total grade: 10.75/12

Q.1

1) a) let $Y=g(X)$, g is monotone & continuously differentiable with non-zero, as Monotone so $h=g^{-1}$

[2/2]

$F_Y(y)$ for g: (strictly Increasing):

$$F_Y(y)=P(Y\le y)=P(g(X)\le y)=P(X\le h(y))=F_X(h(y)),$$

$\because h(y)=g^{-1}(y)$.

Differentiating to get PDF of Y.

$$f_Y(y)=\frac{d}{dy}F_Y(y)=\frac{d}{dy}F_X(h(y))=f_X(h(y))\cdot h'(y).$$

$\because h'(y)=\frac{1}{g'(h(y))}$,

So,

$$f_Y(y)=\frac{f_X(h(y))}{|g'(h(y))|}$$

for g strictly Decreasing:

$$F_Y(y)=P(Y\le y)=P(g(X)\le y)=P(X\ge h(y))=1-F_X(h(y)).$$

Differentiating:

$$f_Y(y)=-f_X(h(y))\cdot h'(y)=\frac{f_X(h(y))}{|g'(h(y))|}$$

So, PDF of $Y=g(X)$ is:

$$f_Y(y)=\frac{f_X(h(y))}{|g'(h(y))|} \quad y\in g(I) \text{ or } \left|\frac{d}{dy}g^{-1}(y)\right|f_X(g^{-1}(y))$$

for Linear Transformation $g(x)=\gamma x+S$.

If $g(x)=\gamma x+S$ $(\gamma>0)$ so, $g^{-1}(y)=\frac{y-S}{\gamma}$. Thus,

$$F_Y(y)=F_X\left(\frac{y-S}{\gamma}\right) \text{ \& } f_Y(y)=\frac{1}{\gamma}f_X\left(\frac{y-S}{\gamma}\right)$$

blind_copy

1b) For Borel subset $A \subset \mathbb{R}^n$ we need to express $P(Y \in A)$ in terms of Prob. of X.
since $Y = g(X)$. defined through bijective.

$$P(Y \in A) = P(g(X) \in A).$$

$B = g^{-1}(A)$ which is preimage of A under g, then

$$P(Y \in A) = P(g(X) \in A) = P(X \in B)$$

Expressing $P(X \in B)$ in terms of f_X.

$$P(X \in B) = \int_B f_X(x)\, dx.$$

$$\int_B f_X(x)\, dx = \int_A f_X(g^{-1}(y)) \left| \det(J_{g^{-1}}(y)) \right| dy.$$

$J_{g^{-1}}(y)$ is Jacobian Matrix of inverse map g^{-1} at y.
$\det J_{g^{-1}}(y)$ is derivative of Jacobian Matrix.
since:

$$P(Y \in A) = \int_A f_Y(y)\, dy \quad \text{by def. of joint density.}$$

$$f_Y(y) = f_X(g^{-1}(y)) \left| \det(J_{g^{-1}}(y)) \right|, \text{ for } y \in Y := g(X)$$

Thus, $Y = g(X)$ is given by.

$$f_Y(y) = \left| \det(J_{g^{-1}}(y)) \right| f_X(g^{-1}(y)), \quad y \in Y := g(X).$$

{1/2}

Q.2

a) $Y := \gamma X + s$

CDF of Y will be $F_Y(y) = P(Y \le y) = P(\gamma X + s \le y)$.

Case 1: $\gamma > 0$.

$$P(\gamma X + s \le y) = P\left(X \le \frac{y-s}{\gamma}\right)$$

So,

$$F_Y(y) = F_X\left(\frac{y-s}{\gamma}\right)$$ ✓

Case 2: $\gamma < 0$

$$P(\gamma X + s \le y) = P\left(X \ge \frac{y-s}{\gamma}\right)$$

$$F_Y(y) = 1 - F_X\left(\frac{y-s}{\gamma}\right)$$ ✓

So CDF of $Y = \gamma X + s$ is:

$$F_Y(y) = \begin{cases} F_X\left(\frac{y-s}{\gamma}\right) & \text{if } \gamma > 0 \\ 1 - F_X\left(\frac{y-s}{\gamma}\right) & \text{if } \gamma < 0 \end{cases}$$

if $\gamma = 0$?

