TAIL SETS AND LEVEL CURVES OF BIVARIATE DISTRIBUTIONS: GEOMETRY AND ESTIMATION

Rishigesh Patgunarajah

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE

GRADUATE PROGRAM IN APPLIED MATHEMATICS

YORK UNIVERSITY
TORONTO, ONTARIO

February 2025

Abstract

By incorporating Gumbel's bivariate exponential (BVE) distribution as the foundation, we aim to minimize losses between two business lines. Motivated by the need for a financial model, we chose Gumbel's BVE for its thin-tailed property, serving as a great foundation for future extensions to multidimensional risk measures. We derive a closed-form expression for $H(x, y)$ using the law of total probability, linking it to the value-at-risk concept represented by set $\mathcal{A}_p$. Recognizing that companies allocate capital near the boundary of $\mathcal{A}_p$, we define set $\mathcal{O}_p$ as the level p curve of optimal values. A convexity analysis via a Hessian matrix and Sylvester's Criterion provides insight into optimal capital allocation, and we apply Lagrange multipliers to prove a loss-minimization theorem. Graphs and tables illustrate our findings. Overall, this research offers practical insights into resource allocation, boosting company growth and financial stability.

Acknowledgements

The author would like to acknowledge Dr. Ricardas Zitikis of the Department of Mathematics for selecting the topic of this thesis and for his invaluable assistance throughout the process of writing it. Another greatly appreciated professor of the Mathematics Department is Dr. Edward Furman; whom also assisted in the preparation of this thesis.

The author is also overly grateful for the funding provided by York University to pursue the research for this project.

I would also like to thank my mother and uncle for encouraging me to continue and helping me find the light at the end of the tunnel during my most difficult times.

Contents

List of Figures

1 Introduction

With the over-saturation of businesses in the current timeline, it has become very crucial for companies operating with multiple business lines to manage their money wisely and allocate their resources effectively. Especially, the ability to minimize losses and maximize profits is an essential aspect in ensuring the companies stay afloat financially and aid in their longevity. This can be done with the help of mathematical models such as the one that will be discussed in this thesis, to capture trends and pinpoint other underlying factors that influence the correlation between certain variables. In particular, we use Gumbel's BVE distribution function to address the optimal allocation of resources, as well as minimize the losses between two business lines.

The properties of the BVE distribution function has important applications used in probability theory and applications. In fact, the tail probability of the function allows us to interpret various forms of data. For instance, one of the derivations of the bivariate exponential distribution function from Gumbel's paper has essential connections in insurance, finance, and economics.

By leveraging Gumbel's work as a foundation, one aim of this paper is to analyze the set $\mathcal{A}_p$ which in finance is called the value-at-risk (VaR). VaR is used as a measure of the risk of loss on a certain group of financial assets, often used as a benchmark for risk management. In particular, its main purpose is to guarantee that the probability of exceeding a certain value deemed dangerous remains within acceptable limits. It is reported as a positive value, which implies that a negative VaR suggests there is a higher probability of making a profit. However, in risk measurement, VaR is considered to be a number rather than a system, where it is used to determine the price movements. Hence, our research extends the application of VaR from a univariate distribution to it's bivariate setting. By closely analyzing this extension, we aim to expand our research even further into it's multidimensional applications.

The other aim of this paper is to determine the optimal value a company needs to minimize the losses between two business lines. By applying the law of total probability, we establish a closed-form expression for the tail probability $\overline{H}(x, y)$, which enables us to depict the relationship between the two business lines and their associated losses. By considering the convexity properties of the sets $\mathcal{A}_p$ and $\mathcal{O}_p$, we can determine the optimal allocation of resources that minimizes losses. This convexity analysis, combined with the exponential nature of the sets, provides valuable insights into the trade-offs involved in resource allocation decisions. Our main result is then obtained by applying the method of Lagrange multipliers, which allows us to derive a theorem that minimizes losses in the two business lines. By applying this theorem, we aim to provide companies with an effective framework in guiding them to allocate their resources in a way that maximizes profit while minimizing losses.

The insights gained from this study offer industries innovative ideas in developing strategies and forming decisions when operating their businesses. Furthermore, our findings provide a solid precursor for the optimization of resource allocation across multiple business lines. Overall, this thesis aims to address the complex challenges faced by companies operating with multiple business lines by presenting a comprehensive mathematical model derived from Gumbel's bivariate exponential distribution function. The subsequent analysis and insights derived from this model offer valuable guidance for companies seeking to make informed resource allocation decisions while minimizing losses, particularly through the lens of value-at-risk in a multidimensional context.

We will explain our result in more detail throughout the subsections of this paper.

2 Theory

This section presents the theoretical foundation of multivariate Value-at-Risk. It begins by utilizing Gumbel's bivariate exponential distribution function to establish a closed form expression for its corresponding tail probability. The discussion also covers Pearson's correlation coefficient, which allows us to compute the value for our own model. Before proceeding further, we also examine the effect our dependent parameter has to the function $\overline{H}(x, y)$ and its relevance to our newfound correlation coefficient.

2.1 Gumbel's Joint Distribution Function

In 1960, E.J. Gumbel published a paper on an in-depth analysis of the bivariate distributions with exponential margins [2]. Initially, Gumbel's distributions were used to model extreme events for weather catastrophes such as floods. Eventually, the insights gained from these applications aided in the development of risk models in insurance and finance of where the distributions would be used to depict extreme losses. The property of the BVE distribution being thin-tailed allowed the model to be a great benchmark in studying and analyzing it's features in a financial risk setting. This paper also influenced a lot of other mathematicians in their research of similar topics. For instance, Marshal and Olkin published a paper in 1967 [1], using the function developed by Gumbel and introduced new parameters and applications to the model. In particular, they provided a derivation of the bivariate exponential distribution by applying a Poisson process with the parameter λ, to depict a "fatal shock" model. They do this by applying Independent Poisson processes $Z_1(t; \lambda_1)$,$Z_2(t; \lambda_2)$,$Z_{12}(t; \lambda_{12})$, which controls the occurrence of shocks. They then denote X and Y as the life of the first and second components of a two component system, and obtain the following result;

$$\overline{F}(s, t) = e^{-\lambda_1 s - \lambda_2 t - \lambda_{12} \max(s,t)}$$

(1967,pg.3), where s and t denote the survival time of the first and second components, respectively. We use a similar derivation to depict our business model.

Hence, we start by defining the joint distribution function presented by Gumbel.

Definition 2.1. Let X and Y be two random variables. Denote their marginal distribution functions by F and G, respectively, and the joint distribution function by H, that is,

$$H(x, y) = \mathbb{P}\big(X \leq x, Y \leq y\big).$$

Furthermore, denote the corresponding tail probability by

$$\overline{H}(x, y) = \mathbb{P}\big(X > x, Y > y\big),$$

which plays an important role in various problems of insurance, finance, and economics.

Definition 2.2. Let X be a continuous random variable. Then, a probability density function (pdf) is a function; f, of X used to determine the probability that X lies within a certain range of values. In particular, a pdf is an integrable function where

$$\int_a^b f(x)dx = \mathbb{P}\big(a < X \leq b\big) \quad \text{such that } a, b \in \mathbb{R},$$

where $\int_a^b f(x)dx = 1$. Furthermore, this function is always non-negative and can be determined by applying the fundamental theorem of calculus and taking the derivative of the cumulative distribution function (cdf); F. Similarly, a bivariate distribution has a density function;

$$f(x,y) = \frac{\partial^2 F}{\partial x \partial y}.$$

Note 2.1. Based off Gumbel's paper [2], we can generalize the bivariate exponential distribution in the following form:

$$H(x,y) = 1 - e^{-\delta_1 x} - e^{-\delta_2 y} + e^{-\delta_1 x - \delta_2 y - \delta_{12} xy}; \qquad \forall x, y \geq 0$$

where $\delta_{12} = \delta\delta_1\delta_2$ for $0 \leq \delta \leq 1$, $\delta_1 > 0$, and $\delta_2 > 0$. The boundary values of $H(x,y)$ are

$$H(x,0) = H(0,y) = H(0,0) = 0; \quad H(\infty,\infty) = 1.$$

Now, we must prove $\delta_1, \delta_2, \delta$, and δ_{12} must satisfy the inequalities mentioned in the distribution function of Note 2.1. From Definition 2.2, the density function $h(x,y)$ is

$$h(x,y) = e^{-\delta_1 x - \delta_2 y - \delta_{12} xy}[(\delta_1 + \delta_{12} y)(\delta_{12} x + \delta_2) - \delta_{12}]$$

with boundary values:

$$h(\infty, y) = h(x, \infty) = 0;\ h(0,0) = \delta_1\delta_2 - \delta_{12}.$$

From the last condition and the non-negativity of a density function, we see that

$$\begin{aligned} &\delta_1\delta_2 - \delta_{12} \geq 0 \\ &\delta_1\delta_2 \geq \delta\delta_1\delta_2 \\ &\delta \leq 1. \end{aligned}$$

Next, by using a property that is true for all bivariate distributions;

$$H(x,y) \leq H_1(x)$$

where $H(x,\infty) = H_1(x)$, we obtain the following

$$\begin{aligned} 1 - e^{-\delta_1 x} - e^{-\delta_2 y} + e^{-\delta_1 x - \delta_2 y - \delta_{12} xy} &\leq 1 - e^{-\delta_1 x} \\ -\delta_1 x - \delta_2 y - \delta_{12} xy &\leq -\delta_2 y \\ -x(\delta_1 + \delta_{12} y) &\leq 0. \end{aligned}$$

In order for this inequality to hold true for all x and y, we must have $\delta_2 \geq 0$ and $\delta \geq 0$. Then, with the inclusion of the boundary condition $h(0,0) = \delta_1\delta_2 - \delta_{12}$, this implies that $\delta_1 \geq 0$ as well. However, if $\delta_1 = 0$, or $\delta_2 = 0$, then $H(x,y) = 0$ which fails to satisfy the condition $H(\infty,\infty) = 1$. Thus, δ_1 and δ_2 must be strictly > 0. By combining the previous mentioned restrictions, the inequalities of the parameters mentioned in Note 3.1 are immediately verified.

