

QRM project. May 2024 Part 2

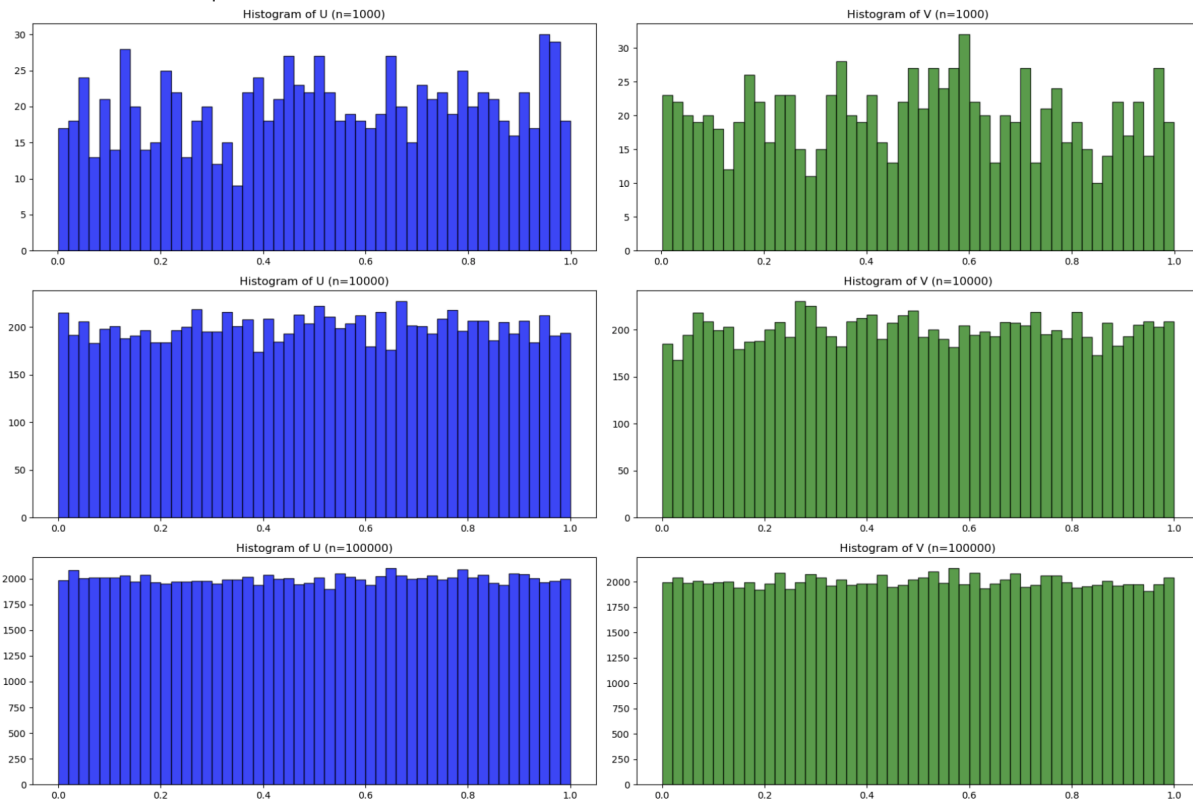
Team 10: Enrico FADDA(B00764165), Linhui Larry SANG(B00809803)
(Python Notebook and R Notebook are attached with submission)

Part II

1.

1.(a)The first step of the project consists in simulating uniform random variables: The `simulate_uniform_rvs` function generates two independent uniform (0,1) random variables U and V of the given sample size.

Then, after defining three samples of size 1000, 10,000 and 100,000 respectively we simulated U and V for each sample size.



b)The next step of our analysis consisted in computing the parameters $\mu(x)$, $\sigma(x)$, $\mu(y)$ and $\sigma(y)$ for the log-normal distributions of X and Y given their expectations and variances.

For a random variable $Z \sim \text{LN}(\mu, \sigma)$:

The mean $E(Z)$ is given by:

$$E(Z) = e^{\mu + \frac{\sigma^2}{2}}$$

While $\text{Var}(Z)$ is computed using:

$$\text{Var}(Z) = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$$

Which solving for $\mu(Z)$ and $\sigma(Z)$ is equivalent to:

$$\mu = \ln(\mathbb{E}[Z]) - \frac{\sigma^2}{2}$$

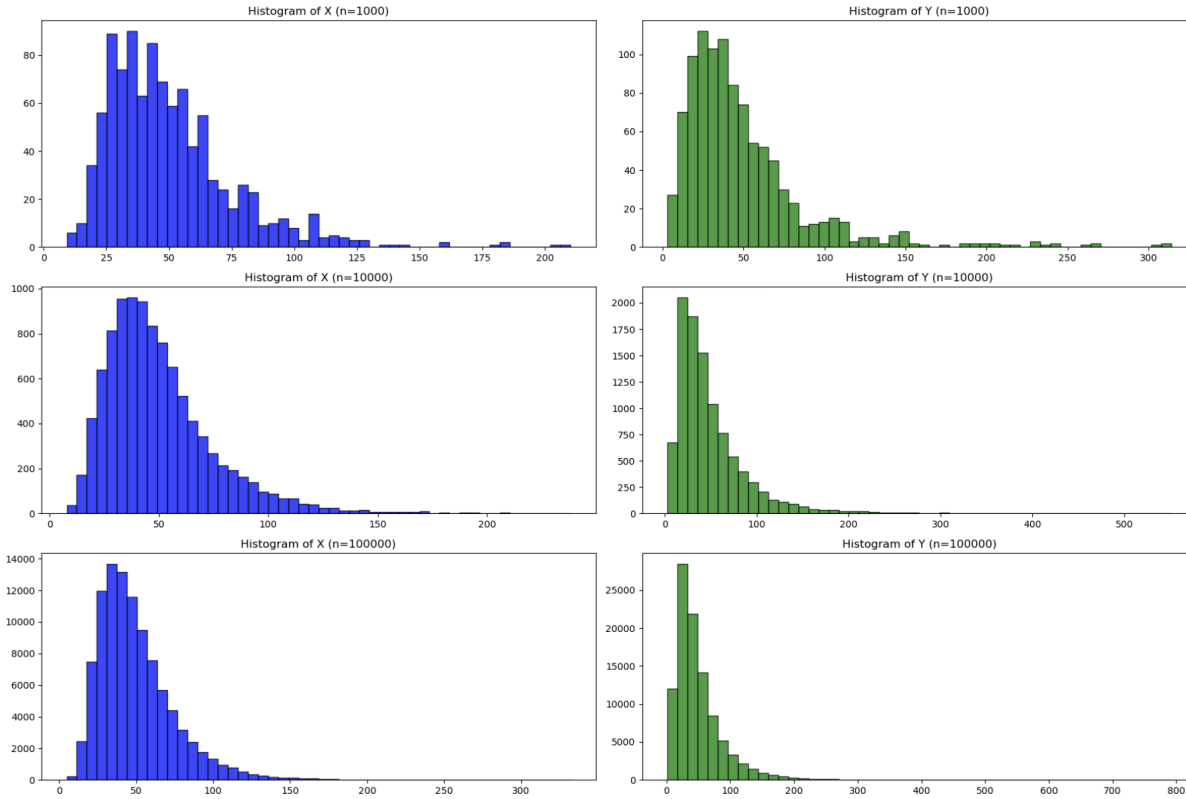
$$\sigma = \sqrt{\ln \left(1 + \frac{\text{Var}[Z]}{\mathbb{E}[Z]^2} \right)}$$

Given $E(X)=50$, $\text{Var}(X)=25^2$, $E(Y)=50$, and $\text{Var}(Y)=40^2$. The following step consisted in solving for X and Y $\mu(x)$, $\sigma(x)$, $\mu(y)$ and $\sigma(y)$.

