

# 10: Introduction to Bayesian Computation

Taylor

University of Virginia

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

Introduction

Accept-Reject Sampling

Importance Sampling

Importance Sampling with Resampling

Sequential Monte Carlo

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

# Introduction

"10"

Taylor

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

This chapter gives an overview for how to approximate intractable quantities such as posterior expectations and predictions.

**Numerical integration** methods approximate integrals.

These methods can be loosely categorized as either **stochastic** or **deterministic** (e.g. quadrature methods).

**Deterministic methods** don't draw samples, and instead approximate the integral by summing up approximate volumes:

$$E[h(\theta) \mid y] = \int h(\theta)p(\theta \mid y)d\theta \approx \frac{1}{S} \sum_{s=1}^S w_s h(\theta^s)p(\theta^s \mid y)$$

# Approximating the posterior on a grid

We used this method earlier for the rat tumor example!

Given that we can evaluate the unnormalized target  $p(\theta, y) = p(y | \theta)p(\theta)$ , we choose a nonrandom grid of points (think seq)  $\theta_1, \dots, \theta_S$ , and then we approximate the the continuous posterior with a discrete random variable with pmf equal to

$$\tilde{p}(\theta_j | y) = \frac{p(y | \theta_j)p(\theta_j)}{\sum_{s=1}^S p(y | \theta_s)p(\theta_s)}$$

for any  $\theta_j \in \{\theta_1, \dots, \theta_S\}$ . Then

$$E[h(\theta) | y] \approx \sum_{j=1}^S h(\theta_j) \tilde{p}(\theta_j | y).$$

[Introduction](#)[Accept-Reject Sampling](#)[Importance Sampling](#)[Importance Sampling with Resampling](#)[Sequential Monte Carlo](#)[References](#)

**Stochastic methods** involve sample averages of simulated draws from some distribution. There are **many** ways to do this, but here are a couple examples that we've seen:

$$E[h(\theta) | y] = \int h(\theta)p(\theta | y)d\theta \approx \frac{1}{S} \sum_{s=1}^S h(\theta^s)$$

with  $\theta^s \sim p(\theta | y)$ , or

$$E[h(\tilde{y}) | y] = \int h(\tilde{y})p(\tilde{y} | y)d\tilde{y} \approx \frac{1}{S} \sum_{s=1}^S h(\tilde{y}^s)$$

with  $\tilde{y} \sim p(\tilde{y} | y)$

## Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

Drawing  $\tilde{y}$  samples can be done in a two-stage way:

1. draw  $\theta^s \sim p(\theta \mid y)$
2. draw  $\tilde{y}^s \sim p(\tilde{y} \mid \theta^s)$



Drawing  $\tilde{y}$  samples can be done in a two-stage way:

1. draw  $\theta^s \sim p(\theta \mid y)$
2. draw  $\tilde{y}^s \sim p(\tilde{y} \mid \theta^s)$

If you can derive  $E(h(\tilde{y}) \mid \theta)$ , you should probably use a Rao-Blackwellized procedure:

$$E[h(\tilde{y}) \mid y] = E[E(h(\tilde{y}) \mid \theta) \mid y] \approx \frac{1}{S} \sum_{s=1}^S E(h(\tilde{y}) \mid \theta^s)$$

with  $\theta^s \sim p(\theta \mid y)$

# The general setup

"10"

Taylor

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

From now on we will write the posterior in terms of an unnormalized density  $q(\theta | y)$ . In other words:

$$p(\theta | y) = \frac{q(\theta | y)}{\int q(\theta | y) d\theta}$$

Most (maybe all) of the sampling techniques will assume that we can't evaluate  $p(\theta | y)$ , but that we can evaluate  $q(\theta | y)$

# Rejection Sampling aka Accept-Reject sampling

"10"

Taylor

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

## Setup

1.  $p(\theta | y)$  the target, posterior
2.  $q(\theta | y) = p(y | \theta)p(\theta)$  the unnormalized target
3.  $g(\theta)$  the “instrumental” or “proposal” distribution
4. need  $q(\theta | y)/g(\theta) \leq M$  uniformly
5. need  $g \gg q$  i.e. the proposal “dominates” your target (won't divide by 0)

We are free to choose our own  $g(\theta)$ . For the time being, we assume that  $\int g(\theta)d\theta = 1$ .

