Simple Application of Approximate Bayesian Computation in Modelling Tumor Growth

Chun-Hei (Sammuel) Lam
Department of Mathematics, Faculty of Natural Science, Imperial College London.

Imperial College London

Scenario

A research [1] suggests that tumor growth is *inhibited* by certain class of drugs. They injected tumor cells in mice and compared the growth of tumor cell with the presence of drug ("experimental condition") / without presence of drug ("control condition")). This project makes use of the data available to perform Bayesian Analysis and justify the conclusion. We try to avoid non-trivial computation other than the Approximate Bayesian Computation (ABC) itself.

Gompertz Model

We assume that the growth of tumor in both conditions follows the deterministic **Gompertz Model**, which is a simple but accurate model [2,3]. If $V(t)$ is the volume (in mm^3) of tumor on the t-th day after injection of cell ($t \geq 0$), then

$$\frac{dV^{(i)}(t)}{dt} = \theta^{(i)} V^{(i)}(t)(\ln V_\infty^{(i)} - \ln V^{(i)}(t)) \quad (1)$$

$$V^{(i)}(t) = V_\infty^{(i)} e^{e^{-\theta^{(i)} t}(\ln V_0^{(i)} - \ln V_\infty^{(i)})} \quad (2)$$

- i is either 1 or 2:, where $i = 1$ refers to the "control condition", and $i = 2$ refers to the "experimental condition".
- $V_0^{(i)}$ is the initial volume of tumor ($t = 0$).
- $V_\infty^{(i)}$ is the volume of tumor in long run ($t \to \infty$).
- $\theta^{(i)}$ is the 'speed of equilibration' (when $\theta^{(i)}$ is large V reaches $V_\infty^{(i)}$ quicker).

We aim to infer the values of those parameters from data $\mathcal{D} = \{\vec{V}^{(1)}, \vec{V}^{(2)}, \vec{t}\}$, where

- $\vec{V}^{(i)} = \{V_1^{(i)}, ..., V_n^{(i)}\}$ is the response under various conditions ($i = 1, 2$), and
- $\vec{t} = \{t_1, ..., t_n\}$ are the time when the measurements are taken.

To simplify our discussion, we further assume that

- All individual observations are **independent** - in such case we can perform ABC separately on $\vec{V}^{(1)}$ and $\vec{V}^{(2)}$.
- The error of measurements are **negligible** - this *oversimplifies* our discussion. However we may use that as a reference to determine the *noise* of likelihood function.

Construction of Bayesian Framework

Let $\vec{\theta}^{(i)} = (V_0^{(i)}, V_\infty^{(i)}, \theta^{(i)})$ be the vector of parameters in various conditions. Bayes' Theorem tells us that [4, 5] the **Posterior** distribution (given data) $f_{\Theta^{(i)}|\mathcal{D}}(\vec{\theta}^{(i)}|\mathcal{D})$ is

$$f_{\Theta^{(i)}|\mathcal{D}}(\vec{\theta}^{(i)}|\mathcal{D}) = \frac{f_{\mathcal{D}|\Theta^{(i)}}(\mathcal{D}|\vec{\theta}^{(i)}) f_{\Theta^{(i)}}(\vec{\theta}^{(i)})}{\int_R f_{\mathcal{D}|\Theta^{(i)}}(\mathcal{D}|\vec{\theta}^{(i)}) f_{\Theta^{(i)}}(\vec{\theta}^{(i)}) d\vec{\theta}^{(i)}} \quad (3)$$

Here $d\vec{\theta}^{(i)}$ means $dV_0^{(i)} dV_\infty^{(i)} d\theta^{(i)}$ is a small region in R. We need to determine the **Prior** distribution $p(\vec{\theta}^{(i)})$ (with support R) and **Likelihood** $p(\mathcal{D}|\vec{\theta}^{(i)})$.

Likelihood

We further assume that

$$p(\mathcal{D}|\vec{\theta}^{(i)}) = \prod_{j=1}^{n} \left(\frac{1}{\sigma^{(i)}\sqrt{2\pi}} e^{-\frac{(V_j - \hat{V}^{(i)}(t_j))^2}{2(\sigma^{(i)})^2}} \right) \quad (4)$$

- $\hat{V}^{(i)}(t_j)$ is obtained by directly substituting t_j into equation (2).
- $\sigma^{(i)}$ = mean of absolute uncertainties in two conditions.

Prior

The prior need not be so accurate for the posterior to give sufficient information. Here are my choice of (marginal) prior:

- $V_0^{(i)} \sim U[0, V^{(i)}(1)]$. by considering that V_0 must lie within the range specified.
- We know that $V_\infty^{(i)}$ should be greater than the last observation taken $V_{\text{last}}^{(i)}$ (which is also maximum among data taken). A guess of prior is $V_\infty^{(i)} \sim U[V_{\text{last}}^{(i)}, 3V_{\text{last}}^{(i)}]$
- We only know that $\theta^{(i)} > 0$. Therefore we guess the prior to be 'modded Cauchy' with pdf

$$f_{\Theta^{(i)}}(\theta^{(i)}) = \begin{cases} \frac{2}{\pi(1+(\theta^{(i)})^2)} & \theta^{(i)} > 0 \\ 0 & \theta^{(i)} \leq 0 \end{cases} \quad (5)$$

- Assume they are independent, so that joint pdf of prior is the product of three pdfs and joint support of prior is the Cartesian product of three supports.

