DOKUZ EYLÜL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

LOCAL DEPENDENCE MEASURES, PROPERTIES AND APPLICATIONS

by
Burcu ÜÇER

November, 2007
İZMİR

LOCAL DEPENDENCE MEASURES, PROPERTIES AND APPLICATIONS

A Thesis Submitted to the
Graduate School of Natural and Applied Sciences of Dokuz Eylül University
In Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Statistics Program

by
Burcu ÜÇER

November, 2007
İZMİR

Ph.D. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"LOCAL DEPENDENCE MEASURES, PROPERTIES AND APPLICATIONS"** completed by **BURCU ÜÇER** under supervision of **PROF. DR. İSMİHAN BAYRAMOĞLU** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Doctor of Philosophy.

Prof. Dr. İsmihan BAYRAMOĞLU

Supervisor

Prof. Dr. Serdar KURT

Thesis Committee Member

Assoc.Prof. Dr. Halil ORUÇ

Thesis Committee Member

Assoc. Prof. Dr. Salih ÇELEBİOĞLU

Examining Committee Member

Asst. Prof. Dr. Esin FİRUZAN

Examining Committee Member

Prof.Dr. Cahit HELVACI

Director

Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENTS

I would like to express my deep and sincere gratitude to my advisor, Prof. Dr. İsmihan Bayramoğlu, for his wisdom, encouragement, advice and patience. His encouragement and instruction lead me to a good start of my research life. I will always be grateful for his guidance. It has been a privilege to work with him.

I would like to thank Prof. Dr. Serdar Kurt, for providing insightful advice on many aspects of my education. I owe him for my being here. I also would like to thank Assoc. Prof. Dr. Halil Oruç for being in my committee and for interesting and helpful comments that he provided.

I would like to thank all my friends in this department for all their supports. Especially, I am very appreciative of the support I received from my friends, Övgü Kınay, Selma Gürler, Evrim Turgutlu and Timur Turgutlu who generously gave their time and encouragement.

I wish to express my great appreciation to my husband, Göksel Üçer, for always being my side, full support, and understanding. His love, support and encouragement of learning and caring in my life have helped immensely. Also I would like to thank for his helpful effort. I would like to thank my parents, Ahmet-Habibe Hüdaverdi and Rabi-Yüksel Üçer for their care and generous support through the years and for their confidence in me. I owe what I achieved to my mother, who has always been a role model for me for her intelligence and hard-working.

Burcu ÜÇER

LOCAL DEPENDENCE FUNCTIONS, PROPERTIES and APPLICATIONS

ABSTRACT

Dependence structure between random variables is generally complex and the single scalar dependence measures cannot be adequate to explain the natural association between them. With this motivation, a new local dependence function characterizing dependence structure between two random variables in an ε-neighbourhood of particular point (x, y) from the domain of underlying bivariate distribution is introduced and its properties are investigated. The local dependence function is constructed by the natural way considering the localized version of the Hoeffding's formula and possesses many of the properties that must satisfy any dependence measure. Examples for the local dependence function of some bivariate distributions are provided.

Dependence maps of the given data sets are constructed by using the natural estimator of the local dependence function. Dependence maps help to determine the dependence structure between random variables visually. Permutation test algorithm is developed for constructing the dependence map. Several examples, concerning inflation rate with stock index rate of return and percentage change in the US dollar and also simulated bivariate normal data are provided.

Also, local numerical characteristics of random variables are introduced and their properties are investigated. Local characteristics of some distributions are also examined.

Keywords: Local dependence function, permutation test, local expected value, local variance, dependence map, correlation.

YEREL BAĞIMLILIK FONKSİYONLARI, ÖZELLİKLERİ ve UYGULAMALARI

ÖZ

Rassal değişkenler arasındaki bağımlılık yapısı genelde karmaşıktır ve skaler bağımlılık ölçüleri rassal değişkenler arasındaki doğal bağımlılık yapısını açıklamada yeterli olmayabilir. Bundan yola çıkararak, iki değişkenli dağılımın tanım aralığındaki (x,y) noktasının ε-komşuluğunda, iki rassal değişken arasındaki bağımlılık yapısını karakterize eden yeni bir yerel bağımlılık fonksiyonu tanımlanmış ve özellikleri incelenmiştir. Yerel bağımlılık fonksiyonu Hoeffding tanımının yerelleştirilmesi sonucunda oluşturulmuştur ve herhangi bir bağımlılık ölçüsünün sağlaması gereken birçok özelliği sağlamaktadır. Bazı iki değişkenli dağılımların yerel bağımlılık fonksiyonları için örnekler verilmiştir.

Çeşitli veri setlerinin bağımlılık haritaları, yerel bağımlılık fonksiyonunun tahmincisi kullanılarak oluşturulmuştur. Bağımlılık haritaları, rassal değişkenler arasındaki bağımlılık yapısını görsel olarak tanımlamaya yardımcı olur. Permütasyon testi algoritması bağımlılık haritalarını oluşturabilmek için geliştirilmiştir. Enflasyon oranı ile hisse senedi getirileri ve ABD dolarındaki yüzde değişim arasında ve türetilen iki değişkenli normal dağılan veri çifti arasında bağımlılık yapısını inceleyen örnekler verilmiştir.

Ayrıca, rassal değişkenlerin yerel sayısal karakteristikleri tanımlanmış ve özellikleri incelenmiştir. Bazı dağılımların yerel karakteristikleri de elde edilmiştir.

Anahtar Sözcükler: Yerel bağımlılık fonksiyonları, permütasyon testi, yerel beklenen değer, yerel varyans, bağımlılık haritası, korelasyon.

CONTENTS

Page

CHAPTER ONE
INTRODUCTION

1.1 Introduction

Dependence relation between random variables is one of the most widely studied topics in probability theory and statistics. Unless specific assumptions are made about the dependence no meaningful statistical model can be constructed. “Statistical dependence” deals with new concepts specific for probability theory which may not coincide with the regular meaning of this concept. Many of the authors, Yule (1897), Lancaster (1963), Lehmann (1966), Esary, Proschan, & Walkup (1967) and Esary & Proschan (1972), initially introduced many concepts about dependence for two variables. These concepts has been extended to a multivariate random vector $(X_1,X_2,\ldots,X_n)$ with $n \geq 2$.

Bivariate dependence can be generalized in another way by considering that the random vector $(X_1,X_2,\ldots,X_n)$ may be split into k sets and we may be more interested in the relationships between k sets than the relationships within each set. Generalizing a bivariate concept of positive dependence leads to a definition of an ordering which compares two bivariate distributions to determine whether one distribution is more concentrated than the other. (Mari & Kotz, 2001)

Basic Concepts of Dependence

Pearson’s correlation and linearity concepts are not the best concepts for non-normal random variables. Some useful concepts of positive and negative dependence are given in Section II. Some of the important uses of dependence concepts are formulated in numerous papers and several books. (See, Joe (1997), Nelsen (1999) and Mari & Kotz (2001))

- Positive quadrant dependence concept which is the basic to the parametric families of copulas in determining whether a multivariate parameter is a dependence parameter.
- Stochastic increasing positive dependence concept which is used in the analysis of the decrease in dependence with lag for stationary Markov chains.
- TP_2 dependence concept which is necessary for constructing families of closed-form copulas with a wide range of dependence.
- Tail dependence concept is important to the construction and analysis of multivariate extreme value distributions and copulas.
- Kendall's tau and Spearman's rho are the summary measures of dependence for bivariate copula; also Kendall's tau is used for compatibility conditions.