After having solved the equation numerically we could proceed by generating samples for X and Y using the uniform random variables U and V simulated previously. The results from this equation are the following: $\mu(x)=2.2753106982399443$, $\sigma(x)=1.8092607922509136$, $\mu(y)=1.8199181847753056$, $\sigma(y)=2.045534072389331$

Given $U \sim \text{Uniform}(0,1)$, a log-normal random variable Z can be generated using the inverse CDF method where $\Phi^{-1}(U)$ is the inverse CDF of the standard normal distribution:

$$Z = e^{\mu + \sigma \Phi^{-1}(U)}$$



c) Given the theoretical values:

For $X \sim \text{LN}(\mu_X, \sigma_X)$:

$$E(X) = e^{\mu_X + \frac{\sigma_X^2}{2}}, \quad \text{Std}(X) = \sqrt{(e^{\sigma_X^2} - 1)e^{2\mu_X + \sigma_X^2}}$$

For $Y \sim \text{LN}(\mu_Y, \sigma_Y)$:

$$E(Y) = e^{\mu_Y + \frac{\sigma_Y^2}{2}}, \quad \text{Std}(Y) = \sqrt{(e^{\sigma_Y^2} - 1)e^{2\mu_Y + \sigma_Y^2}}$$

We used the software to calculate the empirical values and compare them with the actual ones.

	Sample Size	Empirical Mean X	Theoretical Mean X	Empirical Std X	Theoretical Std X	Empirical Mean Y	Theoretical Mean Y	Empirical Std Y	Theoretical Std Y
0	1000	50.542807	50.0	25.920693	25.0	48.549124	50.0	36.705786	40.0
1	10000	49.573720	50.0	24.611950	25.0	49.606330	50.0	39.839675	40.0
2	100000	50.067467	50.0	24.943491	25.0	50.079181	50.0	40.057651	40.0

As we can see, in line with the theoretical framework of the Law of large numbers, as the sample size increases our empirical values approach their true values.

However if the empirical mean or standard deviation significantly were to deviate from the theoretical values, one could:

- Increase Sample Size: Larger sample sizes generally lead to better estimates.
- Multiple Simulations: Perform multiple simulations and average the results to reduce variability.
- Check Randomness: Ensure that the random number generator is properly seeded and functioning correctly.

d) In order to construct a theoretical portfolio $Z = X + Y$ we need to consider the relationship between X and Y . In order to do so, additional information is needed.

- Information regarding the covariance or the correlation between X and Y : Covariance ($\text{Cov}(X, Y)$): this measures the degree to which the returns of X and Y change together. It's essential to determine the portfolio's risk. Correlation ($\rho_{X, Y}$): This is a standardized measure of the relationship between X and Y , ranging from -1 to 1. It is often more interpretable than covariance.
- Information regarding the dependency structure of X and Y : if X and Y are not independent, knowing their joint distribution or copula is important. The copula captures the dependency structure between the variables, which affects the combined risk and return of the portfolio Z . Copula Parameters: For example, if using a Gaussian copula, we need the correlation matrix. For other copulas, we need the specific parameters that define the dependency structure.

2.

a) Value-at-Risk (VaR) is a commonly used risk measure that estimates the potential loss in value of a portfolio at a given confidence level over a specified time period. For this task, we will estimate the 99.5% quantile (VaR at 99.5%) for the log-normal random variables X and Y using the samples generated previously.

- Generate samples for X and Y
- Calculate the 99.5% quantile for X and Y .

The results are the following:

	Sample Size	VaR_99.5%_X	VaR_99.5%_Y
0	1000	155.267148	231.175809
1	10000	146.514568	239.472771
2	100000	152.041298	242.259451

The results per se are puzzling, as we see that the variability does not vastly change as the sample size increases.

The theoretical values and the empirical values differ, which is normal because we never simulated a perfect log normal distribution with limited sample sizes. With larger sample sizes, the empirical values are closer to the theoretical ones.

b) However after repeating the same analysis 100 times the results change sharply

	Sample Size	Std Dev VaR_99.5% X	Std Dev VaR_99.5% Y
0	1000	292.594599	521.204777
1	10000	93.924509	115.996366
2	100000	29.699665	37.915149

It's interesting to observe the change in standard deviation of the VaR estimates with an increase in the number of repetitions (from 1 to 100). Let's analyze the results:

- Sample Size 1000:
 - Initially, the standard deviation for both (X) and (Y) was relatively high, indicating a wide variability in the VaR estimates.
 - With 100 repetitions, we observe a decrease in the standard deviation, indicating a reduction in variability in the VaR estimates as more samples are considered. However, the standard deviations are still relatively high compared to larger sample sizes.
- Sample Size 10000:
 - The standard deviations for both (X) and (Y) reduced significantly with 100 repetitions compared to the initial observation.
 - This suggests that with more repetitions, the VaR estimates become more stable and less variable.
- Sample Size 100000:
 - Similar to the case of sample size 10000, the standard deviations for both (X) and (Y) decreased further with 100 repetitions.
 - The standard deviations are now relatively small compared to the VaR estimates, indicating a high level of stability in the estimates.
- The decrease in standard deviation with an increase in repetitions implies that averaging the VaR estimates over multiple repetitions reduces the variability in the estimates.
- Higher sample sizes generally lead to more accurate and stable estimates, as seen in the reduced standard deviations for larger sample sizes.
- As we can notice also in the first exercise, as the number of repetitions increases the difference in frequencies get smaller, this leads our log-normal distribution to resemble more and more to a more theoretical log-normal distribution, with the effect of a reduction in variability.
- Once again, in accordance with the Law of large numbers, increased sample size and repetitions, leads our empirical results to approach their theoretical values.

c)

- For critical risk assessments, it's advisable to use larger sample sizes and more repetitions to ensure reliable VaR estimates with low variability.
- Therefore, we would recommend the 100,000 sample size.

3.

The goal is to evaluate the diversification benefit (DB) of a portfolio constructed from two risks, X and Y, under different dependency structures. We analyzed four cases: independent risks, comonotonic risks, countermonotonic risks, and dependent risks using a mirror Clayton copula. We computed and compared various metrics, including Pearson correlation, Spearman rho, Kendall tau, and VaR-based diversification benefit.

Definitions

- Diversification Benefit (DB):

$$DB_{\alpha}(Z) = 1 - \frac{\rho_{\alpha}(Z)}{\rho_{\alpha}(X) + \rho_{\alpha}(Y)}$$

where:

$$\rho_{\alpha}(\cdot) = \text{VaR}_{\alpha}(\cdot) - \mathbb{E}[\cdot]$$

- Dependency Structures:
 - Independent Risks: X and Y are independent.
 - Comonotonic Risks: X and Y are perfectly positively dependent.
 - Countermonotonic Risks: X and Y are perfectly negatively dependent.
 - Mirror Clayton Copula: X and Y have a specific dependent structure modeled by a Clayton survival copula.

