

SDS 5531 Homework 1

Gifty Osei

2024-09-11

Remark: If you would like to insert images for your handwritten part into this file, please refer to this article.

Problem 1. Box-Muller transformation

The Box-Muller transformation method simulates random numbers from $N(0, 1)$ as follows.

- Step 1: Generate U_1 and U_2 i.i.d. from $U(0, 1)$.
- Step 2: Let $X_1 = \sqrt{-2 \log U_1} \cos(2\pi U_2)$ and $X_2 = \sqrt{-2 \log U_1} \sin(2\pi U_2)$.

Establish the theoretical validity of the method by proving the following results.

1. (15 points) Use the change-of-variable formula to derive that X_1 and X_2 are two independent draws from $N(0, 1)$.

Solution:

```
knitr::include_graphics(c("D:/WashU/First Year/Sem1/SDS5071_AdvLinearModel/Homework/HW2/Q1.pdf",  
  "D:/WashU/First Year/Sem1/SDS5071_AdvLinearModel/Homework/HW2/Q1b.pdf"))
```

i. To prove that X_1 and X_2 are independent draws from $N(0,1)$

We show that

$$f_{X_1, X_2}(x_1, x_2) = f_{X_1}(x_1) \cdot f_{X_2}(x_2)$$

Using transformation; we recall;

$$X_1 = \sqrt{-2 \log U_1} \cos(2\pi U_2)$$

$$X_2 = \sqrt{-2 \log U_1} \sin(2\pi U_2)$$

find the

$$(U_1, U_2) \longrightarrow (X_1, X_2)$$

using change of variables:

$$\text{Let } \underset{\substack{\uparrow \\ \text{radius}}}{R} = \sqrt{-2 \log U_1} \quad \underset{\substack{\uparrow \\ \text{angle}}}{\Theta} = 2\pi U_2$$

$$f_{U_1, U_2}(u_1, u_2) = 1, \text{ for } u_1 \in (0,1), u_2 \in (0,1)$$

by the inverse transformation;

$$U_1 = e^{-\frac{R^2}{2}} \quad U_2 = \frac{\Theta}{2\pi}$$

by the bivariate transformation;

$$f_{R, \Theta}(r, \theta) = f_{U_1, U_2}(u_1, u_2) \cdot |J|$$

$$|J| = \begin{vmatrix} \frac{\partial X_1}{\partial U_1} & \frac{\partial X_1}{\partial U_2} \\ \frac{\partial X_2}{\partial U_1} & \frac{\partial X_2}{\partial U_2} \end{vmatrix} = \frac{\partial X_1}{\partial U_1} \cdot \frac{\partial X_2}{\partial U_2} - \frac{\partial X_2}{\partial U_1} \cdot \frac{\partial X_1}{\partial U_2}$$

$$f_{R, \Theta}(r, \theta) = f_R(r) \cdot f_{\Theta}(\theta) \\ = \frac{1}{2\pi} e^{-\frac{r^2}{2}}$$

then;

$$f_{X_1, X_2}(x_1, x_2) = \frac{1}{2\pi} e^{-\frac{(x_1^2 + x_2^2)}{2}}$$

$$= g(x_1) \cdot h(x_2)$$

Hence X_1 and X_2 are independent \square .

Figure 1: Problem 1, Question 1

2. (10 points) Show that the polar coordinates $r^2 = X_1^2 + X_2^2 \sim \chi_2^2$, hence $e^{-\frac{r^2}{2}} \sim U(0,1)$, and $\theta = \arctan \frac{X_1}{X_2} \sim U(0, 2\pi)$.

Solution:

```
knitr::include_graphics(c("D:/WashU/First Year/Sem1/SDS5071_AdvLinearModel/Homework/HW2/Q1c.pdf",
"D:/WashU/First Year/Sem1/SDS5071_AdvLinearModel/Homework/HW2/Q1d.pdf"))
```