PDF:

if $\gamma > 0$: $f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X\left(\frac{y-s}{\gamma}\right) = f_X\left(\frac{y-s}{\gamma}\right) \cdot \frac{1}{\gamma}$

if $\gamma < 0$:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy}\left(1 - F_X\left(\frac{y-s}{\gamma}\right)\right) = f_X\left(\frac{y-s}{\gamma}\right) \cdot \frac{1}{|\gamma|}$$

So, PDF of $Y = \gamma X + s$ is:

$$f_Y(y) = \frac{1}{|\gamma|} f_X\left(\frac{y-s}{\gamma}\right)$$

Q.2

b) As X is continuous & differentiable over non-zero so,

$$X = g^{-1}(Y) = e^{Y}$$

$h(y) = g^{-1}(y) = e^{y}$ & it's derivative is also e^{y}.

we know that PDF of $f_Y(y)$ is.

$$f_Y(y) = f_X(h(y)) \left| \frac{d}{dy}(h(y)) \right|$$

$$= f_X(e^{y}) \cdot e^{y}$$

as $f_X(x) = e^{-x}$ for $x > 0$

$$f_X(e^{y}) = e^{-e^{y}}$$

Thus, PDF of $Y = \log(X)$ is:

$$f_Y(y) = e^{-e^{y}} \cdot e^{y} = e^{y} e^{-e^{y}}$$

c) $X \sim \text{Unif}(0,1)$ $\quad Y := \sqrt{X}$

we can write it as $X = g^{-1}(y) = y^2$

derivative of $g^{-1}(y)$ w.r.t y is $\frac{d}{dy} y^2 = 2y$ so

$$\left| \frac{d}{dy} g^{-1}(y) \right| = \frac{1}{2} / \sqrt{y} \quad y \in [0,1]$$

!

Using transformation formula.

$$f_Y(y) = \left| \frac{d}{dy} g^{-1}(y) \right| f_X(g^{-1}(y)) = \cancel{\frac{1}{2\sqrt{y}} \cdot e^{-y^2}}$$

$E[Y]$ using LOTUS:

$$E[Y] = E[\sqrt{X}] = \cancel{\int_0^1 x \cdot f_X(x)\,dx = \int_0^1 x^2\,dx = \frac{1}{3}}$$

or we can also compute $E[Y]$ as.

$$E[Y] = \int_0^1 y \cdot f_Y(y)\,dy = \cancel{\int_0^1 y \cdot \frac{1}{2\sqrt{y}}\,dy = \frac{1}{3}}.$$

the writer has not understood the computation!

Exercice 3

[275/3]

a

$$g(x,y) = (x+y, y) \quad ; \quad \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

inverse Tranformation :

$$g^{-1}(u,v) = (u-v, v)$$

$$J g^{-1}(u,v) = \begin{bmatrix} \frac{\partial}{\partial u} g_1^{-1} & \frac{\partial}{\partial v} g_1^{-1} \\ \frac{\partial}{\partial u} g_2^{-1} & \frac{\partial}{\partial v} g_2^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

- determinant of this matrix :

$$\det\left(J g^{-1}(u,v)\right) = (1)(1) - (-1)(0) = 1$$

- Joint Density :

$$f_{X+Y,Y}(u+v) = \left|\det\left(J g^{-1}(u,v)\right)\right| f_{X,Y}\left(g^{-1}(u,v)\right)$$

$\longrightarrow$ Substituting $g^{-1}(u,v) = (u-v, v)$:

$$= f_{X,Y}(u-v, v)$$

→ Since X and Y are independant, their joint density is the product of their marginals:

$$= f_X(x)\, f_Y(y)$$

Subsituting:

$$f_X(u-v)\, f_Y(v)$$

• Marginal Density of $X+Y$:

– integrate out v

$$f_{X+Y}(u) = \int_{-\infty}^{\infty} f_{X+Y,Y}(u,v)\, dv$$

– Substituting $f_{X+Y,Y}(u,v) = f_X(u-v)\, f_Y(v)$:

$$f_{X+Y}(u) = \int_{-\infty}^{\infty} f_X(u-v)\, f_Y(v)\, dv.$$

⟹ This is exactly the formula we obtained in Exercice 2 Homework 4. ✓

(b) we have:

$$U \sim N(0,2), \qquad V \sim N(0,2)$$

$$g(x,y) = (x+y,\, x-y) \longrightarrow g^{-1}(u,v) = \left(\frac{u+v}{2}, \frac{u-v}{2}\right)$$ ✓

$$J_{g^{-1}}(u,v) = \begin{bmatrix} \frac{\partial}{\partial u} g_1^{-1} & \frac{\partial}{\partial v} g_1^{-1} \\ \frac{\partial}{\partial u} g_2^{-1} & \frac{\partial}{\partial v} g_2^{-1} \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

