Final Project:

2.1: Don't worry about SMC, just ABC from class.

3.3: Model flu outbreaks

q_c: community transmission

q_h: household "

data in a Supplemental file – File006

w_j's give prob. of each cell in that table

then check if simulated table is close to observed data

if so keep it or reject and test new q_c, q_h

Notes on Paper:

. Rather than just using ABC to find Θ, we also want to choose a model m. Goal: calculate $P(m|D_0)$

marginal posterior distribution of m

2 ways to find this:

- Joint Space Based Approach

Define a joint space of model indicators $m = 1, 2, \ldots |M|$
& model parameters Θ

Obtain joint posterior distribution over combined space
$P(\Theta, m|D_0)$

Marginalize over parameters to obtain $P(m|D_0)$

- Marginal Likelihood Based Approach

  "Evidence"

  Estimate marginal likelihoods $P(D_0|m)$ & use

$$P(m|D_0) = \frac{P(D_0|m)\,P(m)}{\sum_{m'} P(D_0|m')\,P(m')}$$

Both are good but computationally costly so paper uses SMC (we won't). Paper prefers Joint Space approach.

Model Selection based on ABC rejection:

particles: (m, θ)

(1) Draw m^* from the prior $P(m)$

(2) Sample θ^* from the prior $P(\theta|m^*)$

(3) Simulate a dataset $D^* \sim f(D|\theta^*, m^*)$

what is f?

(4) Compute distance $\begin{cases} d(D_0, D^*) \leq \varepsilon, & \text{accept } (m^*, \theta^*) \\ \text{else} & \text{reject} \end{cases}$

(5) repeat until N particles are accepted

(6) approximate marginal posterior distribution by:

$$P(m=m'|D_0) \approx \frac{\#\text{ accepted particles of the form } (m', \cdot)}{N}$$

Modifications made for SMC (we won't do this!)

(MS1) Initialize $\varepsilon_1 > \dots > \varepsilon_T$

Set population indicator $t=1$

(MS 2.0) Set particle indicator $i=1$

(MS2.1) · If $t=1$, sample (m^{**}, θ^{**}) from prior $P(m,\theta)$

· If $t>1$, sample m^* with $P_{t-1}(m^*)$ & draw

$m^{**} \sim \underbrace{KM_t(m|m^*)}_{?}$

Sample θ^* from previous population $\{\theta(m^{**})_{t-1}\}$

with weights ω_{t-1} & draw $\theta^{**} \sim KP_{t,m^{**}}(\theta|\theta^*)$

$\begin{cases} P(m^{**}, \theta^{**}) = 0, & \text{return to MS2.1} \\ \text{else: simulate } D^* \sim f(D|\theta^{**}, m^{**}) & \begin{cases} d(D_0, D^*) > \varepsilon_t, \\ \text{else, continue} \end{cases} \end{cases}$

(MS 2.2) Set $(m_t^i, \theta_t^i) = (m^{**}, \theta^{**})$ &

calculate weight of particle via some formula

If $i<N$, set $i=i+1$ & go to (MS2.1)

(MS3) Normalize weights ω_t ← how?

Obtain marginal model probabilities by

$$P_t(m_t = m) = \sum \omega_t^{(i)} (m_t^{(i)}, \theta_t^{(i)})$$

If $t<T$ set $t=t+1$ & go to MS2.0

The goal of this algorithm is to obtain (m_T, θ_T) &
$P_T(m_T)$

In 3.1 choose uniform prior distributions
(discussion of priors at the end of 2.1)

Section 3 Notes

3.3 – Goal: replicate figures 3a & 3c

q_c := probability a susceptible person doesn't get infected from community

q_h := " household

w_{js} := probability that j of s susceptibles in a house get infected
row (j), col (s)

Paper aims to infer q_c & q_h from data on w_{js}
& answer whether the 2 data sets have same q_c & q_h

Consider 2 models:

(1) has q_{h1}, q_{c1}, q_{h2}, q_{c2} – surmises different parameters

(2) has q_h, q_c – assumes same parameters

Prior distributions all assumed to be uniform on $[0,1]$

Distance fxn given on pg. 107