$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ $\int_0^x f(t)\,dt = F_X(x)$

The Acceptance-Rejection Method

Chapter 2 (2)

STATS 102C: Introduction to Monte Carlo Methods

UCLA

CC BY NC SA

Introduction

- This *indirect* method will allow us to simulate virtually any distribution, and only require us to know the form of the *target density* f up to a multiplicative constant.
- The main idea is that we generate a candidate random variable from a simpler density g (called the *instrumental* or *candidate* density) and only accept it subject to passing a test. The only constrains we impose on this candidate density g are that

envelope

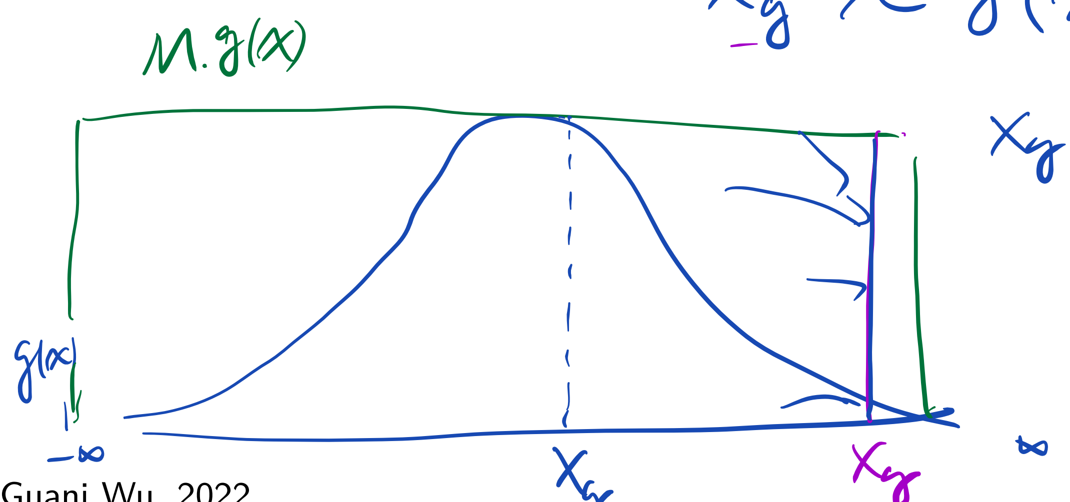
  (i) $f(x) > 0$ and $g(x) > 0$

  (ii) There is a constant M with $\frac{f(x)}{g(x)} \leq M$ for all x.

$u \sim \text{Unif}(0,1)$

$X_g \sim g(x|\theta_g)$

$0 \leq \boxed{\frac{f(x)}{M\,g(x)}} \leq 1$



$X_g \sim \text{Unif}(-\infty, \infty)$

$u \sim \text{Unif}(0,1)$

$u > \frac{f(x_g)}{M.g(x_g)}$ most likely reject X_g

Rejection Method Algorithm

$u \sim Unif(0,1)$

$u < \frac{f(x_g)}{Mg(x_g)}$ accept x_g

$M = \max\left(\frac{f(x)}{g(x)}\right)$

(1) Generate $X_g \sim g(x)$, $U \sim Uniform(0, 1)$;

Find M

(2) Accept X_g if $U < \frac{f(x_g)}{Mg(x_g)}$;