It is also clear that the variables x and y must be greater than or equal to 0, as the function $H(x,y)$ is a probability and thus cannot return a negative value. Hence, we must have that $x, y \in [0,\infty)$. We can observe how the function $H(x,y)$ is depicted in Figure 2.1.

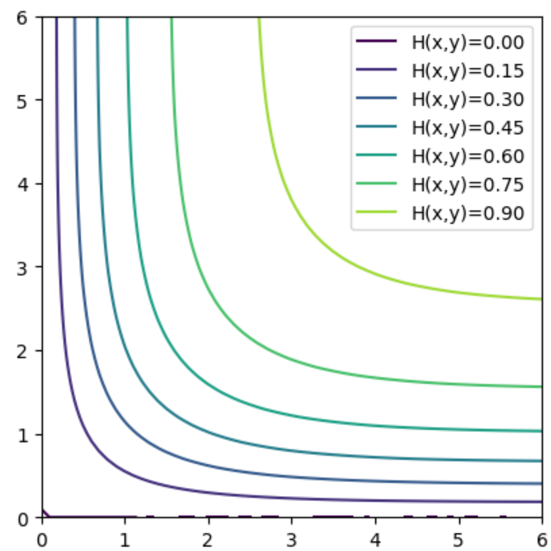

Figure 2.1: The level curves of $H(x,y)$ when δ_1 and $\delta_2 = 0.9$, and $\delta_{12} = 0.729$.

Now, a crucial step in obtaining our main result is beginning by developing a closed form expression for $\overline{H}(x,y)$.

Theorem 2.1. *Let $x, y \in (0,\infty)$ and $\delta_1, \delta_2 > 0, \delta_{12} \geq 0$. If $\overline{H}(x,y) = \mathbb{P}\big(X > x, Y > y\big)$, then $\overline{H}(x,y) = e^{-\delta_1 x - \delta_2 y - \delta_{12} xy}$.*

Proof. First, let $A = X > x$ and $B = Y > y$. Now, we apply the law of total probability to develop a closed form expression for $\overline{H}(x, y)$, as displayed by Grzegorzewski [8]. Hence, since $H(x, y) = \mathbb{P}(X \le x, Y \le y)$, this implies that $H(x, y)$ is the intersection of the complement of the two events A and B. Thus, we have the following:

$$\begin{aligned} H(x, y) &= \mathbb{P}\big(A' \bigcap B'\big) \\ &= 1 - \mathbb{P}\big(A' \bigcap B'\big)' \\ &= 1 - \mathbb{P}\big(A \bigcup B\big) \\ &= 1 - \big(\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \bigcap B)\big). \end{aligned}$$

Similarly, by applying the same rule, we also have that

$$\overline{H}(x, y) = \mathbb{P}\big(A \bigcap B\big).$$

Now since we have:

$$\begin{aligned} H(x, y) &= 1 - (e^{-\delta_1 x} + e^{-\delta_2 y} - e^{-\delta_1 x - \delta_2 y - \delta_{12} xy}) \\ H(x, y) &= 1 - \big(\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \bigcap B)\big), \end{aligned}$$

this implies that $\mathbb{P}(A) = e^{-\delta_1 x}, \mathbb{P}(B) = e^{-\delta_2 y}$ and $\mathbb{P}(A \cap B) = e^{-\delta_1 x - \delta_2 y - \delta_{12} xy}$. This further implies that

$$\overline{H}(x, y) = e^{-\delta_1 x - \delta_2 y - \delta_{12} xy}.$$

□

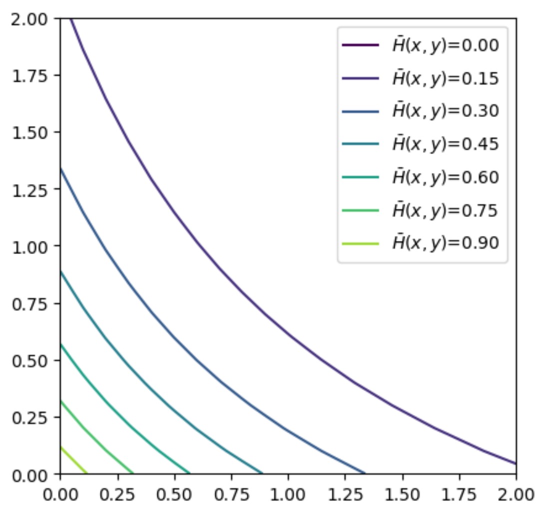

Figure 2.2: The level curves of $\overline{H}(x, y)$ when δ_1 and $\delta_2 = 0.9$ and $\delta_{12} = 0.729$.

2.2 Pearson's Correlation Coefficient

The Pearson's correlation coefficient is a statistic measure used to describe the strength and direction of the linear relationship between two quantitative variables. For instance, a value of $0 < \rho < 1$, would imply that the two variables would move in the same direction. Similarly, a value of $-1 < \rho < 0$, would imply that the two variables move in the opposite direction. A value of $\rho = 0$, would then clearly imply that there is no correlation between the two variables. Hence, the Pearson's correlation coefficient plays an important role in statistics as one can be able to predict how certain parameters can affect the data being collected and how the data will be depicted. For instance, this value is essential in finance as a portfolio manager will monitor the Pearson's correlation coefficient of their companies to limit their investment's volatility and risk. The Pearson's correlation coefficient can be calculated using the formula in the following definition;

Definition 2.3. Let ρ denote the Pearson's correlation coefficient of the two random variables X and Y,

$$\rho(X,Y) = \frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}},$$

where $\mathbb{E}$ is the expected value (or mean) and Var is the variance of either random variable. The expected value is the average number of independent outcomes of a random variable and the variance measures the variability from this value.

Theorem 2.2. *Let $x, y \in (0,\infty)$ and δ_1, $\delta_2 > 0$, $\delta_{12} \geq 0$. Then $\rho(X,Y) = -(e^{\frac{1}{\delta}}/\delta)Ei(-\delta^{-1}) - 1$, where Ei is the exponential integral function defined by $Ei(x) = -\int_{-x}^{\infty} e^{-t}/t dt$.*

Proof. First, we begin by computing the following expectations:

$$\mathbb{E}(X) = \int_0^\infty \mathbb{P}\big(X > x\big)dx = \frac{1}{\delta_1}$$
$$\mathbb{E}(X^2) = \int_0^\infty x^2 d\mathbb{P}\big(X \leq x\big) = \frac{2}{\delta_1^2}.$$

Similarly, we then see that

$$\mathbb{E}(Y) = \frac{1}{\delta_2}$$
$$\mathbb{E}(Y^2) = \frac{2}{\delta_2^2}.$$

As a result, we then obtain the corresponding variances

$$Var(X) = \frac{1}{\delta_1^2}$$
$$Var(Y) = \frac{1}{\delta_2^2}.$$

Next, we compute the expectation $\mathbb{E}(XY)$;

$$\begin{aligned}
\mathbb{E}(XY) &= \int_0^\infty \int_0^\infty \mathbb{P}\big(X > x, Y > y\big) dxdy \\
&= \int_0^\infty \int_0^\infty e^{-\delta_1 x - \delta_2 y - \delta_{12} xy} dxdy \\
&= \int_0^\infty \frac{-e^{-\delta_2 y} e^{-\delta_{12} xy - \delta_1 x}}{\delta_{12} y + \delta_1} |_0^\infty dy \\
&= \int_0^\infty \frac{e^{-\delta_2 y}}{\delta_{12} y + \delta_1} dy \\
&= -\frac{e^{\frac{\delta_1 \delta_2}{\delta_{12}}}}{\delta_{12}} \int_0^\infty -\frac{e^{-(\delta_2 y + \frac{\delta_1 \delta_2}{\delta_{12}})}}{\delta_2 y + \frac{\delta_1 \delta_2}{\delta_{12}}} \delta_2 dy \\
&= \frac{-e^{\frac{\delta_1 \delta_2}{\delta_{12}}} E1(\delta_2 y + \frac{\delta_1 \delta_2}{\delta_{12}})}{\delta_{12}} |_0^\infty \\
&= \frac{e^{\frac{\delta_1 \delta_2}{\delta_{12}}} E1(\frac{\delta_1 \delta_2}{\delta_{12}})}{\delta_{12}} \\
&= \frac{-e^{\frac{\delta_1 \delta_2}{\delta_{12}}} Ei(-\frac{\delta_1 \delta_2}{\delta_{12}})}{\delta_{12}}.
\end{aligned}$$

Now, by Definition 3.3, we receive the following coefficient of correlation:

$$\begin{aligned}
\rho(X, Y) &= \frac{\left(\frac{-e^{\frac{\delta_1 \delta_2}{\delta_{12}}} Ei(-\frac{\delta_1 \delta_2}{\delta_{12}})}{\delta_{12}} - \frac{1}{\delta_1} \frac{1}{\delta_2}\right)}{\frac{1}{\delta_1} \frac{1}{\delta_2}} \\
&= \frac{-\delta_1 \delta_2 e^{\frac{\delta_1 \delta_2}{\delta_{12}}} Ei(-\frac{\delta_1 \delta_2}{\delta_{12}}) - \delta_{12}}{\delta_{12}} \\
&= \frac{-Ei(-\frac{1}{\delta}) e^{\frac{1}{\delta}}}{\delta} - 1.
\end{aligned}$$

Hence, this proves Theorem 3.2. $\square$

By analyzing this equation, we notice that ρ is dependent on the δ parameter, and also must be non-positive values only. Since, if $\delta = 0$, then $\rho = 0$, implying that there is no correlation between the variables X and Y. If $\delta = 1$, then $\rho = -0.40365$, implying that there is a negative correlation between the variables X and Y, as the value of δ increases to 1. This is evident in the graph of the function displayed in Figure 2.3.

