Solution to the Midterm Test

1. [5 marks]

If x and y are positive real numbers and $x \approx y$, but $x \neq y$, $\log_e(x) \approx \log_e(y)$, but $\log_e(x) \neq \log_e(y)$, Hence, the expression

$$\log_e(x) - \log_e(y) \tag{1}$$

suffers from *catastrophic cancellation*. Therefore, you will not be able to evaluate the expression (1) accurately if $x \approx y$, but $x \neq y$.

The question notes that

$$\log_e(x) - \log_e(y) = \log_e(x/y)$$

You might hope that you can evaluate $\log_e(x/y)$ more accurately than you can evaluate $\log_e(x) - \log_e(y)$, since $\log_e(x/y)$ does not involve any cancellation. However, it turns out that both expressions are about equally bad, as I explain below.

The hint suggests that you consider the conditioning of $\log_e(x/y)$ for $x \approx y$, whence $x/y \approx 1$. So, let $u = x/y$ and note that $u \neq 1$, since we assumed above that $x \neq y$. The assumption that $u \neq 1$ is important to avoid a division by zero when we compute the relative error associated with $w = \log_e(u)$. See (3) below.

Now note that, when we compute $u = x/y$, we get

$$\hat{u} = \text{fl}(x/y) = \frac{x}{y}(1+\delta)$$

Assume the relative error

$$\frac{\hat{u} - u}{u} = \delta$$

is small. Now let $w = \log_e(x/y) = \log_e(u)$ and $\hat{w} = \log_e(\hat{u})$. So, using our results from class for the conditioning of evaluating $w = f(u) = \log_e(u)$, we get that the associated condition number is

$$\frac{uf'(u)}{f(u)} = \frac{u\,\frac{d\,\log_e(u)}{du}}{\log_e(u)} = \frac{u\,\frac{1}{u}}{\log_e(u)} = \frac{1}{\log_e(u)} \tag{2}$$

Since $\log_e(u) = \log_e(x/y) \approx 0$ for $u = x/y \approx 1$, the condition number (2) associated with evaluating $\log_e(u) = \log_e(x/y)$ is very large for $u = x/y \approx 1$, whence $x \approx y$. Hence, since

$$\frac{\hat{w} - w}{w} \approx \frac{uf'(u)}{f(u)}\,\frac{\hat{u} - u}{u} = \frac{1}{\log_e(u)}\,\frac{\hat{u} - u}{u} \tag{3}$$

a small relative error in computing $u = x/y$ can be transformed into a very large relative error in computing $w = \log_e(u) = \log_e(x/y)$.

To apply a similar analysis to $\log_e(x) - \log_e(y)$, assume that we change x to $\hat{x} = x(1+\delta)$ for the same δ as above. Then let $g(x) = \log_e(x) - \log_e(y)$, and consider the relative error in $g(x)$ that results from changing x to $\hat{x}$:

$$\frac{g(\hat{x}) - g(x)}{g(x)}$$

where you let x vary, but hold y fixed. In this case, the condition number is

$$\frac{xg'(x)}{g(x)} = \frac{x\,\frac{1}{x}}{\log_e(x) - \log_e(y)} = \frac{1}{\log_e(x/y)} \tag{4}$$

So, you see that the condition number (2) of $f(u) = \log_e(u) = \log_e(x/y)$ is the same as the condition number (4) of $g(x) = \log_e(x) - \log_e(y)$. Hence, if you make a small relative change to x in either $\log_e(x/y)$ or $\log_e(x) - \log_e(y)$ and leave y unchanged, it will produce an equally large relative change in both $\log_e(x) - \log_e(y)$ and $\log_e(x/y)$. That is, both expressions $\log_e(x) - \log_e(y)$ and $\log_e(x/y)$ are equally badly conditioned.

2. [5 marks]

The easiest (and, I think, best) way to generate a pseudo-random variable X with the Cauchy(σ) distribution is to use the inverse CDF method. That is, generate at a pseudo-random Uniform $[0, 1]$ random variable, U, and then let

$$X = F^{-1}(U) \tag{5}$$

Equation (5) is equivalent to solving

$$F(X) = U$$

for X. That is, solve

$$\frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{X}{\sigma}\right) = U \tag{6}$$

for X. The solution of (6) is

$$X = \sigma \tan\left(\pi\left(U - \frac{1}{2}\right)\right)$$

They might be able to use an acceptance-rejection method to compute X, but I think that would be much harder than the solution I gave above. If anyone does this, let me know and I can help you mark their answer.