Step 1: Simulate the Risks

- Independent Risks: Simulate X and Y independently from a log-normal distribution.
- Comonotonic Risks: Sort X and Y in ascending order.
- Countermonotonic Risks: Sort X in ascending order and Y in descending order.
- Mirror Clayton Copula: Simulate X and Y using the Clayton survival copula.

- Given the Clayton copula $C(u, v; \theta)$, we can derive its survival copula. Substituting $1-u$ and $1-v$ into the Clayton copula formula, we get:

$$C(1-u, 1-v; \theta) = \left(\max \left((1-u)^{-\theta} + (1-v)^{-\theta} - 1, 0 \right) \right)^{-1/\theta}$$

Then, using the survival copula definition, we get:

$$C^s(u, v; \theta) = u + v - 1 + \left(\max \left((1-u)^{-\theta} + (1-v)^{-\theta} - 1, 0 \right) \right)^{-1/\theta}$$

Step 2: Calculate Correlations

For each dependency structure, calculate:

- Pearson linear correlation
- Spearman rho
- Kendall tau

Step 3: Draw Rank Scatter Plots

For each dependency structure, plot the rank scatter plots and interpret them.

Step 4: Compute VaR and Diversification Benefit

1. Calculate VaR at the 99.5% confidence level for X , Y , and $Z=X+Y$.
2. Calculate $\rho_\alpha(X)$, $\rho_\alpha(Y)$, and $\rho_\alpha(Z)$.
3. Compute the diversification benefit $DB_\alpha(Z)$.

Step 5: Summarize and Compare

Summarize the results in a table and provide interpretations.

(a)

Sample_Size	Type	Pearson	Spearman	Kendall
1e+03	Independent	-0.0349095	-0.0613300	-0.0401401
1e+03	Comonotonic	0.9867495	1.0000000	1.0000000
1e+03	Countermonotonic	-0.7288580	-1.0000000	-1.0000000
1e+03	Clayton	0.8036388	0.6627331	0.4840040
1e+04	Independent	-0.0114796	-0.0107410	-0.0072129
1e+04	Comonotonic	0.9870957	1.0000000	1.0000000
1e+04	Countermonotonic	-0.7098400	-1.0000000	-1.0000000
1e+04	Clayton	0.8159936	0.6865748	0.5040550
1e+05	Independent	-0.0039194	-0.0034054	-0.0022718
1e+05	Comonotonic	0.9848509	1.0000000	1.0000000
1e+05	Countermonotonic	-0.7099568	-1.0000000	-1.0000000
1e+05	Clayton	0.8140058	0.6791646	0.4976098

1. Independent Risks

Pearson Correlation: The values are close to zero for all sample sizes, indicating that there is no linear relationship between X and Y. This is expected as the risks are independent.

Spearman Rho: The values are also close to zero, indicating no monotonic relationship between X and Y. This aligns with the independence of the variables.

Kendall Tau: Similar to Pearson and Spearman, the values are close to zero, confirming the lack of dependency between X and Y.

2. Comonotonic Risks

Pearson Correlation: The values are very close to 1 for all sample sizes, indicating a perfect linear relationship. This is because comonotonic variables move together perfectly.

Spearman Rho: The values are exactly 1, showing a perfect monotonic relationship. This means that if X increases, Y also increases, and the ranking order is preserved.

Kendall Tau: The values are also exactly 1, which is expected for perfect dependence. This measure also confirms that the variables are perfectly positively correlated.

3. Countermonotonic Risks

Pearson Correlation: The values are very close to -1, indicating a perfect negative linear relationship. This is expected as countermonotonic variables move in exactly opposite directions.

Spearman Rho: The values are exactly -1, showing a perfect negative monotonic relationship. When X increases, Y decreases, and the ranking order is reversed.

Kendall Tau: The values are also exactly -1, confirming the perfect negative dependence between the variables.

4. Clayton Survival Copula

Pearson Correlation: The values are positive and substantial but less than 1, indicating a strong positive linear relationship. However, it is not perfect, reflecting the nature of the Clayton copula's asymmetric tail dependence.