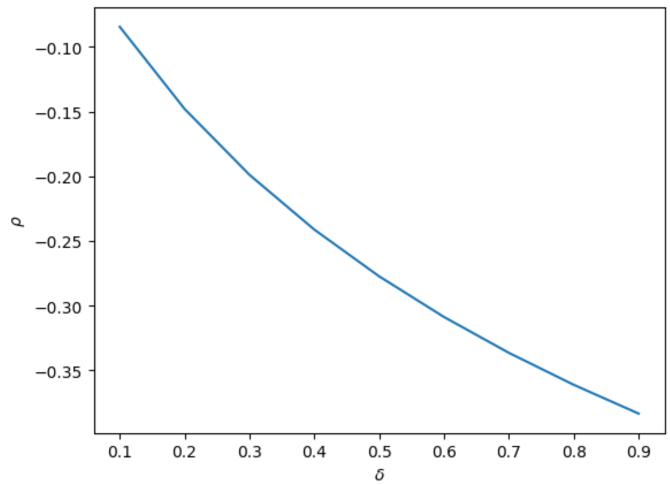

Figure 2.3: The Pearson's correlation coefficient; $\rho(X, Y)$, as a function of δ.

Note 2.2. For our function we observe a Pearson's correlation coefficient that is negative implying that losses between two business lines move in opposite directions. In risk management, this is beneficial because extreme losses in one line are less likely to occur simultaneously with extreme losses in the other. This further implies that losses in one business line may be offset by the gains in the other thus reducing the overall risk. On the other hand, there is also a possibility of having positively correlated random variables. For instance, if we take a look at Marshal and Olkin's BVE model, the common shock parameter induces a positive correlation between X and Y. However, this may not be an ideal model to analyze for risk management as positively correlated exponential random variables affects the optimal capital allocation, as a higher threshold would be required to cover the losses.

Note 2.3. Notice when $\delta = 0$, this implies $\delta_{12} = 0$, hence we have a special case where

$$\overline{H}(x, y) = e^{-\delta_1 x - \delta_2 y} = e^{-\delta_1 x} e^{-\delta_2 y} = \overline{H}(x, 0) \times \overline{H}(0, y).$$

This implies that the two events (i.e the losses) x and y, become independent. This aligns with the observation that when δ equals 0, the Pearson's correlation coefficient also equals 0. The 3-dimensional graphs depicted in Figure 2.4 and Figure 2.5 display how similar the separate functions $\overline{H}(x, 0)$ and $\overline{H}(0, y)$ are. As a result, this gives us an idea of what the overall function $\overline{H}(x, y)$ would appear to be for the case of independence.

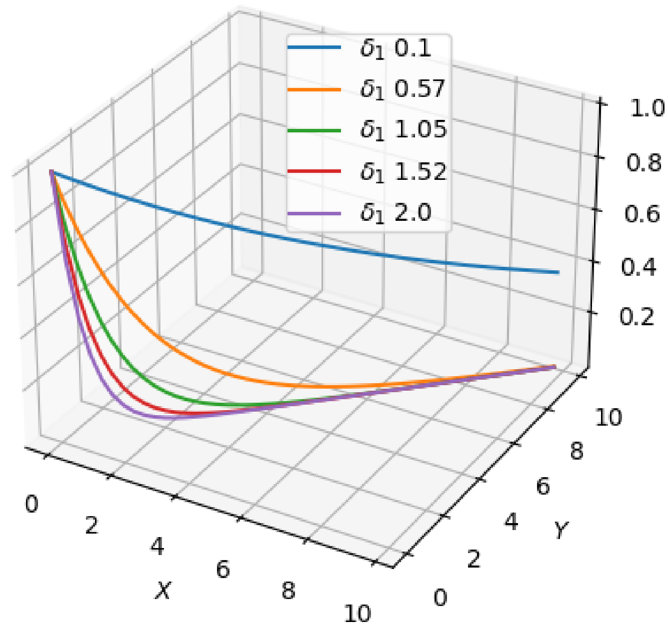

Figure 2.4: Exponential distribution $\overline{H}(x, 0)$ for different values of δ_1.

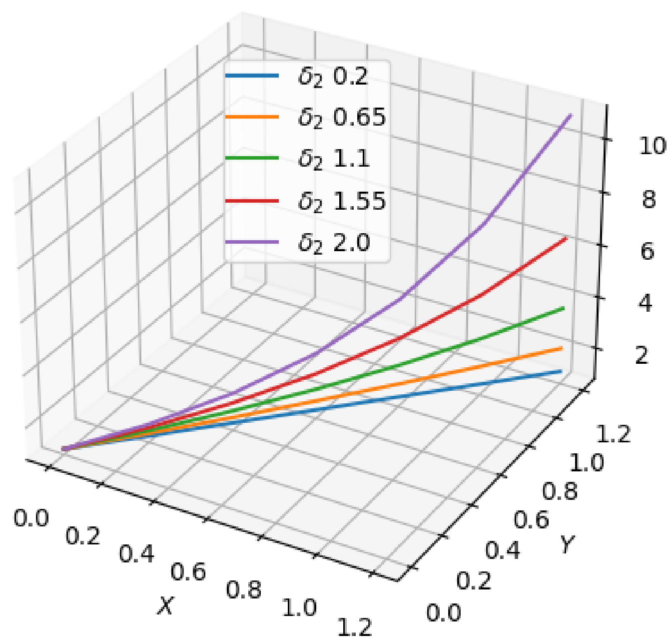

Figure 2.5: Exponential distribution $\overline{H}(0, y)$ for different values of δ_2.

3 Main Analysis and Results

In this section, the bivariate Value-at-risk set is introduced, along with the level p curve of optimal values. The VaR set and the curve are both depicted in the form of a set with an inequality involving the function $\overline{H}(x, y)$. Additionally, this section describes the convexity of these sets. A theorem is then developed to minimize the losses between two business lines. Lastly, a comparison is made between the optimal values obtained in the bivariate setting and the univariate setting.

3.1 Value at Risk

Value at Risk (VaR) has emerged as a crucial element in assessing risk for diverse financial institutions. By offering insights into the likelihood of potential losses, VaR enables businesses to allocate their resources strategically and make informed decisions in real-time, thereby enhancing their overall decision-making process. Furthermore, based on the definition, VaR can also be defined as the tail set of admissible values.

We can depict the Value at Risk, mathematically, by the upcoming definition.

Definition 3.1. Let $\mathcal{A}_p$ represent the tail set of admissible values

$$\mathcal{A}_p = \left\{(x, y) \in \mathbb{R}^2 \;:\; \overline{H}(x, y) \leq 1 - p\right\}, \tag{3.1}$$

where $p \in (0, 1)$ can be any fixed parameter, which in application is usually close to 1. For example, we may think of $p = 0.95$ or $p = 0.99$. In statistics, we call this p value the confidence level.

Note 3.1. To illustrate, think of X and Y as the losses associated with two business lines in a company. Then the set $\mathcal{A}_p$ specifies the amounts x and y that the two business lines should hold so that the two losses would jointly exceed the respective amounts x and y in no more than $(1-p)100\%$ scenarios. In finance, this set is referred to as the value-at-risk at probability level p; $\mathrm{VaR}_p(X, Y)$.

3.2 Curve of Optimal Values

Definition 3.2. Let $\mathcal{O}_p$ represent the level p curve of optimal values.

$$\mathcal{O}_p = \left\{(x, y) \in \mathbb{R}^2 \;:\; \overline{H}(x, y) = 1 - p\right\}. \tag{3.2}$$

It is natural to expect that the company would allocate the capital to the two business lines near the boundary of set (3.1), which in the case of strictly increasing $\overline{H}$ would be **the level p curve**.

Note 3.2. For example, one could think of specifying a point (x^*, y^*) on the curve $\mathcal{O}_p$ that minimizes the sum $x^* + y^*$, but this may not be the best choice from the company's point of view: various considerations play roles when choosing (x^*, y^*) from set $\mathcal{A}_p$.

Theorem 3.1. *Let $p \in (0,1)$ and $\delta_1 > 0$ and $\delta_2 > 0$, and $\delta_{12} \geq 0$. Then the domain of the curve $\mathcal{O}_p$ is [0,${\delta_1}^{-1} \log(1/1-p)$].*

Proof. By applying the definition of the curve $\mathcal{O}_p$ and the closed form expression we derived for $\overline{H}(x,y)$ we have:

$$\overline{H}(x,y) = e^{-\delta_1 x - \delta_2 y - \delta_{12} xy} = 1 - p.$$

Then, based on the fact that the probability cannot be negative, this implies that the function cannot cross the x-axis. Otherwise, we start to receive negative values for the curve $\mathcal{O}_p$. Thus, the curve must be between 0 and the x-intercept. Hence, to determine the x-intercept, we let $y = 0$ and isolate for x:

$$\begin{aligned} e^{-\delta_1 x} &= 1 - p \\ e^{\delta_1 x} &= \frac{1}{1-p} \\ x &= {\delta_1}^{-1} \log(\frac{1}{1-p}). \end{aligned}$$

Hence, the right endpoint of the domain would be the point $x = {\delta_1}^{-1} \log(\frac{1}{1-p})$, and thus this verifies Theorem 3.1. □

The result of this theorem is very important as it provides us with some insight on how the curve $\mathcal{O}_p$ will be depicted before even analyzing the convexity of the function. Specifically, we know that this curve will not exceed the first quadrant of the graph, and the δ_1 and p value plays an important role in determining how far this curve will be extended.