The rest of the thesis is planned as follows: In Section II, general bivariate and multivariate dependence concepts, properties and measures are given. Local dependence measures and properties suggested by Bjerve & Doksum (1993), Holland & Wang (1987) and Bairamov & Kotz (2002) are examined in Section III. Local numerical characteristics of random variables and a new epsilon-local dependence function are introduced in Section IV. Following to this section, in Section V, the local dependence map is constructed by suggested local dependence function for real-data application. Section VI is devoted to conclusion.

CHAPTER TWO
DEPENDENCE MEASURES

Generally, bivariate dependence concepts, properties and measures are given in the probability theory and statistics. In this section, number of dependence concepts are given as mentioned in chapter one.

2.1 Positive quadrant and orthant dependence

Let (X_1, X_2) be a bivariate random vector having distribution function *F*. (X_1, X_2) or *F* is **positive quadrant dependent (PQD)** if

$$P(X_1 > x_1, X_2 > x_2) \geq P(X_1 > x_1)P(X_2 > x_2) \quad \forall x_1, x_2 \in IR \tag{2.1}$$

or equivalently

$$P(X_1 \leq x_1, X_2 \leq x_2) \geq P(X_1 \leq x_1)P(X_2 \leq x_2) \quad \forall x_1, x_2 \in IR \tag{2.2}$$

It is clear that X_1 and X_2 are PQD if the probability that X_1 and X_2 are simultaneously small (or large) is at least great as the probability that they are independent. (Lehmann, 1966)

Similarly, (X_1, X_2) or *F* is **negative quadrant dependent (NQD)** if

$$P(X_1 > x_1, X_2 > x_2) \leq P(X_1 > x_1)P(X_2 > x_2) \quad \forall x_1, x_2 \in IR \tag{2.3}$$

or equivalently

$$P(X_1 \leq x_1, X_2 \leq x_2) \leq P(X_1 \leq x_1)P(X_2 \leq x_2) \quad \forall x_1, x_2 \in IR \tag{2.4}$$

If the probability that X_1 and X_2 are simultaneously small (or large) is at most great as the probability that they are independent, then we say that X_1 and X_2 are NQD.

For the multivariate extension, let $(X_1, X_2, ..., X_m)$ be a random *m*-vector having distribution function *F*. $(X_1, X_2, ..., X_m)$ and *F* is **positive upper orthant dependent (PUOD)** if

$$P(X_1 > x_1, ..., X_m > x_m) \geq \prod_{i=1}^{m} P(X_i > x_i) \;\; \forall (x_1, x_2, ..., x_m) \in IR^m \tag{2.5}$$

Also $(X_1, X_2, ..., X_m)$ or *F* is **positive lower orthant dependent (PLOD)** if

$$P(X_1 \leq x_1, ..., X_m \leq x_m) \geq \prod_{i=1}^{m} P(X_i \leq x_i) \;\; \forall (x_1, x_2, ..., x_m) \in IR^m \tag{2.6}$$

If PUOD and PLOD both occur then, $(X_1, X_2, ..., X_m)$ and *F* are **positive orthant dependent (POD).** Clearly, PUOD means that $(X_1, X_2, ..., X_m)$ are more likely simultaneously to have large values against a vector of independent random variables with the same univariate margins.

Similarly,

$(X_1, X_2, ..., X_m)$ or *F* is **negative upper orthant dependent (NUOD)** if

$$P(X_1 > x_1, ..., X_m > x_m) \leq \prod_{i=1}^{m} P(X_i > x_i) \;\; \forall (x_1, x_2, ..., x_m) \in IR^m \tag{2.7}$$

Also $(X_1, X_2, ..., X_m)$ or *F* is **negative lower orthant dependent (NLOD)** if

$$P(X_1 \le x_1,\ldots,X_m \le x_m) \le \prod_{i=1}^{m} P(X_i \le x_i) \quad \forall (x_1,x_2,\ldots,x_m) \in IR^m \tag{2.8}$$

If NUOD and NLOD both occur then, $(X_1, X_2,\ldots,X_m)$ and F are **negative orthant dependent (NOD).**

2.2 Stochastic Increasing Positive Dependence

Let (X_1, X_2) be a bivariate random vector having c.d.f. $F \in$ F(F_1,F_2) where F is a Frechét classes given the set of margins F_1, F_2. Frechét classes are classes of multivariate distributions with some given margins. Recall that each distribution $F(x, y)$ has upper and lower Frechét bounds, which are $F^+(x, y) = \min(F_1(x), F_2(y))$ and $F^-(x, y) = \max(0, F_1(x) + F_2(y) - 1)$. X_2 is **stochastically increasing (SI)** in X_1 or the conditional distribution $F_{2|1}$ is **stochastically increasing (SI)** if

$$P(X_2 > x_2 \mid X_1 = x_1) = 1 - F_{2|1}(x_2 \mid x_1) \quad \uparrow x_1 \ \forall x_2 \tag{2.9}$$

By reversing the roles of indices of 1 and 2, then X_1 is SI in X_2 or $F_{1|2}$. The reason why this is a positive dependence condition is that X_2 is more likely to take larger values as X_1 increases.

By reversing the direction of monotonicity, the **stochastically decreasing (SD)** condition results.

For the multivariate extensions of SI, there are two dependence concepts. The random vector $(X_1, X_2,\ldots,X_m)$ is **positive dependent through the stochastic ordering (PDS)** if $\{X_i : for\ i \ne j\}$ conditional on $X_j = x$ is increasing stochastically as x increases, for all *j = 1,…,m* and also the random vector $(X_1, X_2, ..., X_m)$ is **conditional increasing in sequence (CIS),** if X_i is stochastically increasing in $X_1, X_2, ..., X_{i-1}$ for *i=1,2,…,m*, i.e. $P(X_i > x_i \mid X_j = x_j, j = 1, ..., i-1)$ is increasing in $(x_1, x_2, ..., x_{i-1})$ for all x_i.

2.3 Right Tail Increasing and Left Tail Decreasing

Let (X_1, X_2) be a bivariate random vector having c.d.f. $F \in$ Ƒ (F_1, F_2) where Ƒ is a Frechét classes given a set of margins. X_2 is **right tail increasing (RTI)** in X_1 if

$$P(X_2 > x_2 \mid X_1 > x_1) = \overline{F}(x_1, x_2) / \overline{F}_1(x_1) \quad \uparrow x_1 \ \forall x_2 \tag{2.10}$$

and X_2 is **left tail decreasing (LTD)** in X_1 if

$$P(X_2 \le x_2 \mid X_1 \le x_1) = F(x_1, x_2) / F_1(x_1) \quad \downarrow x_1 \ \forall x_2 \tag{2.11}$$

RTI and LTD are positive dependence conditions since X_2 is more likely to take on larger values as X_1 for RTI and X_2 is more likely to take on smaller values as X_1 decreases for LTD. For negative dependence conditions, both inequalities are reversed.

As multivariate extension of RTI, $X_i, i \in A^c$, is RTI in $X_j, j \in A$ for $(X_1, X_2, ..., X_m)$, if

$$P(X_i > x_i, i \in A^c \mid X_j > x_j, j \in A) \uparrow x_k, k \in A, \tag{2.12}$$

where A is non-empty subset of $\{1, \ldots, m\}$ and A^c is the complement of A.

2.4 Associated Random Variables

Let $\boldsymbol{X} = (X_1, X_2, ..., X_m)$ be a random vector. $\boldsymbol{X}$ is **positively associated** if

$$E[g_1(\boldsymbol{X})g_2(\boldsymbol{X})] \geq E[g_1(\boldsymbol{X})]E[g_2(\boldsymbol{X})] \tag{2.13}$$

for all real-valued functions g_1 and g_2 which are increasing and these expectations exist. In fact, this is a positive dependence for $\boldsymbol{X}$ since two increasing functions of $\boldsymbol{X}$ have positive covariance whenever the covariance exist.