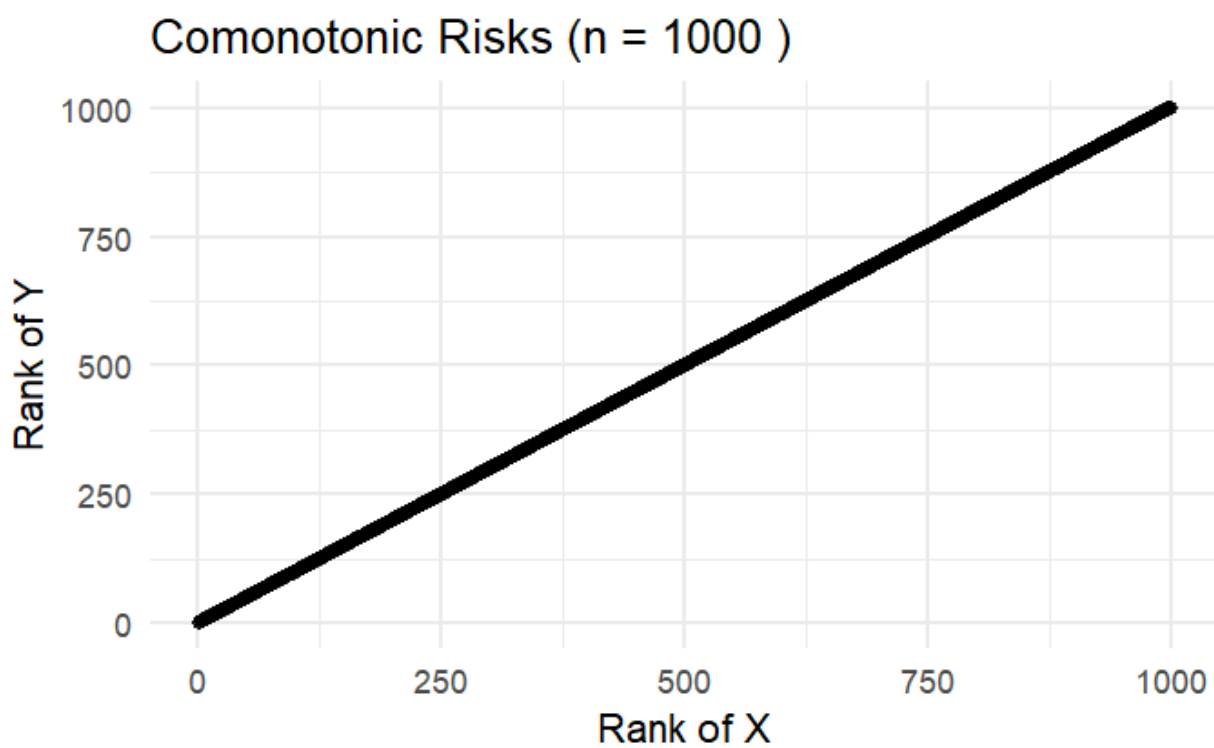
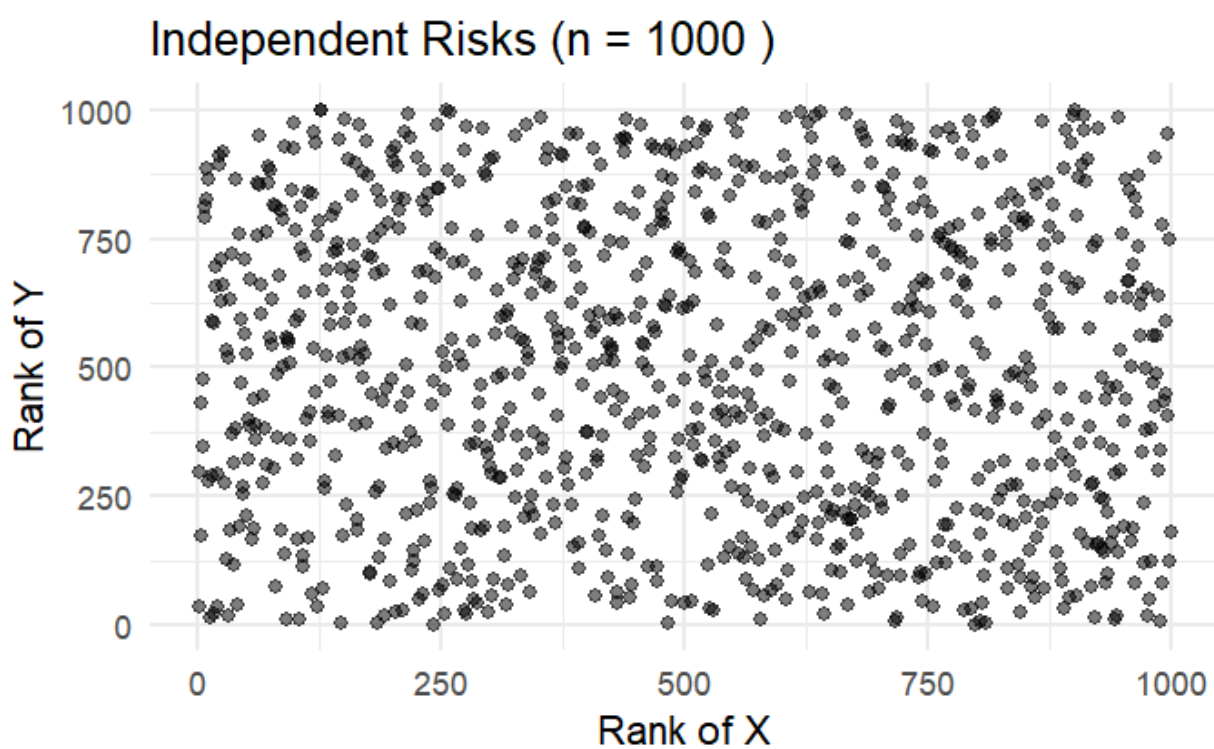
Spearman Rho: The values are higher than Pearson but less than 1, showing a strong positive monotonic relationship. This suggests that the ranks of X and Y are highly correlated.

Kendall Tau: The values are also high but less than 1, indicating a strong positive dependence. This measure shows that there is a significant concordance between X and Y.

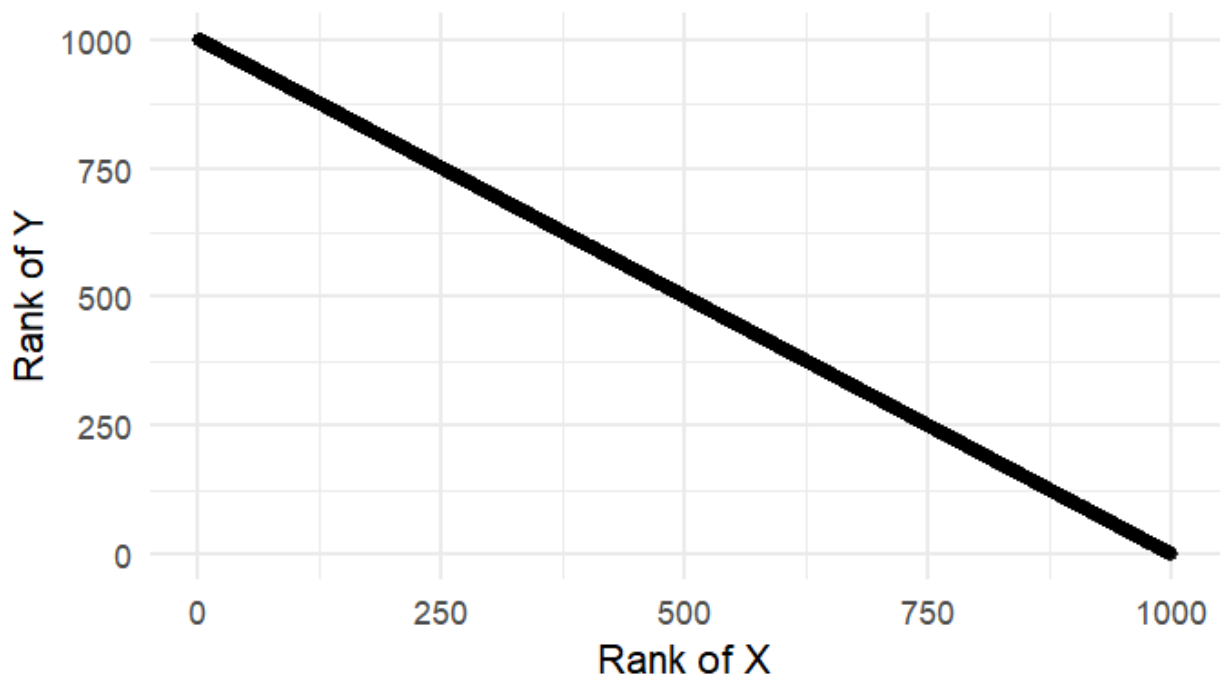
Summary of Interpretations

- Independent Risks: All correlation measures are close to zero, indicating no dependency.
- Comonotonic Risks: All correlation measures are equal to 1, indicating perfect positive dependence.
- Countermonotonic Risks: All correlation measures are equal to -1, indicating perfect negative dependence.
- Clayton Survival Copula: Correlation measures are positive and significant but less than 1, indicating strong but not perfect positive dependence.