3.3 Convexity Analysis

Understanding the behaviour of a function is important to be able to predict a pattern of the received results. Convexity specifically provides us the ability to describe the shape of a function without knowing each and every data point. Furthermore, it is essential in the study of optimization, as a convex function possesses a unique minimum. Thus, the property of being able to minimize functions is why we understand the convexity of our function to solve one of our main objectives of this paper.

In addition, in the field of finance, convexity illustrates how the duration of a bond changes in response to interest rate fluctuations. As a form of loan from an investor, a bond's interest rate represents the percentage of the loan amount. Portfolio managers utilize convexity as a risk-management tool to assess their investments' exposure to interest rate risk. For example, a bond with higher convexity experiences more significant price changes when the interest rate declines compared to a bond with lower convexity.

Hence, prior to the main result of this section (Theorem 3.4), we first recall the definitions of convex sets and functions. As a second step, we define the Hessian matrix as a way to depict the convexity of the sets mentioned. Finally, we state our main result.

Definition 3.3. From the definition provided by Rockafellar [4], a set B is said to be convex if for any two points $x, y \in B$, the line segment

$$c = (1-\lambda)x + \lambda y$$

lies in B, with $\lambda \in [0, 1]$.

Definition 3.4. Again by Rockafellar [4], if we let B be a convex subset of a real vector space and let $f : B \to \mathbb{R}$ be a function. Then f is convex if and only if

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y)$$

such that x,$y \in B$ and $\lambda \in [0, 1]$.

This condition implies that the line segment between $(x, f(x))$ and $(y, f(y))$ must be greater than the line segment between x and y. In general, to be considered a convex function, a straight line between any pair of points on the curve of f must be above the graph of the whole function. In other words, the function at the weighted average of the x values must be less than the weighted average of the function at each x-value. Furthermore, this can be determined by evaluating where the second derivative of a function is non-negative. In particular, if $f''(x)$ is ≥ 0 for all x in an interval I, then f is convex on I. Hence, as suggested by Davydov et al. [7], we must construct a Hessian matrix to determine the convexity of our function.

Definition 3.5. A Hessian matrix is said to be a square matrix consisting of second-order partial derivatives of a function f. Typically, in the 2-D space they would be arranged in the following form:

$$\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}.$$

This is essential in analyzing the local curvature of the function as the second derivatives determine the regions of convexity.

For this paper, we are interested in the analysis of convexity to solve a minimization problem that will be discussed in the upcoming section. Thus, we determine the convexity for the set $\mathcal{A}_p$ and the curve $\mathcal{O}_p$ by the following theorem.

Theorem 3.2. *Let $x, y \in (0, \infty)$ and δ_1, δ_2, and $\delta_{12} > 0$. Then the set $\mathcal{A}_p$ and the curve $\mathcal{O}_p$ are both convex.*

Proof. In order to depict the set $\mathcal{A}_p$ and the curve $\mathcal{O}_p$ we must first calculate the Hessian matrix for $\overline{H}(x,y)$ and then prove that the Hessian is positive semi-definite. Since, a twice continuously differentiable function of several variables is convex if and only if its Hessian matrix of second partial derivatives is positive semidefinite. Hence, we begin by calculating the first derivatives of the function $\overline{H}(x,y)$

$$\frac{d\overline{H}}{dx} = (-\delta_1 - \delta_{12}y)e^{-\delta_1 x - \delta_2 y - \delta_{12}xy}$$
$$\frac{d\overline{H}}{dy} = (-\delta_2 - \delta_{12}x)e^{-\delta_1 x - \delta_2 y - \delta_{12}xy}.$$

Next, we calculate the second partial derivatives of the function to create our Hessian matrix,

$$\frac{d^2\overline{H}}{dx^2} = (-\delta_{12}y - \delta_1)^2 e^{-\delta_1 x - \delta_2 y - \delta_{12}xy}$$
$$\frac{d^2\overline{H}}{dydx} = e^{-\delta_1 x - \delta_2 y - \delta_{12}xy}(\delta_{12}^2 xy + \delta_1\delta_{12}x + \delta_2\delta_{12}y + \delta_1\delta_2 - \delta_{12})$$
$$\frac{d^2\overline{H}}{dy^2} = (-\delta_{12}x - \delta_2)^2 e^{-\delta_1 x - \delta_2 y - \delta_{12}xy}$$
$$\frac{d^2\overline{H}}{dxdy} = e^{-\delta_1 x - \delta_2 y - \delta_{12}xy}(\delta_{12}^2 xy + \delta_1\delta_{12}x + \delta_2\delta_{12}y + \delta_1\delta_2 - \delta_{12}).$$

Thus, we receive the following Hessian matrix:

$$H = \begin{bmatrix} (-\delta_{12}y - \delta_1)^2 e^{-\delta_1 x - \delta_2 y - \delta_{12}xy} & Ae^{-\delta_1 x - \delta_2 y - \delta_{12}xy} \\ Ae^{-\delta_1 x - \delta_2 y - \delta_{12}xy} & (-\delta_{12}x - \delta_2)^2 e^{-\delta_1 x - \delta_2 y - \delta_{12}xy} \end{bmatrix},$$

where $A = \delta_{12}^2 xy + \delta_1\delta_{12}x + \delta_2\delta_{12}y + \delta_1\delta_2 - \delta_{12}$. Next, we apply the Sylvester's criterion to prove the Hessian is positive semidefinite. According to Ghorpade and Limaye [9], the Sylvester's criterion states that the matrix is positive semi-definite if and only if all leading minors are non-negative. Specifically, we must prove $\frac{d^2\overline{H}}{dx^2} \geq 0$, $\frac{d^2\overline{H}}{dy^2} \geq 0$, and the determinant of the whole matrix is non-negative as well. Hence, we see that

$$\frac{d^2\overline{H}}{dx^2} = (-\delta_{12}y - \delta_1)^2 e^{-\delta_1 x - \delta_2 y - \delta_{12}xy} \geq 0$$
$$\frac{d^2\overline{H}}{dy^2} = (-\delta_{12}x - \delta_2)^2 e^{-\delta_1 x - \delta_2 y - \delta_{12}xy} \geq 0.$$

Finally, we calculate the determinant of the whole Hessian;

$$\begin{aligned}\det(H) &= (2xy\delta_{12}^3 + 2x\delta_1\delta_{12}^2 + 2y\delta_2\delta_{12}^2 + 2\delta_{12}\delta_1\delta_2 - \delta_{12}^2)(e^{-2\delta_1 x - 2\delta_2 y - 2\delta_{12}xy}) \\ &= \delta_{12}(2xy\delta_{12}^2 + 2x\delta_1\delta_{12} + 2y\delta_2\delta_{12} + 2\delta_1\delta_2 - \delta_{12})e^{-2\delta_1 x - 2\delta_2 y - 2\delta_{12}xy}.\end{aligned}$$

By definition, we can substitute $\delta_{12} = \delta\delta_1\delta_2$ in the above equation to receive the following:

$$\begin{aligned}&\delta\delta_1\delta_2(2xy(\delta\delta_1\delta_2)^2 + 2x\delta\delta_1^2\delta_2 + 2y\delta\delta_1\delta_2^2 + 2\delta_1\delta_2 - \delta\delta_1\delta_2)e^{-2\delta_1 x - 2\delta_2 y - 2\delta_{12}xy} \\ &= \delta\delta_1^2\delta_2^2(2xy\delta^2\delta_1\delta_2 + 2x\delta\delta_1 + 2y\delta\delta_2 + 2 - \delta)e^{-2\delta_1 x - 2\delta_2 y - 2\delta_{12}xy}.\end{aligned}$$

Now in order for this to be non-negative we must have

$$\delta\delta_1^2\delta_2^2 \geq 0 \quad \text{or} \quad 2xy\delta^2\delta_1\delta_2 + 2x\delta\delta_1 + 2y\delta\delta_2 + 2 - \delta \geq 0.$$

The first case is clearly non-negative as by definition we know that $0 \leq \delta \leq 1$ and since $\delta_1, \delta_2 > 0$, the overall result must be 0 or greater. Then, for our second case, if we let $\delta = 0$, then

$$2xy\delta^2\delta_1\delta_2 + 2x\delta\delta_1 + 2y\delta\delta_2 + 2 - \delta = 2$$

which is clearly non-negative. When $\delta = 1$, we receive

$$2xy\delta_1\delta_2 + 2x\delta_1 + 2y\delta_2 + 1$$

which are all positive parameters and variables, and thus must be ≥ 0. Thus, we can conclude that

$$\delta\delta_1^2\delta_2^2(2xy\delta^2\delta_1\delta_2 + 2x\delta\delta_1 + 2y\delta\delta_2 + 2 - \delta)e^{-2\delta_1 x - 2\delta_2 y - 2\delta_{12}xy} \geq 0,$$

implying that the the determinant of the Hessian matrix is non-negative. Hence, this also proves that all the leading minors are non-negative which implies that the Hessian matrix is positive semi-definite by the Sylvester's criterion. Thus, this proves that the set $\mathcal{A}_p$ and the curve $\mathcal{O}_p$ are both convex.