# Rejection Sampling aka Accept-Reject sampling

"10"

Taylor

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

To (potentially) produce one draw:

1. propose the draw  $\theta^s \sim g(\theta)$
2. accept  $\theta^s$  with probability  $\frac{q(\theta^s|y)}{g(\theta^s)M}$

# Rejection Sampling aka Accept-Reject sampling

"10"

Taylor

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

To (potentially) produce one draw:

1. propose the draw  $\theta^s \sim g(\theta)$
2. accept  $\theta^s$  with probability  $\frac{q(\theta^s|y)}{g(\theta^s)M}$

Note this is the same as

1. propose the draw  $\theta^s \sim g(\theta)$
2. draw  $U \sim \text{Uniform}(0, 1]$
3. accept  $\theta^s$  if  $U < q(\theta^s | y) / \{g(\theta^s)M\}$

$$\begin{aligned}P\left(\theta \leq t \mid U \leq \frac{q(\theta | y)}{Mg(\theta)}\right) &= \frac{P\left(\theta \leq t, U \leq \frac{q(\theta | y)}{Mg(\theta)}\right)}{P\left(U \leq \frac{q(\theta | y)}{Mg(\theta)}\right)} \\&= \frac{\int_{-\infty}^t \int_0^{\frac{q(\theta | y)}{Mg(\theta)}} g(\theta) 1 du d\theta}{\int_{-\infty}^{\infty} \int_0^{\frac{q(\theta | y)}{Mg(\theta)}} g(\theta) 1 du d\theta} \\&= \frac{\int_{-\infty}^t g(\theta) \frac{q(\theta | y)}{Mg(\theta)} d\theta}{\int_{-\infty}^{\infty} g(\theta) \frac{q(\theta | y)}{Mg(\theta)} d\theta} \\&= \frac{\int_{-\infty}^t q(\theta | y) d\theta}{\int_{-\infty}^{\infty} q(\theta | y) d\theta} \\&= P(\theta \leq t | y).\end{aligned}$$

# Example 1

"10"

Taylor

Assume  $y \sim \text{Normal}(\theta, 1)$ , and  $p(\theta) = \frac{1}{\pi(1+\theta^2)}$ . Our goal is to draw from

$$\begin{aligned} p(\theta | y) &\propto q(\theta | y) \\ &= p(y | \theta)p(\theta) \\ &= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(y - \theta)^2\right] \frac{1}{\pi(1 + \theta^2)} \\ &\propto \exp\left[-\frac{(\theta - y)^2}{2} - \log(1 + \theta^2)\right], \end{aligned}$$

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

## Example 1

"10"

Taylor

Let's assume that we want to use our prior distribution as a proposal:  $g(\theta) = p(\theta)$ . Then we have to find  $M$ :

$$\begin{aligned}\frac{q(\theta | y)}{g(\theta)} &= \frac{p(y | \theta)p(\theta)}{p(\theta)} \\ &= p(y | \theta) \\ &= \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2}(y - \theta)^2 \right] \\ &\leq \frac{1}{\sqrt{2\pi}} \stackrel{\text{def}}{=} M\end{aligned}$$

Our acceptance probability for draw  $\theta^s$  is then

$$q(\theta^s | y) / \{g(\theta^s)M\} = p(y | \theta^s) / M = \exp \left[ -\frac{1}{2}(y - \theta^s)^2 \right]$$

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References



# Rejection Sampling aka Accept-Reject sampling

"10"

Taylor

Generally a good strategy is work with logarithms, and then exponentiate as late as possible.

```
y <- 2 # fake data
num_trials <- 1000
theta_proposals <- rt(num_trials, 1)
us <- runif(num_trials, min = 0, max = 1)
log_accept_prob <- function(theta){
  -.5*(y - theta)^2
}
probs <- exp(log_accept_prob(theta_proposals))
accepts <- ifelse(us < probs, TRUE, FALSE)
hist(theta_proposals[accepts]) # only the accepted draws
#hist(theta_proposals) # all draws!
```

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

# Importance Sampling

"10"

Taylor

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

**importance sampling** also involves ratios like the previous algorithm. However, instead of using those ratios to either accept or discard samples, it uses the ratios to weight samples.

# Importance Sampling

"10"

Taylor

Introduction

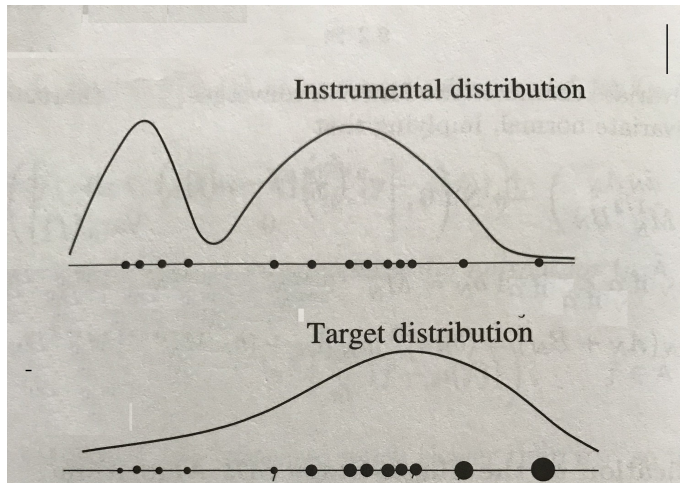
Accept-Reject  
Sampling

**Importance  
Sampling**

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References



Original, unedited image is from  
<https://www.springer.com/us/book/9780387402642>