2.5 Total Positivity of Order 2

Let b be a non-negative function on A^2 where $A \subset IR$, then for all $x_1 < x_2$, $y_1 < y_2$ with $x_i, y_j \in A$, b is **totally positive of order 2 ($\mathbf{TP_2}$)** if it satisfies the condition

$$b(x_1, y_1)b(x_2, y_2) \geq b(x_1, y_2)b(x_2, y_1) \tag{2.14}$$

Also it can be expressed as the non-negativity of the determinant of the square matrix of order 2,

$$\begin{vmatrix} b(x_1, y_1) & b(x_1, y_2) \\ b(x_2, y_1) & b(x_2, y_2) \end{vmatrix} \geq 0. \tag{2.15}$$

Total positivity of higher orders involves the non-negativity of determinants of larger square matrices. If the inequality in 2.14 is reversed then b is **reverse rule of order 2 ($\mathbf{RR_2}$).**

Let F be a bivariate distribution function having density function f. There are three notions of positive dependence:

i) f is TP_2.

ii) F is TP_2.

iii) $\overline{F}$ is TP_2.

In fact, (i) means that both F and $\overline{F}$ are TP_2. An explanation of (ii) as a positive dependence condition is shown below.

$$F(x_1, y_1)F(x_2, y_2) - F(x_1, y_2)F(x_2, y_1) \geq 0, \qquad \forall x_1 < x_2,\ y_1 < y_2. \tag{2.16}$$

This inequality means that

$$\begin{aligned} F(x_1, y_1)&\left[F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1)\right] \\ &- \left[F(x_1, y_2) - F(x_1, y_1)\right]\left[F(x_2, y_1) - F(x_1, y_1)\right] \geq 0 \end{aligned} \tag{2.17}$$

$$\forall x_1 < x_2,\ y_1 < y_2$$

If $(X_1, Y_1) \sim F$, then the inequality can be written as

$$\begin{aligned} P(X_1 \leq x_1, Y_1 \leq y_1)&P(x_1 \leq X_1 \leq x_2, y_1 \leq Y_1 \leq y_2) \\ &- P(X_1 \leq x_1, y_1 \leq Y_1 \leq y_2)P(x_1 \leq X_1 \leq x_2, Y_1 \leq y_1) \geq 0 \end{aligned} \tag{2.18}$$

$$\forall x_1 < x_2,\ y_1 < y_2$$

Also, survival function $\overline{F}$ satisfies the above condition when it is reversed with F for (iii).

A multivariate extension of TP_2 can be expressed as the following. Suppose that X is a random m-vector having density $\boldsymbol{f}$. $\boldsymbol{X}$ or $\boldsymbol{f}$ is **multivariate totally positive of order 2 (MTP_2)** if it satisfies the condition

$$f(x \vee y)f(x \wedge y) \geq f(x)f(y) \tag{2.19}$$

for all $x, y \in IR$ where

$$x \vee y = \left(\max\{x_1, y_1\}, \max\{x_2, y_2\}, \ldots, \max\{x_m, y_m\}\right)$$
$$x \wedge y = \left(\min\{x_1, y_1\}, \min\{x_2, y_2\}, \ldots, \min\{x_m, y_m\}\right)$$

One important property of MTP_2 is that if a density is MTP_2 then so are all of its marginal densities of order 2 and higher. (Joe, 1997)

When the inequality 2.19 is reversed, then *f* is **multivariate reverse rule of order 2 (MRR_2).** This is a weak dependence concept, because the property MRR_2 is not closed under taking the margins.

2.6 Positive function dependence

When all univariate margins are the same, then there is a special concept of dependence called **positive function dependence**, i.e. for $F(F_1,\ldots,F_m)$ with $F_1 = \ldots = F_m$ For the bivariate case, let (X_1, X_2)~*F*. Then (X_1, X_2) or *F* is **positive function dependent (PFD)** if

$$Cov\left[h(X_1), h(X_2)\right] \geq 0, \ \forall \text{ real-valued } h \tag{2.20}$$

such that the covariance exists. The multivariate extension with $(X_1, X_2, \ldots, X_m)$ is **positive function dependent** if

$$E\left[\prod_{i=1}^{m} h(X_i)\right] \geq \prod_{i=1}^{m} E\left[h(X_i)\right] \qquad \forall \text{ real-valued } h \tag{2.21}$$

such that the expectations exist. If *m* is odd, then another restriction is that *h* is non-negative.

2.7 Relationships among dependence properties

Relationships among dependence properties are summarized in this part. The proofs of these relationships can be seen in Joe, 1997.

Theorem: All of the properties given above are invariant with respect to strictly increasing transformations on the components of the random vector. This means that if (X_1, X_2) is PQD, then $(g_1(X_1), g_2(X_2))$ for strictly increasing functions g_1 and g_2.

Theorem: The bivariate relations are:

(i) TP_2 density $\Rightarrow$ SI $\Rightarrow$ LTD, RTI;
(ii) LTD or RTI $\Rightarrow$ association $\Rightarrow$ PQD;
(iii) TP_2 density $\Rightarrow$ TP_2 c.d.f. and TP_2 survival function;
(iv) TP_2 c.d.f. $\Rightarrow$ LTD and TP_2 survival function $\Rightarrow$ RTI.

Theorem: The multivariate relations are:

(i) a random subvector of an associated random vector is associated;
(ii) association $\Rightarrow$ PUOD and PLOD
(iii) PDS $\Rightarrow$ PUOD and PLOD
(iv) CIS $\Rightarrow$ association

2.8 Concordance

Let (x_i, y_i) and (x_j, y_j) be two observations from a vector (X, Y) of continuous random variables. If $x_i < x_j$ and $y_i < y_j$, or if $x_i > x_j$ and $y_i > y_j$, then (x_i, y_i) and (x_j, y_j) are concordant; and also when $x_i < x_j$ and $y_i > y_j$ or when $x_i > x_j$ and $y_i < y_j$, then similarly we say that (x_i, y_i) and (x_j, y_j) are discordant. And also

(x_i, y_i) and (x_j, y_j) are concordant if $(x_i - y_i)\,(x_j - y_j) > 0$, and discordant if $(x_i - y_i)\,(x_j - y_j) < 0$.

If large values of one of the pair of random variables tend to be associated with large values of the other, and also small values of one of the pair of random variables tend to be associated with small values of the other, then it is said these random variables are concordant.

Kendall’s tau and Spearman’s rho are bivariate measures of dependence for continuous variables. They are

(i) invariant with respect to strictly increasing transformations,

(ii) equal to 1 for the bivariate Frechét upper bound (one variable is an increasing transform of the other),

(iii) equal to -1 for the bivariate Frechét lower bound (one variable is a decreasing transform of the other).

These properties do not hold for Pearson’s correlation, so that τ and ρ_s are more desirable as measures of association for multivariate non-normal distributions.

2.8.1 Kendall’s Tau

Let (X, Y) be a random vector and $(X_1, Y_1), (X_2, Y_2)$ be i.i.d. copies of (X, Y). One of the measure of association known as Kendall’s tau is defined in terms of concordance as follows:

$$\tau = \tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]; \qquad (2.22)$$

i.e., Kendall's tau is defined as the probability of concordance minus the probability of discordance.