$\square$

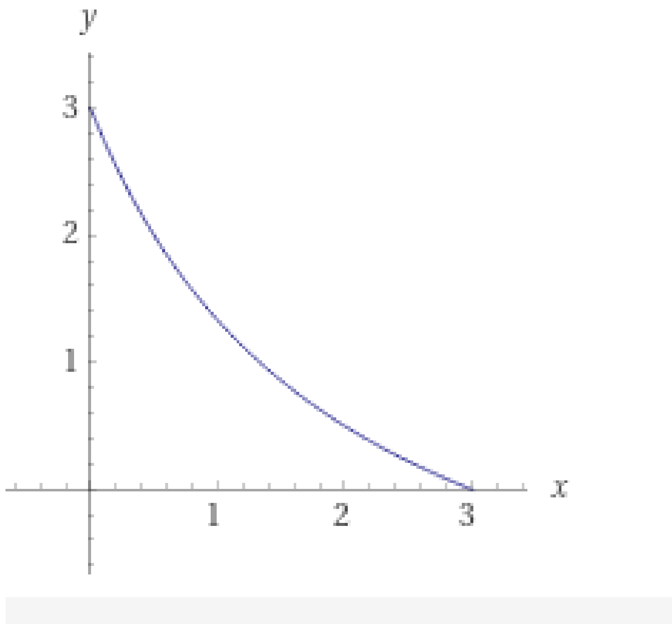

Figure 3.1: The curve $\mathcal{O}_p$ when $\delta_1, \delta_2 = 1$ and $\delta_{12} = 0.5$ and $p = 0.95$.

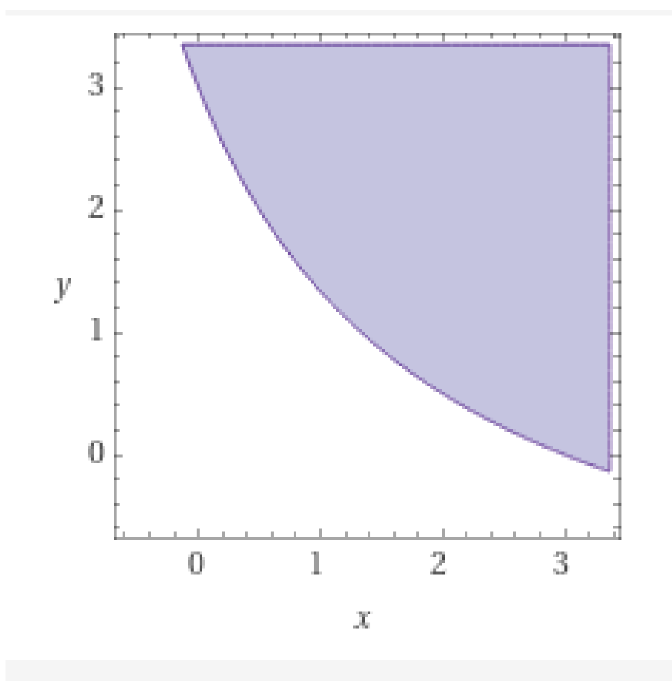

Figure 3.2: The set $\mathcal{A}_p$ when $\delta_1, \delta_2 = 1$ and $\delta_{12} = 0.5$ and $p = 0.95$.

Note 3.3. The convexity of the exponential function allow us to get a better understanding of what values the two business lines should hold to minimize their losses. For instance, if a company invests a low amount x in their first business line, then they should invest a higher amount y in their second business line. Similarly, if a company invests a higher amount x in the first business line, then a lower amount y should be invested in the second business line. This inverse relationship is a clear implication from the negative Pearson's correlation coefficient calculated from section 2.2. Hence, observing this relationship allows a company to make smarter decisions in allocating their money to either business lines, to make the most profit.

Furthermore, when analyzing the convexity of a function, it may not be convex everywhere or all the time. For instance, some functions may oscillate between convex and non-convex portions of the graph, hence resulting in being an overall non-convex function. Then, there are linear functions that are considered to be both convex and concave, due to the fact that they possess characteristics of both. Hence, the following theorem illustrates why the set $\mathcal{A}_p$ and the curve $\mathcal{O}_p$ are not strictly convex all the time.

Theorem 3.3. *Let $x, y \in (0, \infty)$ and δ_1 and $\delta_2 > 0$, and $\delta_{12} = 0$. Then the set $\mathcal{A}_p$ and the curve $\mathcal{O}_p$ are both convex and concave.*

Proof. If $\delta_{12} = 0$, then the set

$$\mathcal{A}_p = \left\{(x, y) \in \mathbb{R}^2 \ : \ e^{-\delta_1 x - \delta_2 y} \leq 1 - p\right\}$$

which implies

$$\mathcal{A}_p = \left\{(x, y) \in \mathbb{R}^2 \ : \ y \geq \frac{-\delta_1 x - \log(1-p)}{\delta_2}\right\}.$$

We can observe that this is a function in the form of $y = mx + b$. As, we can see $m = -\delta_1/\delta_2$ and $b = -\log(1-p)/\delta_2$. Similarly, for the curve $\mathcal{O}_p$, we can see

$$\mathcal{O}_p = \left\{(x, y) \in \mathbb{R}^2 \ : \ y = \frac{-\delta_1 x - \log(1-p)}{\delta_2}\right\}. \tag{3.3}$$

Hence, both functions are linear and thus are both convex and concave when $\delta_{12} = 0$. $\square$

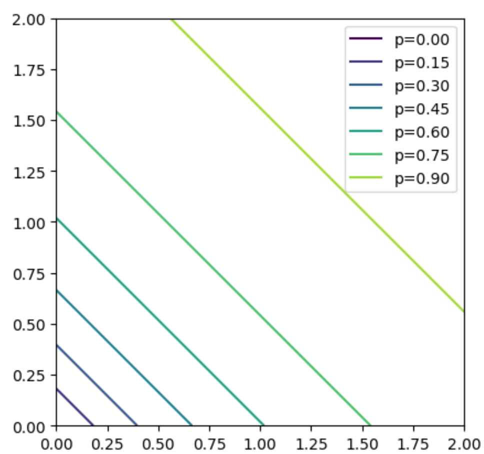

Figure 3.3: The level curves of $\mathcal{O}_p$ when $\delta_{12} = 0$, and when δ_1 and $\delta_2 = 0.9$.

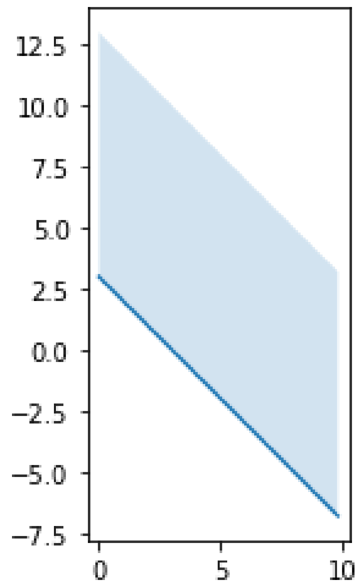

Figure 3.4: The set $\mathcal{A}_p$ when $\delta_{12} = 0$, $p = 0.95$, $\delta_1 = 1$ and $\delta_2 = 1$.

Note 3.4. From note 2.3, we know that $\overline{H}(x, 0) \times \overline{H}(0, y) = \overline{H}(x, y)$ when $\delta_{12} = 0$, which also implies $\rho(X, Y) = 0$, as mentioned in section 2.2. Hence, the two losses no longer depend on each other, and thus curve $\mathcal{O}_p$ becomes a constant decreasing function. This implies that there is a constant decrease in the losses as the amount invested in x increases at a $(1-p)100\%$ tail probability.

3.4 Minimization

Now that we have established the convexity of the set $\mathcal{A}_p$ and the curve $\mathcal{O}_p$, we are now able to solve one of our one main objectives of this paper, that is, minimizing the losses between two business lines. This can be done by applying the method of Lagrange multipliers. Lagrange multipliers are used to minimize a convex function subject to a certain constraint. In our case, the inequalities within the set $\mathcal{A}_p$ and the curve $\mathcal{O}_p$ serve as a constraint for our minimization problem. Thus, we obtain the following theorem to solve the problem mentioned previously in Note 3.2.

Theorem 3.4. *Let $p \in (0, 1)$, and let $\delta_1 > 0$ and $\delta_2 > 0$, and $\delta_{12} > 0$. Then, if α and β are parameters such that α ,$\beta \neq 0$, then the minimum of the sum $\alpha x^* + \beta y^*$ is the following point:*

$$(x^*, y^*) = ((-k\delta_2 + \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)})/k\delta_{12}, (-\delta_1 + \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)})/\delta_{12}),$$

where $k = \alpha/\beta$, on the curve $\mathcal{O}_p$.

Proof. First, with the help of Rockafellar's research in page 8 of [4],we apply the method of Lagrange multipliers to minimize the function $f(x, y) = \alpha x + \beta y$ subject to the constraint $\mathrm{e}^{-\delta_1 x - \delta_2 y - \delta_{12} xy} =$

$1-p$. Hence, for the method of Lagrange multipliers, the constraint

$$g(x,y) = e^{-\delta_1 x - \delta_2 y - \delta_{12} xy} - (1-p) = 0.$$

is used to obtain the Lagrange function

$$\begin{aligned} L(x,y,\lambda) &= f(x,y) + \lambda g(x,y) \\ &= \alpha x + \beta y + \lambda(e^{-\delta_1 x - \delta_2 y - \delta_{12} xy} - (1-p)). \end{aligned}$$

Next, we calculate the gradient

$$\begin{aligned} &\nabla_{x,y,\lambda} L(x,y,\lambda) = (\frac{dL}{dx}, \frac{dL}{dy}, \frac{dL}{d\lambda}) \\ &= (\alpha + \lambda(-\delta_1 - \delta_{12} y)e^{-\delta_1 x - \delta_2 y - \delta_{12} xy}, \beta + \lambda(-\delta_2 - \delta_{12} x)e^{-\delta_1 x - \delta_2 y - \delta_{12} xy}, e^{-\delta_1 x - \delta_2 y - \delta_{12} xy} - (1-p)). \end{aligned}$$