## Setup

1.  $p(\theta | y)$  the target, posterior
2.  $q(\theta | y) = p(y | \theta)p(\theta)$  the unnormalized target
3.  $g(\theta)$  the “instrumental” or “proposal” distribution
4.  $g \gg q$  i.e. the proposal dominates your target

We are free to choose our own  $g(\theta)$ . For the time being, we assume that  $\int g(\theta)d\theta = 1$ .

# Importance Sampling

"10"

Taylor

Introduction

Accept-Reject  
Sampling

**Importance  
Sampling**

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

Algorithm: for each iteration  $s$

1. draw  $\theta^s \sim g(\theta)$
2. calculate unnormalized weight  $\tilde{w}(\theta^s) = \frac{q(\theta^s|y)}{g(\theta^s)}$
3. calculate normalized weights  $w(\theta^s) = \tilde{w}(\theta^s) / \sum_r \tilde{w}(\theta^r)$

Final calculation:

$$E_q[h(\theta) | y] \approx \sum_s w(\theta^s) h(\theta^s)$$

# Importance Sampling

"10"

Taylor

Motivation:

$$\begin{aligned} E_q[h(\theta) | y] &= \int h(\theta) p(\theta | y) d\theta \\ &= \frac{\int h(\theta) q(\theta | y) d\theta}{\int q(\theta | y) d\theta} \\ &= \frac{\int h(\theta) \frac{q(\theta|y)}{g(\theta)} g(\theta) d\theta}{\int \frac{q(\theta|y)}{g(\theta)} g(\theta) d\theta} \end{aligned}$$

Introduction

Accept-Reject  
Sampling

**Importance  
Sampling**

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

# Importance Sampling

"10"

Taylor

Motivation:

$$\begin{aligned} E_q[h(\theta) | y] &= \int h(\theta) p(\theta | y) d\theta \\ &= \frac{\int h(\theta) q(\theta | y) d\theta}{\int q(\theta | y) d\theta} \\ &= \frac{\int h(\theta) \frac{q(\theta | y)}{g(\theta)} g(\theta) d\theta}{\int \frac{q(\theta | y)}{g(\theta)} g(\theta) d\theta} \end{aligned}$$

So first:

$$\frac{1}{S} \sum_{s=1}^S \frac{q(\theta^s | y)}{g(\theta^s)} \rightarrow E_g \left[ \frac{q(\theta | y)}{g(\theta)} \right] = \int \frac{q(\theta | y)}{g(\theta)} g(\theta) d\theta = \int q(\theta | y) d\theta$$

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

# Importance Sampling

"10"

Taylor

Introduction

Accept-Reject  
Sampling

**Importance  
Sampling**

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

$$1. E_q[h(\theta) | y] = \frac{\int h(\theta)q(\theta|y)d\theta}{\int q(\theta|y)d\theta}$$

$$2. \frac{1}{S} \sum_{s=1}^S \frac{q(\theta^s|y)}{g(\theta^s)} \rightarrow \int q(\theta | y)d\theta \text{ ( for the denominator)}$$

And second:

$$\frac{1}{S} \sum_{s=1}^S h(\theta^s) \frac{q(\theta^s | y)}{g(\theta^s)} \rightarrow E_g \left[ h(\theta) \frac{q(\theta | y)}{g(\theta)} \right] = \int h(\theta)q(\theta | y)d\theta$$

which converges to the numerator



# Importance Sampling

"10"

Taylor

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

$$1. E_q[h(\theta) | y] = \frac{\int h(\theta) q(\theta|y) d\theta}{\int q(\theta|y) d\theta}$$

$$2. \frac{1}{S} \sum_{s=1}^S \frac{q(\theta^s|y)}{g(\theta^s)} \rightarrow \int q(\theta | y) d\theta$$

$$3. \frac{1}{S} \sum_{s=1}^S h(\theta^s) \frac{q(\theta^s|y)}{g(\theta^s)} \rightarrow \int h(\theta) q(\theta | y) d\theta$$