Let $R = \sqrt{-2 \log U_1}$ and $\Theta = 2\pi U_2$
 then: $X_1 = R \cos \Theta$ and $X_2 = R \sin \Theta$
 $X_1, X_2 \sim N(0,1)$; we can show that
 $X_1^2 + X_2^2 = r^2$ and
 From 1, we show that $X_1, X_2 \sim N(0,1)$ then recall that the sum
 of 2 independent standard normal random variables follows
 a chi-squared with 2 degrees of freedom so;
 $r^2 = X_1^2 + X_2^2 \sim \chi_2^2$
 Now; $U_1 = e^{-\frac{r^2}{2}}$ then $r^2 = -2 \log U_1$ and $U_1 \sim U(0,1)$ so;
 $\therefore P(U_1 \leq v) = P[e^{-\frac{r^2}{2}} \leq v]$
 $= P[r^2 \leq -2 \log v]$
 $= \int_{-2 \log v}^{\infty} \frac{1}{2} e^{-\frac{x}{2}} dx$ b/c pdf of χ_2^2
 $P[U_1 \leq v] = e^{-\frac{-2 \log v}{2}} = v \quad \square$

$\therefore F_{U_1}(u) = u$ then it follows that $U_1 \in (0,1)$ hence
 $U_1 \sim \text{Uniform}(0,1)$

Now given ratio $\frac{X_1}{X_2}$, $X_1, X_2 \sim N(0,1)$

Θ is the angle coordinate for an iid standard normal
 random variables so;

$$f_{\Theta}(\theta) = \frac{1}{2\pi}, \quad \theta \in [0, 2\pi]$$

recall that $\frac{X_1}{X_2} = \frac{r \cos \theta}{r \sin \theta}$

$$\frac{X_1}{X_2} = \tan \theta$$

$$\therefore \theta = \arctan\left(\frac{X_1}{X_2}\right)$$

Figure 2: Problem 1, Question 2

Problem 2. Generate Cauchy random numbers

The Cauchy distribution is Student's t -distribution with 1 degree of freedom.

1. (10 points) Derive an algorithm to simulate random numbers from the Cauchy distribution using the inverse cdf approach. (Hint: Show the Cauchy cdf is $F(x) = \tan^{-1}(x)/\pi$.)

Solution:

```
knitr::include_graphics("D:/WashU/First Year/Sem1/SDS5071_AdvLinearModel/Homework/HW2/Q2a.pdf")
```

Problem 2.

$$\begin{aligned} 1. \quad F(x) &= \int_{-\infty}^x f(t) dt, \quad \text{recall that } f(t) = \frac{1}{\pi(1+t^2)} \\ &= \int_{-\infty}^x \frac{1}{\pi(1+t^2)} dt \\ \therefore F(x) &= \frac{1}{\pi} \tan^{-1}(x) + \frac{1}{2} \quad \text{b/c } \int \frac{1}{1+t^2} = \tan^{-1}(t) \end{aligned}$$

Using the inverse cdf method:

$$\text{Let } U = F(x) \quad ; \quad U = \frac{1}{2} + \frac{\tan^{-1}(x)}{\pi}$$

$$U - \frac{1}{2} = \frac{\tan^{-1}(x)}{\pi} \quad \text{making } x \text{ the subject}$$

$$x = \tan \left[\pi \left(U - \frac{1}{2} \right) \right]$$

\therefore Algorithm:

- Generate a random number U from uniform $(0,1)$
- Compute and simulate the Cauchy random numbers from the inverse cdf as:

$$x = \tan \left[\pi \left(\underset{U(0,1)}{U} - \frac{1}{2} \right) \right]$$

This is a fast approach!

Figure 3: Problem 2, Question 1

2. (10 points) Alternatively, one can simulate from the Cauchy distribution by computing the ratio $\frac{X_1}{X_2}$, where X_1 and X_2 are two independent $N(0,1)$ random variables. Explain why this works.