Hence, we receive the following system of equations:

$$\begin{aligned} \nabla_{x,y,\lambda} L(x,y,\lambda) = 0 \iff \alpha + \lambda(-\delta_1 - \delta_{12} y)e^{-\delta_1 x - \delta_2 y - \delta_{12} xy} &= 0 \\ \beta + \lambda(-\delta_2 - \delta_{12} x)e^{-\delta_1 x - \delta_2 y - \delta_{12} xy} &= 0 \\ e^{-\delta_1 x - \delta_2 y - \delta_{12} xy} - (1-p) &= 0. \end{aligned}$$

The last equation implies that $\mathrm{e}^{-\delta_1 x - \delta_2 y - \delta_{12} xy} = 1-p$. Hence, if we substitute this in the first and second equation we get the following two equations respectively:

$$\begin{aligned} \alpha + \lambda(-\delta_1 - \delta_{12} y)(1-p) &= 0 \qquad (1) \\ \beta + \lambda(-\delta_2 - \delta_{12} x)(1-p) &= 0 \qquad (2). \end{aligned}$$

This implies that

$$\begin{aligned} \alpha &= \lambda(\delta_1 + \delta_{12} y)(1-p) \qquad (3) \\ \beta &= \lambda(\delta_2 + \delta_{12} x)(1-p) \qquad (4). \end{aligned}$$

By dividing the third equation by the fourth, we then receive

$$\frac{\alpha}{\beta} = \frac{\delta_1 + \delta_{12} y}{\delta_2 + \delta_{12} x}$$

which implies

$$y^* = \frac{k(\delta_2 + \delta_{12} x) - \delta_1}{\delta_{12}} \quad \text{where} \quad k = \frac{\alpha}{\beta}.$$

Next, we substitute this back in our third equation and receive the following:

$$k{\delta_{12}}^2x^2 + 2k\delta_2\delta_{12}x - \delta_2\delta_1 + k{\delta_2}^2 + \delta_{12}\log(1-p) = 0.$$

Now, we notice that this equation can be written in the form of $ax^2 + bx + c$, as we can assume $a = k{\delta_{12}}^2$, $b = 2k\delta_2\delta_{12}$, and $c = -\delta_2\delta_1 + k{\delta_2}^2 + \delta_{12}\log(1-p)$. Hence, we can apply the quadratic formula to solve for the value of x,

$$\begin{aligned} x^* &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\ &= \frac{-k\delta_2 \pm \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)}}{k\delta_{12}}. \end{aligned}$$

If $x^* = (-k\delta_2 - \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)})/k\delta_{12}$, then we receive an x^* value that is negative and hence outside our domain which is a contradiction. Thus, it must be that $x^* = (-k\delta_2 + \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)})/k\delta_{12}$. Now, we can substitute this back in $y^* = [k(\delta_2 + \delta_{12}x) - \delta_1]/\delta_{12}$ to receive the following y^* value:

$$y^* = \frac{-\delta_1 + \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)}}{\delta_{12}}.$$

Hence, this implies that

$$(x^*, y^*) = ((-k\delta_2 + \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)})/k\delta_{12}, (-\delta_1 + \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)})/\delta_{12})$$

is the point on the curve $\mathcal{O}_p$ that minimizes the sum $\alpha x^* + \beta y^*$. □

Corollary 3.1. *If $\delta_{12} = 0$, then $(x^*, y^*) = (-\frac{\log(1-p)}{\delta_1}, 0)$ when $\frac{\alpha}{\delta_1} \leq \frac{\beta}{\delta_2}$ and $(x^*, y^*) = (0, -\frac{\log(1-p)}{\delta_2})$ when $\frac{\beta}{\delta_2} < \frac{\alpha}{\delta_1}$. Hence, the minimum sum is then $\min\{\alpha\frac{-\log(1-p)}{\delta_1}, \beta\frac{-\log(1-p)}{\delta_2}\}$.*

Proof. From Theorem 3.3, we analyzed the case of when $\delta_{12} = 0$, giving us a linear curve of optimal values, depicted by equation (3.3). For a linear function with a linear constraint, the minimum or maximum must occur on one of the two extremes. Hence, the problem we have here is minimizing the sum $\alpha x + \beta y$ subject to the constraint $y = \frac{-\delta_1 x - \log(1-p)}{\delta_2}$. Now knowing that the optimal points occur at the two extreme values, we first let y^*=0, which implies that $x^* = -\frac{\log(1-p)}{\delta_1}$, implying that the minimum sum is then $-\alpha\frac{\log(1-p)}{\delta_1}$. For the second optimal point we let $x^* = 0$, implying that $y^* = -\frac{\log(1-p)}{\delta_2}$. Hence, the minimum sum value in this case would be $-\beta\frac{\log(1-p)}{\delta_2}$. Hence, the minimum value of $\alpha x^* + \beta y^*$ when $\delta_{12} = 0$ is $\min\{\alpha\frac{-\log(1-p)}{\delta_1}, \beta\frac{-\log(1-p)}{\delta_2}\}$. Thus, if $\frac{\alpha}{\delta_1} \leq \frac{\beta}{\delta_2}$ then $(x^*, y^*) = (\frac{-\log(1-p)}{\delta_1}, 0)$ with a minimum sum value of $\alpha\frac{-\log(1-p)}{\delta_1}$. Conversely, if $\frac{\beta}{\delta_2} < \frac{\alpha}{\delta_1}$ then $(x^*, y^*) = (0, \frac{-\log(1-p)}{\delta_2})$ with a minimum sum value of $-\beta\frac{\log(1-p)}{\delta_2}$. □

Note 3.5. From here we notice that both the x and y values consist of parameters within a square root. Knowing that a non-negative value must be obtained to give any meaningful implications in this problem, only certain values can be assigned to α and β. Specifically, we can see that

$$
\begin{aligned}
&k\delta_2\delta_1 - k\delta_{12}\log(1-p) \geq 0 \\
\Rightarrow\ &k(\delta_2\delta_1 - \delta_{12}\log(1-p)) \geq 0.
\end{aligned}
$$

Since, δ_1, δ_2, and δ_{12} are non-negative parameters, and $p \in (0,1)$, $\log(1-p)$ must be negative, which implies that k is multiplied by a positive value. Hence, for the above inequality to be true, it must be that $k \geq 0$. Since $k = \alpha/\beta$, this implies that $\beta \neq 0$, and α and β must be both positive values, or both negative values to return a co-ordinate that lies on the curve $\mathcal{O}_p$. However, in terms of interpretation, for both values to be negative, this would imply there to be negative risk which doesn't have any real implications. Hence, α and β must both obtain positive values, as they represent the percentage of loss each business is willing to take. Which further depicts why $\alpha x + \beta y$ must be a convex combination, with $\alpha + \beta = 1$.

Corollary 3.2. *If $\alpha = \beta$, then $((-\delta_2 + \sqrt{\delta_2\delta_1 - \delta_{12}\log(1-p)})/\delta_{12}, (-\delta_1 + \sqrt{\delta_2\delta_1 - \delta_{12}\log(1-p)})/\delta_{12})$ is the point on curve $\mathcal{O}_p$ that minimizes the sum $\alpha x^* + \beta y^*$ with a value of $(-\delta_2 - \delta_1 + 2\sqrt{\delta_2\delta_1 - \delta_{12}\log(1-p)})/\delta_{12}$.*

Proof. By virtue of Theorem 3.4, we know that

$$
\begin{aligned}
x^* &= \frac{-k\delta_2 + \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)}}{k\delta_{12}} \\
y^* &= \frac{-\delta_1 + \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)}}{\delta_{12}},
\end{aligned}
$$

where $k = \alpha/\beta$. If $\alpha = \beta$, this implies that $k = 1$. As a result, we have

$$
\begin{aligned}
x^* &= \frac{-\delta_2 + \sqrt{\delta_2\delta_1 - \delta_{12}\log(1-p)}}{\delta_{12}} \\
y^* &= \frac{-\delta_1 + \sqrt{\delta_2\delta_1 - \delta_{12}\log(1-p)}}{\delta_{12}}.
\end{aligned}
$$

Hence, by adding these two values together we obtain our minimized sum $\alpha x^* + \beta y^*$ when $\alpha = \beta$, to be $(-\delta_2 - \delta_1 + 2\sqrt{\delta_2\delta_1 - \delta_{12}\log(1-p)})/\delta_{12}$. □

Note 3.6. When α and β obtain the same values, the minimized sum just becomes a multiple of the $x^* + y^*$ value. As we can see here, if $\alpha = \beta$, then $\alpha x^* + \beta y^* = \alpha x^* + \alpha y^* = \alpha(x^* + y^*)$.

Corollary 3.3. *If $\delta_1 = \delta_2$ and $\alpha = \beta$, then $x^* = y^* = (-\delta_1 + \sqrt{{\delta_1}^2 - \delta_{12}\log(1-p)})/\delta_{12}$ is the point on curve $\mathcal{O}_p$ that minimizes the sum $\alpha x^* + \beta y^*$. Hence, the minimized sum $\alpha x^* + \beta y^*$ would be $\alpha((-\delta_1 + \sqrt{{\delta_1}^2 - \delta_{12}\log(1-p)}/\delta_{12})$.*

Proof. By virtue of Theorem 3.4, we know that,

$$x^* = \frac{-\delta_2 + \sqrt{\delta_2\delta_1 - \delta_{12}\log(1-p)}}{\delta_{12}}$$
$$y^* = \frac{-\delta_1 + \sqrt{\delta_2\delta_1 - \delta_{12}\log(1-p)}}{\delta_{12}}.$$

Hence, if $\delta_1 = \delta_2$ then,

$$x^* = \frac{-\delta_1 + \sqrt{{\delta_1}^2 - \delta_{12}\log(1-p)}}{\delta_{12}}$$
$$y^* = \frac{-\delta_1 + \sqrt{{\delta_1}^2 - \delta_{12}\log(1-p)}}{\delta_{12}}.$$

Which implies that $x^* = y^*$. Hence, the minimized sum $x^* + y^*$ would be

$$\alpha(x^* + x^*) = \alpha x^* = \alpha[(-\delta_1 + \sqrt{{\delta_1}^2 - \delta_{12}\log(1-p)})/\delta_{12}].$$

□

Note 3.7. When δ_1 and δ_2 obtain the same values, and $\alpha = \beta$, this implies that the x and y variables vary by the same amount. Hence, to minimize the sum $x^* + y^*$ on the curve $\mathcal{O}_p$, x^* and y^* must obtain the same value as well.

