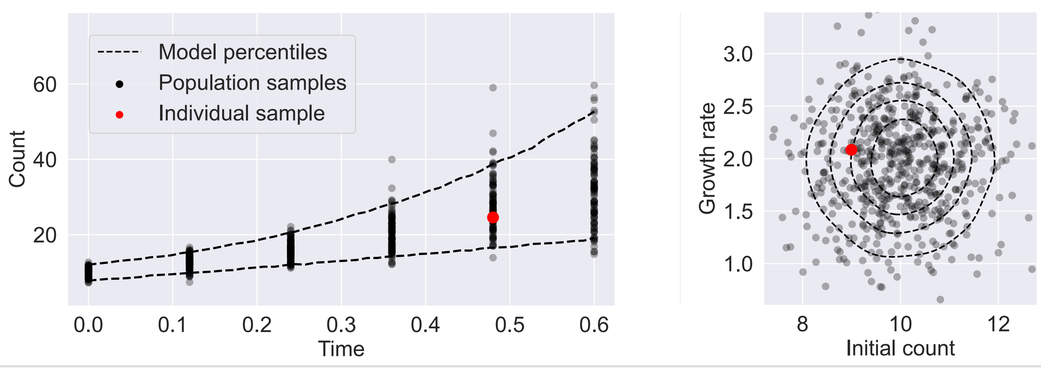
$$\psi = (y_0, \lambda, \sigma), \quad \theta = (\mu_{y_0}, \sigma_{y_0}, \mu_\lambda, \sigma_\lambda, \theta_\sigma)$$



What are snapshot time series data?

Snapshot time series data are measurements over time where individuals in the population are only measured once.

Snapshot data: Early cancer growth across patients



How do we infer the model parameters?

The model defines a joint distribution, $p(y, \psi|\theta, t)$, which we can use to define a (hierarchical) log-likelihood for the parameters given the data

$$\log p(\mathcal{D}, \Psi|\theta) = \sum_{ij} \log p(y_{ij}|\psi_i, t_{ij}) + \sum_i \log p(\psi_i|\theta)$$

with data $\mathcal{D} = \{(y_{ij}, t_{ij})\}$ and individual-level parameters $\Psi = \{\psi_i\}$.

Using Bayes' rule, we can derive a posterior distribution for the parameters

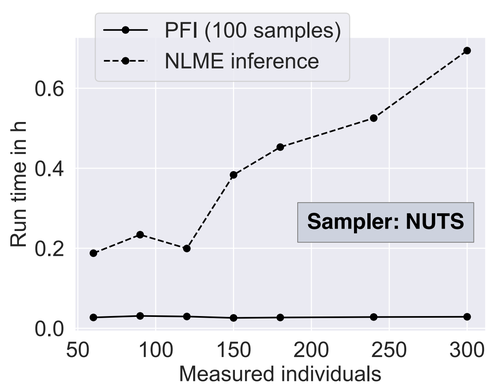
$$\log p(\theta, \Psi|\mathcal{D}) = \log p(\mathcal{D}, \Psi|\theta) + \log p(\theta) + \text{const.},$$

where $p(\theta)$ is the prior distribution of the population parameters. This posterior distribution can be inferred using MCMC sampling.

NLME inference is not tractable for snapshot data

1. Posterior evaluations become expensive.
2. Posterior dimension becomes large.
3. Local changes in posterior curvature become extreme.

Scaling: Early cancer growth across patients



Population filter inference (PFI)

PFI uses population filters to do approximate population-level inference. Population filters summarise the population measurement distribution using virtual measurements, $\tilde{Y}_j$, of the model