3. [5 marks]

Since both MC_1 and MC_2 use the same number of iterations, N, the computational work required to evaluate them will be about the same. Actually, the computational work required to evaluate MC_2 will be a little more than the computational work required to evaluate MC_1, since MC_2 requires two independent random variables, $X_{n,1} \sim F$ and $X_{n,2} \sim F$, per iteration, while MC_1 requires only one, $X_n \sim F$. However, this will be a minor difference between the two simulations. The bigger difference is that the variance associated with MC_1 should be much smaller than the variance associated with MC_2, as we explain below. Hence, the confidence interval for MC_1 should be much smaller than the confidence interval for MC_2.

To be more specific,

$$\text{Var}[\text{MC}_1] = \frac{1}{N}\Big(\text{Var}\left[g(X)\right] - 2\text{Cov}\left[g(X), h(X)\right] + \text{Var}\left[h(X)\right]\Big) \tag{7}$$

while

$$\text{Var}[\text{MC}_2] = \frac{1}{N}\Big(\text{Var}\left[g(X)\right] + \text{Var}\left[h(X)\right]\Big) \tag{8}$$

where, in both (7) and (8), $X \sim F$.

I'll show below a derivation of both (7) and (8), but first note that, from a comparison of (7) and (8), we see immediately that we should expect

$$\text{Var}[\text{MC}_1] < \text{Var}[\text{MC}_2]$$

since we should expect that

$$\text{Cov}\left[g(X), h(X)\right] > 0$$

because of our assumption that $g(x) \approx h(x)$ for all x. Indeed, we should expect that

$$\text{Var}[\text{MC}_1] \ll \text{Var}[\text{MC}_2]$$

since, in addition, we should expect that

$$\text{Cov}\left[g(X), h(X)\right] \approx \text{Var}\left[g(x)\right] \approx \text{Var}\left[h(x)\right]$$

because of our assumption that $g(x) \approx h(x)$ for all x.

I give a justification of (7) and (8) below. My justification is much longer than I expect the students' justifications will be. They can leave out any details from my justification that you think are reasonably obvious.

To begin, for a random variable $X \sim F$, let

$$\begin{aligned}
\mu_g &= \mathbb{E}[g(X)] \\
\mu_h &= \mathbb{E}[h(X)] \\
\mu &= \mathbb{E}[g(X) - h(X)] = \mathbb{E}[g(X)] - \mathbb{E}[h(X)] = \mu_g - \mu_h
\end{aligned} \tag{9}$$

First note that

$$\mathbb{E}[\text{MC}_1] = \mathbb{E}[\text{MC}_2] = \mu \tag{10}$$

They could just state (10), without proving it, since we proved it in class. However, to make this solution self-contained, I show below that (10) holds.

To this end, note that, since $X_n \sim F$ for $n = 1, 2, \ldots, N$,

$$\mathbb{E}[g(X_n) - h(X_n)] = \mu \qquad \text{for } n = 1, 2, \ldots, N$$

Therefore,

$$\begin{aligned}
\mathbb{E}[\text{MC}_1] &= \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\Big(g(X_n) - h(X_n)\Big)\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}[g(X_n) - h(X_n)] \\
&= \frac{1}{N}\sum_{n=1}^{N}\mu \\
&= \mu
\end{aligned}$$

Similarly, since $X_{n,1} \sim F$ and $X_{n,2} \sim F$ for $n = 1, 2, \ldots, N$,

$$\begin{aligned}
\mathbb{E}[g(X_{n,1})] &= \mu_g \qquad \text{for } n = 1, 2, \ldots, N \\
\mathbb{E}[h(X_{n,2})] &= \mu_h \qquad \text{for } n = 1, 2, \ldots, N \\
\mu &= \mu_g - \mu_h
\end{aligned}$$

from (9). Therefore,

$$\begin{aligned}
\mathbb{E}[\text{MC}_2] &= \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}g(X_{n,1}) - \frac{1}{N}\sum_{n=1}^{N}h(X_{n,2})\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}[g(X_{n,1})] - \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}[h(X_{n,2})] \\
&= \frac{1}{N}\sum_{n=1}^{N}\mu_g - \frac{1}{N}\sum_{n=1}^{N}\mu_h \\
&= \mu_g - \mu_h \\
&= \mu
\end{aligned}$$

Now note that

$$\text{Var}[\text{MC}_1] = \frac{1}{N}\text{Var}[g(X) - h(X)] \tag{11}$$

where $X \sim F$. Again, the students could just state (11) without proof, since we proved it in class. However, to make this solution self-contained, I show below that (11) holds.