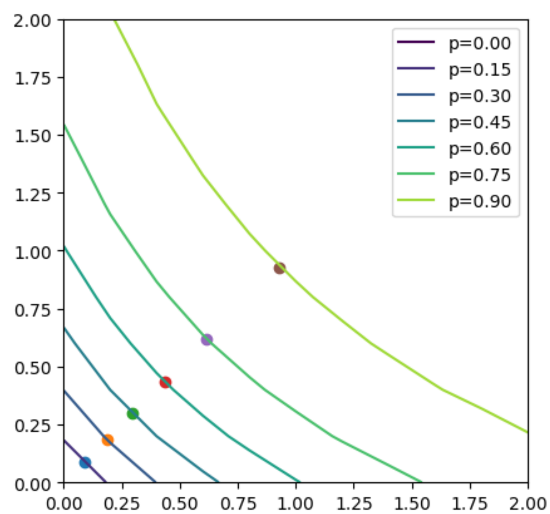

Figure 3.5: The curve $\mathcal{O}_p$ for different values of p, when δ_1 and $\delta_2 = 0.9$, and $\delta_{12} = 0.729$, with the minimum point; $((-\delta_2 + \sqrt{\delta_2\delta_1 - \delta_{12}\log(1-p)})/\delta_{12}, (-\delta_1 + \sqrt{\delta_2\delta_1 - \delta_{12}\log(1-p)})/\delta_{12})$.

Having illustrated the graph in Figure 3.5, we can confidently affirm the accuracy of our devised calculations and theorems. Since, our proof supports the notion that the minimum sum of two values occurs at the midpoint of an exponential function when the δ_1 and δ_2 parameters are equal.

Note 3.8. When α and β obtain different values then we can expect the x and y values to vary differently as well. For instance, when $\alpha > \beta$, then $k > 1$, which implies that we can expect to see a lower x-value compared to the y-value, due to the effects k has on the x and y values separately. This implies that a lower amount should be invested in the first business line, x, than the second to gain the most profit. Evidently, in Figure 3.6, we see that the minimum point is located near the upper half of the curve $\mathcal{O}_p$.

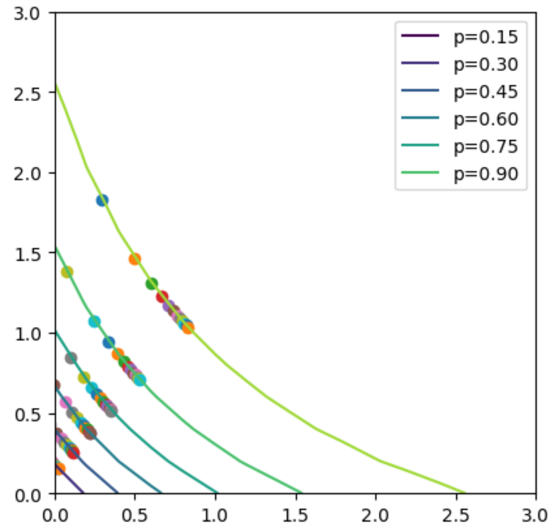

Figure 3.6: The curve $\mathcal{O}_p$ for different values of p,α, and β, when δ_1 and $\delta_2 = 0.9$, and $\delta_{12} = 0.729$, with the minimum point; $((-k\delta_2 + \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)})/k\delta_{12}, (-\delta_1 + \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)})/\delta_{12})$.

Similarly, if $\alpha < \beta$, then $k < 1$, implying that $x > y$, which we can also see is depicted in the Figure 3.7, where the minimum point is located near the lower half of the curve. Which also implies that a lower amount should be invested in the second business line, y, than the first to receive the least losses.

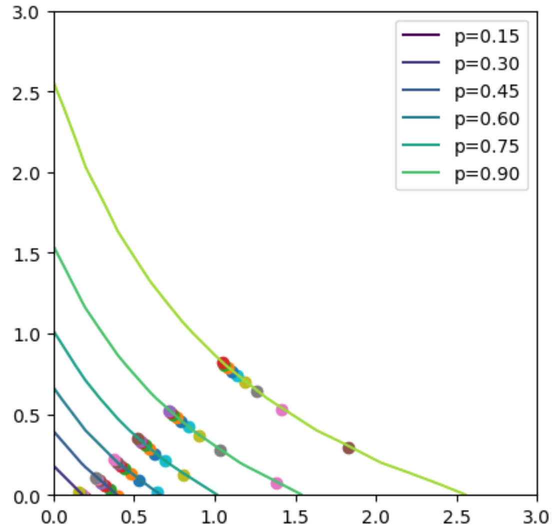

Figure 3.7: The curve $\mathcal{O}_p$ for different values of p,α, and β, when δ_1 and $\delta_2 = 0.9$, and $\delta_{12} = 0.729$, with the minimum point; $((-k\delta_2 + \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)})/k\delta_{12}, (-\delta_1 + \sqrt{k\delta_2\delta_1 - k\delta_{12}\log(1-p)})/\delta_{12})$.

Moreover, by varying the values of δ_1 and δ_2, we notice the respective x and y values are positively proportional, regardless of the α or β values. For instance, increasing the value of δ_1 results in an increase in the x-value. Similarly, an increase in the value of δ_2 results in an increase in the y-value.

Corollary 3.4. *i)When $p = 0.95$,$\delta_1 = 1$,$\delta_2 = 1$, $\delta_{12} = 0.5$, and $\alpha = \beta = 0.5$, then the curve $\mathcal{O}_p$ has the point (1.16,1.16) that minimizes the sum $\alpha x^* + \beta y^*$ with a value of 2.32.*
ii)When $p = 0.99$,$\delta_1 = 1$,$\delta_2 = 1$, $\delta_{12} = 0.5$, and $\alpha = \beta = 0.5$, then the curve $\mathcal{O}_p$ has the point (1.63,1.63) that minimizes the sum $\alpha x^* + \beta y^*$ with a value of 3.26.
iii) When $p = 0.95$,$\delta_1 = 2$,$\delta_2 = 1$, $\delta_{12} = 1$, $\alpha = 0.7$ and $\beta = 0.3$, then the curve $\mathcal{O}_p$ has the point (0.464,1.412) that minimizes the sum $\alpha x^* + \beta y^*$ with a value of 0.7484.

Proof. i)Given that $p = 0.95$,$\delta_1 = 1$,$\delta_2 = 1$, and $\delta_{12} = 0.5$, then from Theorem 3.4 and Collorary 3.2, we obtain the following x and y value:

$$x^*, y^* = \frac{-1 + \sqrt{(1)^2 - 0.5\log(1-0.95)}}{0.5} = 1.16.$$

Hence, this implies that (1.16,1.16) is the point on the curve $\mathcal{O}_p$ that minimizes the sum $0.5x^*+0.5y^*$ with a value of 1.16+1.16=2.32.
ii)Given that $p = 0.99$, then we obtain the following x^* and y^* value:

$$x^*, y^* = \frac{-1 + \sqrt{(1)^2 - 0.5\log(1-0.99)}}{0.5} = 1.63.$$

Hence, this implies that (1.63,1.63) is the point on the curve $\mathcal{O}_p$ that minimizes the sum $0.5x^*+0.5y^*$ with a value of 1.63+1.63=3.26.

iii) Given that $\alpha = 0.7$ and $\beta = 0.3$, we see that $k = 2.33$. Also, from given $p = 0.95$,$\delta_1 = 2$,$\delta_2 = 1$, $\delta_{12} = 1$, we receive the following values for x and y respectively:

$$x^* = \frac{-2.33 + \sqrt{(2.33)(1)(2) - (2.33)(1)\log(1-0.95)}}{2.33} = 0.464$$
$$y^* = \frac{-2 + \sqrt{(2.33)(1)(2) - (2.33)(1)\log(1-0.95)}}{1} = 1.412.$$

Hence, (0.464,1.412) would be the point on the curve $\mathcal{O}_p$ that minimizes the sum $0.7x^*+0.3y^*$, with a value of 0.7(0.464)+0.3(1.412)=0.7484. □

3.5 Comparison of Bivariate and Univariate VaR values

We notice that curve $\mathcal{O}_p$ is also the bivariate- VaR curve induced by the set (3.1). Hence, to display as a comparison, we may observe the Value-at-Risk curve in the univariate setting. Various details have already been compared between the two settings by Torres et al. [6] and Xu and Chen [10]. For instance, it was said that the univariate Value-at-Risk models are depicted to be simpler to construct but do not consider certain factors such as the time-varying covariance or correlation between financial assets. On the other hand, multivariate VaR approaches are not well-developed compared to univariate ones. Multivariate VaR estimates consider the correlation between asset returns, providing a more comprehensive assessment of risk by incorporating more information and risk factors. Both univariate and multivariate VaR models have their advantages and disadvantages, with multiple calculation methods available for both. However, the addition of risk factors in the multivariate VaR estimation, makes it a much more conservative measurement for portfolio VaR compared to univariate VaR models.