So finally

$$\sum_{i=1}^S w(\theta^s) h(\theta^s) = \frac{\sum_{s=1}^S h(\theta^s) \frac{q(\theta^s|y)}{g(\theta^s)}}{\sum_{r=1}^S \frac{q(\theta^r|y)}{g(\theta^r)}} = \frac{\frac{1}{S} \sum_{s=1}^S h(\theta^s) \frac{q(\theta^s|y)}{g(\theta^s)}}{\frac{1}{S} \sum_{r=1}^S \frac{q(\theta^r|y)}{g(\theta^r)}} \rightarrow E[h(\theta) | y]$$

where  $w(\theta^s) = \frac{q(\theta^s|y)}{g(\theta^s)} \bigg/ \sum_{r=1}^S \frac{q(\theta^r|y)}{g(\theta^r)}$  are the self-normalized weights

## Example 2

"10"

Taylor

Assume  $y \sim \text{Normal}(\theta, 1)$ , and  $p(\theta) = \frac{1}{\pi(1+\theta^2)}$ . Approximate  $E_q[\theta \mid y]$  using proposal  $g(\theta) = p(\theta)$ .

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

## Example 2

"10"

Taylor

Assume  $y \sim \text{Normal}(\theta, 1)$ , and  $p(\theta) = \frac{1}{\pi(1+\theta^2)}$ . Approximate  $E_q[\theta \mid y]$  using proposal  $g(\theta) = p(\theta)$ .

If we sample from  $g(\theta) = p(\theta) = \frac{1}{\pi(1+\theta^2)}$  then the unnormalized weights are

$$\begin{aligned}\tilde{w}(\theta^s) &= \frac{q(\theta^s \mid y)}{g(\theta^s)} \\ &= p(y \mid \theta^s) \\ &= \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2}(y - \theta^s)^2 \right]\end{aligned}$$

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

## Example 2

"10"

Taylor

Assume  $y \sim \text{Normal}(\theta, 1)$ , and  $p(\theta) = \frac{1}{\pi(1+\theta^2)}$ . Approximate  $E_q[\theta \mid y]$  using proposal  $g(\theta) = p(\theta)$ .

If we sample from  $g(\theta) = p(\theta) = \frac{1}{\pi(1+\theta^2)}$  then the unnormalized weights are

$$\begin{aligned}\tilde{w}(\theta^s) &= \frac{q(\theta^s \mid y)}{g(\theta^s)} \\ &= p(y \mid \theta^s) \\ &= \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2}(y - \theta^s)^2 \right]\end{aligned}$$

then normalize these...

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

## Example 2

"10"

Taylor

```
y <- 2 # fake data
num_samples <- 1000000
theta_draws <- rt(num_samples , 1)
log_unnorm_weight <- function(theta){
  # sqrt(2pi) because it will cancel out
  -.5*(y - theta)^2
}
lunws <- log_unnorm_weight(theta_draws)
norm_weights <- exp(lunws)/sum(exp(lunws))
sum(norm_weights * theta_draws)
hist(norm_weights)
```

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

## Example 2

"10"

Taylor

The choice of proposal is very important. In a HW problem, you will derive the following using the delta method:

$$\text{Var}_g \left( \sum_{s=1}^S w(\theta^s) h(\theta^s) \right) \approx \frac{1}{S} E_g \left[ \underbrace{\left( \frac{\tilde{w}(\theta)}{E_g[\tilde{w}(\theta)]} \right)^2}_{!} (h(\theta) - E_q[h(\theta)])^2 \right]$$

Beware of proposals that have tails that are thin relative to the target!

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

## Example 2

"10"

Taylor

Introduction

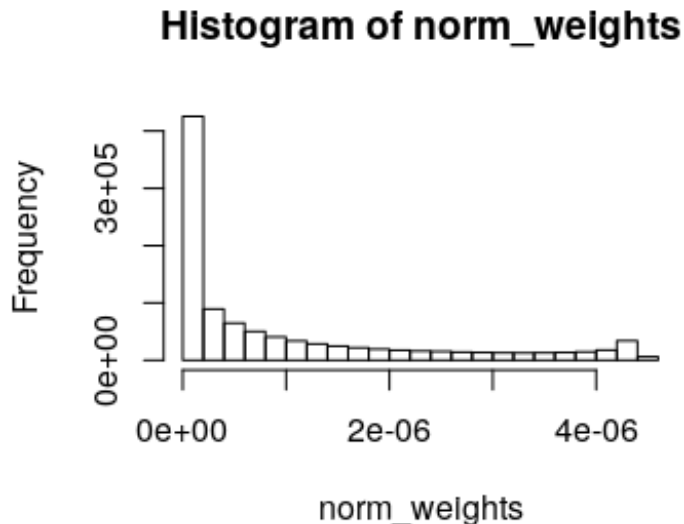
Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References



## Example 2

"10"

Taylor

Beware of bad proposal distributions!