To this end, recall that $X_n \sim F$ for $n = 1, 2, \ldots, N$ and the X_n are independent. Hence, if $m \neq n$, then

$$\begin{aligned} &\mathbb{E}\left[\Big(g(X_m) - h(X_m) - \mu\Big)\Big(g(X_n) - h(X_n) - \mu\Big)\right] \\ &\quad = \mathbb{E}\left[\Big(g(X_m) - h(X_m) - \mu\Big)\right] \mathbb{E}\left[\Big(g(X_n) - h(X_n) - \mu\Big)\right] \\ &\quad = 0 \end{aligned} \tag{12}$$

since X_m and X_n are independent for $m \neq n$, whence the random variables $Y = g(X_m) - h(X_m) - \mu$ and $Z = g(X_n) - h(X_n) - \mu$ are also independent for $m \neq n$, so $\mathbb{E}[YZ] = \mathbb{E}[Y]\,\mathbb{E}[Z]$. In addition, from (9),

$$\begin{aligned} \mathbb{E}\left[\Big(g(X_m) - h(X_m) - \mu\Big)\right] &= 0 \\ \mathbb{E}\left[\Big(g(X_n) - h(X_n) - \mu\Big)\right] &= 0 \end{aligned}$$

Therefore,

$$
\begin{aligned}
\mathrm{Var}[\mathrm{MC}_1] &= \mathrm{Var}\left[\frac{1}{N}\sum_{n=1}^{N}\Big(g(X_n)-h(X_n)\Big)\right] \\
&= \mathbb{E}\left[\left(\left(\frac{1}{N}\sum_{n=1}^{N}\Big(g(X_n)-h(X_n)\Big)\right)-\mu\right)^2\right] \\
&= \mathbb{E}\left[\left(\frac{1}{N}\sum_{n=1}^{N}\Big(g(X_n)-h(X_n)-\mu\Big)\right)^2\right] \\
&= \frac{1}{N^2}\,\mathbb{E}\left[\left(\sum_{n=1}^{N}\Big(g(X_n)-h(X_n)-\mu\Big)\right)\left(\sum_{n=1}^{N}\Big(g(X_n)-h(X_n)-\mu\Big)\right)\right] \\
&= \frac{1}{N^2}\,\mathbb{E}\left[\left(\sum_{m=1}^{N}\Big(g(X_m)-h(X_m)-\mu\Big)\right)\left(\sum_{n=1}^{N}\Big(g(X_n)-h(X_n)-\mu\Big)\right)\right] \\
&= \frac{1}{N^2}\,\mathbb{E}\left[\sum_{m=1}^{N}\sum_{n=1}^{N}\Big(\Big(g(X_m)-h(X_m)-\mu\Big)\Big(g(X_n)-h(X_n)-\mu\Big)\Big)\right] \\
&= \frac{1}{N^2}\,\sum_{m=1}^{N}\sum_{n=1}^{N}\mathbb{E}\left[\Big(g(X_m)-h(X_m)-\mu\Big)\Big(g(X_n)-h(X_n)-\mu\Big)\right] \\
&= \frac{1}{N^2}\left(\sum_{n=1}^{N}\mathbb{E}\left[\Big(g(X_n)-h(X_n)-\mu\Big)\Big(g(X_n)-h(X_n)-\mu\Big)\right]\right. \\
&\qquad\left. +\sum_{m=1}^{N}\sum_{\substack{n=1\\ n\neq m}}^{N}\mathbb{E}\left[\Big(g(X_m)-h(X_m)-\mu\Big)\Big(g(X_n)-h(X_n)-\mu\Big)\right]\right) \\
&= \frac{1}{N^2}\,\sum_{n=1}^{N}\mathbb{E}\left[\Big(g(X_n)-h(X_n)-\mu\Big)\Big(g(X_n)-h(X_n)-\mu\Big)\right] \\
&= \frac{1}{N^2}\,\sum_{n=1}^{N}\mathbb{E}\left[\Big(g(X_n)-h(X_n)-\mu\Big)^2\right] \\
&= \frac{1}{N^2}\,\sum_{n=1}^{N}\mathrm{Var}\left[g(X_n)-h(X_n)\right] \\
&= \frac{1}{N^2}\,\sum_{n=1}^{N}\mathrm{Var}\left[g(X)-h(X)\right] \\
&= \frac{1}{N}\,\mathrm{Var}\left[g(X)-h(X)\right]
\end{aligned}
\tag{13}
$$