2.8.1.1 *Kendall's τ and TP_2 property*

From the definition of Kendall's τ, it can be shown that

$$\tau = 2P\left[(X_1 - X_2)(Y_1 - Y_2) \geq 0\right] - 1. \tag{2.23}$$

Let (X,Y) be a random vector and $(X_1,Y_1),(X_2,Y_2)$ be i.i.d. copies of (X,Y). If (X,Y) are varies over the whole plane, then

$$P(X_2 < X_1, Y_2 < Y_1) = \int FdF$$

and therefore,

$$\tau = 4\int FdF - 1. \tag{2.24}$$

Estimation of τ

Let (X,Y) be the pair having sample size of n, then

$$\hat{\tau} = \frac{\#(concordantpairs) - \#(discordantpairs)}{\binom{n}{2}}.$$

2.8.2 Spearman's Rho

Like Kendall's tau, the measure of association known as Spearman's rho is also based on concordance and discordance. Let (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) be three independent random vectors with a common joint distribution function *F* and copula *C*.

The Spearman's rho is defined to be proportional to the probability of concordance minus the probability of discordance for the two vectors (X_1, Y_1) and (X_2, Y_3)

$$\rho_{X,Y} = 3\{P[(X_1 - X_2)(Y_1 - Y_3) > 0] - [(X_1 - X_2)(Y_1 - Y_3) < 0]\} \; . \tag{2.25}$$

Here, the joint distribution function of (X_1, Y_1) is *F*(*x,y*) and the joint distribution function of (X_2, Y_3) is $F_1(x)F_2(y)$ since they are independent.

2.8.2.1 Spearman's ρ and PQD concept

Spearman's ρ can be defined as the linear correlation coefficient between two uniform random variables as $U = F_1(X)$ and $V = F_2(Y)$, where F_1 and F_2 are distribution function of *X* and *Y*.

$$\begin{aligned}
\rho_s(X,Y) &= \frac{E(UV) - \frac{1}{4}}{\frac{1}{12}} \\
&= 12E(UV) - 3 \\
&= 12\int_0^1\int_0^1 UVdudv - 3 \\
&= 12\left(\int_0^1\int_0^1 UVdudv - \int_0^1 Udu\int_0^1 Vdv\right)
\end{aligned} \tag{2.26}$$

where $E(U) = E(V) = 1/2$ and $Var(U) = Var(V) = 1/12$, since U and V are uniformly distributed.

Also, ρ_s can be defined as

$$\rho_s = 12\iint_{R^2} \left(F(x,y) - F_1(x)F_2(y)\right)dF_1.dF_2 \tag{2.27}$$

Estimator of ρ_s

$$\hat{\rho}_s = 12\sum_{i,j}(\hat{F}(i,j) - \hat{F}_1(i)\hat{F}_2(j))\hat{p}_i\hat{p}_j \tag{2.28}$$

Renyi (1959), Schweizer & Wolff (1981) and Lancaster (1963) considered measures of association which are commonly called “measures of dependence”. They discussed a minimal set of desirable for a nonparametric measure of dependence.

Definition: A numeric measure κ of association between two continuous random variables X and Y whose copula is C is a measure of concordance if it satisfies the following properties:

1. κ is defined for every pair X, Y of continuous random variables;
2. $-1 \le \kappa_{X,Y} \le 1$, $\kappa_{X,X} = 1$ $\kappa_{X,-X} = -1$;
3. $\kappa_{X,Y} = \kappa_{Y,X}$;
4. If X and Y are independent, then $\kappa_{X,Y} = 0$;
5. $\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$;
6. if C_1 and C_2 are copulas such that $C_1 < C_2$, then $\kappa_{C_1} < \kappa_{C_2}$;

7. If $\{(X_n, Y_n)\}$ is a sequence of continuous random variables with copulas C_n, and if $\{C_n\}$ converges pointwise to C, then $\lim_{n\to\infty} \kappa_{C_n} = \kappa_C$.

2.9 Tail Dependence

Bivariate tail dependence concerns with the amount of dependence in the upper-quadrant tail or lower-quadrant tail of a bivariate distribution. It is a concept that is relevant to dependence in extreme values that mainly depends on the tails. Because of invariance to increasing transformations, the definition will be given in terms of copulas.

2.10 Measures of Dependence

Seven properties for one class of measures of association known as "measures of concordance" are presented. Measures of concordance are interested in how large values of one variable tend to be associated with large values of the other, and consequently they attain their extreme values when the copula of the random variables where the measures is +1 or -1.

On the other hand, a measure of dependence indicates how closely X and Y are related with extremes at mutual independence and (monotone) dependence. (Lancaster, 1963)

.

Definition: A numeric measure δ of association between two continuous random variables X and Y whose copula is C is a "measure of dependence" if it satisfies the following properties:

1. δ is defined for every pair of continuous random variables X and Y;
2. $\delta_{X,Y} = \delta_{Y,X}$;
3. $0 \leq \delta_{X,Y} \leq 1$;
4. $\delta_{X,Y} = 0$ if and only if X and Y are independent;

5. $\delta_{X,Y} = 1$ if and only if each of X and Y is almost surely a strictly monotone function of the other;
6. If α and β are almost surely strictly monotone functions on Ran$X$ and Ran$Y$, respectively, then $\delta_{\alpha(X),\beta(Y)} = \delta_{X,Y}$;
7. If $\{(X_n, Y_n)\}$ is a sequence of continuous random variables with copulas C_n, and if $\{C_n\}$ converges pointwise to C, then $\lim_{n\to\infty} \delta_{C_n} = \delta_C$.

2.11 Correlation and Dependence

2.11.1 Independence

Many statisticians interested in correlation and dependence concepts and give some verbal definitions of dependence and independence. Bayes (1763) concern with the notion of independence and defined it as follows:

"Events are independent when the happening of any one of them does neither increase nor abate the probability of the rest."

P.S Laplace (1810) presents the earliest mathematical description of independence pointing out that if $\{E_i\}$ is a sequence of independent simple events with $p_i = P(E_i)$, then $P(E_1,...,E_n) = \prod_{i=1}^{n} p_i$ and then say that in the case of simple events that the occurrence of the first E_1 affects the probability of the occurrence of the second E_2, i.e., we have $P(E_1E_2) = P(E_2|E_1)P(E_1)$.

2.11.2 Zero Correlation versus Dependence

Correlation is one of the basic concept in statistics, familiar to researchers in natural, social and medical sciences, and of course economics. In mathematics, especially in measure theory, the "independence" notion is investigated. Probability theory, which was developed as a branch of measure theory transmitted this concept to statistical methodology. So we have two concepts *–independence (and dependence)* and *correlation (lack of correlation)* utilized in statistical sciences. (Mari & Kotz, 2001) As simplification in statistical methodology, many authors do not distinguish between independence and lack of correlation. Also some readers of these writings often establish practical independence, or absence of meaningful relationships among the variables; it suffices to verify that the correlation coefficients are zero; this may cause considerable degree of harm.

2.11.2.1 A Technical Discussion

A technical discussion between zero correlation (uncorrelatedness) and dependence is proceeded in this section. If the covariance between two random variables *X* and *Y*, i.e.,

$$Cov(X,Y) = E(XY) - E(X)E(Y) \tag{2.29}$$

is zero, i.e.;

$$E(X)E(Y) = E(XY) \tag{2.30}$$

the variables are said *uncorrelated.*

If two variables are *independent,* i.e.

$$F_x(x)F_Y(y) = F_{X,Y}(x,y) \qquad \forall x, \forall y \tag{2.31}$$

or, if derivatives exist, equivalently,

$$f_X(x)f_Y(y) = f_{X,Y}(x,y) \qquad \forall x, \forall y \tag{2.32}$$

then, they are uncorrelated.

The variables which are not independent-that is they are *dependent*- can also be uncorrelated. This is

$$E(X|Y = y) = E(X), \quad \forall y \tag{2.33}$$

or

$$E(Y|X = x) = E(Y), \quad \forall x \tag{2.34}$$

Independence is more stringent than uncorrelatedness. The conditions (2.31) and (2.32) require the equality of function for all x and y, while (2.29) requires the equality of a function of expected value and the product moment of X and Y. This means that the arithmetic means of two sets of numbers can be equal without necessarily implying that each number in one set is equal to the corresponding number in the second set.