$\det(J_{g^{-1}}) = -\frac{1}{2}$ ✓

So :

$$f_{U,V}(u,v) = \left|\det(J_{g^{-1}}(u,v))\right| f_{X,Y}(g^{-1}(u,v))$$

$$= \frac{1}{2} f_{X,Y}\left(\frac{u+v}{2}, \frac{u-v}{2}\right)$$

$$= \frac{1}{2} f_X\left(\frac{u+v}{2}\right) f_Y\left(\frac{u-v}{2}\right)$$

$$= \frac{1}{2} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{u+v}{2}\right)^2\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{u-v}{2}\right)^2\right)$$

$$= \frac{1}{\sqrt{2}\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{u}{\sqrt{2}}\right)^2\right) \frac{1}{\sqrt{2}\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{v}{\sqrt{2}}\right)^2\right).$$

$$= f_U(u) f_V(v)$$ ✓

$\Longrightarrow$ we conclude that u and v are independant.

(c) we have : ✗ Comparison to HW3, Ex. 5

$$\bar{X} = \frac{X_1 + X_2}{2}$$

$$S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$$

$$= \frac{1}{2-1}\left(\left(\frac{X_1 - X_2}{2}\right)^2 + \left(\frac{X_2 - X_1}{2}\right)^2\right)$$

$$= \frac{1}{2}(X_1 - X_2)^2$$

we have also:

$$\mathbb{P}(\bar{X} \leq x, S^2 \leq s) = \mathbb{P}\left(\frac{X_1+X_2}{2} \leq x, \frac{1}{2}(X_1-X_2)^2 \leq s\right)$$

$$= \mathbb{P}\left(X_1+X_2 \leq 2x, -\sqrt{2s} \leq (X_1-X_2) \leq \sqrt{2s}\right)$$

$$= \mathbb{P}(X_1+X_2 \leq 2x) \cdot \mathbb{P}(-\sqrt{2s} \leq (X_1-X_2) \leq \sqrt{2s})$$

using Part b:

$$= \mathbb{P}(\bar{X} \leq x) \cdot \mathbb{P}(S^2 \leq s)$$

$\Rightarrow$ we can conclude that $\bar{X}$ and S^2 are independent.

Exercice 4 [2/2]

a X_n is defined as:

$$X_n = \frac{1}{n} \sum_{i=1}^{n} X_i$$

- Expectation is:

$$E[X_n] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \frac{1}{n} \sum_{i=1}^{n} \mu_i \quad \checkmark$$

- Variance:

$$Var(X_n) = Var\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right) = \frac{1}{n^2} \sum_{i=1}^{n} Var(X_i)$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \sigma_i^2. \quad \checkmark$$

- Applying Chebyshev's inequality:

$$\mathbb{P}(|Y - E[Y]| > a) \leq \frac{Var(Y)}{a^2}$$

– in our case $Y = X_n$:

$$\mathbb{P}(|X_n - E[X_n]| > a) \leq \frac{Var(X_n)}{a^2}$$

- Substituting the Variance into the Chebyshev's inequality:

$$\mathbb{P}(|X_n - E[X_n]| > a) \leq \frac{\frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2}{a^2}$$

$$\leq \frac{\sum_{i=1}^{n} \sigma_i^2}{n^2 a^2}$$

- Expectation into the inequality:

$$\mathbb{P}\left(\left|X_n - \frac{1}{n}\sum_{i=1}^{n} \mu_i\right| > a\right) \leq \frac{\sum_{i=1}^{n} \sigma_i^2}{n^2 a^2}$$

(b) we have from (a):

$$\mathbb{P}\left(\left|X_n - \frac{1}{n}\sum_{i=1}^{n} \mu_i\right| > a\right) \leq \frac{\sum_{i=1}^{n} \sigma_i^2}{n^2 a^2}$$

- and we have that $0 \leq \sigma_i^2 \leq M$ for all $i \geq 1$
  where $M < \infty$, so:

$$\sum_{i=1}^{n} \sigma_i^2 \leq nM$$

- substitute $\sum \sigma_i^2 \leq nM$ into the inequality:

$$\mathbb{P}\left(\left|X_n - \frac{1}{n}\sum_{i=1}^{n} \mu_i\right| > a\right) \leq \frac{nM}{n^2 a^2}.$$

$$\leq \underbrace{\frac{M}{na^2}}$$

- as $n \to \infty$, the right hand side $\frac{M}{na^2} \longrightarrow 0$ so:

$$\lim_{n\to\infty} \mathbb{P}\left(\left|X_n - \frac{1}{n}\sum_{i=1}^{n} \mu_i\right| > a\right) = 0$$

