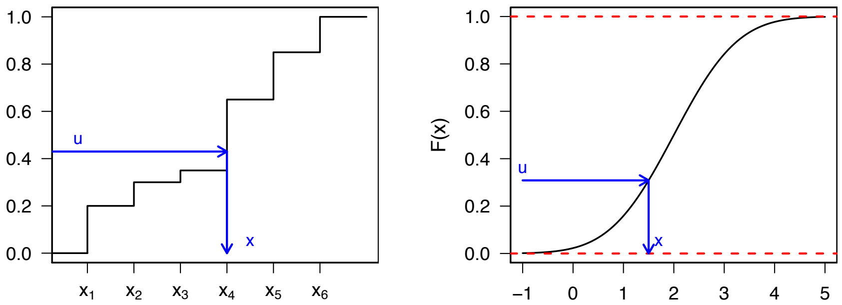
Summary of TMA4300

What did we do?

We had three blocks:

- Simulation
- Markov chain Monte Carlo and INLA
- Bootstrap and EM-algorithm

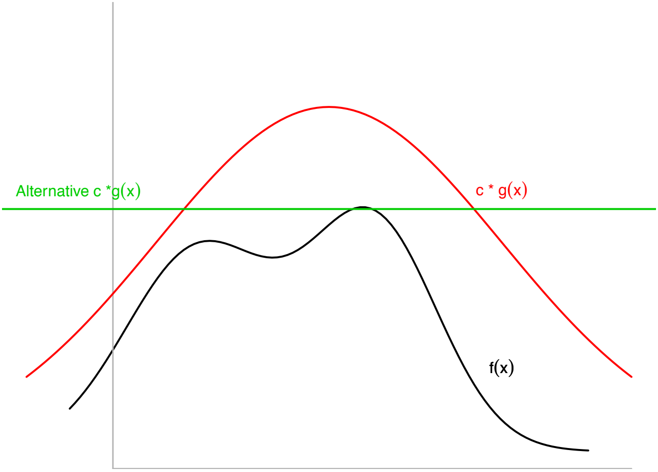
Block 1



What else ...?

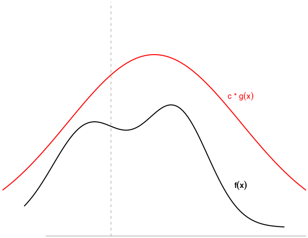
- Variable transformation
  - Univariate (e.g. location and scale transformation)
  - Bivariate (e.g. the Box-Muller algorithm)
- Ratio-of-uniforms method
- Methods based on mixtures

Rejection Sampling: Do you remember this figure?



Refinements: Make the envelope adaptive (different approaches)

Rejection Sampling: When do we accept?

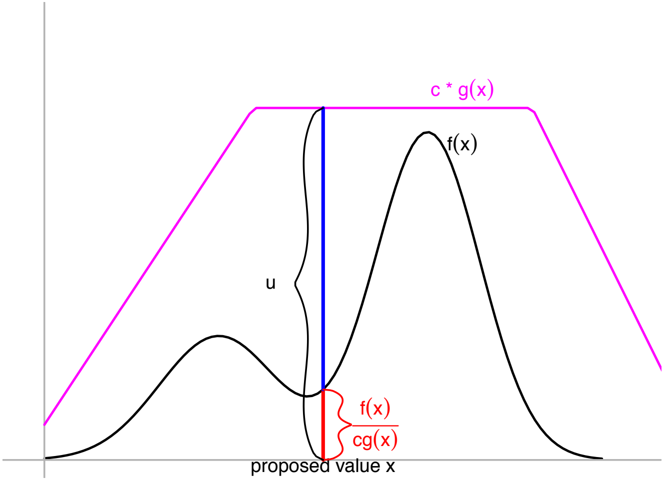


Let $x \sim g(x)$ and $u \sim \text{Unif}(0, 1)$ the

$$\begin{cases} u \leq \frac{1}{c}\frac{f(x)}{g(x)} & \text{keep} \\ u > \frac{1}{c}\frac{f(x)}{g(x)} & \text{reject} \end{cases}$$

- What is the overall acceptance probability?
- Do we need to know the normalizing constant of $f(\cdot)$?

Rejection sampling



Why do we want samples?

Often we would like to approximate a statistic that is difficult to compute directly.

Keywords:

- Monte Carlo integration
- Importance sampling
  - What is it?
  - When is it particularly usefull?

