ORIE 4580/5580: Simulation Modeling and Analysis

ORIE 5581: Monte Carlo Simulation

Unit 5: generating non-uniform random variables

Sid Banerjee

School of ORIE, Cornell University

random variable generation

random number: a sample from $U[0,1]$

modern PRNGs (like `np.random.rand()`) are

- random enough for your simulation to be correct
- deterministic enough (by setting seed) for your simulation to be repeatable

the 'fundamental theorem' of simulation

can 'transform' a stream of i.i.d. $U[0,1]$ into

– a random variable with any given cdf

– a random vector with any given correlation matrix

– any stochastic process

generating rvs with arbitrary distributions

aim: "transform" $U[0,1]$ rv to another rv with given probability distribution.

monte carlo sampling techniques

basic methods

- inversion
- acceptance-rejection
- distribution-specific techniques (Box-Muller for Gaussians)
- advanced techniques (adaptive rejection sampling, SIR)

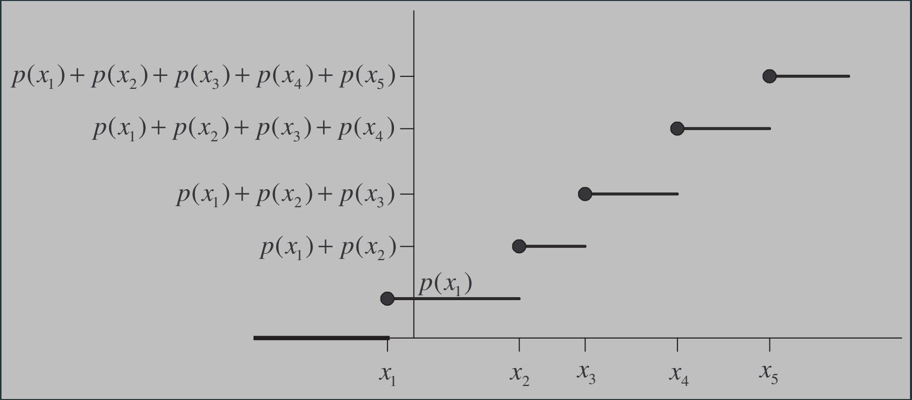
markov-chain monte carlo (MCMC)

- Metropolis-Hastings
- Gibbs sampling
- advanced techniques (slice sampling, Hamiltonian MC)

inversion

warm-up: simulating discrete rv

X takes values $x_1 \leq x_2 \leq \ldots \leq x_5$, $\mathbb{P}[X = x_i] = p(x_i)$



the inversion method

X continuous r.v. with pdf f and c.d.f. $F(\cdot)$

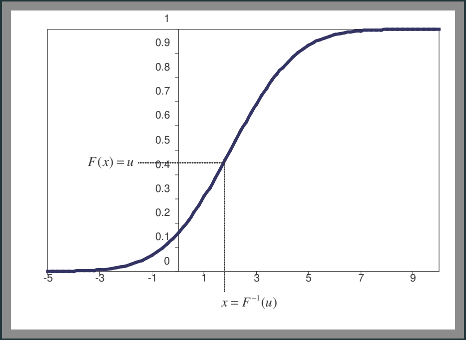
- want to generate samples of X.
- $F(\cdot)$ non-decreasing $\implies$ can define inverse $F^{-1}(\cdot)$
- $F(x) = u \iff F^{-1}(u) = x$

inversion method

given desired cdf F (continuous, increasing), generate sample $X_0 \sim F$ as:

1. generate $U \sim U[0,1]$.
2. return $X_o = F^{-1}(U)$.

intuition/proof for inversion method



example

example – the pdf of X is given by

$$f(x) = \begin{cases} 1/3 & \text{if } 0 \leq x \leq 1 \\ 2/3 & \text{if } 1 < x \leq 2 \\ 0 & \text{otherwise.} \end{cases}$$

develop an inversion method to generate samples of X.

example (contnd)

example (exponential rv)

generate samples of an exponential r.v. with parameter λ, with cdf

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

example (geometric distribution)

generate samples of a geometric distribution with pdf

$$p(k) = (1-p)^{k-1}p \qquad \text{for } k = 1, 2, \ldots$$

drawback of inversion method

- inversion method may be computationally expensive.
- computing $F^{-1}(\cdot)$ may require numerical search.

example – the pdf of X is given by

$$f(x) = \begin{cases} 60x^3(1-x)^2 & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

$$F(x) = 15x^4 - 24x^5 + 10x^6 \qquad \text{for } 0 \le x \le 1.$$

generate samples of X by using the inversion method.