where we used (12) in the fifth equation from the bottom in the sequence of equations (13) and

$$\text{Var}\left[g(X_n) - h(X_n)\right] = \text{Var}\left[g(X) - h(X)\right] \tag{14}$$

in the second equation from the bottom in the sequence of equations (13). Note that (14) holds because both $X_n \sim F$ and $X \sim F$.

Now, using (9), we can expand (11) to get

$$\begin{aligned}
\text{Var}[\text{MC}_1] &= \frac{1}{N}\text{Var}[g(X) - h(X)] \\
&= \frac{1}{N}\mathbb{E}\left[(g(X) - h(X) - \mu)^2\right] \\
&= \frac{1}{N}\mathbb{E}\left[(g(X) - h(X) - (\mu_g - \mu_h))^2\right] \\
&= \frac{1}{N}\mathbb{E}\left[((g(X) - \mu_g) - (h(X) - \mu_h))^2\right] \\
&= \frac{1}{N}\mathbb{E}\left[(g(X) - \mu_g)^2 - 2(g(X) - \mu_g)(h(X) - \mu_h) + (h(X) - \mu_h))^2\right] \\
&= \frac{1}{N}\Big(\mathbb{E}\left[(g(X) - \mu_g)^2\right] - 2\mathbb{E}\left[(g(X) - \mu_g)(h(X) - \mu_h)\right] + \mathbb{E}\left[(h(X) - \mu_h)^2\right]\Big) \\
&= \frac{1}{N}\Big(\text{Var}\left[g(X)\right] - 2\text{Cov}\left[g(X), h(X)\right] + \text{Var}\left[h(X)\right]\Big)
\end{aligned}$$

Therefore, we have verified (7).

Equation (8) is a little different from any results that we proved in class, but it is not too different. Therefore, the students should give a little justification of why (8) is true, but their justification need not be as detailed as the one I give below.

To verify (8), assume $X_{n,1} \sim F$, $X_{n,2} \sim F$ for $n = 1, 2, \ldots, N$ and also assume that all the $\{X_{n,1}, X_{n,2} : n = 1, 2, \ldots, N\}$ are independent. Hence,