- Checking the weak LLN:

for random variables:

$\mu_i = \mu$ for all i so: $\frac{1}{n}\sum_{i=1}^{n} \mu_i = \mu$

and $\sigma_i^2 = \sigma^2$ for all i so:

$$\sum_{i=1}^{n} \sigma_i^2 = n\sigma^2$$

substituting these into the result from Exercice4:

$$\mathbb{P}(|X_n - \mu| > \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$$

as $n \to \infty$, $\frac{\sigma^2}{n\varepsilon^2} \longrightarrow 0$ so:

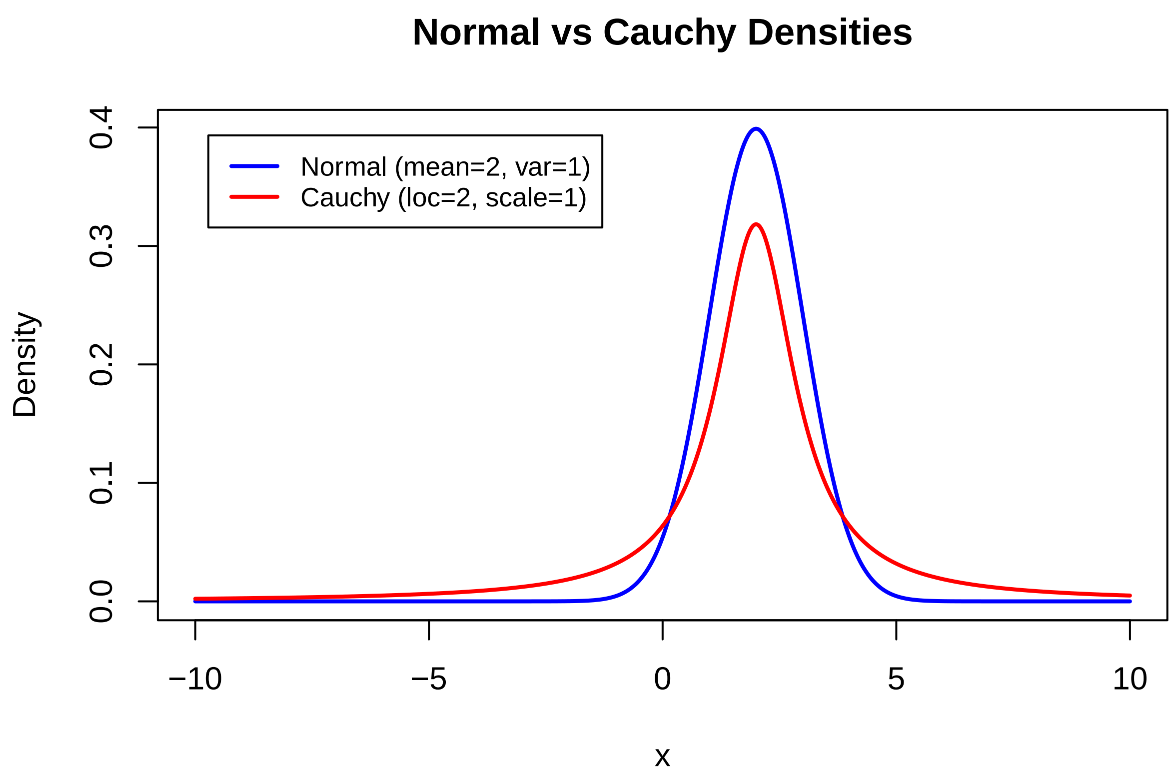
$$\lim_{n\to\infty} \mathbb{P}(|X_n - \mu| > \varepsilon) = 0$$

- The general result applies to independant but not identically variables, while the weak LLN assumes that the random variables are independant and identically distributed, which is the specific case of constant mean ($\mu_i = \mu$) and ($\sigma_i^2 = \sigma^2$).