Mari & Kotz (2001) gave a simple example of two dependent but uncorrelated variables:

$$P(X = Y = 0) = P(X = Y = 1) = P(X = -1, Y = 1) = \frac{1}{3}$$

Marginal distributions are

$$P(X = -1) = P(X = 0) = P(X = 1) = \frac{1}{3};$$

$$P(Y=0)=\frac{1}{3};P(Y=1)=\frac{2}{3}$$

Hence:

$$E(X)=\frac{1}{3}(-1+0+1)=0$$
$$E(Y)=\frac{1}{3}\times 0+\frac{2}{3}\times 1=\frac{2}{3}$$

and

$$E(XY)=\frac{1}{3}(0\times 0+1\times 1+(-1\times 1))=0$$

and $Cov(X,Y)=0\times\frac{2}{3}-0=0$

This means that X and Y are uncorrelated, but they are clearly dependent since $P(X=1|Y=0)=0$ but $P(X=1|Y=1)=\frac{1}{2}$

In this case, note that $E(X|Y=y)$ does not depend on y:

$$E(X|Y=0)=E(X|Y=1)=0.$$

Another example,

Let X and Y be independent random variables and let

$E(X)=0$, $E(Y)=0$, $E(X^2)<0$,

Let $\delta_1 = X$ and $\delta_2 = XY$

and $E(\delta_1\delta_2) = E(X^2Y) = E(X^2)E(Y) = 0$.

In this case, δ_1 and δ_2 are uncorrelated, but clearly they are dependent.

2.11.3 Linear correlation

Correlation is only one particular measure of dependency; a linear correlation cannot capture non-linear dependence relationships. The linear correlation coefficient between *X* and *Y* is

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \tag{2.35}$$

where $Cov(X,Y)$ is the covariance between the random variables *X* and *Y* which is

$$Cov(X,Y) = E(XY) - E(X)E(Y)$$

also, σ_X and σ_Y denote the standard deviations of *X* and *Y*.

When *X, Y* and the product *X.Y* are integrable functions, covariance function can be used to measure the dependence between random variables. For a larger class of random variables, Hoeffding (1940) extended covariance concept as given below:

$$Cov(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\big(F(x,y) - F(x)F(y)\big)dxdy \tag{2.36}$$

When the distribution of (X,Y) is a standardized bivariate normal,

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(2\rho xy - x^2 - y^2)}, \quad -\infty < x < \infty,\ -\infty < y < \infty,\ -1 < \rho < 1 \tag{2.37}$$

then the conditional density of $(Y|X = x)$ is:

$$f(y \mid X = x) = \frac{1}{\sqrt{2\pi(1-\rho^2)}} e^{-\frac{(y-\rho x)^2}{2(1-\rho^2)}} \tag{2.38}$$

so, expected value and variance of the conditional random variable

$$E(Y \mid X = x) = \rho x \tag{2.39}$$

and

$$Var(Y \mid X = x) = 1 - \rho^2 \tag{2.40}$$

Note that ρ characterizes the distribution and when ρ = 0 then it is clear that (X,Y) are independent, but the equivalence is not valid if the distribution is not normal.

The relation between $\mu(x) = E(Y \mid X = x)$ and x may not be linear and the variance $\sigma^2(x) = Var(Y \mid X = x)$ may not be constant for other distributions different from normal distribution. So it can be thought that ρ is not necessarily an adequate index to explain the relationship between X and Y. Bjerve & Doksum (1993) has characterized these connections for non-normal distributions by defining a local linear correlation coefficient.

ρ and concepts of dependence

If the distribution of (X,Y) variables satisfies any of the concepts of positive dependence, even weak concepts of positive dependence, then ρ will be positive. Note that if ρ is positive and the random variables (X,Y) is bivariate normal then they satisfy a stronger condition of TP_2. (Mari & Kotz, 2001)

Here are some properties of $\rho(X,Y)$,

(i) $-1 \leq \rho(X,Y) \leq 1$

(ii) If X and Y are independent, then $\rho(X,Y)=0$

(iii) Let $\widetilde{X} = aX + b$ and $\widetilde{Y} = cY + d$, then $\rho(\widetilde{X},\widetilde{Y}) = sign(ab)\rho(X,Y)$ for $\forall a,b \in IR \setminus \{0\}$, $b,d \in IR$

(iv) Let $Y = aX + b$ for $\forall a \in IR \setminus \{0\}$, and $b \in IR$, then $\rho(X,Y) = \pm 1$, i.e. there occurs perfect dependence. This means that linear correlation is invariant under positive transformations.

Correlation coefficient has many popular properties in application explained in the following;

(i) Naturalness of correlation coefficient as a measure of dependence in multivariate normal distributions or more generally spherical and elliptical distributions makes it a more popular dependence measure.

(ii) Correlation coefficient is easy to manipulate under linear operations.

(iii) For many bivariate distributions, it is simple to calculate second moments and so to derive the correlation coefficient.

While working with models other than multivariate normal, one must be careful in using correlation. Elliptical distributions make someone have true interpretations. Unfortunately, most dependent random variables do not appear to have elliptical distributions. If we restrict our attention to multivariate normal distributions or elliptical distributions, marginal distributions and correlation coefficient determine the joint distribution. (Yule, 1897) For example, if we know that X and Y have standard normal distributions with a given correlation coefficient, there is only one bivariate elliptical distribution which fits with this information. Otherwise, there are many other distributions that fit with this information. Also, the coefficient tells us nothing about the degree of dependence in the tail of underlying distributions.

For elliptical distribution family given marginal distributions F_1 and F_2 for all X and Y, all linear correlations between -1 and +1 can be attained through suitable specification of the joint distribution F.

Imperfections of correlation

Matteis (2001) summarizes the cases of imperfections of correlations. In these cases, it is assumed that X and Y are real-valued random variables.

(i) Independence of two random variables implies that they are uncorrelated, but zero correlation does not generally imply independence. However in the case of multivariate normal distribution, it is admissible to interpret uncorrelatedness as independence.

(ii) If linear correlation is invariant under nonlinear strictly increasing transformations $T : IR \to IR$, then it has the serious deficiency.

(iii) Since single observation can even have an arbitrarily high influence on the linear correlation, linear correlation is not a robust measure.

2.12 Copula

Understanding relationships among bivariate and multivariate outcomes is a basic problem in statistical science. *Copulas* focus on the simple problem of understanding the distribution of several outcomes, multivariate distribution. The study of copulas and their applications are modern phenomenon in statistics; most of the statistical applications have arisen in the last twenty years. However, copulas have been studied in the probability literature for about 50 years. This chapter introduces the concept of copula functions.

Copulas are function that join or "couple" multivariate distribution functions to their one-dimensional marginal distribution functions. (Nelsen, 1999) Also, as another definition; copulas are multivariate distribution functions whose one dimensional margins are uniform on the interval (0,1). The reason why statisticians are interested in copula is that firstly, it is a scale-free measure of dependence; secondly, it is a starting point for constructing families of bivariate distributions.

The word *copula* is a Latin noun which means "a link, tie, bond". The word copula was first employed in mathematical or statistical sense by Abe Sklar (1959). Also, Frechét, Dall'Aglio, Feron and many others studied multivariate distributions with fixed univariate marginal distributions. Indeed, many of the basic results about copulas can be seen in the work of Wassily Hoeffding (1940).