$$\begin{aligned}
\mathrm{Var}[\mathrm{MC}_2] &= \mathrm{Var}\left[\frac{1}{N}\sum_{n=1}^{N} g(X_{n,1}) - \frac{1}{N}\sum_{n=1}^{N} h(X_{n,2})\right] \\
&= \mathbb{E}\left[\left(\left(\frac{1}{N}\sum_{n=1}^{N} g(X_{n,1}) - \frac{1}{N}\sum_{n=1}^{N} h(X_{n,2})\right) - \mu\right)^2\right] \\
&= \mathbb{E}\left[\left(\left(\frac{1}{N}\sum_{n=1}^{N} g(X_{n,1}) - \frac{1}{N}\sum_{n=1}^{N} h(X_{n,2})\right) - (\mu_g - \mu_h)\right)^2\right] \\
&= \mathbb{E}\left[\left(\left(\frac{1}{N}\sum_{n=1}^{N}(g(X_{n,1}) - \mu_g)\right) - \left(\frac{1}{N}\sum_{n=1}^{N}(h(X_{n,2}) - \mu_h)\right)\right)^2\right] \\
&= \mathbb{E}\left[\left(\left(\frac{1}{N}\sum_{n=1}^{N}(g(X_{n,1}) - \mu_g)\right) - \left(\frac{1}{N}\sum_{n=1}^{N}(h(X_{n,2}) - \mu_h)\right)\right)\right. \\
&\qquad \left. \times \left(\left(\frac{1}{N}\sum_{n=1}^{N}(g(X_{n,1}) - \mu_g)\right) - \left(\frac{1}{N}\sum_{n=1}^{N}(h(X_{n,2}) - \mu_h)\right)\right)\right] \\
&= \mathbb{E}\left[\left(\left(\frac{1}{N}\sum_{m=1}^{N}(g(X_{m,1}) - \mu_g)\right) - \left(\frac{1}{N}\sum_{m=1}^{N}(h(X_{m,2}) - \mu_h)\right)\right)\right. \\
&\qquad \left. \times \left(\left(\frac{1}{N}\sum_{n=1}^{N}(g(X_{n,1}) - \mu_g)\right) - \left(\frac{1}{N}\sum_{n=1}^{N}(h(X_{n,2}) - \mu_h)\right)\right)\right] \\
&= \mathbb{E}\left[\left(\frac{1}{N}\sum_{m=1}^{N}(g(X_{m,1}) - \mu_g)\right)\left(\frac{1}{N}\sum_{n=1}^{N}(g(X_{n,1}) - \mu_g)\right)\right. \\
&\qquad - \left(\frac{1}{N}\sum_{m=1}^{N}(g(X_{m,1}) - \mu_g)\right)\left(\frac{1}{N}\sum_{n=1}^{N}(h(X_{n,2}) - \mu_h)\right) \\
&\qquad - \left(\frac{1}{N}\sum_{m=1}^{N}(h(X_{m,2}) - \mu_h)\right)\left(\frac{1}{N}\sum_{n=1}^{N}(g(X_{n,1}) - \mu_g)\right) \\
&\qquad \left. + \left(\frac{1}{N}\sum_{m=1}^{N}(h(X_{m,2}) - \mu_h)\right)\left(\frac{1}{N}\sum_{n=1}^{N}(h(X_{n,2}) - \mu_h)\right)\right]
\end{aligned} \tag{15}$$

Carrying on from (15) above, we get

$$
\begin{aligned}
\operatorname{Var}[\mathrm{MC}_2] &= \mathbb{E}\Bigg[\frac{1}{N^2}\sum_{m=1}^{N}\sum_{n=1}^{N}(g(X_{m,1})-\mu_g)(g(X_{n,1})-\mu_g) \\
&\qquad - \frac{1}{N^2}\sum_{m=1}^{N}\sum_{n=1}^{N}(g(X_{m,1})-\mu_g)(h(X_{n,2})-\mu_h) \\
&\qquad - \frac{1}{N^2}\sum_{m=1}^{N}\sum_{n=1}^{N}(h(X_{m,2})-\mu_h)(g(X_{n,1})-\mu_g) \\
&\qquad + \frac{1}{N^2}\sum_{m=1}^{N}\sum_{n=1}^{N}(h(X_{m,2})-\mu_h)(h(X_{n,2})-\mu_h)\Bigg] \\
&= \frac{1}{N^2}\Bigg(\sum_{m=1}^{N}\sum_{n=1}^{N}\mathbb{E}[(g(X_{m,1})-\mu_g)(g(X_{n,1})-\mu_g)] \\
&\qquad - \sum_{m=1}^{N}\sum_{n=1}^{N}\mathbb{E}[(g(X_{m,1})-\mu_g)(h(X_{n,2})-\mu_h)] \\
&\qquad - \sum_{m=1}^{N}\sum_{n=1}^{N}\mathbb{E}[(h(X_{m,2})-\mu_h)(g(X_{n,1})-\mu_g)] \\
&\qquad + \sum_{m=1}^{N}\sum_{n=1}^{N}\mathbb{E}[(h(X_{m,2})-\mu_h)(h(X_{n,2})-\mu_h)]\Bigg)
\end{aligned}
\tag{16}
$$

Now note that, since $X_{m,1}$ and $X_{n,2}$ are independent for all $m = 1, 2, \ldots, N$ and $n = 1, 2, \ldots, N$, $(g(X_{m,1}) - \mu_g)$ and $(h(X_{n,2}) - \mu_h)$ are also independent for all $m = 1, 2, \ldots, N$ and $n = 1, 2, \ldots, N$. In addition, since $X_{m,1} \sim F$ and $X_{n,2} \sim F$ for all $m = 1, 2, \ldots, N$ and $n = 1, 2, \ldots, N$, it follows from (9) that

