Controlled error for combinatorial structures

Better estimation through Gibbs

Mark Huber
Fletcher Jones Associate Professor of Mathematics and Statistics and George R. Roberts Fellow
Claremont McKenna College

Estimate integrals/sums

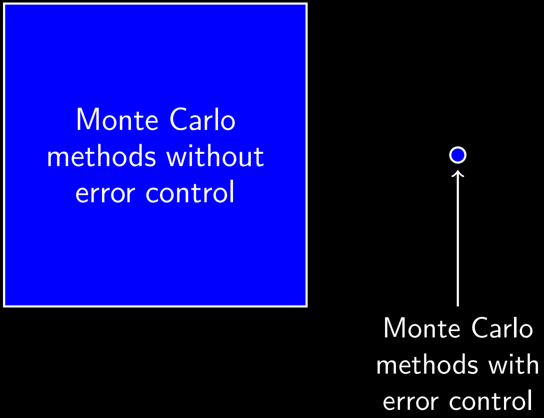
$$Z = \sum_{i \in \Omega} f(i)$$

to within relative error ϵ with

$$O(\ln(Z)\ln(\ln(Z))\epsilon^{-2})$$

samples

Error control


Monte Carlo methods without error control
Monte Carlo methods with error control

The goal is to estimate

$$Z = \int_{\Omega} f(x)\ d\mu$$

Lebesgue measure

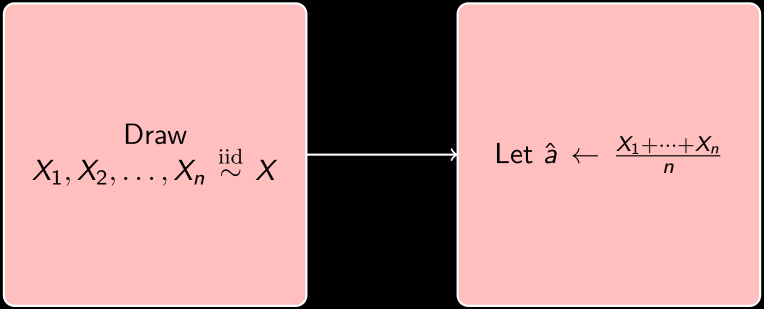
$$\int_{\Omega} f(x)\ d\mathbb{R}^n$$

counting measure

$$\sum_{i \in \Omega} f(i)$$

Basic Monte Carlo method

To estimate a, create r.v. X with $\mathbb{E}[X] = a$



Error in $\hat{a}$ is $\mathrm{SD}(X)/\sqrt{n}$

The problem

We do not know $\mathrm{SD}(X)$!

Typically forced to estimate

Estimation does not really work

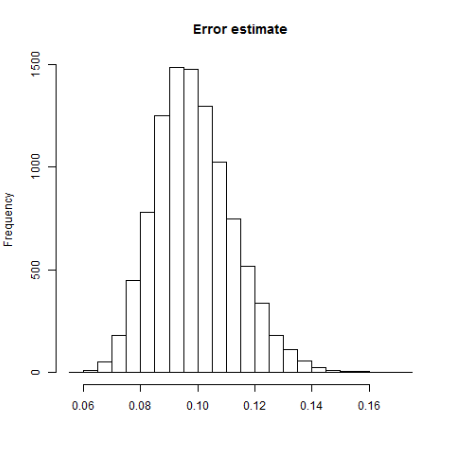
A simple example

$$Z = \int_0^\infty x \exp(-x) \; dx$$

$$X_1, X_2, \ldots, X_{100} \leftarrow \mathrm{Exp}(1)$$

True error in estimate is 0.1

Histogram of error results



Chance that error estimate below 0.1 is $\approx 57\%$

How to get around this problem

1. Convert to estimating partition function of a Gibbs distribution
2. Create well balanced cooling schedule
3. Then product estimator from importance sampling automatically has small variance

Result, # of samples needed is

$$O(\ln(Z)\ln(\ln(Z)))$$

Just to compare...