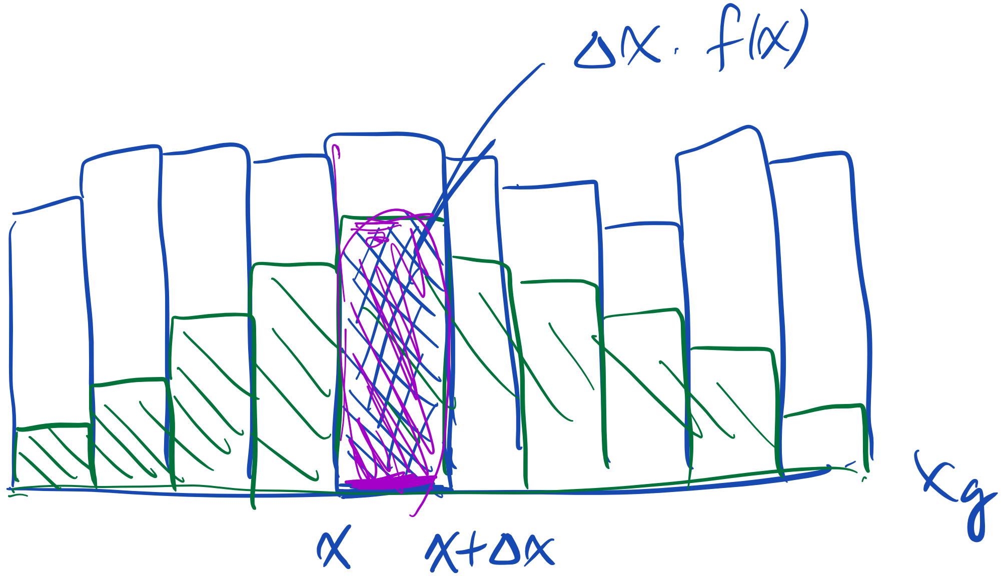
$X_g^{(accepted)} \sim f(x)$

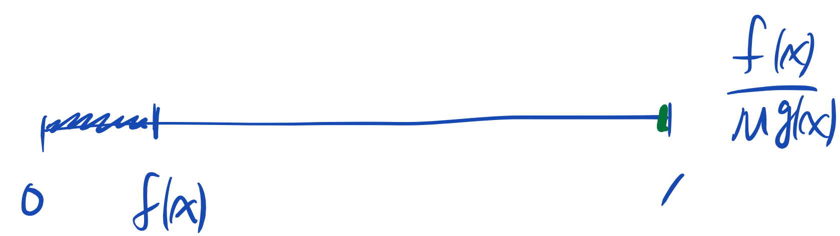
(3) Return to (1) otherwise.

Can we prove that the accepted points $x \sim f(x)$ as desired?

$x_1, x_2, \dots, x_N \sim g(x \mid \theta)$

$u \sim Unif(0,1)$





$= \Delta x \cdot f(x)$

$$f(x) \simeq \frac{\text{prob. of } x \text{ accepted in } (x, x+\Delta x)}{\Delta x} = \frac{P(x \in (x, x+\Delta x))}{\Delta x}$$

$$\underset{x \in (x,\, x+\Delta x)}{p(\text{event})} = \frac{\#\text{ of event}}{\text{sample space}} = \frac{\frac{N}{M}\Delta x\, f(x)}{\frac{N}{M}} = \Delta x \cdot f(x)$$

$$N * g(x) * \Delta x$$

$$\#\text{ of data points accepted in } (x, x+\Delta x) = N \cdot g(x)\, \Delta x\, \frac{f(x)}{M \cdot g(x)}$$

$$= \text{total \# of points in } (x, x+\Delta x) * \underbrace{p(X \text{ being accepted})}_{\frac{f(x)}{M \cdot g(x)}} = \frac{N}{M}\Delta x\, f(x)$$

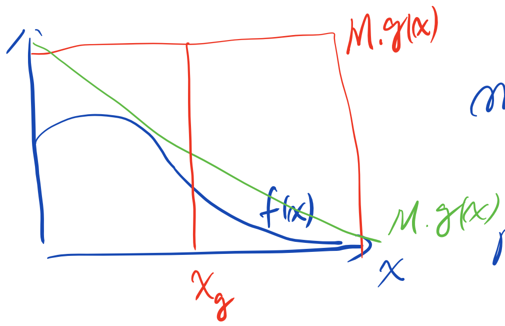
Across the all bins the # of points accepted

$$\sum_{\text{bins}} \frac{N}{M}\Delta x\, f(x) = \frac{N}{M}\sum_{\text{bin}} \Delta x\, f(x) = \frac{N}{M}$$

$$\text{The acceptance rate} = \frac{\#\text{ of data points accepted}}{\text{total data points}} = \frac{N/M}{N} = \frac{1}{M}$$

$$\frac{1}{M}\uparrow \Rightarrow \underline{M \text{ must be small}}$$

$$0 \le \frac{f(x)}{M \cdot g(x)} \le 1 \Rightarrow \frac{f(x)}{g(x)} \le M \quad \forall x$$



$$\max \frac{f(x)}{g(x)} \le M$$

$$M = \max\left[\frac{f(x)}{g(x)}\right]$$

Toy example

X	X_1	X_2
$P(X)$	0.6	0.4

X	X_1	X_2
$g(X)$	0.5	0.5

$$M = \max\left(\frac{P(X)}{g(X)}\right) \qquad \frac{0.6}{0.5} = 1.2 \qquad \frac{0.4}{0.5}$$