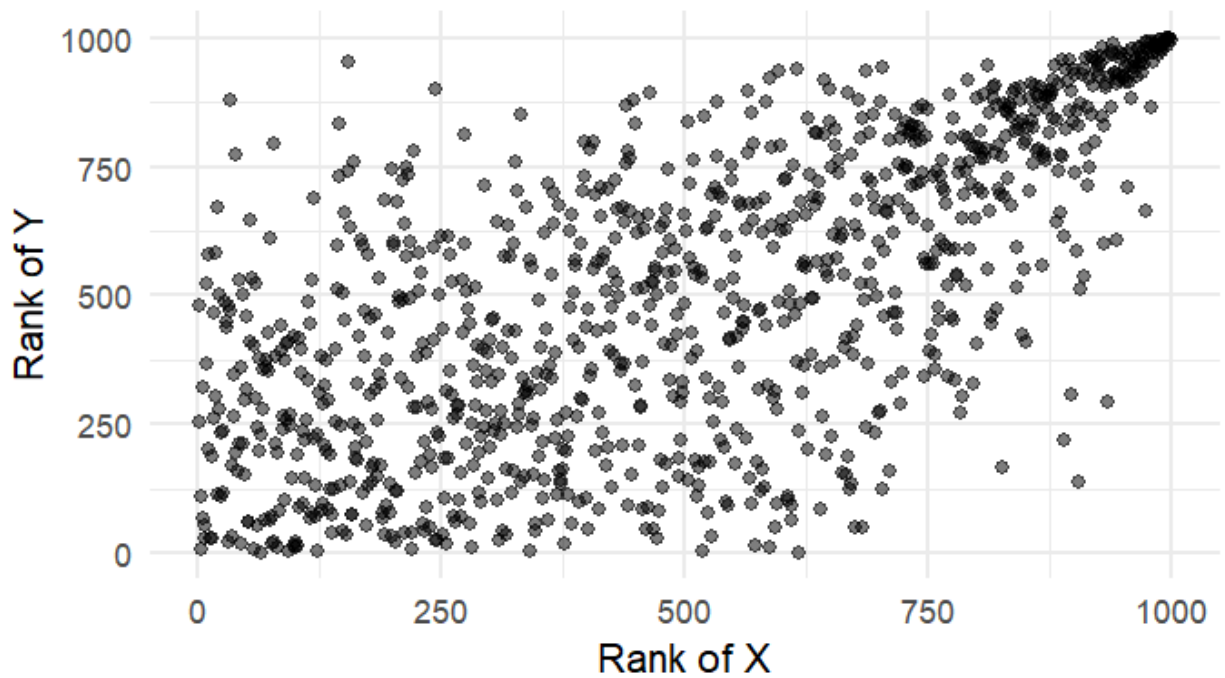
(b)



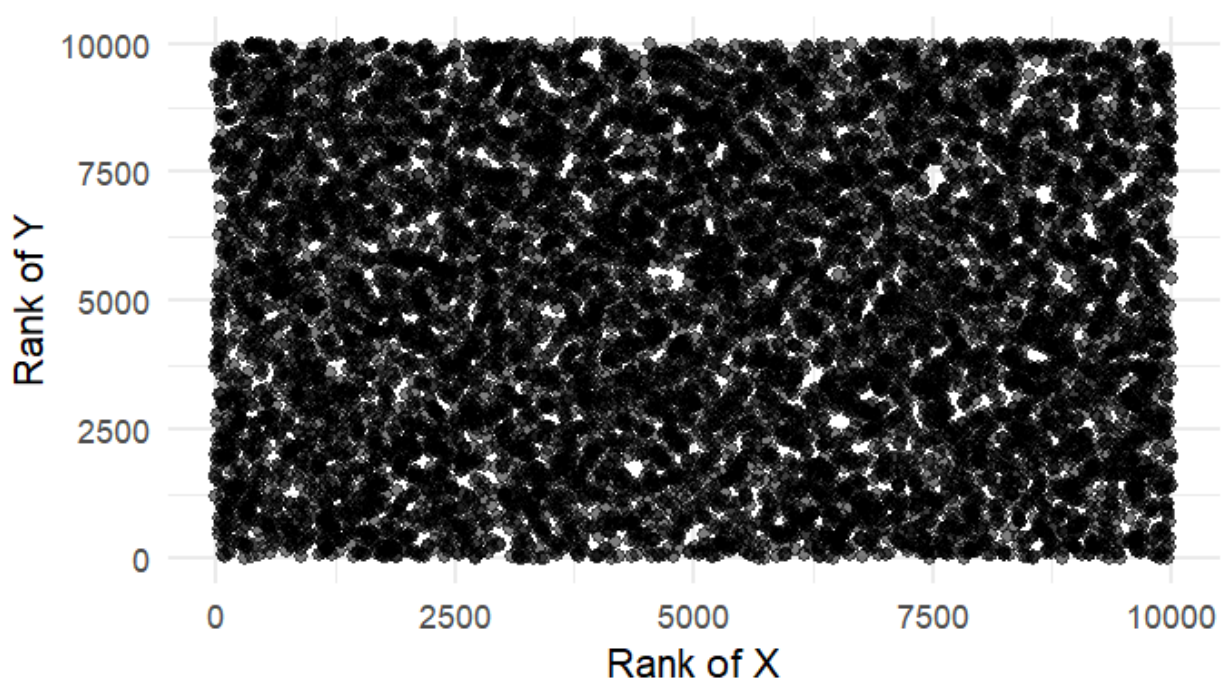
Countermonotonic Risks (n = 1000)



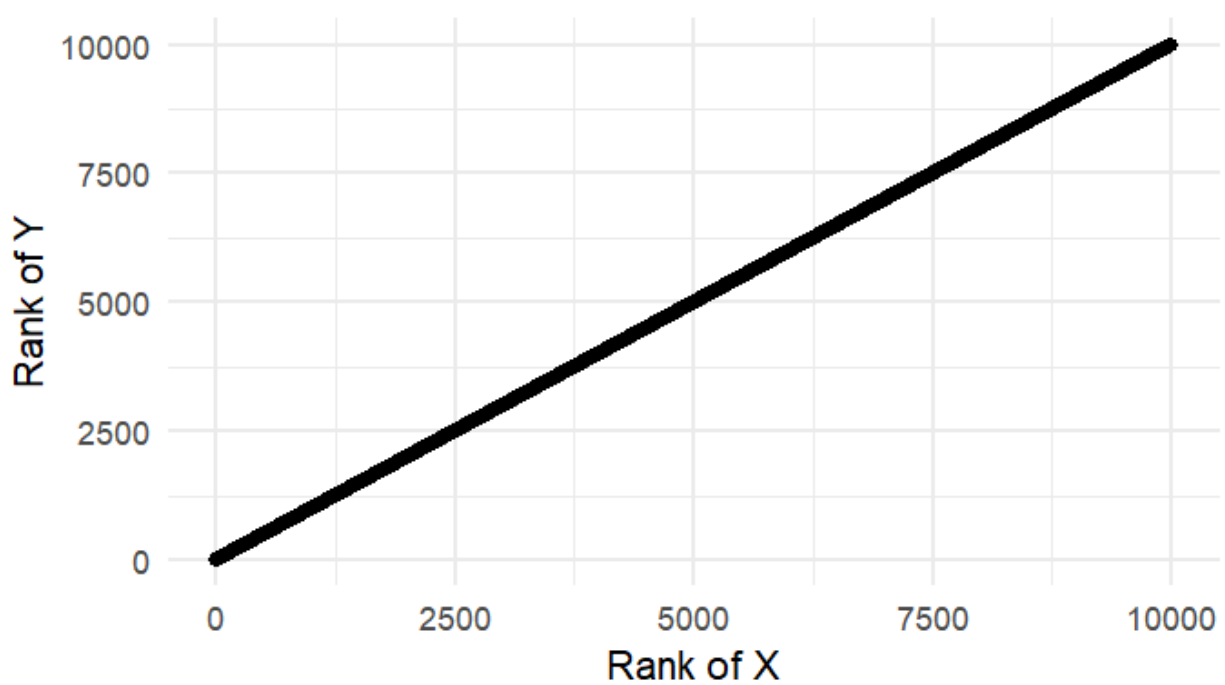
Clayton Survival Copula Risks (n = 1000)