[1.5/1.5]

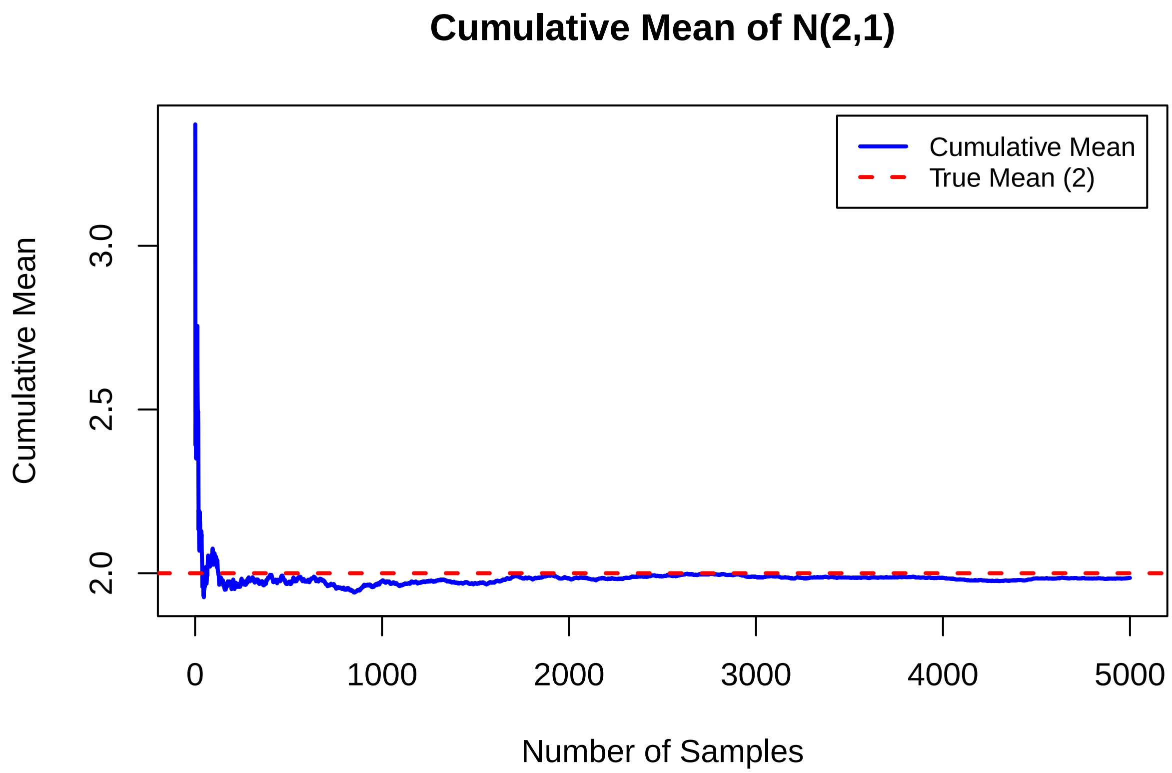
R Notebook

```r
#Q5(a): Plotting Densities of Normal and Cauchy Distributions

# Set seed for reproducibility
set.seed(42)

# Define the x-axis range
x <- seq(-10, 10, length.out = 1000)

# Compute densities
normal_density <- dnorm(x, mean = 2, sd = 1)
cauchy_density <- dcauchy(x, location = 2, scale = 1)

# Plot the normal density
plot(x, normal_density, type = "l", col = "blue", lwd = 2,
     main = "Normal vs Cauchy Densities",
     ylab = "Density", xlab = "x", xlim = c(-10, 10))

# Add the Cauchy density
lines(x, cauchy_density, col = "red", lwd = 2)

# Add a legend with adjusted position and size
legend("topleft", inset = 0.05,
       legend = c("Normal (mean=2, var=1)", "Cauchy (loc=2, scale=1)"),
       col = c("blue", "red"), lty = 1, lwd = 2, cex = 0.8)
```