① Generate 1000 random samples

$$\frac{500 \quad X_1}{500 \quad X_2}$$

For X_1, $500 * \frac{P(X_1)}{M \cdot g(X_1)} = 500 * \frac{0.6}{1.2(0.5)}$

$= 500$

For X_2, $500 * \frac{P(X_2)}{M \cdot g(X_2)} = 500 \frac{0.4}{(1.2)(0.5)}$

$= 333$

For our sample prob. function

$$P(X_1) = \frac{500}{500+333} = \frac{500}{833} \simeq 0.6$$

$$P(X_2) = \frac{333}{833} \simeq 0.4$$

$$M = \max\left(\frac{f(x)}{g(x)}\right)$$

$X_g \sim g(x) = \text{Unif}(0, \pi/2)$

$U_g \sim \text{Unif}(0,1)$

$X_g = U_g \frac{\pi}{2}$

$$M = \max\left(\frac{f(X_g)}{g(X_g)}\right)$$

$u \sim \text{Unif}(0, 1)$

If $u \le \frac{f(X_g)}{M\, g(X_g)}$ accept X_g

Example 1: $f(x) = \sin x$

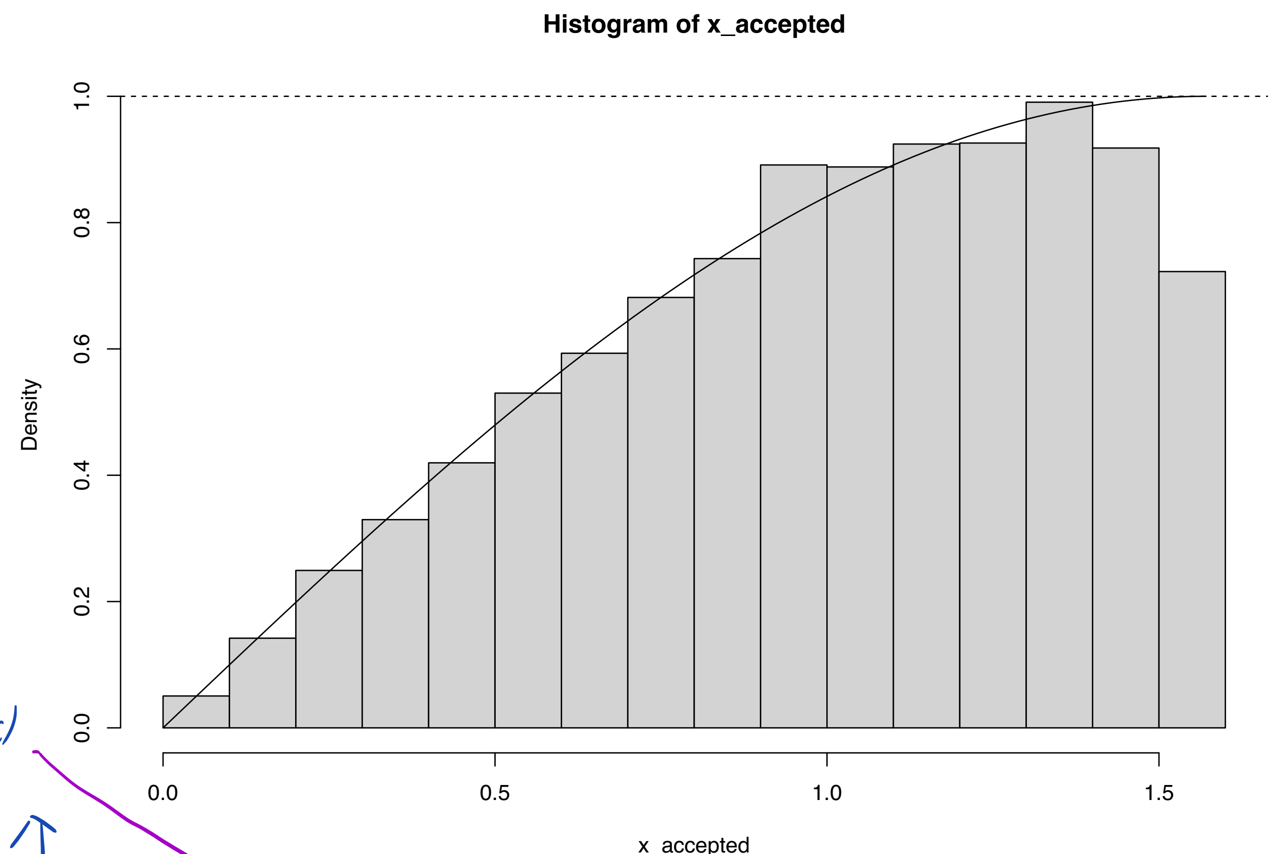
- Let $f(x) = \sin(x)$, for $x \in [0, \pi/2]$.
- $\int_0^{\pi/2} \sin(x)dx = 1$, so $f(x)$ is a PDF on $[0, \pi/2]$

The R code to generate samples from $f(x)$ using Rejection method.