A sample estimator of this approximate variance is

$$\sum_{s=1}^S w(\theta^s) \left( h(\theta^s) - \hat{E}[h(\theta)] \right)^2$$

where  $\hat{E}[h(\theta)] = \sum_s w(\theta^s) h(\theta^s)$ .

Note the weights aren't the same for each sample, like a "standard" estimation of the sample variance.

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References



## Example 2

"10"

Taylor

```
y <- 2 # fake data
log_unnorm_weight <- function(theta){
  # sqrt(2pi) because it will cancel out
  -.5*(y - theta)^2 }
getISEstimator <- function(num_samples){
  theta_draws <- rt(num_samples , 1)
  lunws <- log_unnorm_weight(theta_draws)
  norm_weights <- exp(lunws)/sum(exp(lunws))
  estimator <- sum(norm_weights * theta_draws)
  list("estimate" = estimator,
       "approx_var" =
         sum( norm_weights*(theta_draws - estimator)^2) ) }
# two ways to calculate standard errors
num_samps_per_estimate <- 10
sqrt(getISEstimator(num_samps_per_estimate)$approx_var)
sd(replicate(1000,
  getISEstimator(num_samps_per_estimate)$estimate))
```

[Introduction](#)

[Accept-Reject  
Sampling](#)

[Importance  
Sampling](#)

[Importance  
Sampling with  
Resampling](#)

[Sequential Monte  
Carlo](#)

[References](#)

# Effective Sample Size

"10"

Taylor

Effective sample size calculations use a second order delta method.

One form is

$$\text{ESS} = \frac{S}{1 + \text{Var}_g(W)}.$$

and the other form is

$$\text{ESS} = \frac{1}{\sum_{i=1}^S w_i^2}.$$

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

Importance Sampling gives you weighted draws  $(\theta^1, w(\theta^1)), (\theta^2, w(\theta^2)), \dots$

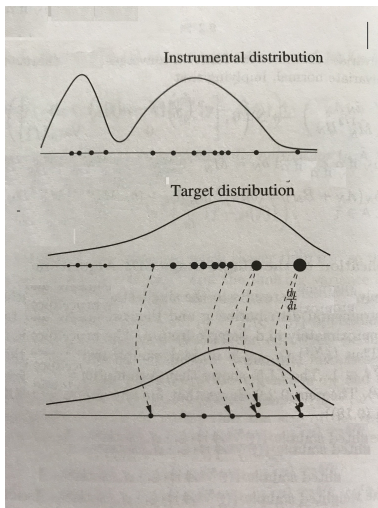
You can draw from these, with replacement. At the expense of more variance, it will give you unweighted draws from your target distribution:  $\tilde{\theta}^1, \tilde{\theta}^2, \dots$

This is known as **factored sampling** or **importance sampling with resampling** or **sampling importance resampling** (SIR).

# Adding Resampling

"10"

Taylor



Original, unedited image is from

<https://www.springer.com/us/book/9780387402642>

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

Stage 1: do importance sampling to get  $\{\theta^i, w(\theta^i)\}_{i=1}^S$ .

Stage 2: for  $i = 1, \dots, S$ , select

$$P[\tilde{\theta}^i = \theta^j \mid \theta^1, w(\theta^1), \dots, \theta^S, w(\theta^S)] = \frac{w(\theta^j)}{\sum_k w(\theta^k)}.$$

Another way to write it:

Stage 1: do importance sampling to get  $\{\theta^i, w(\theta^i)\}_{i=1}^S$ .

Stage 2: for  $i = 1, \dots, S$ , select indexes

$$P[I^i = j \mid \theta^1, w(\theta^1), \dots, \theta^S, w(\theta^S)] = \frac{w(\theta^j)}{\sum_k w(\theta^k)}$$

and then set

$$\tilde{\theta}^i = \theta^{I^i}.$$

## Example 2 revisited

"10"

Taylor

```
y <- 2 # fake data
log_unnorm_weight <- function(theta){
  # sqrt(2pi) because it will cancel out
  -.5*(y - theta)^2 }
num_samples <- 10000
theta_draws <- rt(num_samples , 1)
lunws <- log_unnorm_weight(theta_draws)
# note: prob arg automatically normalizes
random_indexes <- sample(x = num_samples,
                        size = num_samples,
                        replace = T,
                        prob = exp(lunws))
sort(random_indexes) # repeats!
resampled_draws <- theta_draws[random_indexes]
hist(resampled_draws) # can't do this unless we resample
```

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

# Going sequential

"10"

Taylor

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

Resampling adds variance, so why do it?

It throws away bad samples, and duplicates promising ones.  
When you're looking at a sequence of distribution targets,  
this can have a good effect on future samples.