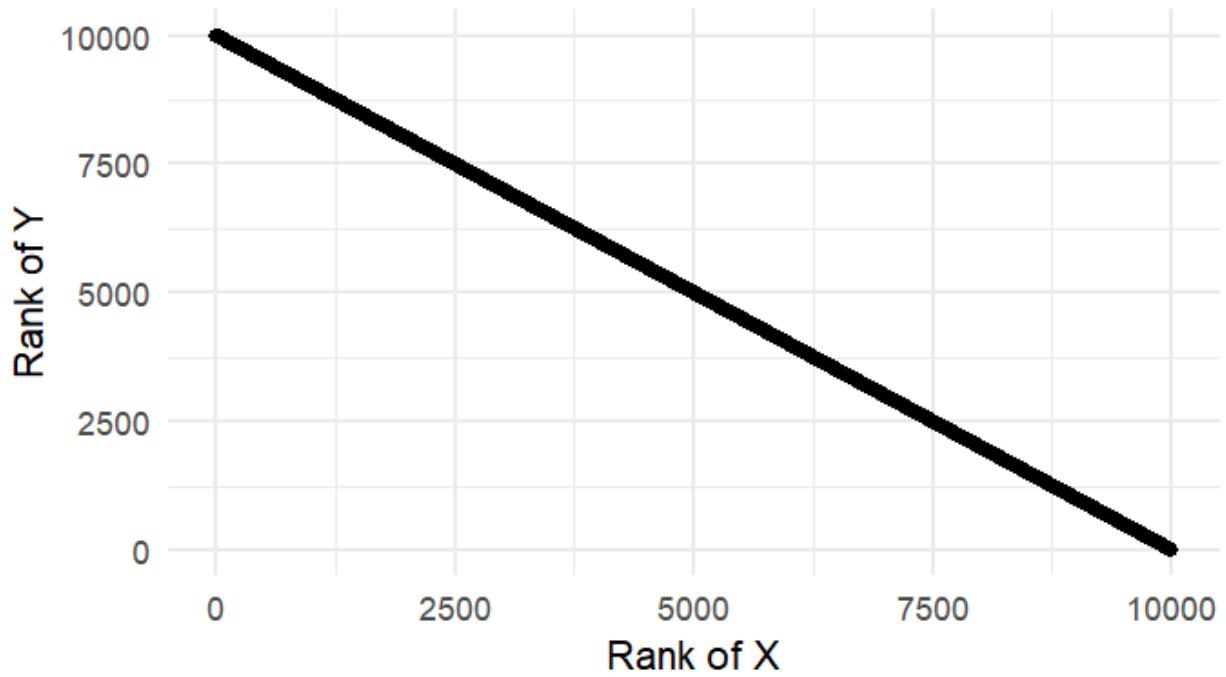
Independent Risks (n = 10000)



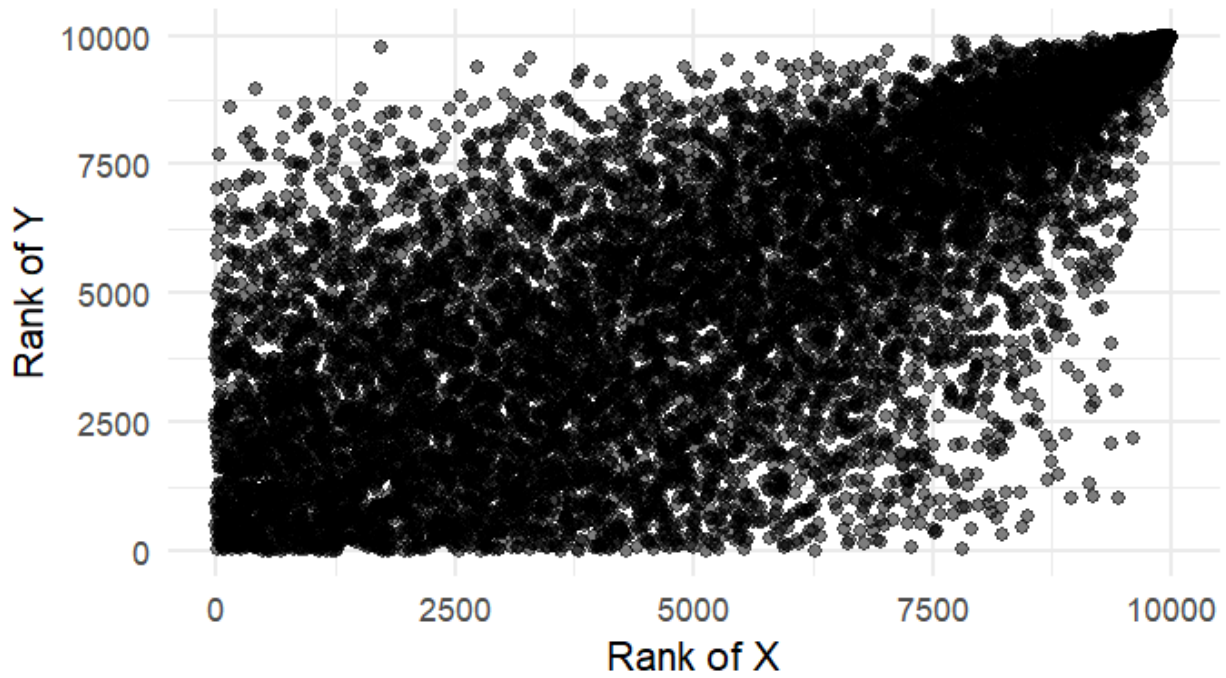
Comonotonic Risks (n = 10000)