$$
\begin{aligned}
\mathbb{E}[(g(X_{m,1})-\mu_g)] &= 0 \qquad \text{for all } m = 1, 2, \ldots, N \\
\mathbb{E}[(h(X_{n,2})-\mu_h)] &= 0 \qquad \text{for all } n = 1, 2, \ldots, N
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\mathbb{E}[(g(X_{m,1})-\mu_g)(h(X_{n,2})-\mu_h)] \\
&\quad = \mathbb{E}[(g(X_{m,1})-\mu_g)]\,\mathbb{E}[(h(X_{n,2})-\mu_h)] \\
&\quad = 0
\end{aligned}
$$

for all $m = 1, 2, \ldots, N$ and $n = 1, 2, \ldots, N$, Similarly,

$$
\mathbb{E}[(h(X_{m,2})-\mu_h)(g(X_{n,1})-\mu_g)] = 0
$$

for all $m = 1, 2, \ldots, N$ and $n = 1, 2, \ldots, N$. Therefore, carrying on from (16) above, we get

$$\begin{aligned} \mathrm{Var}[\mathrm{MC}_2] = \frac{1}{N^2} \Bigg(& \sum_{m=1}^{N} \sum_{n=1}^{N} \mathbb{E}[(g(X_{m,1}) - \mu_g)(g(X_{n,1}) - \mu_g)] \\ & + \sum_{m=1}^{N} \sum_{n=1}^{N} \mathbb{E}[(h(X_{m,2}) - \mu_h)(h(X_{n,2}) - \mu_h)] \Bigg) \end{aligned} \tag{17}$$

Since $X_{m,1}$ and $X_{n,1}$ are independent for $m \neq n$, it follows from an argument similar to the one above that

$$\mathbb{E}[(g(X_{m,1}) - \mu_g)(g(X_{n,1}) - \mu_g)] = 0$$

for $m \neq n$. Similarly,

$$\mathbb{E}[(h(X_{m,2}) - \mu_h)(h(X_{n,2}) - \mu_h)] = 0$$

for $m \neq n$. Hence, (17) reduces to

$$\begin{aligned} \mathrm{Var}[\mathrm{MC}_2] &= \frac{1}{N^2} \Bigg(\sum_{n=1}^{N} \mathbb{E}[(g(X_{n,1}) - \mu_g)(g(X_{n,1}) - \mu_g)] \\ &\qquad + \sum_{n=1}^{N} \mathbb{E}[(h(X_{n,2}) - \mu_h)(h(X_{n,2}) - \mu_h)] \Bigg) \\ &= \frac{1}{N^2} \Bigg(\sum_{n=1}^{N} \mathbb{E}[(g(X_{n,1}) - \mu_g)^2] \\ &\qquad + \sum_{n=1}^{N} \mathbb{E}[(h(X_{n,2}) - \mu_h)^2] \Bigg) \\ &= \frac{1}{N^2} \left(\sum_{n=1}^{N} \mathrm{Var}[g(X_{n,1})] + \sum_{n=1}^{N} \mathrm{Var}[h(X_{n,2})] \right) \\ &= \frac{1}{N^2} \left(\sum_{n=1}^{N} \mathrm{Var}[g(X)] + \sum_{n=1}^{N} \mathrm{Var}[h(X)] \right) \\ &= \frac{1}{N} \Big(\mathrm{Var}[g(X)] + \mathrm{Var}[h(X)] \Big) \end{aligned} \tag{18}$$

where we used in the second to last equation in the sequence of equations (18) that

$$\begin{aligned} \mathrm{Var}[g(X_{n,1})] &= \mathrm{Var}[g(X)] \\ \mathrm{Var}[h(X_{n,2})] &= \mathrm{Var}[h(X)] \end{aligned}$$

since $X_{n,1} \sim F$, $X_{n,2} \sim F$ and $X \sim F$. Therefore, we have verified that (8) holds.