Solution:

```
knitr::include_graphics("D:/WashU/First Year/Sem1/SDS5071_AdvLinearModel/Homework/HW2/Q2b.pdf")
```

2. We know that $X_1, X_2 \sim N(0,1)$ and we have proved that X_1 and X_2 are independent so
Let

$Y_1 = \frac{X_1}{X_2}$ & $Y_2 = X_2$ Using transformation then;

$$f_{Y_1, Y_2}(y_1, y_2) = f_{X_1, X_2}(y_1 y_2, y_2) \cdot |J|$$

$$= \frac{1}{2\pi} \exp\left[-\frac{y_2^2}{2} (y_1^2 + 1)\right] |y_2|$$

Finding the marginal of Y_1 which is the ratio;

$$f_{Y_1}(y_1) = \int_{-\infty}^{\infty} f_{Y_1, Y_2}(y_1, y_2) dy_2$$

$$= 2 \int_{-\infty}^{\infty} \frac{y_2}{2\pi} \exp\left[-\frac{y_2^2}{2} (y_1^2 + 1)\right] dy_2$$

$$= 2 \left[\frac{1}{2\pi} \cdot \frac{1}{1+y_1^2} \right]$$

$f_{Y_1}(y_1) = \frac{1}{\pi(1+y_1^2)}$ which is the pdf of a Cauchy distribution but hence, this is why simulating the ratio of 2 independent $N(0,1)$ works!.

Figure 4: Problem 2, Question 2

3. (20 points) Implement these two methods in R (or Python). Then compare their computing time and efficiency by simulating n Cauchy random numbers. (Choose n to be reasonably large to be able to tell the time difference in running the two methods. The exact choice of n depends on your hardware and implementation.)

Solution:

1. Implement the two methods. (Make sure your implementation is efficient. For example, avoid loops i

Part 1

Using the Inverse CDF to generate Cauchy random numbers

```
inv_cdf_cauc <- function(n) {
```

```

# Generate n uniform random numbers between 0 and 1
u <- runif(n)
# Using inverse CDF formula
x <- tan(pi * (u - 0.5))
return(x)
}

# 2. Simulate n Cauchy random numbers and compare the execution time of the two methods. (You can find

# Part 2

# Ratio of 2 standard Normals to generate Cauchy RV.
ratio_of_norms_cauc <- function(n) {
# Generate n N(0,1) numbers as x1
X1 <- rnorm(n)

# Generate n N(0,1) numbers as x2
X2 <- rnorm(n)
# ratio of the two
x <- X1 / X2

return(x)
}

# Part 3

## Comparing time to compute
library(microbenchmark)

# set a reasonable n
n <- 10000

# Compare the two methods using micro benchmark

benchmark_results <- microbenchmark(
  inverse_cdf = inv_cdf_cauc(n),

  ratio_of_stan_normals = ratio_of_norms_cauc(n),

  times = 10
)

## Summarize
kable(summary(benchmark_results), caption = "Time to Compute",
      latex_options = c("hold_position",
                        "striped" ))

```

Table 1: Time to Compute

expr	min	lq	mean	median	uq	max	neval
inverse_cdf	198.0	200.7	445.25	205.90	329.0	2242.9	10
ratio_of_stan_normals	637.1	661.0	942.42	905.85	963.3	2005.6	10

```
## We can see that inverse CDF is faster
```

Table 1: From the descriptive summary, we can see that inverse CDF is doing much better in time to compute and generate the Cauchy random variables

```
N <- 2000
system.time(for (i in 1:N) inv_cdf_cauc(n))

system.time(for (i in 1:N) ratio_of_norms_cauc(n))

# 3. For both methods, draw the empirical cdf of your simulated numbers and see how close it is to the
## Plotting All 3 plots (Cauchy, Ratio, Inverse CDF)

## Empirical Values from inverse cdf and ratio
cauchy_samples <- list(
  inverse = inv_cdf_cauc(10000),
  ratio = ratio_of_norms_cauc(10000)
)

## X-values
x_values <- seq(-5, 5, length.out = 10000)

# Calculate the true Cauchy CDF values
true_cdf <- pcauchy(x_values)

## Create Data frame

empirical_cdfs <- map(cauchy_samples, function(samples) {
  data.frame(
    x = sort(samples),
    cdf = seq(1/n, 1, length.out = n)
  )
})

data_true <- data.frame(x = x_values, cdf = true_cdf)

#data_inverse <- data.frame(x = empirical_cdfs$inverse, cdf = seq(1/n, 1, length.out = n))

#data_ratio <- data.frame(x = empirical_cdfs$ratio, cdf = seq(1/n, 1, length.out = n))

ggplot() +
  geom_line(data = data_true, aes(x = x, y = cdf), color = "blue",
    size = 1, linetype = "solid", alpha = 0.8) +
```

```

geom_line(data = empirical_cdfs$inverse ,
          aes(x = x, y = cdf), color = "red", size = 1,
          linetype = "dashed", alpha = 0.5) +
geom_line(data = empirical_cdfs$ratio,
          aes(x = x, y = cdf), color = "green",
          size = 1, linetype = "dotted", alpha = 0.5) + # Restrict x-axis range
xlim(-10, 10) +
ggtitle("Comparison of Simulated Cauchy Distribution CDFs")+
xlab("x")+
ylab("CDF")+
labs(caption = "Blue: True Cauchy CDF,
              Red: Inverse CDF Method, Green: Ratio of Normals Method"
) + theme_bw()