Resampling adds variance, so why do it?

It throws away bad samples, and duplicates promising ones. When you're looking at a sequence of distribution targets, this can have a good effect on future samples.

**sequential monte carlo** or **particle filtering** methods are basically doing SIR over and over again.

# Examples of sequences of distributions

"10"

Taylor

## Data annealing<sup>1</sup>

$$p(\theta), p(\theta \mid y_1), p(\theta \mid y_{1:2}), \dots, p(\theta \mid y_{1:n}),$$

## Temperature annealing<sup>2</sup>

$$p(y \mid \theta)^{a_0} p(\theta), p(y \mid \theta)^{a_1} p(\theta), \dots p(y \mid \theta)^{a_n} p(\theta)$$


with  $0 = a_0 < a_1 < \dots < a_n = 1$ .

## filtering and smoothing in state space models

$$p(x_1 \mid y_1, \theta), \dots, p(x_n \mid y_{1:n}, \theta)$$

---

<sup>1</sup>Chopin, *A Sequential Particle Filter Method for Static Models*.

<sup>2</sup>Neal, "Annealed Importance Sampling". 

Introduction

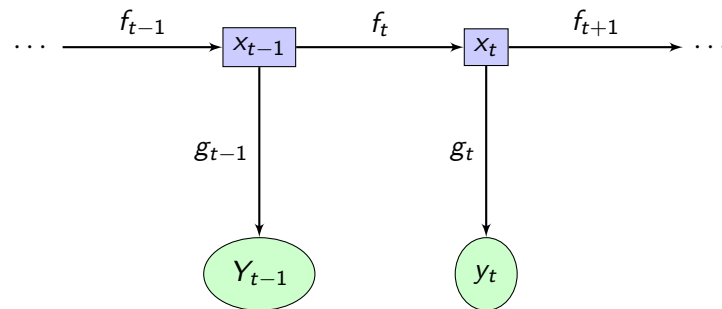
Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References



$$p(x_{1:T}, y_{1:T} \mid \theta)$$

$$= g(y_1 \mid x_1, \theta) f(x_1 \mid \theta) \prod_{t=2}^T g(y_t \mid x_t, \theta) f(x_t \mid x_{t-1}, \theta)$$

# Example: filtering in state space models

Here's an example of a state space model.  $y_t$  is a univariate time series, and  $x_t$  is a hidden/unobserved/latent time series.

$$y_t = \exp(x_t/2)\epsilon_t \quad (1)$$

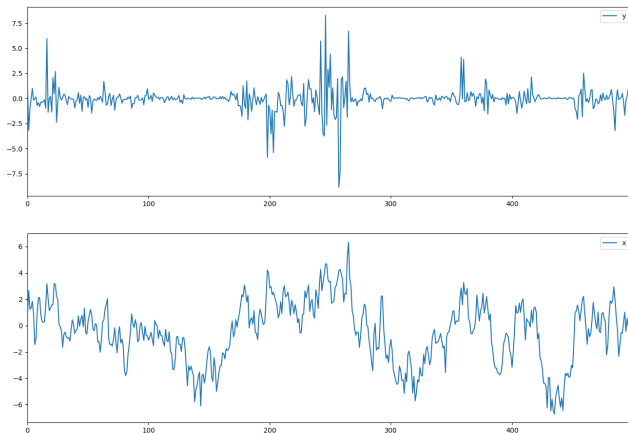
$$x_t = c + \phi x_{t-1} + v_t \quad (2)$$

We sometimes refer to (1) as  $g(y_t | x_t, \theta)$  or the observation equation, and (2) as the state transition equation or  $f(x_t | x_{t-1}, \theta)$ .

[Introduction](#)[Accept-Reject Sampling](#)[Importance Sampling](#)[Importance Sampling with Resampling](#)[Sequential Monte Carlo](#)[References](#)

# Example: filtering in state space models

$y_{1:t}$  observed,  $x_{1:t}$  hidden. Goal:  $p(x_t | y_{1:t})$  in real-time.



