Background: ABC

The posterior for θ given observed data x_{obs}:

$$\pi(\theta \mid x_{\text{obs}}) = \frac{f(x_{\text{obs}} \mid \theta)\pi(\theta)}{\int f(x_{\text{obs}} \mid \theta)\pi(\theta)d\theta} = \frac{f(x_{\text{obs}} \mid \theta)\pi(\theta)}{f(x_{\text{obs}})}$$

Approximate Bayesian Computation

"Likelihood-free" approach to approximating $\pi(\theta \mid x_{\text{obs}})$

$\longrightarrow f(x_{\text{obs}} \mid \theta)$ not specified

Proceeds via simulation of the forward process

Why would we not know $f(x_{\text{obs}} \mid \theta)$?

1. Physical model too complex
2. Strong dependency in data
3. Observational limitations

Basic ABC algorithm

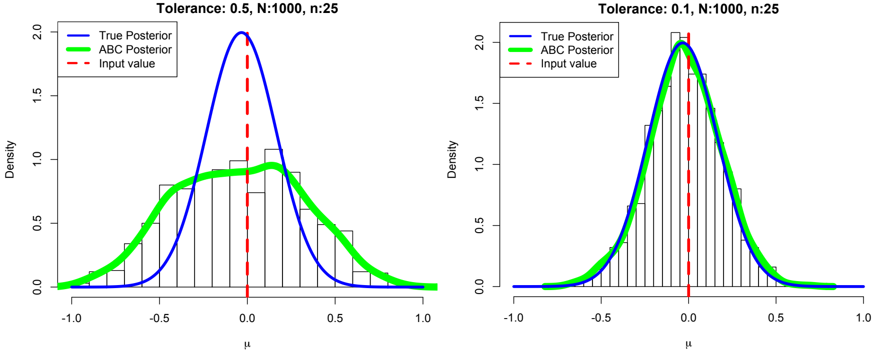
For the observed data x_{obs} and prior $\pi(\theta)$:

1. Sample θ_{prop} from prior $\pi(\theta)$
2. Generate x_{prop} from forward process $f(x \mid \theta_{\text{prop}})$
3. Accept θ_{prop} if $x_{\text{obs}} = x_{\text{prop}}$
4. Return to step 1

*Introduced in Pritchard et al. (1999) (population genetics)

Gaussian illustration

- Data x_{obs} consists of 25 iid draws from Normal$(\mu, 1)$
- Summary statistics $S(x) = \bar{x}$
- Distance function $\Delta(S(x_{prop}), S(x_{obs})) = |\bar{x}_{prop} - \bar{x}_{obs}|$
- Tolerance $\epsilon = 0.50$ and 0.10
- Prior $\pi(\mu) =$ Normal(0,10)


Tolerance: 0.5, N:1000, n:25
True Posterior
ABC Posterior
Input value
Density
μ
Tolerance: 0.1, N:1000, n:25
True Posterior
ABC Posterior
Input value
Density
μ

For observations x_{obs}, distance function Δ, and (small) tolerance ϵ

Algorithm 1 Basic ABC Algorithm

```
for i = 1 to N do
    while Δ(S(x_prop), S(x_obs)) > ε do
        Propose θ_prop by drawing θ_prop from prior π(θ)
        Generate x_prop from forward process F(x | θ_prop)
        Calculate summary statistics {S(x_prop), S(x_obs)}
    end while
    θ^(i) ← θ_prop
end for
```

- ABC posterior based on $\{\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(N)}\} = \{\theta^{(i)}\}_{i=1}^{N}$
- $\{\theta^{(i)}\}_{i=1}^{N}$ are often referred to as *particles*

Introduced in Pritchard et al. (1999), Rubin (1984) (conceptually)

How to pick a tolerance, ϵ?

Sequential ABC

Main idea
Instead of starting the ABC algorithm over with a smaller tolerance (ϵ), use the already sampled particle system as a proposal distribution *rather* than drawing from the prior distribution.

Particle system:

(1) retained sampled values, (2) importance weights

Beaumont et al. (2009); Moral et al. (2011); Bonassi and West (2004)

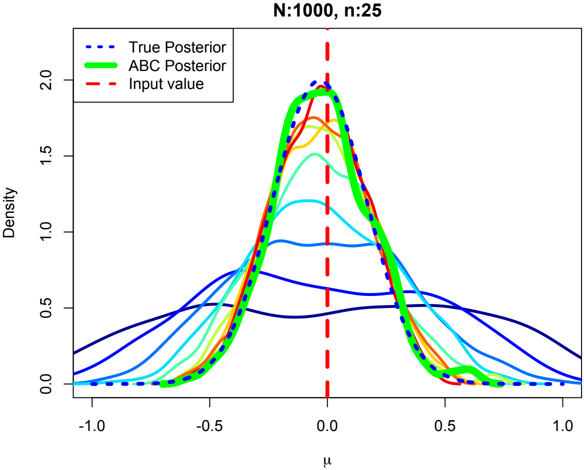
Algorithm 2 ABC - Population Monte Carlo algorithm*

At iteration $t = 1$
Algorithm 1: Basic ABC sampler to obtain $\{\theta_1^{(i)}\}_{i=1}^N$
Set importance weights $W_1^{(i)} = 1/N$ for $i = 1, \dots, N$
for $t = 2$ to T **do**
    Set $\tau_t^2 = 2 \cdot \text{var}\left(\{\theta_{t-1}^{(i)}, W_{t-1}^{(i)}\}_{i=1}^N\right)$
    for $i = 1$ to N **do**
        while $\rho(S(y_{1:n}), S(x_{1:n})) > \epsilon_t$ **do**
            Draw θ_0 from $\{\theta_{t-1}^{(i)}\}_{i=1}^N$ with probabilities $\{W_{t-1}^{(i)}\}_{i=1}^N$
            Propose $\theta^* \sim N(\theta_0, \tau_t^2)$
            Generate $x_{1:n}$ from $F(x \mid \theta^*)$
            Calculate summary statistics $\{S_y, S_x\}$
        end while
        $\theta_t^{(i)} \leftarrow \theta^*$
        $\widetilde{W}_t^{(i)} \leftarrow \dfrac{\pi\left(\theta_t^{(i)}\right)}{\sum_{j=1}^N W_{t-1}^{(j)} \phi\left[\tau_t^{-1}(\theta_t^{(i)} - \theta_{t-1}^{(j)})\right]}$
    end for
    $\{W_t^{(i)}\}_{i=1}^N \leftarrow \{\widetilde{W}_t^{(i)}\}_{i=1}^N / \sum_{i=1}^N \widetilde{W}_t^{(i)}$
end for

Decreasing tolerances $\epsilon_1 \geq \cdots \geq \epsilon_T$, $\phi(\cdot)$ is the density function of a $N(0, 1)$

*From Beaumont et al. (2009)

Gaussian illustration: sequential posteriors



Tolerance sequence, $\epsilon_{1:10}$:

1.00 0.75 0.53 0.38 0.27 0.19 0.15 0.11 0.08 0.06