```

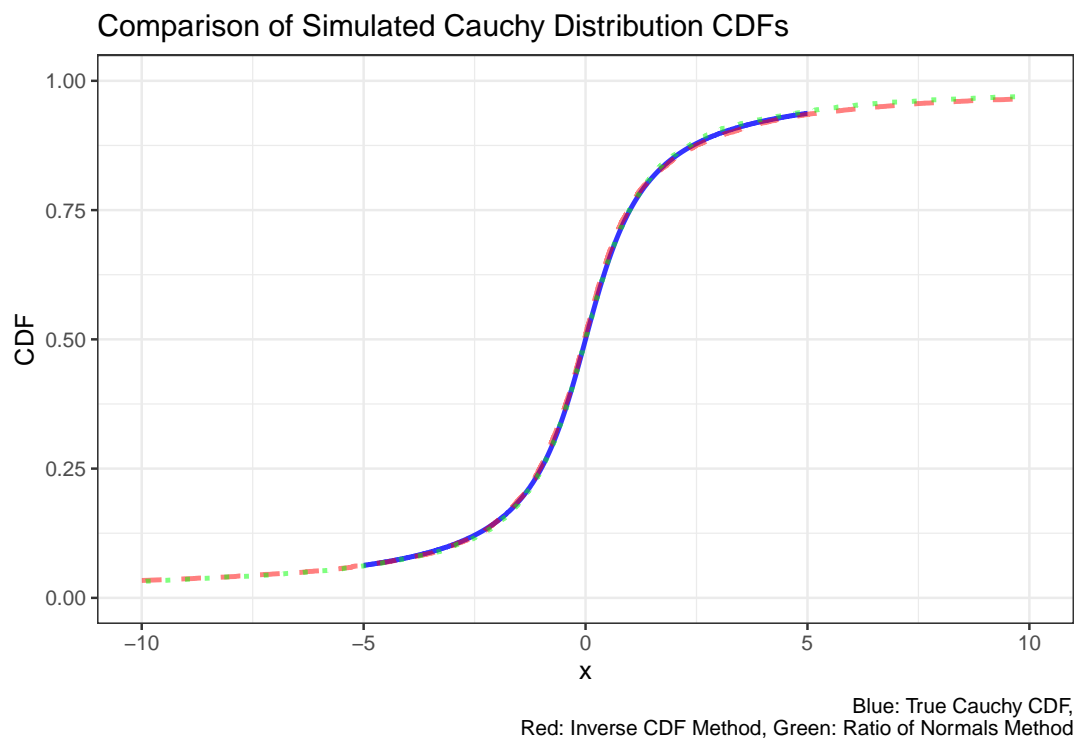


Figure 5: Plot showing the true, inverse cdf and ratio of normals approach of simulating cauchy random number. As n gets bigger over time, both simulation approach converges to the true cdf from a cauchy distribution.

Problem 3. Accept-Reject sampling

Consider simulating from $N(0, 1)$ using the accept-reject sampling. Pretend you do not know the normalizing constant of the pdf, so $f(x) = e^{-\frac{x^2}{2}}$. First, consider using the standard Cauchy distribution as an envelope distribution. Let $g(x) = \frac{1}{1+x^2}$. (Note we have dropped the normalizing constant in the Cauchy pdf.)

1. (10 points) Show that the ratio

$$\frac{f(x)}{g(x)} = (1+x^2)e^{-\frac{x^2}{2}} \leq \frac{2}{\sqrt{e}},$$

with the equality attained at $x = \pm 1$.