4. [5 marks]

As noted in the question, the option price at time $t = 0$ is

$$P_0 = \mathbb{E}[\mathrm{e}^{-rT} h(S_T^{(1)}, S_T^{(2)})] \tag{19}$$

where

$$h(S_T^{(1)}, S_T^{(2)}) = \max(S_T^{(1)} - S_T^{(2)} - \hat{K}, 0) \tag{20}$$

for some constant $\hat{K}$. So, combining (19) and (20), we get

$$P_0 = \mathbb{E}[\mathrm{e}^{-rT} \max(S_T^{(1)} - S_T^{(2)} - \hat{K}, 0)] \tag{21}$$

Also, the question notes that

$$\begin{aligned} S_T^{(1)} &= S_0^{(1)} \mathrm{e}^{(r-\sigma_1^2/2)T + \sigma_1 W_T^{(1)}} \\ S_T^{(2)} &= S_0^{(2)} \mathrm{e}^{(r-\sigma_2^2/2)T + \sigma_2 W_T^{(2)}} \end{aligned}$$

and the correlated Brownian motions $W_T^{(1)}$ and $W_T^{(2)}$ satisfy

$$\begin{aligned} W_T^{(1)} &= \sqrt{T}\left(\sqrt{1-\rho^2} Z^{(1)} + \rho Z^{(2)}\right) \\ W_T^{(2)} &= \sqrt{T} Z^{(2)} \end{aligned}$$

where $Z^{(1)} \sim N(0,1)$, $Z^{(2)} \sim N(0,1)$ and $Z^{(1)}$ and $Z^{(2)}$ are independent. Therefore,

$$\begin{aligned} S_T^{(1)} &= S_0^{(1)} \mathrm{e}^{(r-\sigma_1^2/2)T + \sigma_1 \sqrt{T}\left(\sqrt{1-\rho^2} Z^{(1)} + \rho Z^{(2)}\right)} \\ S_T^{(2)} &= S_0^{(2)} \mathrm{e}^{(r-\sigma_2^2/2)T + \sigma_2 \sqrt{T} Z^{(2)}} \end{aligned} \tag{22}$$

where $Z^{(1)} \sim N(0,1)$, $Z^{(2)} \sim N(0,1)$ and $Z^{(1)}$ and $Z^{(2)}$ are independent. Substituting the $S_T^{(1)}$ and $S_T^{(2)}$ from (22) into (21), we get

$$\begin{aligned} P_0 = \mathbb{E}[\mathrm{e}^{-rT} \max(&S_0^{(1)} \mathrm{e}^{(r-\sigma_1^2/2)T + \sigma_1 \sqrt{T}\left(\sqrt{1-\rho^2} Z^{(1)} + \rho Z^{(2)}\right)} \\ &- S_0^{(2)} \mathrm{e}^{(r-\sigma_2^2/2)T + \sigma_2 \sqrt{T} Z^{(2)}} - \hat{K}, 0)] \end{aligned} \tag{23}$$

Now we can apply conditional expectation to (23) to get

$$\begin{aligned} P_0 = \mathbb{E}_{Z^{(2)}}[\mathbb{E}_{Z^{(1)}}[\mathrm{e}^{-rT} \max(&S_0^{(1)} \mathrm{e}^{(r-\sigma_1^2/2)T + \sigma_1 \sqrt{T}\left(\sqrt{1-\rho^2} Z^{(1)} + \rho Z^{(2)}\right)} \\ &- S_0^{(2)} \mathrm{e}^{(r-\sigma_2^2/2)T + \sigma_2 \sqrt{T} Z^{(2)}} - \hat{K}, 0) | Z^{(2)}]] \end{aligned} \tag{24}$$

where the outer expectation is with respect to $Z^{(2)}$ and the inner expectation is with respect to $Z^{(1)}$. To clarify (24) a little, we introduce a deterministic variable z_2 and re-write (24) as

$$\begin{aligned} P_0 = \mathbb{E}_{Z^{(2)}}[\mathbb{E}_{Z^{(1)}}[\mathrm{e}^{-rT} \max(&S_0^{(1)} \mathrm{e}^{(r-\sigma_1^2/2)T + \sigma_1 \sqrt{T}\left(\sqrt{1-\rho^2} Z^{(1)} + \rho z_2\right)} \\ &- S_0^{(2)} \mathrm{e}^{(r-\sigma_2^2/2)T + \sigma_2 \sqrt{T} z_2} - \hat{K}, 0) | z_2 = Z^{(2)}]] \end{aligned} \tag{25}$$