acceptance-rejection

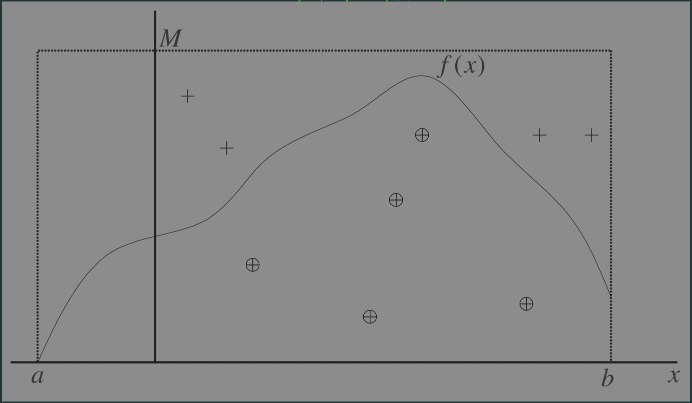
acceptance-rejection

want to generate samples of a rv X

- pdf $f(\cdot)$ of X takes positive values only over $[a, b]$
- M is an upper bound on pdf of X, i.e., $M \geq \max_{x\in[a,b]} f(x)$

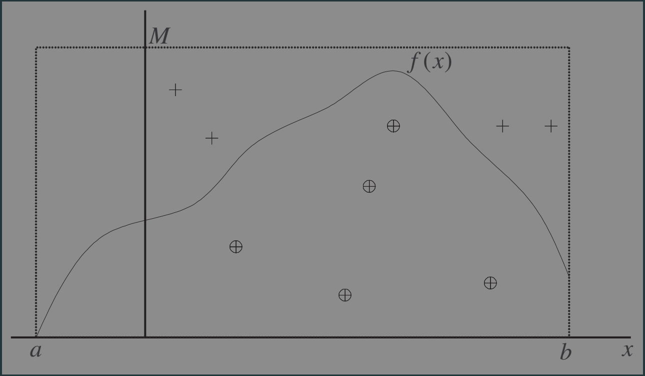
  $\Rightarrow$ can enclose pdf in the rectangle

$$[a, b] \times [0, M]$$



acceptance-rejection

want samples of a rv $X \in [a, b]$, with pdf $f(x) \leq M$



acceptance-rejection sampling

1. generate $U_1, U_2 \sim U[0,1]$
2. set $Z_1 = a + (b-a)U_1$, $Z_2 = MU_2$
3. if $Z_2 \leq f(Z_1)$, return $X_o = Z_1$; else, reject and repeat

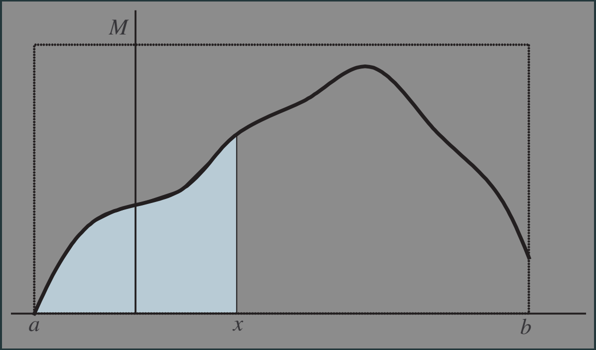
AR sampling: proof of correctness

let X_o denote the output of the AR method for cdf F

- $F_{X_o}(x) = \mathbb{P}[X_o \leq x] =$

AR sampling: proof of correctness

observe: $\mathbb{P}[Z_1 \leq x,\ Z_2 \leq f(Z_1)] = \mathbb{P}[(Z_1, Z_2) \in \text{shaded region in figure}]$



AR sampling: running time

how many $U[0,1]$ samples do we need for one sample of X?

example: X has pdf

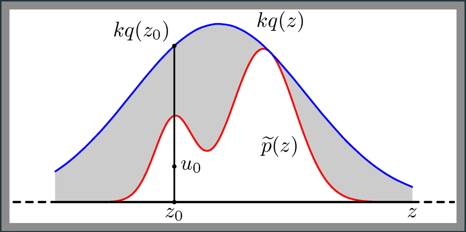
$$f(x) = \begin{cases} 6x(1-x) & \text{if } 0 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

for rejection sampling, we choose

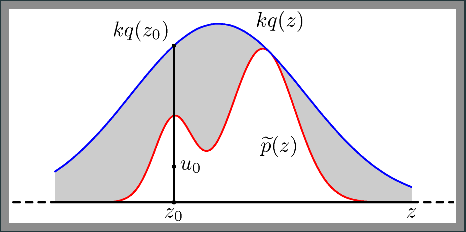
$a =$ $b =$ $M =$

on average, per sample we require

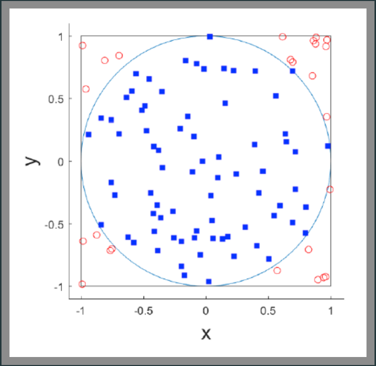
generalized AR sampling



generalized AR sampling



(even more) generalized AR sampling



AR sampling: challenges in high dimensions