2.12.1 Definitions and Basic Properties

Copula functions are tools for extracting a kind of transformation invariant description of the dependence structure from the joint distribution function. It is firstly concerned with "2-increasing" function- a two-dimensional analog of a nondecreasing function of one variable.

Definition: Let S_1, S_2 be nonempty subsets of IR where IR denotes the extended real line $[-\infty, \infty]$. Let *H* be a real function of *two* variable such that Dom(*H*) $= S_1 \times S_2$, where Dom is the domain of the function. Let $B = [x_1, x_2] \times [y_1, y_2]$ be a rectangle all of whose vertices are in Dom(*H*). Then the *H*-volume of *B* is given by

$$V_H(B) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1). \tag{2.41}$$

In the beginning of defining a copula, it is considered two standard uniform random variable X_1 and X_2. It is not assumed that X_1 and X_2 are independent, they may be related. This dependence between real-valued random variables X_1 and X_2 is described by their joint distribution function

$$F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2). \tag{2.42}$$

If we know the joint distribution function *F*, then everything is known about random variables X_1 and X_2. Marginal behaviour can be known and also conditional probabilities can be evaluated:

$$F_{1|2}(x_1, x_2) = P(X_1 \leq x_1 | X_2 \leq x_2) \tag{2.43}$$

The dependence structure of the random variables is contained within *F*. The idea of separating *F* into a part which describes the dependence structure and parts which describe the marginal behaviour only had led to the concept of copula. (Matteis, 2001) In other words, copulas are multivariate uniform distributions and they represent the way of extracting the dependence structure from the joint distribution and separating dependence and marginal behaviour. Conventional methods like correlation summarize dependence but a copula gives a model for the dependence structure.

Definition: (Copula) A two-dimensional copula whose domain is $\mathbf{I}^2$ is a function C from $\mathbf{I}^2$ to $\mathbf{I}$ with the following properties:

.

(i) Dom $C = S_1 \times S_2$, where S_1 and S_2 are subsets of $\mathbf{I}$ containing 0 and 1;

(ii) C is increasing in each component.

(iii) For every u,v in **I,**

$$C(u,0) = C(0,v) = 0$$

and

$$C(u,1) = u \qquad and \qquad C(1,v) = v\,.$$

(iv) For every u_1, u_2, v_1, v_2 in **I** such that $u_1 \le u_2 \quad and \quad v_1 \le v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \ge 0\,.$$

Proposition: Let *X* be a random variable with distribution function *F*. Let F^{-1} be the quantile function of F, i.e.

$$F^{-1}(\alpha) = \inf\{x | F(x) \ge \alpha,\} \qquad \alpha \in (0,1) \tag{2.44}$$

Then

(i) For any standard uniformly distributed U~U(0,1), we have F^{-1}(U)~ *F*.

(ii) If F is continuous then the random variable *F*(*X*) is standard uniformly distributed, i.e. *F*(*X*) ~U(0,1),

Theorem: (Sklar,1959) Let F be a joint distribution with margins F_1 and F_2 (not necessarily continuous). Then there exists a copula $C:[0,1]^2 \rightarrow [0,1]$ such that for all x_1 and x_2 in $\overline{IR}$,

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)). \tag{2.45}$$

If the margins are continuous then C is unique; otherwise C is uniquely determined on $RanF_1 \times RanF_2$ where $Ran\,F_1$ and $Ran\,F_2$ are the ranges of marginals. Conversely if C is a copula, and F_1 and F_2 are distribution functions, and the function F is a joint distribution function with marginals F_1 and F_2.

Theorem: (Frechét -Hoeffding Bounds) Let C be a copula for all u,v in **I**, then the Frechét -Hoeffding bounds satisfy the following inequalities,

$$W(u,v) = \max(u + v - 1,0) \le C(u,v) \le \min(u,v) = M(u,v). \tag{2.46}$$

As a consequence of Sklar's theorem, if X and Y are random variables with joint distribution function $F(x,y)$ and margins F and G, respectively, then for all IR,

$$\max(F(x) + G(y) - 1,0) \le F(x,y) \le \min(F(x), G(y)) . \tag{2.47}$$

Since M and W are copulas, the above bounds are joint distribution functions and they are called the Frechét-Hoeffding bounds for joint distribution functions with margins F and G.

2.12.2 Properties of Copula

2.12.2.1 Invariance

Copulas are invariant under strictly monotone increasing transformations of random variables. Many of the useful properties of copula are based on invariance property.

Proposition: Let (X_1, X_2) be a vector continuously distributed random variables with copula *C*. Also let T_1, T_2 be strictly increasing transformation functions. Then $(T_1(X_1), T_2(X_2))$ also has the same copula.

Proof. Let F_i denote distribution function of X_i and let T_i^{-1} denote the inverse of T_i. The transformed random variable $T_i(X_i)$ has continuous distribution function $\widetilde{F} = F_i \circ T_i^{-1}$. It is known that $F_i^{-1}(U) \sim F_i$ and $F_i^{-1}(u_i) = x_i$

$$\begin{aligned} C(u_1, u_2) &= P\left[X_1 \leq F_1^{-1}(u_1), X_2 \leq F_2^{-1}(u_2)\right] \\ &= P\left[T_1(X_1) \leq T_1 \circ F_1^{-1}(u_1), T_2(X_2) \leq T_2 \circ F_2^{-1}(u_2)\right] \\ &= P\left[T_1(X_1) \leq \widetilde{F}_1^{-1}(u_1), T_2(X_2) \leq \widetilde{F}_2^{-1}(u_2)\right] \end{aligned} \tag{2.48}$$

So X_1 and X_2 move together by copula regardless of the scale in each variable measured. In the case of transformation of random variables, copulas remain the same.

2.12.2.2 Comonotonicity and Countermonotonicity

Firstly it is necessary to give the definitions of these two concepts. The upper boundary represents a situation where X_1 and X_2 are perfectly positively dependent and it is said that (X_1, X_2) are 'comonotonic'. The lower boundary represents a situation where X_1 and X_2 are perfectly negatively dependent and it is said that (X_1, X_2) are 'countercomonotonic'. Comonotonicity and countermonotonicity are the strongest types of concordance and discordance.

As mentioned before Frechét -Hoeffding bounds are universal bounds for any copula C for all x_1, x_2 in [0,1], that is

$$W(x_1,x_2) = \max(x_1 + x_2 - 1,0) \le C(x_1,x_2) \le \min(x_1,x_2) = M(x_1,x_2) \tag{2.49}$$

In order to obtain the copula M, firstly assume that the random variables X_1 and X_2 are continuous and $X_1 \sim \mathrm{F}_1$ and $X_2 \sim \mathrm{F}_2$. If it is assumed that $X_2 = T(X_1)$ where T is strictly increasing, then a strong dependence is occurred. So this kind of dependence represents the desired form of M. In proving this, it is considered that $F_2 = T^{-1}(F_1)$ and $F_1(X_1) \sim \mathrm{U}(0,1)$, then

$$\begin{aligned} C(x_1,x_2) &= P\left[F_1(X_1) \le x_1, F_2(X_2) \le x_2\right] \\ &= P\left[F_1(X_1) \le x_1, F_1 \circ T^{-1}(X_2) \le x_2\right] \\ &= P\left[F_1(X_1) \le x_1, F_1(X_1) \le x_2\right] \\ &= P\left[F_1(X_1) \le \min\{x_1,x_2\}\right] \\ &= \min\{x_1,x_2\} \\ &= M(x_1,x_2) \end{aligned} \tag{2.50}$$

This is a new copula-based definition for perfect positive dependence, which is also known as comonotonicity.

Analogously, if the random variables X_1 and X_2 have the Fréchet-Hoeffding lower bound, then they are called countermonotonic. The proof is similar to that given above.