Let $q = \ln(Z)$, main factor for # of samples

1. Stefankovič, Vempala, Vigoda 2009

$$10^{10} q \ln(q)^5 + O(q)$$

2. TPA, H. & Schott 2010

$$2q^2$$

3. Today

$$10.7 q \ln(q) + O(q)$$

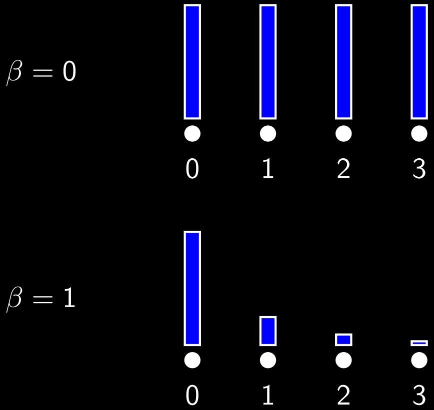
Gibbs distributions

X has a *Gibbs distribution* with parameter β if there exists a function H where

$$\mathbb{P}(X = x) = \frac{\exp(-\beta H(x))}{Z(\beta)}.$$

Graphically

$H(x) = x$



Finding a ratio...

Usually know either $Z(0)$ or $Z(\infty)$, so want to estimate

$$\frac{Z(\infty)}{Z(0)}$$

Importance sampling

For $\beta_a < \beta_b$, X drawn with parameter β_a, set

$$W = \exp(-H(X)(\beta_b - \beta_a))$$

Then

$$\mathbb{E}[W] = \frac{Z(\beta_b)}{Z(\beta_a)}$$

so unbiased estimate of ratio

The relative variance is

$$\frac{\mathbb{V}(W)}{\mathbb{E}[W]^2} = \frac{Z(\beta_b + (\beta_b - \beta_a))Z(\beta_a)}{Z(\beta_b)^2}$$

1. Involves $Z(\beta)$ function for β not in $[\beta_a, \beta_b]$
2. Can be very large!

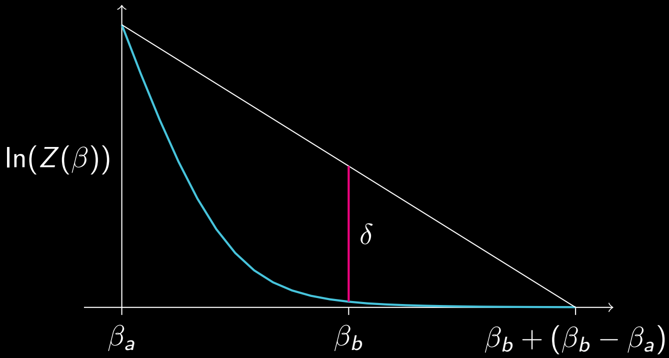
Need three ideas to make scheme work

1. Put in extra β values (a *cooling schedule*):

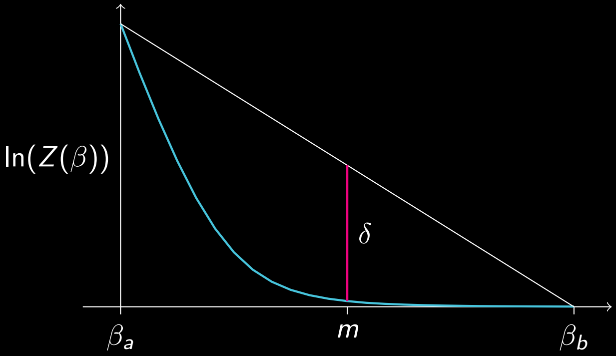
$$\frac{Z(\infty)}{Z(0)} = \frac{Z(\beta_1)}{Z(0)} \frac{Z(\beta_2)}{Z(\beta_1)} \cdots \frac{Z(\infty)}{Z(\beta_{\ell-1})}$$

   (Valleau and Card 1972: Multistage sampling
   Gelman and Meng 1998: Bridge sampling)
2. Well balanced cooling schedule [H. 2012]
3. Paired estimator [H. 2012]

Graphical representation of relative variance



Paired Estimator



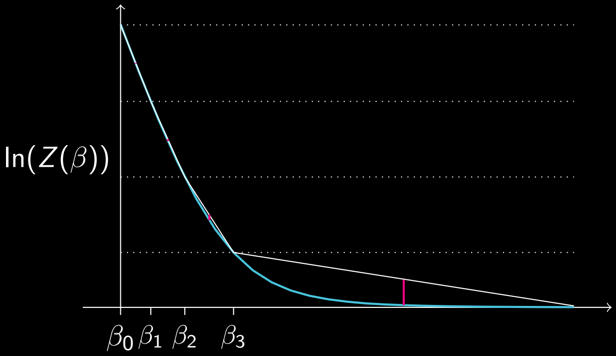
Estimate both $Z(\beta_a)/Z(m)$ and $Z(\beta_b)/Z(m)$
Relative variance δ for both!