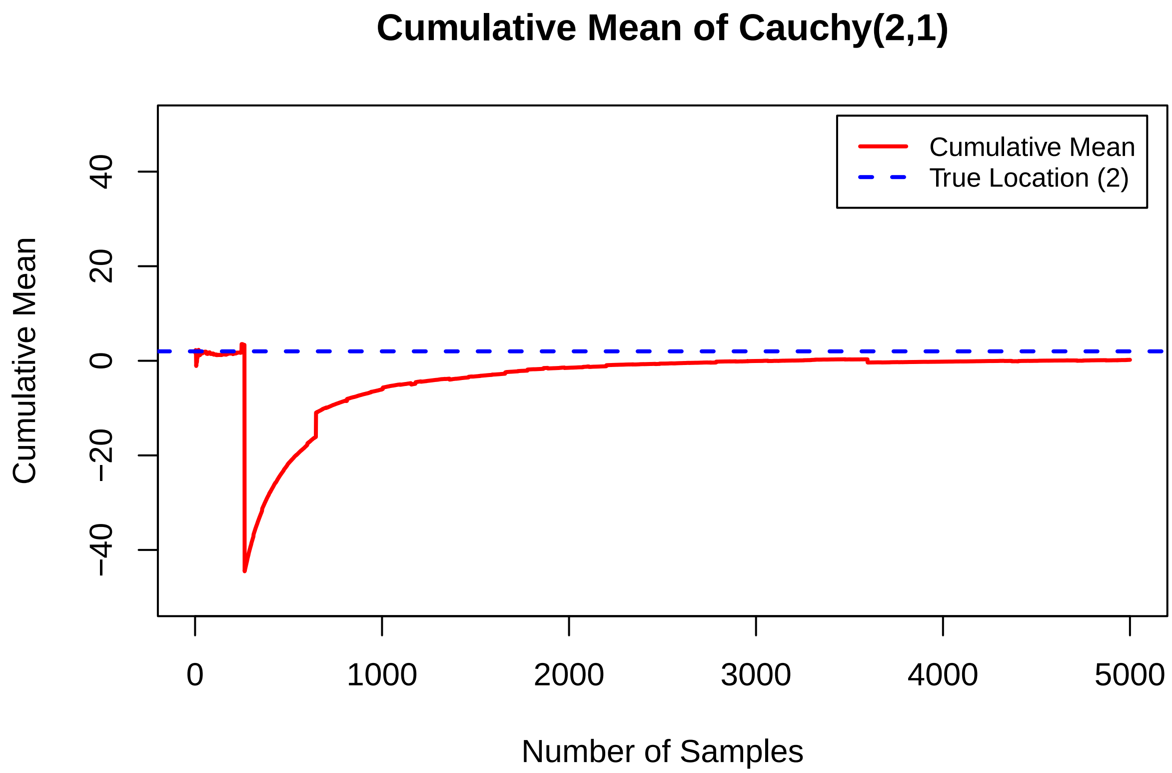


Cauchy distribution has heavier tails compared to the normal distribution due to the behavior of its PDF at large values of x.

```r
#Q.5(b)
# Set seed for reproducibility
set.seed(42)

# Generate samples from N(2, 1)
n <- 5000
normal_sample <- rnorm(n, mean = 2, sd = 1)

# Compute cumulative mean
cumulative_mean <- cumsum(normal_sample) / (1:n)

# Plot cumulative mean
plot(cumulative_mean, type = "l", col = "blue", lwd = 2,
     main = "Cumulative Mean of N(2,1)",
     xlab = "Number of Samples", ylab = "Cumulative Mean")

# Add true mean line
abline(h = 2, col = "red", lwd = 2, lty = 2)

# Add legend
legend("topright", inset = 0.02,
       legend = c("Cumulative Mean", "True Mean (2)"),
       col = c("blue", "red"), lty = c(1, 2), lwd = 2, cex = 0.8)
```




```r
#Q.5(c)
# Set seed for reproducibility
set.seed(42)

# Generate samples from Cauchy(loc = 2, scale = 1)
cauchy_sample <- rcauchy(n, location = 2, scale = 1)

# Compute cumulative mean
cumulative_mean_cauchy <- cumsum(cauchy_sample) / (1:n)

# Plot cumulative mean
plot(cumulative_mean_cauchy, type = "l", col = "red", lwd = 2,
     main = "Cumulative Mean of Cauchy(2,1)",
     xlab = "Number of Samples", ylab = "Cumulative Mean", ylim = c(-50, 50))

# Add true location line
abline(h = 2, col = "blue", lwd = 2, lty = 2)

# Add legend
legend("topright", inset = 0.02,
       legend = c("Cumulative Mean", "True Location (2)"),
       col = c("red", "blue"), lty = c(1, 2), lwd = 2, cex = 0.8)
```