Importance sampling

We are interested in

$$\mu = E_f(h(x)) = \int h(x) f(x) dx$$

- If possible compute it analytically!
- If we can sample from $f(x)$ we can use Monte Carlo integration
- Possible alternative: Importance sampling
  - sample from ausiliary distribution $g(x)$ and re-weight
  - can be used as variance-reduction technique

Importance sampling Algorithm

Let $x_1, \ldots, x_n \sim g(x)$, and let $w(x_i) = \frac{f(x_i)}{g(x_i)}$, $i = 1, \ldots, n$ then

$$\hat{\mu}_{IS} = \frac{\sum h(x_i) w(x_i)}{n}$$

- Unbiased
- Consistent
- Need to know the normalizing constant

$$\tilde{\mu}_{IS} = \frac{\sum h(x_i) w(x_i)}{\sum w(x_i)}$$

- Biased for finite n
- Consistent
- Self-normalizing

Bayesian inference

Basics:

- Posterior $\propto$ Likelihood $\times$ Prior
- Prior Choice
- Bayesian hierarchical models:
  - Observation $\pi(\boldsymbol{y}|\boldsymbol{x})$
  - Latent process $\pi(\boldsymbol{x}|\boldsymbol{\theta})$
  - Hyperpriors $\pi(\boldsymbol{\theta})$
- Full-conditional distributions

Block 2: Two big topics

- Markov chain Monte Carlo
- Integrated Nested Laplace Approximarion (INLA)

Markov chain Monte Carlo:

- What is the idea? Can we generate any Markov chain?
  - Construct a Markov chain that converges to the distribution of interest.
- Why do we not use an approach from block 1?
- What kind of different MCMC techniques have we seen?
- Is the algorithm working at all?

Elements of a MCMC algorithm

- Target distribution $\pi(x)$: Given by the problem
- Proposal distribution $Q(y|x)$: Chosen by the user
- Acceptance probability $\alpha(y|x)$: Derived in order to fullfill the detailed balance condition

Some keywords

detailed balance condition, Metropolis-within-Gibbs, random-walk proposal, burn-in, convergence diagnostics, mixing, effective sample size, ...

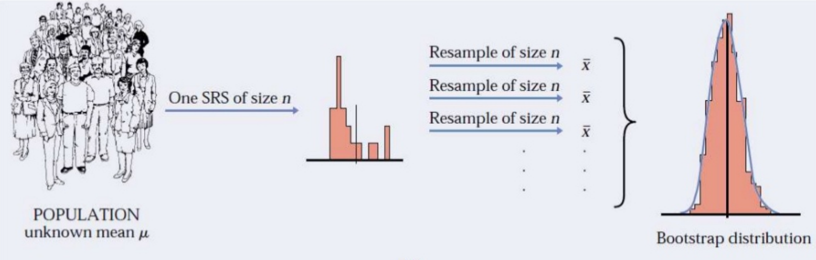
Integrated nested Laplace approximations

- What is the idea?
- For which models does it work?
  - Latent Gaussian Models (with Markov dependence structure)
  - Deterministic approximation instead of simulations
  - Focus on marginal posterior instead of joint posterior
- What are the main "ingredients"
  - Conditional probability

$$\pi(x|z) = \frac{\pi(x,z)}{\pi(z)} \Rightarrow \pi(z) = \frac{\pi(x,z)}{\pi(x|z)}$$

  - Laplace Approximation
    - 2 order Taylor approximarion of $\log g(x) \rightarrow g(x) \approx \mathcal{N}(\hat{x}, \hat{\sigma^2})$
- Potential advantages over MCMC

Bootstrap



One SRS of size n
Resample of size n
Resample of size n
Resample of size n
x̄
x̄
x̄
POPULATION
unknown mean μ
Bootstrap distribution

(c)

Bootstrap

Some keywords:

- Empirical distribution function
- Plug-in principle
- Bootstrap sample
  - Non-parametric
  - Parametric
- Bootstrap Regression
  - Bootstrap residuals
  - Paired-Bootstrap
- Bootstrap time series
  - Model based bootstrap
  - BLock-bootstrap

EM-algorithm

- Goal? Basic idea? What are the steps?
- Field of applications: mixture models, censored data, missing data, hidden models, ...

The exam - 05.06.2016 - 15:00

Permitted aids:

- Calculator HP30S, CITIZEN SR-270X, CITIZEN SR-270X College, Casio fx-82ES PLUS with empty memory.
- Tabeller og formler i statistikk, Tapir forlag.
- K. Rottman: Matematisk formelsamling.
- A dictionary in any language.
- One yellow, stamped A5 sheet with your own handwritten formulas and notes (on both sides).