Well balanced

A cooling schedule is *well balanced* when there are constants $0 < a < b < 1$ such that for all i,

$$\frac{Z(\beta_{i+1})}{Z(\beta_i)} \in [a, b]$$

With cooling schedule



Standard dev. of product estimator controlled by sum of δ_i.

Theorem [H. 2012]
Let $z(\beta) = \ln(Z(\beta))$.
For a cooling schedule over $[0, \beta]$ with $z(\beta_{i+1}) - z(\beta_i) \leq \eta$ for all i

$$\mathbb{V}_{\text{relative}}(W) \leq 2e^{\eta}[2z'(\beta)]^{\eta/2}.$$

In particular, for $\eta = 2/(2 + \ln(2\mathbb{E}[H(X)]))$,

$$\mathbb{V}_{\text{relative}}(W) \leq 2e.$$

How do we get a well balanced schedule?

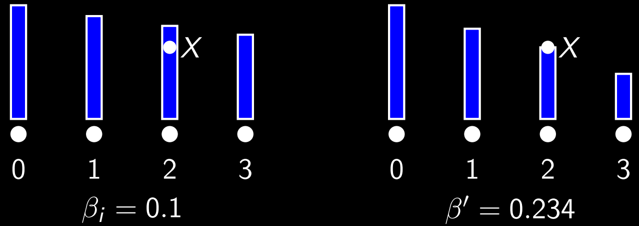
1. Need to draw from many different values of β
2. TPA (H. & Schott 2010) gives such a schedule automatically
3. (Woodard, Schmidler, & H. 2009) Also good for sampling
4. Parallel tempering

Key insight

Currently at β_i

Draw X uniformly from blue region

Let β' be smallest β parameter such that X still in β region



Cool fact [H. & Schott 2012]

When you obtain β' from β_i in this fashion

$$\frac{Z(\beta')}{Z(\beta_i)} \sim \text{Unif}([0,1])$$

Draw several β' use median gets

$$\frac{Z(\beta_{i+1})}{Z(\beta_i)} \approx \frac{1}{2}$$

[Even better, use TPA.]

With these three elements...

1. Put in extra β values (a *cooling schedule*):

$$\frac{Z(\infty)}{Z(0)} = \frac{Z(\beta_1)}{Z(0)} \frac{Z(\beta_2)}{Z(\beta_1)} \cdots \frac{Z(\infty)}{Z(\beta_{\ell-1})}$$

   (Valleau and Card 1972: Multistage sampling)
2. Well balanced cooling schedule [H. 2012]
3. Paired estimator [H. 2012]

Number of samples

$$10.7q \ln(q) + O(q)$$

Example: Linear Extensions

Example: Linear extensions of a poset

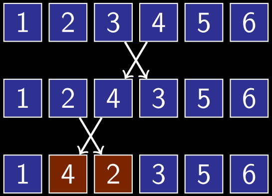
- Linear extensions = permutations with consistent restrictions
- Ex: 2 preceeds 4, 4 preceeds 6
- (Couldn't add 6 preceeds 2)

1	2	3	4	5	6	Is a linear extension
1	4	3	2	5	6	Is not a linear extension (violates 2 before 4)

- Morton, Pacheter, Shiu, Sturmfels, Wienand (2008): All convex rank tests reduce to counting linear extensions

Simple Metroplis-Hastings chain mixes rapidly

- Kharzanov-Khachiyan chain
- Pick random position transpose item and next item...
- ...unless violate restrictions



Accept move-new state valid

Reject move-new state invalid

This chain fairly fast

- Wilson showed this chain rapidly mixing
- Can perfectly sample uniformly from linearly extensions in $O(n^3)$ time (H. 2006)

To continuize...

- Create *home* linear extension
- Example: home $= 123456$
- let $\beta \in [0, 1]$
- Given linear extension X,

$$[Y|X] \sim \text{Unif}([0, \beta^{\#\text{number of items to right of home}}])$$

- 124356 has 3 to the right of its home, so $Y \sim \text{Unif}([0, \beta])$
- When $\beta = 1$, uniform over linear extensions
- when $\beta = 0$, only home position has mass

Delicate situation

- Simple approach does *not* work
- let $\beta \in [0, 1]$
- Given linear extension X,

$$[Y|X] \sim \text{Unif}([0, \beta^{\#\text{number of items to right or left of home}}])$$

Conclusions

- Many problems can be embedded in Gibbs framework
- Care must be taken to make sure fast chains...
- ...stay fast after adding in β