Now, by first analyzing the level p curve of optimal values, and by Cousin and Di Bernardino's VaR function [3], we can depict the univariate Value-at-Risk as a comparison for our paper. Hence, in the univariate setting, the level p curve of optimal values for X and Y can be displayed in the following way:

$$\mathcal{O}_p(X) = \inf\{x \in \mathbb{R} : \overline{H}(x) = 1-p\} \tag{3.4}$$

where $\overline{H}(x) = e^{-\delta_1 x}$, and

$$\mathcal{O}_p(Y) = \inf\{y \in \mathbb{R} : \overline{H}(y) = 1-p\} \tag{3.5}$$

where $\overline{H}(y) = e^{-\delta_2 y}$. We then see that the univariate VaR of X and Y can be viewed as the boundary of the set $\{x \in \mathbb{R} : \overline{H}(x) \leq 1-p\}$ and $\{y \in \mathbb{R} : \overline{H}(y) \leq 1-p\}$ respectively. Now, since the x and y variables take the same values, this implies that $\mathcal{O}_p(X)$=$\mathcal{O}_p(Y)$, which is also similar to the results obtained by Arbia in page 237 of [5]. By observing the equation within the set represented by $\mathcal{O}_p(X)$, we see that

$$e^{-\delta_1 x} = 1 - p$$
$$\Rightarrow x^* = -\frac{\log(1-p)}{\delta_1}$$

which provides us with some insight on the behavior of curve $\mathcal{O}_p(X)$. Since, we can see that curve $\mathcal{O}_p(X)$ will be a logarithmic function in the first quadrant, gradually increasing from $x = 0$, with a vertical assymptote at $p = 1$. In order to illustrate the graph of curve Op, we set δ_1 to 1 and verify that our interpretations of the equation align with our expectations, as evidenced in Figure 3.8.

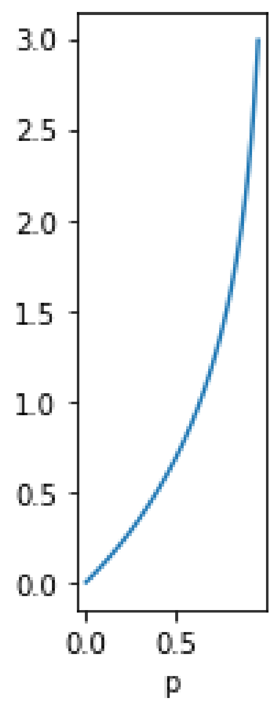

Figure 3.8: The univariate curve $O_p(X)$ when $\delta_1 = 1$.

To closely analyze the difference in value-at-risk values induced by the bivariate curve and the univariate curve, we may compare the pair $\mathcal{O}_p(X, Y)$ to $(\mathcal{O}_p(X), \mathcal{O}_p(Y))$ coming from their respective curves. To compute $\mathcal{O}_p(X, Y)$, we apply Theorem 3.4 and equation 3.2, and for computing the vector $(\mathcal{O}_p(X), \mathcal{O}_p(Y))$, we apply the equations 3.4 and 3.5. In both cases, we let α and $\beta = 0.5$. We are then able to clearly depict our data in a table of values to compare the two curves.

			$(\mathcal{O}_p(X), \mathcal{O}_p(Y))$	$\mathcal{O}_p(X,Y)$		
p	δ_1	δ_2		$\delta_{12} = 0.1$	$\delta_{12} = 0.5$	$\delta_{12} = 1.0$
0.95	1	1	(3.00,3.00)	(1.40,1.40)	(1.16,1.16)	(0.999,0.999)
0.5	1	1	(0.693,0.693)	(0.341,0.341)	(0.321,0.321)	(0.301,0.301)
0.2	1	1	(0.223,0.223)	(0.111,0.111)	(0.109,0.109)	(0.106,0.106)

Table 3.1: Comparison of optimal values for the Univariate and Bivariate function when $k = 1$.

Based on the results from Table 3.1, we can see that the Value-at-Risk pair coming from the univariate curve is always greater than the values coming from the bivariate curve, regardless of the δ_{12} value. Which is quite typical, due to the functions negative dependence as discovered earlier in Section 2.2. Next, we can observe the differences in the values when varying the δ_1 and δ_2 parameters. Next, if we let $\alpha = 0.7$ and $\beta = 0.3$, we can observe a similar pattern of results when $k > 1$ in Table 3.2.

			$(\mathcal{O}_p(X), \mathcal{O}_p(Y))$	$\mathcal{O}_p(X,Y)$		
p	δ_1	δ_2		$\delta_{12} = 1.0$	$\delta_{12} = 1.5$	$\delta_{12} = 2.0$
0.95	2	1	(1.498,3.000)	(0.464,1.412)	(0.446,1.260)	(0.426,1.158)
0.5	2	1	(0.347,0.693)	(0.075,0.505)	(0.095,0.441)	(0.103,0.404)
0.3	2	1	(0.178,0.357)	(0.006,0.343)	(0.029,0.287)	(0.040,0.257)

Table 3.2: Comparison of optimal values for the Univariate and Bivariate function when $k > 1$.

Now, suppose $\alpha = 0.3$ and $\beta = 0.7$, then we can observe the results for when $k < 1$ in Table 3.3.

			$(\mathcal{O}_p(X), \mathcal{O}_p(Y))$	$\mathcal{O}_p(X,Y)$		
p	δ_1	δ_2		$\delta_{12} = 3.0$	$\delta_{12} = 4.0$	$\delta_{12} = 5.0$
0.95	2	5	(1.498,0.599)	(0.551,0.285)	(0.560,0.268)	(0.526,0.255)
0.5	2	5	(0.347,0.139)	(0.102,0.092)	(0.114,0.085)	(0.121,0.081)
0.3	2	5	(0.178,0.071)	(0.027,0.060)	(0.040,0.054)	(0.048,0.050)

Table 3.3: Comparison of optimal values for the Univariate and Bivariate function when $k < 1$.

Not only do these tables help to display a comparison within the univariate and bivariate functions, but it is now evident to see the difference in x^* and y^*-values for both cases of $k > 1$ and $k < 1$ as mentioned in Note 3.8. The only anomalies happen to be for when $p = 0.3$ corresponding to the bivariate function when $k < 1$. However, this is expected as with higher p-values, and higher δ_2 values, the x^* and y^*-values become almost identical.

4 Conclusion and Further Research

In this paper, we have successfully applied Gumbel's probabilistic model to depict and solve challenges in the world of finance. In particular, we analyzed the set of admissible values, $\mathcal{A}_p$, and the level p curve of optimal values, $\mathcal{O}_p$, which provided valuable insights into the allocation of capital for two business lines within a company. By applying optimization techniques, we were then able to depict how our new found theorems and other discoveries allows businesses to strategically invest their resources to minimize losses.

When we set our parameters to certain values such as $\delta_1 = 1$, $\delta_2 = 1$, $\delta_{12} = 0.5$, and $\alpha = \beta = 0.5$, our findings reveal a minimized sum of 2.32 resulted in the most profitable outcome at a 95 percentile level. This highlights one instance where the discoveries made from this paper help obtain results where actual numerical values are involved.

Furthermore, we were able to a draw a crucial connection between our findings and the financial concept of value-at-risk (VaR). In particular, we were able to depict VaR at probability p by the set $\mathcal{A}_p$, in a bivariate setting, by incorporating a function we derived from the original distribution noted by Gumbel (Bivariate exponential distribution, 1960). This allowed us to project our new found knowledge into the form of innovative theorems and graphs, hence deepening our understanding of how risk management principles can be integrated into resource allocation decisions.

In conclusion, by utilizing Gumbel's BVE as a framework for the results in our thesis, we have been able to make significant progress in the correlations this model has in the field of finance. In particular, our ability to connect the set of admissible values to the concept of value-at-risk, enabled us to develop a theorem to depict and solve the optimal value to minimize the losses between two business lines. These probabilities can serve as an essential risk assessment tool, enabling companies to make well-informed decisions regarding the products they would invest into.

Moving forward, we hope our research can be extended in further dimensions such as $\mathbb{R}^3$, $\mathbb{R}^4$, and eventually to $\mathbb{R}^n$ to be applicable for multivariate scenarios. Of which, we can delve into more complex form of mathematics, as it may consist of functions with marginals of fatter tails. This way, companies with multiple business lines will have an easy and efficient way of allocating their resources, resulting in financial stability and the longevity of a company.

References

[1] Marshall, A.W. and Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association*, 62, 30-44,

[2] Gumbel, E.J. (1960). Bivariate exponential distribution. *Journal of the American Statistical Association*, 55(292), 698-707,

[3] Cousin, A. and Di Bernardino, E. (2013). On multivariate extensions of Value-at-Risk. *Journal of multivariate analysis*, 119, 32–46,

[4] Rockafellar,R.T. (1993). Lagrange Multipliers and Optimality. *Society for Industrial and Applied Mathematics*, 35, 183–238,

[5] Arbia,G. (2002). Bivariate Value-at-Risk. *Statistica*, 2, 231–247,

[6] Torres,R., Lillo, R.E., and Laniado, H. (2015). A directional multivariate value at risk. *Insurance:Mathematics and Economics*, 65, 111-123,

[7] Davydov,Y., Moldavskaya,E., Zitikis,R. (2018). Searching for, and quantifying, non-convexity of functions. *arXiv: Functional Analysis*, 1-16,

[8] Grzegorzewski, P. (2013). On some basic concepts in probability of IF-events. *Information Sciences*, 232, 411–418,

[9] Ghorpade, R.S. and Limaye, V.B (2007). Sylvester's Minorant Criterion, Lagrange-Beltrami Identity, and Nonnegative Definiteness. *The Mathematics Student. Special Centenary Volume*, 1,123-130,

[10] Xu, C. and Chen, H. (2012). Measuring Portfolio Value at Risk. *School of Economics and Management, Lund University*,