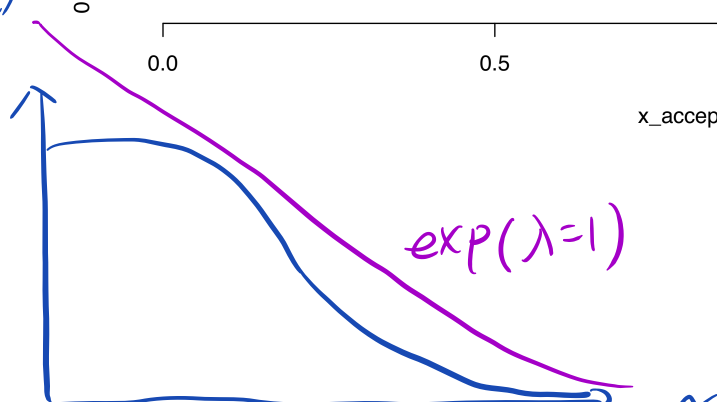
```
set.seed(9999)
n <- 10000
u1 <- runif(n)
x <- u1 * (pi / 2)   这是个uniform dist的 inverse cdf
f_x <- sin(x)
M <- max(f_x) / (2 / pi))
                g(x)
u <- runif(n)
x_accepted <- x[which(u <= f_x / (M * (2 / pi)))]
```

Example 1: $f(x) = \sin x$ (Cont.)

```
hist(x_accepted, probability = T)
curve(sin, from = 0, to = pi/2, add = T)
abline(h = 1, lty = 2)
```



$f(x)$



$g(x) = e^{-x}$

Example 2: Folded Normal Distribution

- Let $Z \sim N(0,1)$ with pdf
$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad \text{for } z \in (-\infty, \infty).$$
- Suppose we want to sample from $X = |Z|$, which has PDF
$$f_X(x) = 2 \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} = \sqrt{\frac{2}{\pi}} e^{-x^2/2}, \quad \text{for } x \geq 0.$$
X has a **folded normal distribution**.
- How can we use rejection sampling to sample from X?

$$M = \max\left(\frac{f(x)}{g(x)}\right) = \max \frac{\sqrt{\frac{2}{\pi}}\, e^{-x^2/2}}{e^{-x}} = \max_x \sqrt{\frac{2}{\pi}}\, e^{-(x^2/2 - x)}$$

$$M = \frac{\sqrt{\frac{2}{\pi}}\, e^{-x^2/2}}{e^{-x}} = 1.315$$

$$\min_x \left(x^2/2 - x\right)$$

$$\Rightarrow x^* = 1$$

R Code to sample from Folded Normal Distribution

写出 F, G

```
f <- function(x){
    return(sqrt(2/pi) * exp(-x^2/2))
}
g <- function(x, lambda){
  return(lambda * exp(-lambda * x))
}

n <- 20000
v <- runif(n, 0, 1)
x <- -log(v)
M <- 1.315
u <- runif(n)
x_accepted <- x[which(u < f(x) / (M * g(x, 1)))]
```

用inverse cdf先找出gx的x

判断 是否小于fx/Mgx

$g(x) = \exp(\lambda = 1)$

$= e^{-x}$

$\min_x \left(\frac{x^2}{2} - x \right)$

$x^* = 1$

$M = \max \left(\frac{f(x)}{g(x)} \right)$

$\frac{\partial}{\partial x} \frac{f(x)}{g(x)} = 0$

R Code to sample from Folded Normal Distribution (Cont.)

```
hist(x_accepted, prob = TRUE, las = 1,
     main = "", ylim = c(0, M))
curve(f(x),  col = "blue", lwd = 2, add = T)
curve(M * g(x, 1), lty = 2, lwd = 2, add = TRUE)
legend("topright", c("f(x)", "M g(x)"),
       inset = 0.1, lty = 1:2,
       lwd = 2, col=c("blue", "black"))
```