Solution:

```
knitr::include_graphics("D:/WashU/First Year/Sem1/SDS5071_AdvLinearModel/Homework/HW2/Q3a.pdf")
```

Problem 3

1. $f(x) = e^{-\frac{x^2}{2}}$ $g(x) = \frac{1}{1+x^2}$ the Envelop: Cauchy dist.

ratio: $\frac{f(x)}{g(x)} = \frac{e^{-\frac{x^2}{2}}}{\frac{1}{1+x^2}}$

$= (1+x^2)e^{-\frac{x^2}{2}} \stackrel{?}{\leq} \frac{2}{\sqrt{e}}$ at $x = \pm 1$

Now let $h(x) = (1+x^2)e^{-\frac{x^2}{2}}$ and find the 1st derivative

$\frac{d[h(x)]}{dx} = \frac{d}{dx} (1+x^2) \cdot e^{-\frac{x^2}{2}} + (1+x^2) \cdot \frac{d[e^{-\frac{x^2}{2}}]}{dx}$

$h'(x) = e^{-\frac{x^2}{2}} [2x - x(1+x^2)]$

$h'(x) = x e^{-\frac{x^2}{2}} [1-x^2]$

Now to find the maximum point, set $h'(x) = 0$

$x e^{-\frac{x^2}{2}} [1-x^2] = 0$ but recall that $e^{\oplus} > 0$ so;

$x[1-x^2] = 0$ $x = 0$ or $x = \pm 1$

when $x = 0$ $h(0) = (1+0^2)e^{-\frac{0^2}{2}} = 1$

when $x = 1$ $h(1) = (1+1^2)e^{-\frac{1^2}{2}} = 2 \cdot e^{-\frac{1}{2}}$

the maximum value of $h(x)$

$h(x) = \frac{2}{\sqrt{e}}$ at $x = \pm 1$

$\therefore (1+x^2)e^{-\frac{x^2}{2}} \leq \frac{2}{\sqrt{e}} \quad \square$

Figure 6: Problem 3, Question 1

2. (15 points) Show that the probability of acceptance is $\sqrt{\frac{e}{2\pi}} \approx 0.66$. Also run an empirical evaluation of the probability of acceptance.

Solution:

```
knitr::include_graphics("D:/WashU/First Year/Sem1/SDS5071_AdvLinearModel/Homework/HW2/Q3b.pdf")
```

2. Probability of acceptance = $P[X \in A]$
 Given as: $P[X \leq x] = P[-1 \leq x | v < f(y)]$

$$= \frac{1}{M} \int_{-\infty}^{\infty} f(x) dx$$

 M is the ratio

$$= \frac{\int_{-\infty}^{\infty} f(x) dx}{M \int_{-\infty}^{\infty} g(x) dx} = \frac{\sqrt{2\pi} \cdot \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} dx}{M \int_{-\infty}^{\infty} \frac{1}{1+x^2} dx \cdot \pi}$$

$$\therefore P[X \leq x] = \frac{\sqrt{2\pi}}{M \cdot \pi} = \frac{\sqrt{2\pi}}{\frac{2}{\sqrt{e}} \cdot \pi}$$

$$= \frac{\sqrt{2\pi}}{\frac{2\pi}{\sqrt{e}}}$$

$$= \frac{\sqrt{2\pi}}{\frac{2\pi\sqrt{e}}{e}}$$

$$P[X \leq x] = \frac{\sqrt{e}}{\sqrt{2\pi}} = \sqrt{\frac{e}{2\pi}}$$

$$\therefore \text{Probability of acceptance} = \sqrt{\frac{e}{2\pi}} \approx 0.66$$

Figure 7: Problem 3 Question 2

```
# 1. implement your algorithm and record the number of acceptances or
# rejections 2. Simulate n=1000 random numbers, and find the empirical
# proportion of acceptances or rejections

set.seed(300)

# Number of simulations
n <- 10000

# Generate n samples from the standard Cauchy distribution using vectorized
```

```

# operations
x <- rcauchy(n)

# target (unnormalized standard normal)
f_x <- exp(-x^2/2)

# envelope density (unnormalized standard Cauchy)
g_x <- 1/(1 + x^2)

# ratio M
M <- 2/sqrt(exp(1))

# Acceptance probability
acceptance_prob <- f_x/g_x

# simulate uniform random numbers for comparison
u <- runif(n)

# Accept if u is less or equal to acceptance_prob / M
accepted <- u <= (acceptance_prob/M)

# empirical proportion of acceptances
acceptance_rate <- mean(accepted)

# acceptance_rate gives the proportion of acceptance printed below