"10"

Taylor

[Introduction](#)

[Accept-Reject  
Sampling](#)

[Importance  
Sampling](#)

[Importance  
Sampling with  
Resampling](#)

[Sequential Monte  
Carlo](#)

[References](#)

# Example: filtering in state space models

"10"

Taylor

Introduction

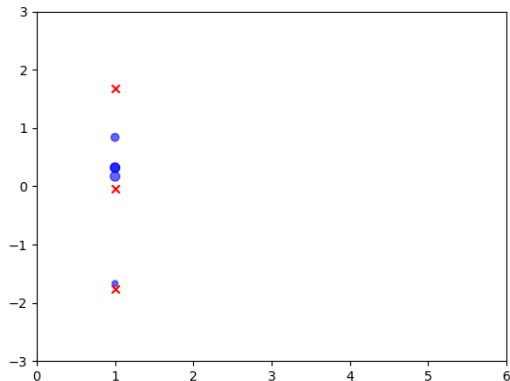
Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

**Sequential Monte  
Carlo**

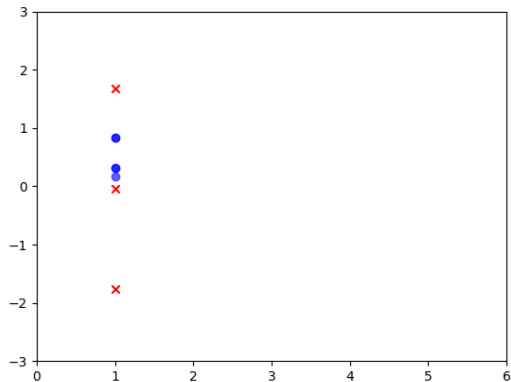
References



# Going sequential

"10"

Taylor



Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

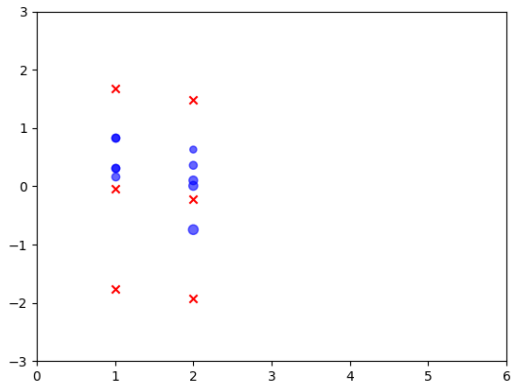
**Sequential Monte  
Carlo**

References

# Going sequential

"10"

Taylor



Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

**Sequential Monte  
Carlo**

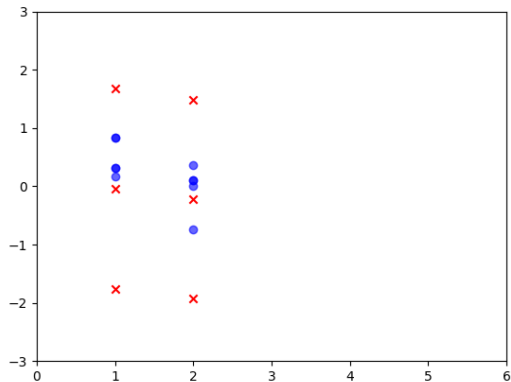
References



# Going sequential

"10"

Taylor



Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

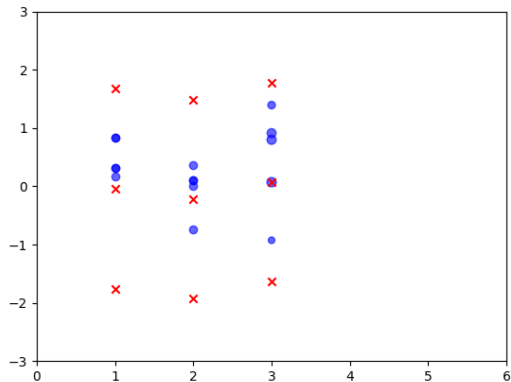
**Sequential Monte  
Carlo**

References

# Going sequential

"10"

Taylor



Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

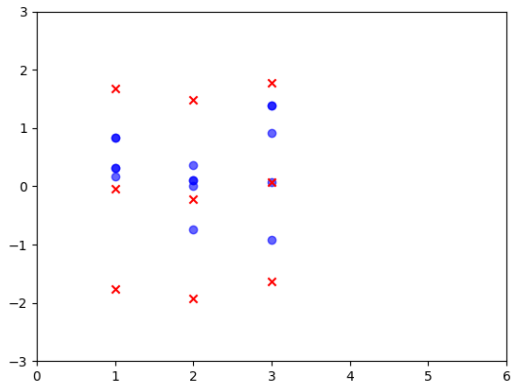
**Sequential Monte  
Carlo**

References

# Going sequential

"10"

Taylor



Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

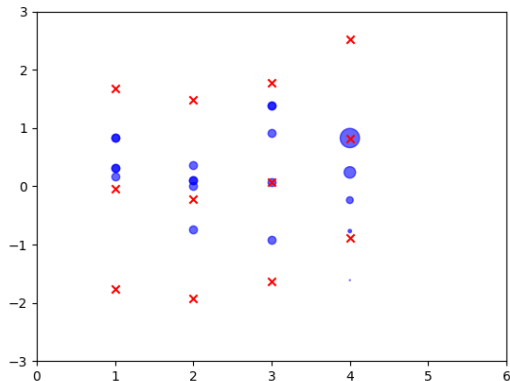
**Sequential Monte  
Carlo**

References

# Going sequential

"10"