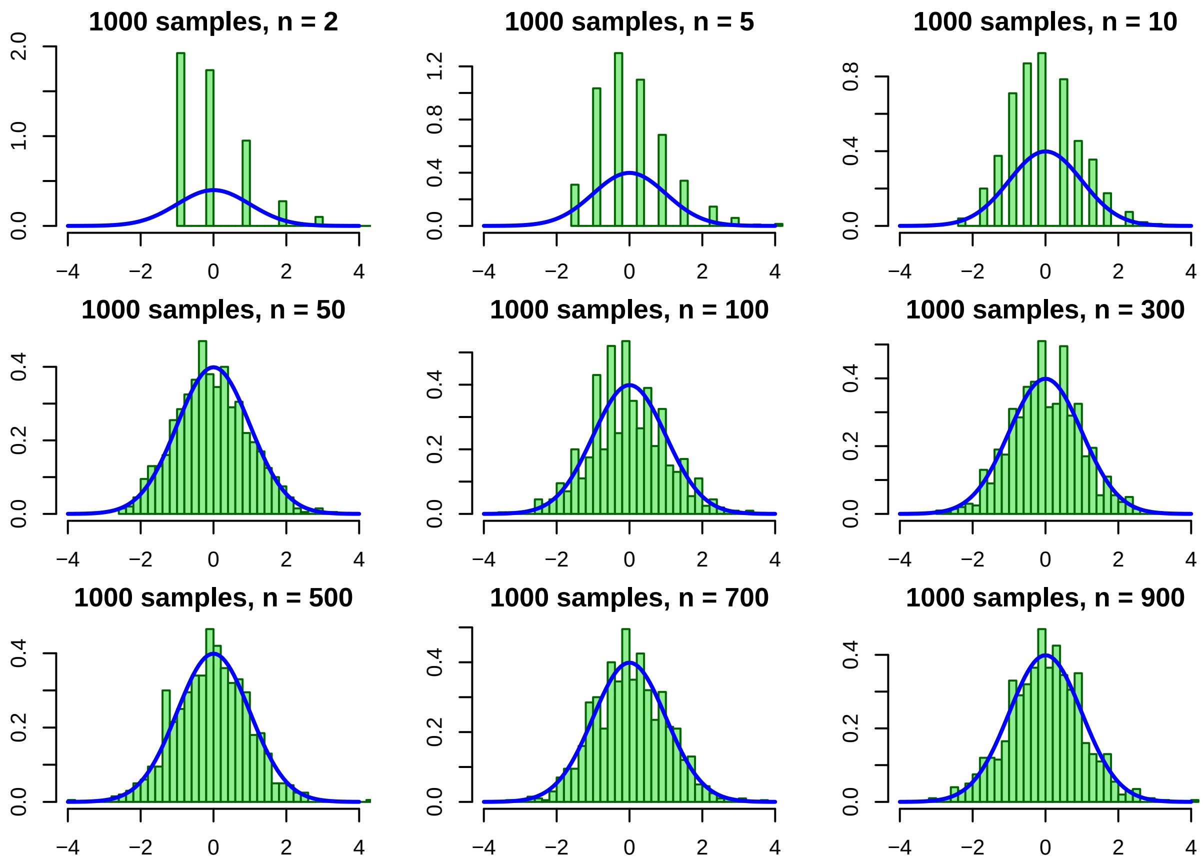


Q.5(c) Explanation:

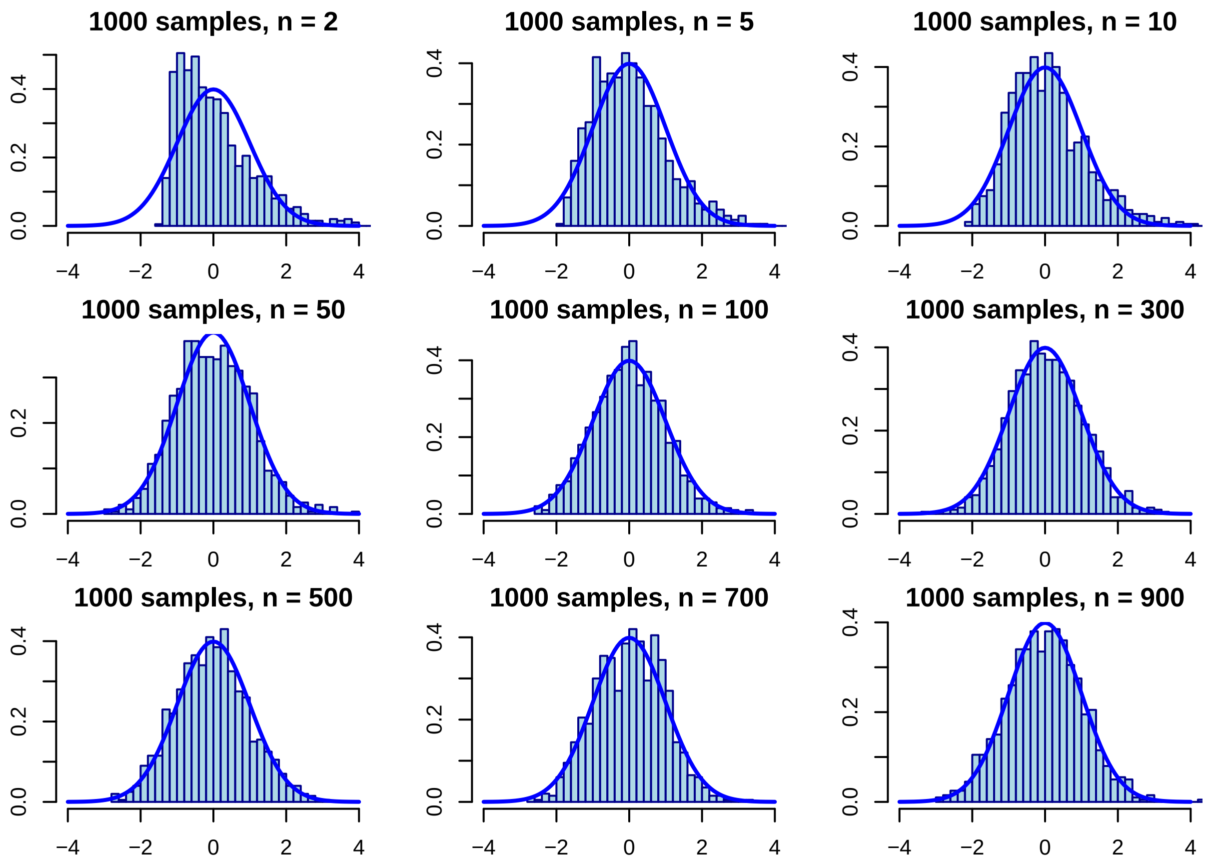
No we can't observe a similar convergence in the case of the Cauchy distribution as we did with the Normal distribution.the Law of Large Numbers (LLN) ensures that the sample mean converges to the true mean (2) because the Normal distribution has a finite variance. In contrast, the Cauchy distribution has infinite variance (its tails are so heavy that the variance diverges), which means the conditions required for the LLN are not satisfied.

```r
##Q.6
set.seed(42)
# Number of samples
K <- 1000

# Poisson distribution parameters
lambda_pois <- 0.5

# Exponential distribution parameters
lambda_exp <- 2

# Function for standardizing the sum of Poisson random variables
rp_pois <- function(n) {
  Zn <- rep(NA, K)
  for (k in 1:K) {
    s <- rpois(n, lambda = lambda_pois)  # Poisson random variables
    Zn[k] <- (sum(s) - n * lambda_pois) / (sqrt(n) * sqrt(lambda_pois))  # Standardized sum
  }
  return(Zn)  # Length K = 1000
```

```r
}

# Function for standardizing the sum of Exponential random variables
rp_exp <- function(n) {
  Zn <- rep(NA, K)
  for (k in 1:K) {
    s <- rexp(n, rate = lambda_exp)  # Exponential random variables
    Zn[k] <- (sum(s) - n * (1 / lambda_exp)) / (sqrt(n) * sqrt(1 / (lambda_exp^2)))  # Standardized sum
  }
  return(Zn)  # Length K = 1000
}

# Function to plot histogram and normal distribution overlay for Poisson
plot_hist_pois <- function(n) {
  Zn <- rp_pois(n)
  hist(Zn, breaks = 30, prob = TRUE, xlim = c(-4, 4), main = sprintf("%s samples, n = %s", K, n), col =
  curve(dnorm(x), col = "blue", lwd = 2, add = TRUE, yaxt = "n")  # Overlay normal distribution curve
}

# Function to plot histogram and normal distribution overlay for Exponential
plot_hist_exp <- function(n) {
  Zn <- rp_exp(n)
  hist(Zn, breaks = 30, prob = TRUE, xlim = c(-4, 4), main = sprintf("%s samples, n = %s", K, n), col =
  curve(dnorm(x), col = "blue", lwd = 2, add = TRUE, yaxt = "n")  # Overlay normal distribution curve
}

# Set plotting area to show 9 plots in a 3x3 grid
par(mfrow = c(3, 3), mar = c(2, 2, 2, 2))

# Run and plot Poisson distribution simulations for n = 3, 5, 7, ..., 512
n_values <- c(2, 5, 10, 50, 100, 300, 500,700,900)  # Modified n values
for (n in n_values) {
  plot_hist_pois(n)  # Poisson distribution plots
}
```




```
# Run and plot Exponential distribution simulations for n = 3, 5, 7, ..., 512
par(mfrow = c(3, 3), mar = c(2, 2, 2, 2))  # Reset to 3x3 grid
for (n in n_values) {
  plot_hist_exp(n)  # Exponential distribution plots
}
```



Final result does not depend on distribution of Xi as n goes at certain value(large) and in both cases the distribution converges to the standard normal distribution but it depends on distributon and n how long it takes to converge.

6
[1.5/1.5]