```

Based on the simulation example above, we can show that the empirical proportion of acceptance for a unnormalized Standard Normal and Cauchy is 0.66

3

. (10 points) Now, consider using a scaled Cauchy distribution as the envelope distribution, i.e. $g_\sigma(x) = \frac{1}{\pi\sigma(1+\frac{x^2}{\sigma^2})}$. Find the upper bound for $\frac{f(x)}{g_\sigma(x)}$ and the value of σ that minimizes this bound.

Solution:

```
knitr::include_graphics(c("D:/WashU/First Year/Sem1/SDS5071_AdvLinearModel/Homework/HW2/Q3c.pdf",
"D:/WashU/First Year/Sem1/SDS5071_AdvLinearModel/Homework/HW2/Q3d.pdf"))
```

<p>3. Using a Scaled Cauchy as the envelop distribution;</p> $g_\sigma(x) = \frac{1}{\pi\sigma(1+\frac{x^2}{\sigma^2})}$ <p>ratio: $\frac{f(x)}{g_\sigma(x)}$</p> <p>recall that $f(x) = e^{-\frac{x^2}{2}}$</p> $\frac{f(x)}{g_\sigma(x)} = \frac{e^{-\frac{x^2}{2}}}{\frac{1}{\pi\sigma(1+\frac{x^2}{\sigma^2})}} = \pi\sigma\left(1+\frac{x^2}{\sigma^2}\right)e^{-\frac{x^2}{2}}$ <p>Let $h_\sigma(x) = \pi\sigma\left(1+\frac{x^2}{\sigma^2}\right)e^{-\frac{x^2}{2}}$ and find $h'_\sigma(x)$ as</p> $h'_\sigma(x) = \frac{d}{dx} \left[\pi\sigma\left(1+\frac{x^2}{\sigma^2}\right)e^{-\frac{x^2}{2}} \right]$ $h'_\sigma(x) = \pi\sigma e^{-\frac{x^2}{2}} \left[\frac{x}{\sigma^2} (2 - \sigma^2 - x^2) \right]$ <p>set $h'_\sigma(x) = 0$</p> <p>but recall $e^x > 0$ so;</p> $0 = \frac{x}{\sigma^2} (2 - \sigma^2 - x^2)$ $x = 0 \quad \text{or} \quad 2 - \sigma^2 - x^2 = 0$ $x^2 = 2 - \sigma^2$ <p>the Critical points are $x = 0$ and $x = \pm \sqrt{2 - \sigma^2}$, $\sigma^2 < 2$</p>	<p>when $x = 0$</p> $h_\sigma(0) = \pi\sigma(1+0) = \pi\sigma$ <p>at $x = \sqrt{2 - \sigma^2}$</p> $h_\sigma(x) = \pi\sigma \left[1 + \frac{2 - \sigma^2}{\sigma^2} \right] e^{-\frac{(2 - \sigma^2)}{2}}$ $= \pi\sigma \left[\frac{2}{\sigma^2} \right] e^{-\frac{(2 - \sigma^2)}{2}}$ <p>to find the σ that minimize the upper bound;</p> $h_\sigma(0) = h_\sigma(\sqrt{2 - \sigma^2})$ $\pi\sigma = \frac{2\pi}{\sigma} e^{-\frac{(2 - \sigma^2)}{2}}$ $\sigma^2 = 2 e^{-\frac{(2 - \sigma^2)}{2}}$ <p>after evaluating $\sigma^2 = 1$</p> <p>So $\frac{f(x)}{g_\sigma(x)} = \pi\sigma \left[\frac{2}{\sigma^2} \right] e^{-\frac{(2 - \sigma^2)}{2}}$ is minimized at</p> $\sigma^2 = 1 \quad \text{or} \quad \sigma = \pm \sqrt{1}$
---	---

Figure 8: Problem 3, Question 3