Addressing Situation

We need to determine whether $V_\infty^{(1)} - V_\infty^{(2)} > 0$ (reduction of **maximum size**) and $\theta^{(1)} - \theta^{(2)} > 0$ (reduction of **growth rate**).
A rather simple way is to perform the Monte Carlo Hypothesis Test [7] with $\alpha = 0.1$ as followed.

Hypotheses	Result
$H_0 : V_\infty^{(1)} - V_\infty^{(2)} \leq 0$ $H_1 : V_\infty^{(1)} - V_\infty^{(2)} > 0$	reject H_0, $(p \approx 0.07)$
$H_0 : \theta^{(1)} - \theta^{(2)} \leq 0$ $H_1 : \theta^{(1)} - \theta^{(2)} > 0$	accept H_0, $(p \approx 0.24)$

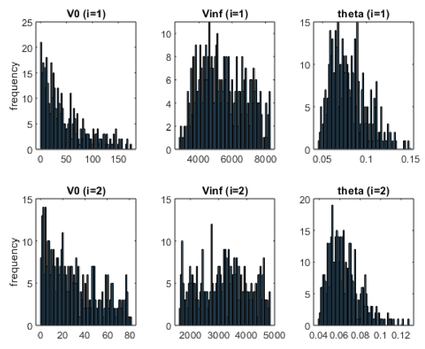
Unfortunately, I had made a rather unreasonable assumption that the variables in posterior distributions are **independent**. Moreover, p just represents the percentile of statistic under H_0 in the list of samples (which is different from p in Frequentists' view). Nevertheless it provides a hint that the **volume of tumor in long run** is reduced.

Possible Further Investigation

There are sophisticated methods to perform ABC, including the **ABC-Monte Carlo Markov Chain (MCMC)** (ABC and Metropolitan-Hasting) and **ABC-Sequential Monte Carlo (SMC)**. [6]. In general, the algorithms are more computationally efficient. We may further use **Bayes' Factor** to perform hypothesis testing. These techniques, although not included here, may be used in the future.

Results of ABC

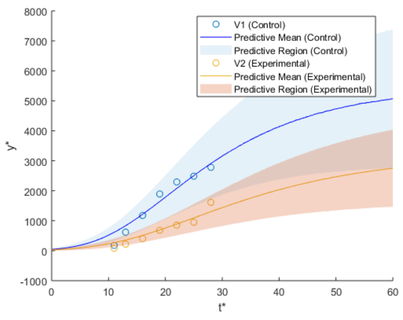
Joint Posterior Distribution: We perform the simplest **rejection-based ABC** as described in [6]. Here are the histograms of samples (size = 400) from Marginal Distributions.



Predictive Distribution: Given $t = t^*$, then the probability density function of $V^{(i)}(t^*) = V^*$ is: [5]

$$f(V^*|t^*, \mathcal{D}) = \int_R f(V^*|t^*, \vec{\theta}^{(i)}, \mathcal{D}) f_{\Theta^{(i)}|\mathcal{D}}(\vec{\theta}^{(i)}|\mathcal{D}) d\vec{\theta}^{(i)} \quad (6)$$

Below is the plot of raw data, means of predictive distributions and predictive bands.



References

[1] ZANFARDINO M, SPAMPANATO C, DE CICCO R, BUOMMINO E, DE FILIPPIS A, BAIANO S, et al. Simvastatin reduces melanoma progression in a murine model. International Journal of Oncology. 2013; 43 (6): 1763-1770. Available from: doi: 10.3892/ijo.2013.2126 Available from: https://www.ncbi.nlm.nih.gov/pubmed/24101161.

[2] Dennis B, Ponciano JM. Density-dependent state-space model for population-abundance data with unequal time intervals. Ecology. 2014; 95 (8): 2069-2076. Available from: doi: 10.1890/13-1486.1 Available from: https://www.jstor.org/stable/43494714.

[3] Retsky MW, Swartzendruber DE, Wardwell RH, Bame PD. Is Gompertzian or exponential kinetics a valid description of individual human cancer growth? Medical Hypotheses. 1990; 33 (2): 95-106. Available from: doi: 10.1016/0306-9877(90)90186-I Available from: https://www.sciencedirect.com/science/article/pii/030698779090186I.

[4] Ben Calderhead, Mark Girolami, Desmond J. Higham. Is it safe to go out yet? Mathematical Modelling of Zombies. University of Ottawa Press; 2014. pp. 129.

[5] Girolami M. Bayesian inference for differential equations. Theoretical computer science. 2008; 408 4-16. Available from: doi: 10.1016/j.tcs.2008.07.005 Available from: https://www.sciencedirect.com/science/article/pii/S030439750800501X.

[6] Beaumont MA. Approximate Bayesian Computation in Evolution and Ecology. Annual Review of Ecology, Evolution, and Systematics. 2010; 41 (1): 379-406. Available from: doi: 10.1146/annurev-ecolsys-102209-144621 Available from: https://www.jstor.org/stable/27896228.

[7] Adery C. A. Hope. A Simplified Monte Carlo Significance Test Procedure. Journal of the Royal Statistical Society. Series B (Methodological). 1968; 30 (3): 582-598. Available from: doi: 10.1111/j.2517-6161.1968.tb00759.x Available from: https://www.jstor.org/stable/2984263.

I further acknowledge Dr. B. Calderhead and his students for suggesting sources [4] and [5], as well as Dr. A. Duncan for his supervision.