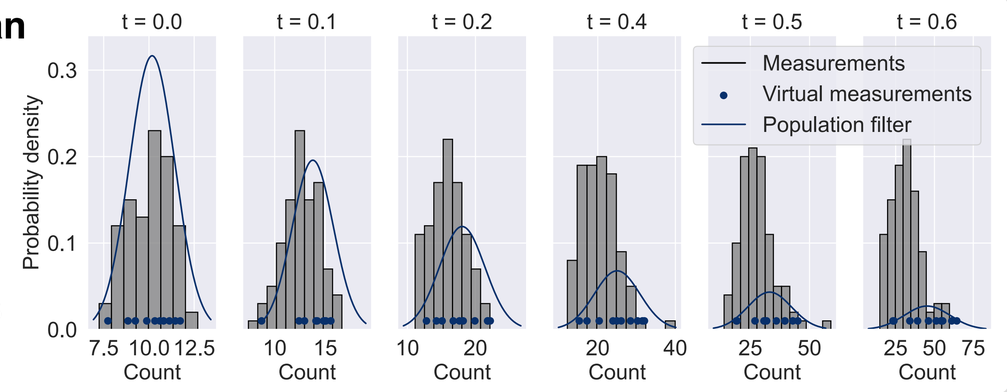
$$p(y|\tilde{Y}_j), \quad \tilde{Y}_j = \{\tilde{y}_{sj}\} \quad \text{for each measurement time } t_j.$$

The simplest population filter is a Gaussian population filter

$$p(y|\tilde{Y}_j) = \mathcal{N}(y|\tilde{\mu}_j, \tilde{\sigma}^2_j)$$

where the mean and variance of the Gaussian are estimated from the virtual measurements. Other choices are possible, e.g. KDE-based filters.

Gaussian Filters: Early cancer growth across patients

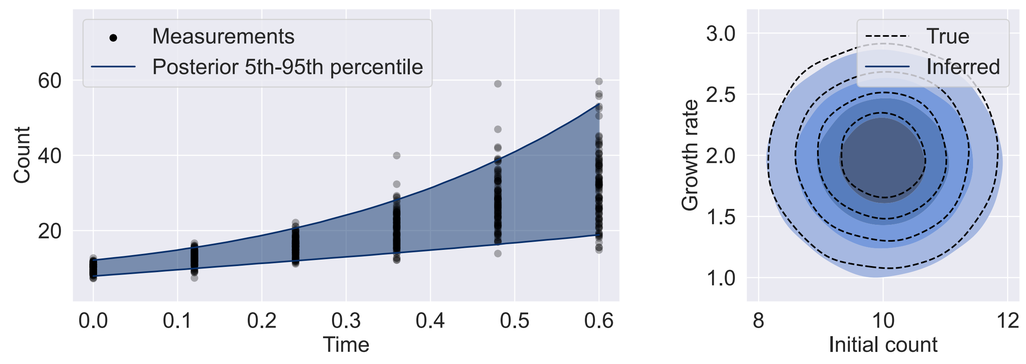


With population filters we can define an approximate population-level posterior

$$\log p(\theta, \Psi, \tilde{Y}|\mathcal{D}) = \sum_{ij} \log p(y_{ij}|\tilde{Y}_j) + \log p(\tilde{\mathcal{D}}, \Psi|\theta) + \log p(\theta) + \text{const.}$$

whose **evaluation cost no longer scales with the number of measured individuals**. Instead, the evaluation cost can be tuned with the number of measured virtual individuals in the virtual dataset, $\tilde{\mathcal{D}} = \{(\tilde{y}_{sj}, t_j)\}$.

PFI results: Early cancer growth across patients



Summary

Population filter inference enables NLME inference from snapshot data by:

1. Decoupling the inference cost from the number of measured individuals;
2. Decoupling the dimensionality of the posterior from the number of measured individuals;
3. Eliminating shrinkage by fitting to data on a population level.

Limitations:

1. Population filters can result in information loss which would translate into widened posterior distributions.
2. Noise and inter-individual variability are no longer distinguishable on a population level. Strong priors on error model parameters are required to reduce the uncertainty of population parameter estimates.

References

Preprint will be soon available on bioRxiv. This work generalises and extends earlier work by

Hasenauer et al., BMC Bioinformatics 12, 125, 2011.

Contact and open-source software

david.augustin@cs.ox.ac.uk | davaug.github.io | DavAug

Dose Response Modelling

www.github.com/DavAug/chi