Taylor



Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References

# Going sequential

"10"

Taylor

Introduction

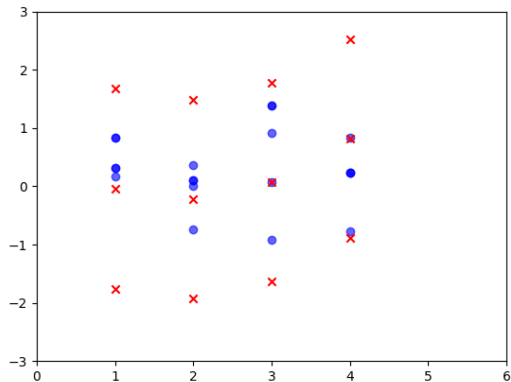
Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

**Sequential Monte  
Carlo**

References



# Going sequential

"10"

Taylor

Introduction

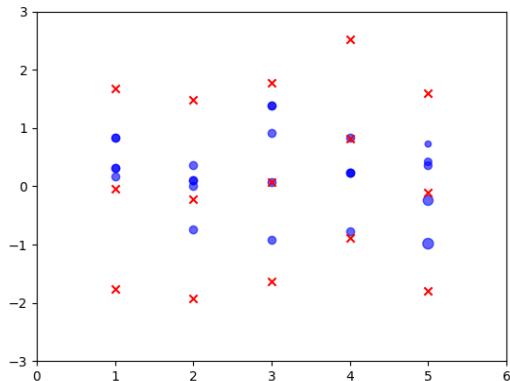
Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

**Sequential Monte  
Carlo**

References



# Going sequential

"10"

Taylor

Introduction

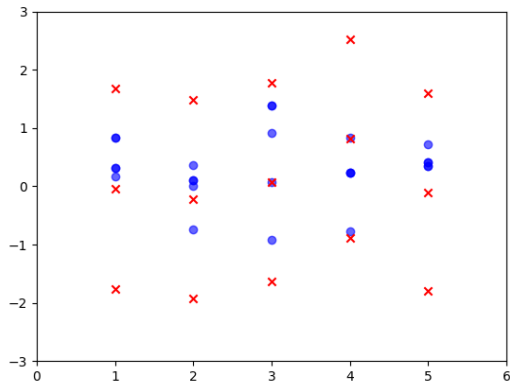
Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References



# Filtering recursions

Drop dependence on  $\theta$  from the notation...

$$\begin{aligned}
 p(x_{1:t}|y_{1:t}) &= C_t^{-1} p(x_t, y_t \mid x_{t-1}) p(x_{1:t-1} \mid y_{1:t-1}) \\
 &= C_t^{-1} \frac{p(x_t, y_t \mid x_{t-1})}{q_t(x_t \mid y_t, x_{t-1})} \times \\
 &\quad q_t(x_t \mid y_t, x_{t-1}) p(x_{1:t-1} \mid y_{1:t-1}) \\
 &= C_t^{-1} \frac{g(y_t|x_t)f(x_t|x_{t-1})}{q_t(x_t|x_{t-1}, y_t)} \times \\
 &\quad q_t(x_t \mid x_{t-1}, y_t) p(x_{1:t-1} \mid y_{1:t-1})
 \end{aligned}$$

Repeat through time:

1. start with samples from  $p(x_{1:t-1} \mid y_{1:t-1})$
2. mutate/propagate/extend using  $q_t(x_t \mid x_{t-1}, y_t)$
3. adjust weights by multiplying by  $\frac{g(y_t|x_t)f(x_t|x_{t-1})}{q_t(x_t|x_{t-1}, y_t)}$
4. resample, giving you particles distributed as  $p(x_{1:t} \mid y_{1:t})$

Introduction

Accept-Reject  
Sampling

Importance  
Sampling

Importance  
Sampling with  
Resampling

Sequential Monte  
Carlo

References



[Introduction](#)[Accept-Reject  
Sampling](#)[Importance  
Sampling](#)[Importance  
Sampling with  
Resampling](#)[Sequential Monte  
Carlo](#)[References](#)

Chopin, Nicolas. *A Sequential Particle Filter Method for Static Models*. 2000.

Neal, Radford M. "Annealed Importance Sampling". In: *Statistics and Computing* 11.2 (Apr. 2001), pp. 125–139. ISSN: 0960-3174. DOI: 10.1023/A:1008923215028. URL: <https://doi.org/10.1023/A:1008923215028>.