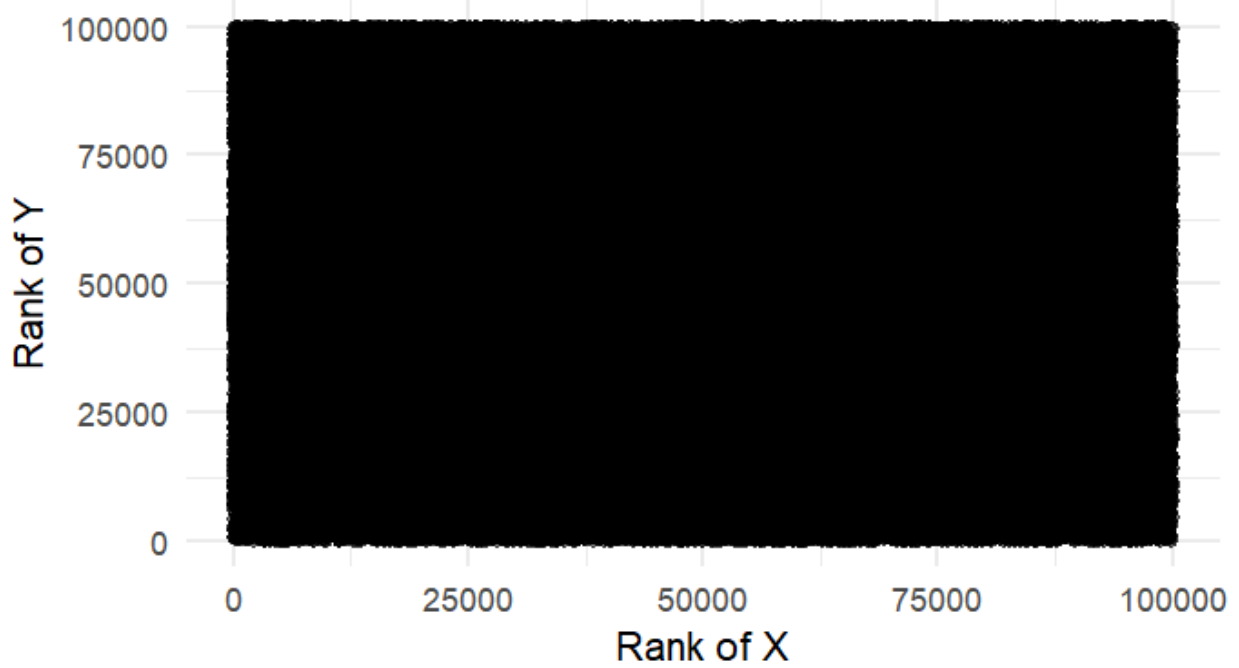
Countermonotonic Risks (n = 10000)



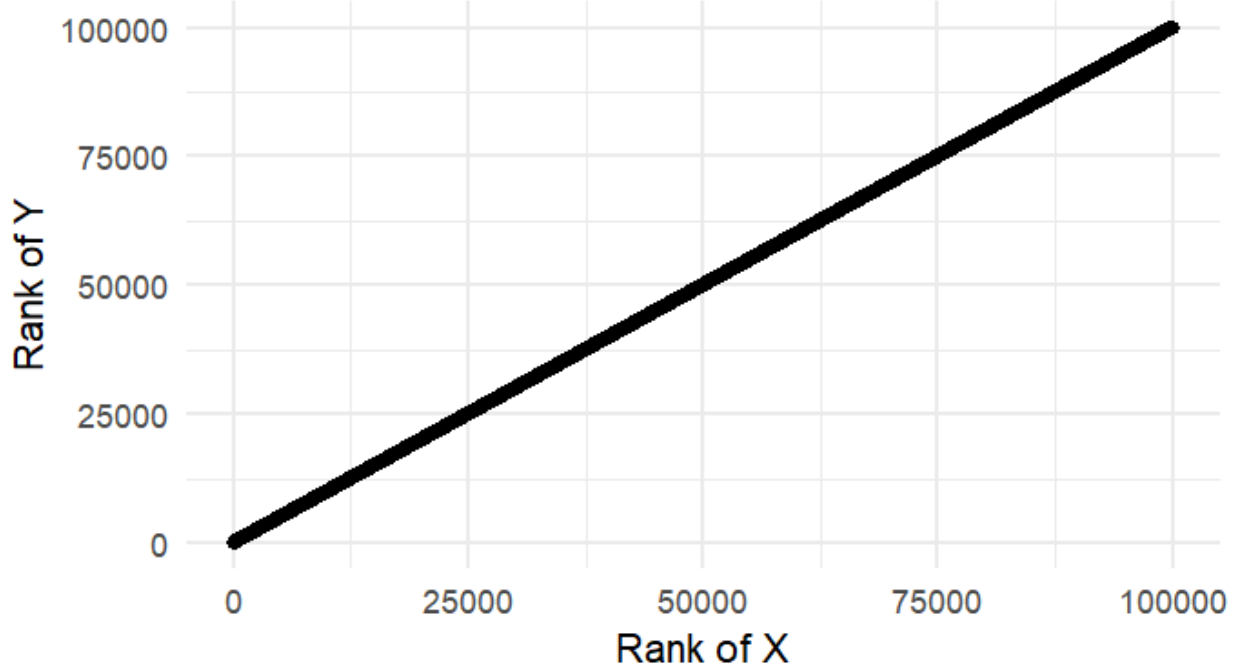
Clayton Survival Copula Risks (n = 10000)