2.12.2.3 Tail Dependence

Copulas are suitable functions for modelling tail dependence. Let λ_r and λ_l be the coefficients of right tail and left tail dependence. Assume that F_1 and F_2 are continuous, the coefficients of tail dependence can be verified

$$\begin{aligned}\lambda_r &= \lim_{\alpha\to 1-} P\left[X_2 > F_2^{-1}(\alpha) \middle| X_1 > F_1^{-1}(\alpha)\right] \\ &= \lim_{\alpha\to 1-} \frac{P\left[X_2 > F_2^{-1}(\alpha), X_1 > F_1^{-1}(\alpha)\right]}{P(X_1 > F_1^{-1}(\alpha))} \\ &= \lim_{\alpha\to 1-} \frac{\overline{C}(\alpha,\alpha)}{1-\alpha}\end{aligned} \tag{2.51}$$

where $\overline{C}(u,u) = 1 - 2u + C(u,u)$ and C is unique.

Similarly,

$$\begin{aligned}\lambda_l &= \lim_{\alpha\to 0+} P\left[X_2 \le F_2^{-1}(\alpha) \middle| X_1 \le F_1^{-1}(\alpha)\right] \\ &= \lim_{\alpha\to 0+} \frac{P\left[X_2 \le F_2^{-1}(\alpha), X_1 \le F_1^{-1}(\alpha)\right]}{P(X_1 \le F_1^{-1}(\alpha))} \\ &= \lim_{\alpha\to 0+} \frac{C(\alpha,\alpha)}{\alpha}\end{aligned} \tag{2.52}$$

So, tail dependence is best understood as an asymptotic property of copula if the limit exists.

2.12.2.4 Rank Correlation with Copulas

Working with multivariate distributions creates moments that may be difficult to determine in the case of calculation. Schweizer & Wolff (1981) showed that the Spearman's ρ and Kendall's τ measures could be expressed in terms of copula function.

Copulas play an important role in concordance and measures of association such as Kendall's tau. Firstly, a concordance function τ is defined as the difference of the probabilities of concordance and discordance between two vectors (X_1, Y_1) and (X_2, Y_2) of continuous random variables with different joint distributions $F_1(x, y)$ and $F_2(x, y)$ with margins F_1 and F_2.

Theorem: Let (X_1, Y_1) and (X_2, Y_2) be independent vectors of continuous random variables with joint distribution functions $F_1(x, y)$ and $F_2(x, y)$, respectively, with common margins F_1 (of X_1 and X_2) and F_2 (of Y_1 and Y_2). Let C_1 and C_2 denote the copulas of (X_1, Y_1) and (X_2, Y_2) respectively, so that $F_1(x, y) = C_1(F_1(x), F_2(y))$ and $F_2(x, y) = C_2(F_1(x), F_2(y))$. Let Q denote the difference between the probabilities of concordance and discordance of (X_1, Y_1) and (X_2, Y_2), i.e., let

$$\tau = \tau_{X,Y} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] \tag{2.53}$$

Proposition. $\tau = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1$

Proof. In order to prove the formula, recall that

$$P[(X_1 - X_2)(Y_1 - Y_2) > 0] + P[(X_1 - X_2)(Y_1 - Y_2) < 0] = 1$$

$$
\begin{aligned}
\tau(X,Y) &= 2P\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] - 1 \\
&= 2E\left[I_{(X_1-X_2)(Y_1-Y_2)>0}\right] - 1 \\
&= 2E\left[I_{(X_1-X_2)>0} I_{(Y_1-Y_2)>0} + I_{(X_1-X_2)<0} I_{(Y_1-Y_2)<0}\right] - 1 \\
&= 2(\int\int\int\int I_{(X_1-X_2)>0} I_{(Y_1-Y_2)>0} dF(x_2,y_2) dF(x_1,y_1) \\
&\quad + \int\int\int\int I_{(X_1-X_2)<0} I_{(Y_1-Y_2)<0} dF(x_2,y_2) dF(x_1,y_1)) - 1 \\
&= 2.2.(\int\int\int\int I_{(X_1-X_2)>0} I_{(Y_1-Y_2)>0} dF(x_2,y_2) dF(x_1,y_1)) - 1 \\
&= 4\int\int F(x_1,y_1) dF(x_1,y_1) - 1 \\
&= 4\int_0^1\int_0^1 C(u,v) dC(u,v) - 1
\end{aligned}
$$

This equation is the expected value of the function $C(U,V)$ of uniform distribution (0,1) random variables with joint distribution function C, i.e.,

$$
\tau = 4E(C(U,V)) - 1 \tag{2.54}
$$

Corollary 1: Let C_1, C_2 and τ be given as in the foregoing theorem. Then

1. τ is symmetric in its arguments: $\tau\,(C_1,C_2) = \tau\,(C_2,C_1)$.
2. τ is nondecreasing in each argument: if $C_1 < C_1'$ and $C_1 < C_2'$ for all (u,v) in $\mathbf{I}^2$, then $\tau\,(C_1,C_2) \le \tau(C_1',C_2')$.
3. Copulas can be replaced by survival copulas in τ, that is $\tau\,(C_1,C_2) = \tau(\hat{C}_1,\hat{C}_2)$.

Proposition. $\rho_s = 12\int_0^1\int_0^1 \{C(u,v) - uv\} dudv$ (2.55)

Proof. Let $(F_1(X), F_2(Y))$ have joint distribution C and also $F_1(X), F_2(Y)$ are standard uniformly distributed. So,

$$\begin{aligned}
\rho_s(X,Y) &= \rho(F_1(X),F_2(Y)) \\
&= \frac{Cov(F_1(X),F_2(Y))}{\sqrt{\sigma^2 F_1(X)\sigma^2 F_1(Y)}} \\
&= \frac{Cov(F_1(X),F_2(Y))}{\sqrt{\frac{1}{12}.\frac{1}{12}}} \\
&= 12Cov(F_1(X),F_2(Y)) \\
&= 12\int_0^1\int_0^1 \{C(x,y) - F_{F_1(X)}(x)F_{F_2(Y)}(y)\}dxdy \\
&= 12\int_0^1\int_0^1 \{C(x,y) - xy\}dxdy
\end{aligned}$$

2.13 Notions of Aging

The notions of dependence between two variables in bivariate distributions are useful concepts in the analysis of reliability theory. A notion of aging helps us to understand which life distributions are important in reliability models. Aging conveniently studied in terms of the failure rate function. A constant failure rate is obtained in the exponential distribution where there is no aging. (Barlow & Proschan, 1975).

The reliability (or survival probability) of a fresh unit corresponding to a mission of duration x is, $\overline{F}(x) = 1 - F(x)$, where F is the life distribution of the unit.

Let T be a random lifetime with continuous distribution on IR^+, then

$$\lim_{x\to 0}\frac{1}{x}P\{t \le T \le t+x | T > t\} = \lim_{x\to 0}\frac{1}{x}\frac{P\{t \le T \le t+x\}}{P(T>t)} = \frac{1}{P(T>t)}\lim_{x\to 0}\frac{P\{t \le T \le t+x\}}{x}$$

$$= \frac{f(t)}{1-F(t)} = \frac{f(t)}{\overline{F}(t)} \tag{2.56}$$

Here, the function $r(t) = \dfrac{f(t)}{\overline{F}(t)}$ is called failure rate function, when $f(t)$ exists and $\overline{F}$ (t) > 0. Useful identities are readily obtained by integrating both sides of $r(t)$.

$$\int_0^x r(t)dt = -\ln(1 - F(x)), \tag{2.57}$$

then exponentiating gives,

$$F(x) = 1 - e^{-\int_0^x r(t)dt}, \quad x \geq 0 \tag{2.58}$$

and hence

$$\overline{F}(x) = e^{-\int_0^x r(t)dt} \quad . \tag{2.59}$$

The cumulative failure rate, $R(x) = \int_0^x r(t)dt$ is referred to as the hazard function. $R(x)$ gives a useful theoretical representation of reliability as a function of failure rate. An alternate representation is

$$\overline{F}(x) = e^{-R(x)}, \tag{2.60}$$

which gives reliability in terms of hazard.

