Lecture 4 Classical Simulation

Fei Tan

Department of Economics
Chaifetz School of Business
Saint Louis University

E6930 Applied Bayesian Statistics

December 26, 2022

The Road Ahead...

- Probability integral transform method
- Composition method
- Accept-reject method
- Importance sampling method
- Using simulated output

Probability Integral Transform

Algorithm 1

Step 1: draw $u \sim \mathbb{U}(0, 1)$

Step 2: return $y = F^{-1}(u)$ as a draw from $f(y)$

- Represent $f(y)$ with $\mathbb{P}(Y \leq y) = F(y)$ by simulating *independent* samples from uniform distribution
  - useful for sampling from truncated $F(y)$: $\frac{F(y)-F(c_1)}{F(c_2)-F(c_1)}$ for $c_1 \leq y \leq c_2$
  - not applicable for multivariate as F is not injective
- Example: $f(y) = \frac{3}{8}y^2$ for $0 \leq y \leq 2$ and 0 otherwise
  - compute $F(y) = \frac{1}{8}y^3$ for $0 \leq y \leq 2$
  - draw $u \sim \mathbb{U}(0, 1) \Rightarrow y = 2u^{\frac{1}{3}} \sim f(y)$

Composition

Algorithm 2

Step 1: draw $y \sim h(y)$

Step 2: draw $x \sim g(x|y) \Rightarrow x \sim f(x) = \int g(x|y)h(y)dy$

- Example: sample regression error $u|\sigma^2 \sim t_\nu(0, \sigma^2)$

$$f(u|\sigma^2) = \int \underbrace{g(u|\lambda, \sigma^2)}_{\mathbb{N}(u|0,\sigma^2/\lambda)} \underbrace{h(\lambda)}_{\mathbb{G}(\lambda|\nu/2,\nu/2)} d\lambda$$

- Finite mixture distribution

$$f(x) = \sum_{i=1}^{K} p_i f_i(x), \qquad \sum_{i=1}^{K} p_i = 1$$

Accept-Reject

Algorithm 3

Step 1: draw $y \sim g(y)$

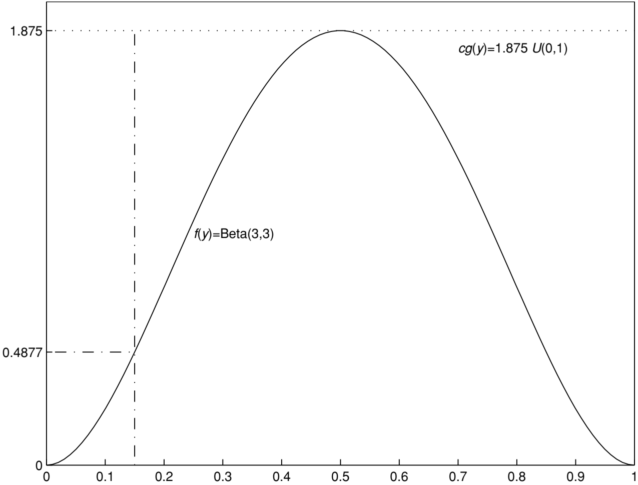
Step 2: draw $u \sim \mathbb{U}(0,1)$

Step 3: accept y as a draw from $f(y)$ if $u \leq \frac{f(y)}{cg(y)}$;

otherwise reject and return to step 1

- Represent target $f(y)$ by simulating *independent* samples from proposal $g(y)$ with $f(y) \leq cg(y)$ for some $c \geq 1$
  - $1/c$ = probability of acceptance $\Rightarrow$ choose small c
  - difficult to find proposal in multivariate case
- Example: sample $y \sim \mathbb{B}(3,3)$
  - choose proposal $\mathbb{U}(0,1)$
  - set $c = f(.5)/g(.5) = 1.875$

Accept-Reject (Cont'd)



- Efficient sampler tailors proposal to mimic target

Importance Sampling

Algorithm 4

$$\mathbb{E}[g(X)] \approx \frac{1}{G}\sum_{g=1}^{G} g(x^{(g)})\underbrace{f(x^{(g)})/h(x^{(g)})}_{\text{importance weight}}, \quad \{x^{(g)}\}_{g=1}^{G} \sim h(x)$$

- Monte Carlo integration: estimate $\mathbb{E}[g(X)] = \int g(x)f(x)dx$ by simulating *independent* samples from proposal $h(x)$
  - efficiency requires tailoring $h(x)$ to $f(x)$
  - why Gaussian is not suitable for $h(x)$? (thin tails)
- Example: $\mathbb{E}[(1+x^2)^{-1}]$, $x \sim$ Exponential(1) truncated to $[0,1]$
  - step 1: sample $\{x^{(g)}\}_{g=1}^{G} \sim \mathbb{B}(2,3)$
  - step 2: compute $\frac{1}{G}\sum_{g=1}^{G} \frac{1}{1+(x^{(g)})^2} \frac{e^{-x^{(g)}}}{1-e^{-1}} \frac{\mathbb{B}(2,3)}{x^{(g)}(1-(x^{(g)})^2)}$

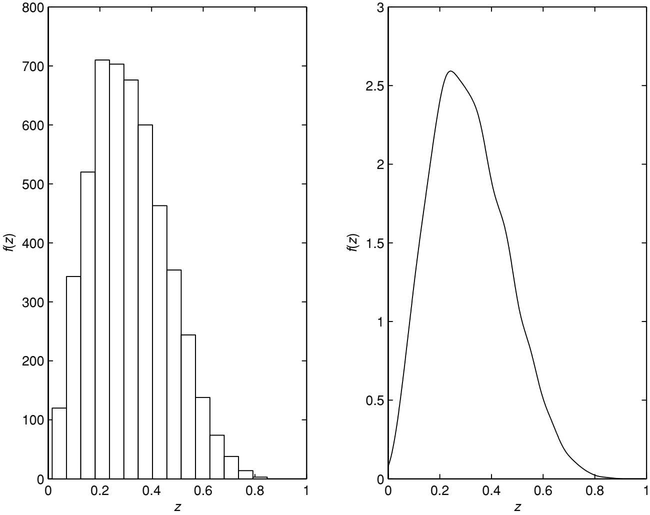
Using Simulated Output

- Use $\{y^{(g)}\}_{g=1}^{G} \sim f(y)$ to investigate properties of $f(y)$, e.g.
  - approximate distribution of $X = h(Y)$, e.g. moments; numerical standard error (n.s.e.) = $\sqrt{\mathbb{V}(X)/G}$
  - 90% credible set: $0.05G$-th & $0.95G$-th ordered $y^{(g)}$
  - marginal (column) vs. joint (row) distribution

$$\{\theta^{(g)}\}_{g=1}^{G} = \begin{bmatrix} \theta_1^{(1)} & \theta_2^{(1)} & \cdots & \theta_d^{(1)} \\ \theta_1^{(2)} & \theta_2^{(2)} & \cdots & \theta_d^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_1^{(G)} & \theta_2^{(G)} & \cdots & \theta_d^{(G)} \end{bmatrix}$$

- Example: interested in learning distribution of $Z = XY$, $X \sim \mathbb{B}(3,3)$ and $Y \sim \mathbb{B}(5,3)$ are independent
  - sample $\{x^{(g)}\}_{g=1}^{G}$, $\{y^{(g)}\}_{g=1}^{G}$, compute $z^{(g)} = x^{(g)}y^{(g)}$
  - $\{z^{(g)}\}_{g=1}^{G}$ represent distribution of Z

Using Simulated Output (Cont'd)



- Histogram (left) vs. kernel-smoothed density (right)

Readings

- To be added...