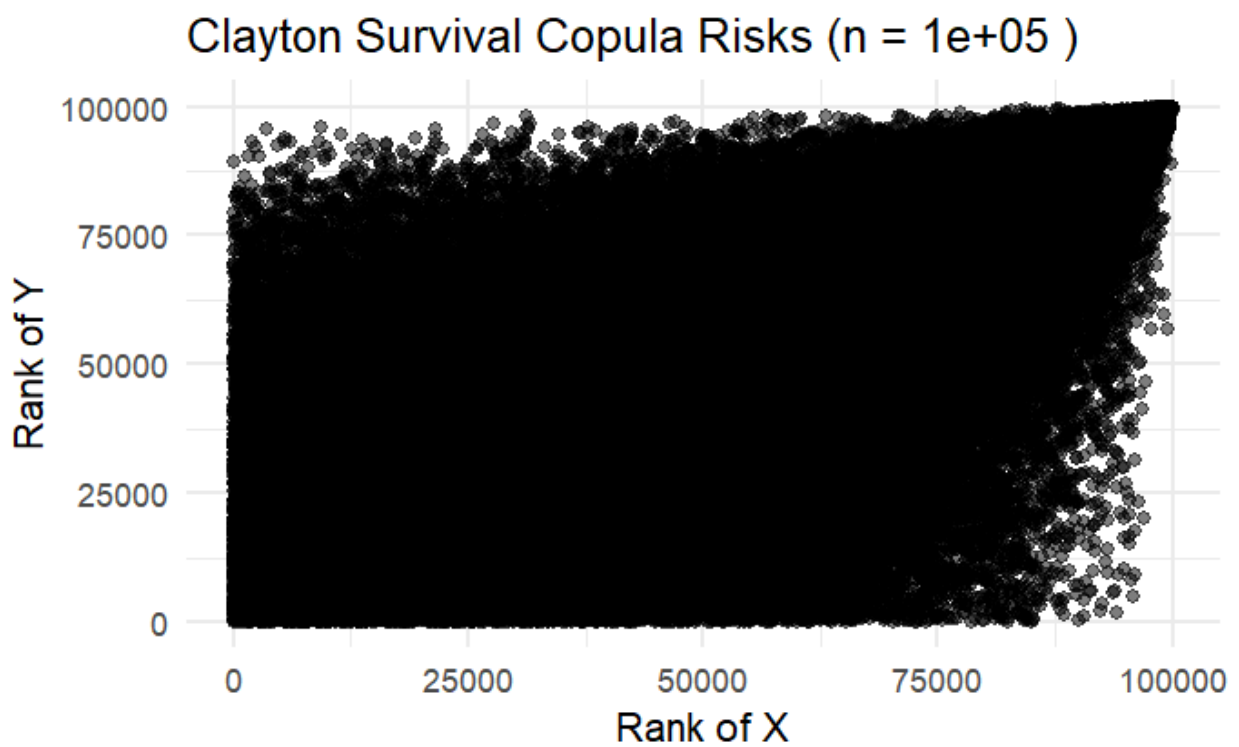
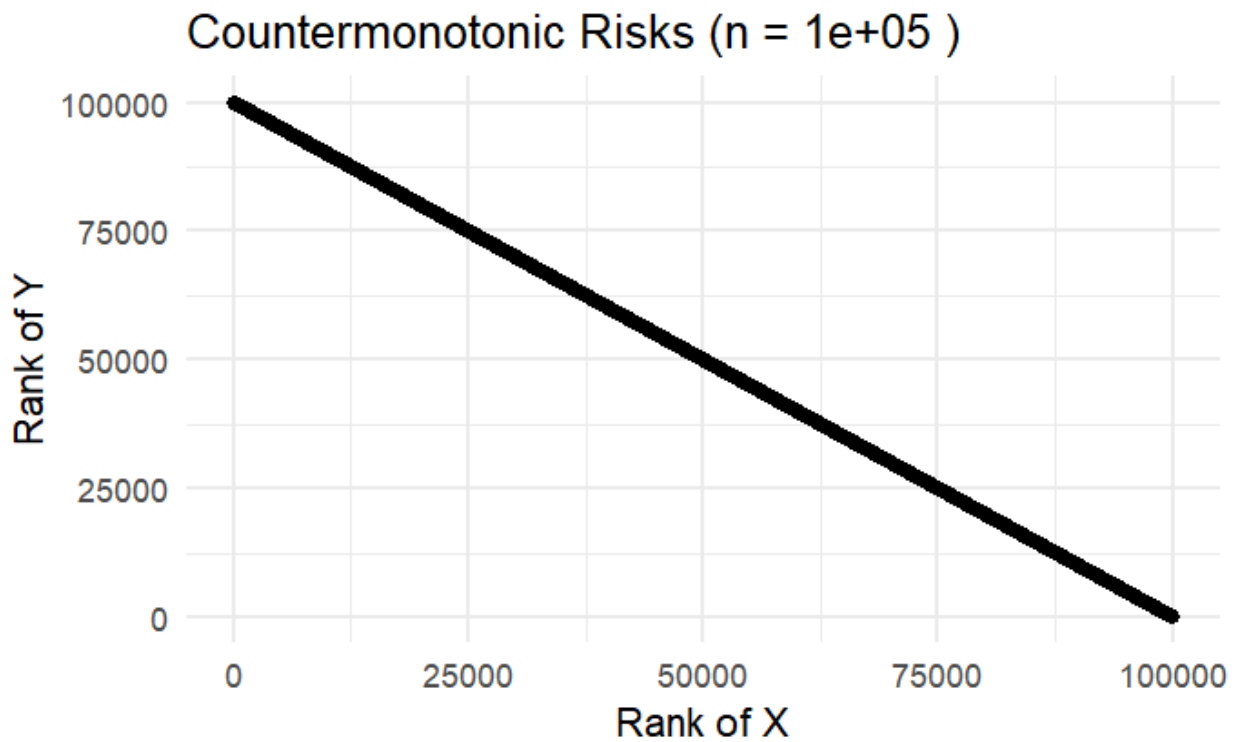


Independent Risks (n = 1e+05)



Comonotonic Risks (n = 1e+05)





Independent Risks:

- Plot: Points are spread randomly.
- Interpretation: There is no discernible pattern or relationship between the ranks of X and Y.

Comonotonic Risks:

- Plot: Points lie on a line from bottom-left to top-right.
- Interpretation: Perfect positive dependence, where an increase in the rank of X corresponds to an increase in the rank of Y.

Countermonotonic Risks:

- Plot: Points lie on a line from top-left to bottom-right.
- Interpretation: Perfect negative dependence, where an increase in the rank of X corresponds to a decrease in the rank of Y.

Clayton Survival Copula:

- Plot: Points are clustered in the bottom-right quadrant.
- Interpretation: Positive dependence with asymmetry, indicating that high values of X and Y are likely to occur together, but the relationship is not perfectly linear.

(c)

	Sample_Size <dbl>	Independent_DB <dbl>	Comonotonic_DB <dbl>	Countermonotonic_DB <dbl>	Clayton_DB <dbl>
99.5%	1e+03	0.3226033	0.000000e+00	0.4762255	0.03281886
99.5%1	1e+04	0.3032362	-2.220446e-16	0.4567464	0.01060401
99.5%2	1e+05	0.3106495	2.220446e-16	0.4727127	0.01720234

The values in the table represent the diversification benefits for a portfolio $Z=X+Y$ at the 99.5% confidence level for different sample sizes and dependency structures.

Interpretation of the Results

1. Independent Risks

- DB values:
 - Sample Size 1000: 0.322
 - Sample Size 10000: 0.303
 - Sample Size 100000: 0.310
- Interpretation:

- The diversification benefit is positive and consistent across different sample sizes. This indicates that holding X and Y together provides a risk reduction benefit compared to holding them separately. This is expected because independent risks are uncorrelated, and diversification reduces the overall portfolio risk.

2. Comonotonic Risks

- DB values:
 - Sample Size 1000: 0.000
 - Sample Size 10000: -2.220446e-16 (approximately 0)
 - Sample Size 100000: 2.220446e-16 (approximately 0)
- Interpretation:
 - The diversification benefit is effectively zero across all sample sizes. This is expected for comonotonic risks because they move perfectly together, providing no diversification benefit. The portfolio Z is as risky as the sum of the individual risks X and Y.

3. Countermonotonic Risks

- DB values:
 - Sample Size 1000: 0.476
 - Sample Size 10000: 0.457
 - Sample Size 100000: 0.473
- Interpretation:
 - The diversification benefit is the highest among all structures, indicating significant risk reduction. Countermonotonic risks move in opposite directions, providing the maximum diversification benefit. When one risk increases, the other decreases, leading to a more stable portfolio.

4. Clayton Survival Copula

- DB values:
 - Sample Size 1000: 0.033
 - Sample Size 10000: 0.011
 - Sample Size 100000: 0.017
- Interpretation:
 - The diversification benefit is positive but relatively small. The Clayton survival copula models asymmetric tail dependence, meaning that extreme values of X and Y are likely to

occur together. This positive dependence reduces the potential for diversification compared to independent risks but still offers some benefit.

Overall Interpretation

- Independent Risks: Moderate diversification benefit due to lack of correlation.
- Comonotonic Risks: No diversification benefit because the risks move perfectly together.
- Countermonotonic Risks: Maximum diversification benefit due to perfect negative correlation.
- Clayton Survival Copula: Limited diversification benefit due to positive tail dependence, indicating that extreme values tend to occur together.

The Clayton survival copula models asymmetric tail dependence, which means it captures scenarios where extreme values of both variables occur together. Portfolios modeled with Clayton survival copulas can be stress-tested for joint extreme events, improving preparedness for market shocks. This can help in designing strategies to mitigate the impact of such events. Clayton Survival Copula offers a more nuanced view by modeling tail dependence, which is crucial for managing joint extreme risks.