Now note that

$$
\begin{aligned}
&S_0^{(1)}\mathrm{e}^{(r-\sigma_1^2/2)T+\sigma_1\sqrt{T}\left(\sqrt{1-\rho^2}Z^{(1)}+\rho z_2\right)}\\
&\quad= S_0^{(1)}\mathrm{e}^{(r-\sigma_1^2(1-\rho^2)/2)T-\sigma_1^2\rho^2T/2+\sigma_1\sqrt{1-\rho^2}\sqrt{T}Z^{(1)}+\sigma_1\rho\sqrt{T}z_2}\\
&\quad= S_0^{(1)}\mathrm{e}^{-\sigma_1^2\rho^2T/2+\sigma_1\rho\sqrt{T}z_2}\mathrm{e}^{(r-\sigma_1^2(1-\rho^2)/2)T+\sigma_1\sqrt{1-\rho^2}\sqrt{T}Z^{(1)}}\\
&\quad= S_0(z_2)\mathrm{e}^{(r-\sigma^2/2)T+\sigma\sqrt{T}Z^{(1)}}
\end{aligned}
$$

where

$$
\begin{aligned}
S_0(z_2) &= S_0^{(1)}\mathrm{e}^{-\sigma_1^2\rho^2T/2+\sigma_1\rho\sqrt{T}z_2}\\
\sigma &= \sigma_1\sqrt{1-\rho^2}
\end{aligned}
\tag{26}
$$

In addition, let

$$
K(z_2) = S_0^{(2)}\mathrm{e}^{(r-\sigma_2^2/2)T+\sigma_2\sqrt{T}z_2} + \hat{K} \tag{27}
$$

So, we can re-write the inner expectation in (25) as

$$
\mathbb{E}_{Z^{(1)}}[\mathrm{e}^{-rT}\max(S_0(z_2)\mathrm{e}^{(r-\sigma^2/2)T+\sigma\sqrt{T}Z^{(1)}} - K(z_2), 0)] \tag{28}
$$

Note that (28) is the price of a "vanilla" call option with strike price $K(z_2)$, expiry $t = T$ and underlying asset S_t that starts from $S_0(z_2)$ at time $t = 0$ and evolves in time according to the SDE

$$
dS_t = rS_t dt + \sigma S_t dW_t
$$

Therefore,

$$
\text{Call} = \mathbb{E}_{Z^{(1)}}[\mathrm{e}^{-rT}\max(S_0(z_2)\mathrm{e}^{(r-\sigma^2/2)T+\sigma\sqrt{T}Z^{(1)}} - K(z_2), 0)]
$$

where

$$
\texttt{[Call, Put] = blsprice(}S_0(z_2)\texttt{, }K(z_2)\texttt{, }r\texttt{, }T\texttt{, }\sigma\texttt{)}
$$

and $\sigma = \sigma_1\sqrt{1-\rho^2}$. So, write a little wrapper function `blspriceCall` such that

$$
\texttt{Call = blspriceCall(}S_0(z_2)\texttt{, }K(z_2)\texttt{, }r\texttt{, }T\texttt{, }\sigma\texttt{)}
$$

where `Call` is given above by `blsprice` with the same parameters. Therefore,

$$
\begin{aligned}
&\mathbb{E}_{Z^{(1)}}[\mathrm{e}^{-rT}\max(S_0(z_2)\mathrm{e}^{(r-\sigma^2/2)T+\sigma\sqrt{T}Z^{(1)}} - K(z_2), 0)]\\
&\quad= \texttt{blspriceCall(}S_0(z_2)\texttt{, }K(z_2)\texttt{, }r\texttt{, }T\texttt{, }\sigma\texttt{)}
\end{aligned}
\tag{29}
$$

So, we can use (29) to rewrite (25) as

$$
P_0 = \mathbb{E}_{Z^{(2)}}[\texttt{blspriceCall(}S_0(Z_2)\texttt{, }K(Z_2)\texttt{, }r\texttt{, }T\texttt{, }\sigma\texttt{)}] \tag{30}
$$

We can use (30) as the basis for the Monte Carlo simulation below to price this *exchange spread option.*

(a) For $n = 1, 2, \ldots, N$,

i. Generate $Z_n \sim N(0, 1)$ using a function such as MatLab's `randn`

ii. Let

$$Y_n = \texttt{blspriceCall(}S_0(Z_n)\texttt{, } K(Z_n)\texttt{, } r\texttt{, } T\texttt{, } \sigma\texttt{)}$$

where $S_0(z_2)$ and σ are given in (26) and $K(z_2)$ is given in (27).

(b) Approximate the option price P_0 by

$$\hat{P}_0 = \frac{1}{N} \sum_{n=1}^{N} Y_n$$