Considering device which does not age stochastically; that is its survival probability over an additional period of duration x is the same regardless of its present age t.

Symbolically,

$$\overline{F}(x \mid t) = \overline{F}(x) \text{ for all } x \text{ and } t \geq 0 \tag{2.61}$$

Conversely, a distribution with constant failure rate is $\overline{F}(x) = e^{-\lambda x}$, $\lambda > 0$ and $x \geq 0$ which is the exponential survival probability.

Another property of the exponential distribution which makes it especially important in reliability theory and application is that the remaining life of a used component is independent of its initial age (the memoryless property)

The conditional reliability of a unit of age t is

$$\overline{F}(x \mid t) = \frac{\overline{F}(t+x)}{\overline{F}(t)} \quad \text{if } \overline{F}(t) > 0 \tag{2.62}$$

Similarly, the conditional probability of failure during the next interval of duration x of a unit of age t is

$$\overline{F}(x \mid t) = \frac{F(t+x) - F(t)}{\overline{F}(t)} = 1 - F(x \mid t) \tag{2.63}$$

The reliability, failure rate, residual reliability and cumulative hazard functions are defined respectively,

$$\forall t \geq 0, \qquad \overline{F}(t) = P\,(T > t) \tag{2.64}$$

$$\forall t \geq 0, \qquad \lim_{x \to 0} \frac{1}{x} P\{t \leq T \leq t + x | T > t\} = \frac{f(t)}{\overline{F}(t)} \tag{2.65}$$

$$\forall t \geq 0, \forall x \geq 0, \quad \overline{F}(x \mid t) = P(T > x + t \mid T > t) = \frac{\overline{F}(x+t)}{\overline{F}(t)} \tag{2.66}$$

$$\forall x \geq 0, \qquad \mathrm{R}(x) = \int_0^x r(t)dt = -\ln \overline{F}(x) \quad . \tag{2.67}$$

Suppose now that the unit ages in the sense that the conditional survival probability is a decreasing function of age:

$$\overline{F}(x \mid t) = \frac{\overline{F}(t+x)}{\overline{F}(t)} \text{ is decreasing in } -\infty < t < \infty \text{ for each } x \geq 0. \tag{2.68}$$

As a consequence we obtain,

$$r(t) = \lim_{x \to 0} \frac{1}{x} \left[1 - \frac{\overline{F}(t+x)}{\overline{F}(t)} \right] \text{ is increasing in } t \geq 0 \text{, when the density } f(t) \text{ exists.}$$

Conversely, $r(t)$ is increasing implies

$$\overline{F}(x \mid t) = e^{-\int_t^{t+x} r(u)du} \tag{2.69}$$

is decreasing in $t \geq 0$ for each $x \geq 0$. Thus when the density exists, $\overline{F}(x \mid t) = \dfrac{\overline{F}(t+x)}{\overline{F}(t)}$ is equivalent to the failure rate $r(t)$ increasing.

2.13.1 Increasing Failure Rate (IFR)

T is said to be IFR if and only if, equivalently

- *r*(*t*) is an increasing function.
- $\forall x \geq 0$, $\overline{F}(x|t)$ is a decreasing function in *t*.
- $\overline{\mathrm{F}}$ is a Pólya frequency function of order 2 (PF_2), i.e.

$$\forall x_1, x_2, t_1, t_2 \in IR^+, x_1 < x_2, t_1 < t_2, \quad \begin{vmatrix} \overline{F}(x_1 - t_1) & \overline{F}(x_1 - t_2) \\ \overline{F}(x_2 - t_1) & \overline{F}(x_2 - t_2) \end{vmatrix} \geq 0$$

- $\ln \overline{F}(t)$ is a concave function.

If a real valued function *f* satisfies $f(x\text{-}y) = K(x,y)$ where *K* is TP_2, the function *f* is said to be Pólya function of order two. In that case $-log(f)$ is convex. This property is used in case of product convolution.

Example: Bivariate Failure Rate of Basu for Lomax Distribution

The joint survival function is given by

$$\overline{F}(x, y) = (1 + ax + by + \theta xy)^{-c}, \quad 0 \leq \theta \leq (c+1)ab : a,b,c > 0 \tag{2.70}$$

with p.d.f. (Lai, Xie. & Bairamov, 2001),

$$f(x,y) = \frac{c\left[c(b + \theta x)(a + \theta y) + ab - \theta\right]}{(1 + ax + by + \theta xy)^{c+2}}. \tag{2.71}$$

There are several versions of bivariate failure rate, one is defined by Basu:

$$r(x,y) = \frac{f(x,y)}{\overline{F}(x,y)}, \tag{2.72}$$

Thus, bivariate failure rate for the bivariate Lomax is

$$r(x,y) = \frac{f(x,y)}{\overline{F}(x,y)} = \frac{c[c(b+\theta x)(a+\theta y)+ab-\theta]}{(1+ax+by+\theta xy)^2},\ 0 \le \theta \le (c+1)ab \tag{2.73}$$

If $\theta = ab$, then $r(x,y)$ becomes the product of two marginal failure rate functions.

It is said that F is BIFR (BDFR) if $r(x,y)$ is increasing (decreasing) in both x and y. Note that since $\frac{\partial r(x,y)}{\partial x} = \frac{\partial \log r(x,y)}{\partial x} \times r(x,y)$, then it follows that

$$\frac{\partial \log r(x,y)}{\partial x} \ge 0 \Rightarrow \frac{\partial r(x,y)}{\partial x} \ge 0$$

2.13.2 Increasing Failure Rate In Average (IFRA)

T is said to be IFRA if and only if, equivalently

- $[\overline{F}\ (t)]^{1/2}$ is a decreasing function.
- $\frac{R(t)}{t}$ is an increasing function.

  Note that $R(t) = -\ln \overline{F}\ (t)$.

2.13.3 New Better Than Used (NBU)

T is said to be NBU, if and only if, equivalently

- $\overline{F}(x \mid t) \leq \overline{F}(x),\ \forall t \geq 0, \forall x \geq 0$
- $R(t+x) \geq R(t) + R(x),\ \forall t \geq 0, \forall x \geq 0$.

Note that $R(t)$ = -ln$\overline{F}$ (t).

2.13.4 Relationships between Basic Aging Notions

In continuous time, it is easy to show that (Barlow & Proschan,1975):

IFR $\Rightarrow$ IFRA $\Rightarrow$NBU

2.13.5 Mean Residual Life Function (MRLF)

The mean residual life function (MRLF) of a random variable T is defined as

$$\mu(t) = E(T - t \mid T > t) = \frac{1}{\overline{F}(t)} \int_t^{\infty} x dF(x) - t \tag{2.74}$$

2.13.6 Dependent Components

It is generally assumed that there are some forms of positive dependence among components in many reliability situations. This positive dependence among component life lengths arises from common environmental stresses and shocks, from components depending on common sources of power. (Bracquemond, Gaudoin, Roy, & Xie, 2001).

(T_1,T_2) will be jointly distributed with bivariate distribution $F(t_1,t_2)$ and marginals $F_1(t_1),F_2(t_2)$ if all properties hold below,

$$\begin{aligned} &F(t_1,\infty) = F_1(t_1) \\ &F(\infty,t_2) = F_2(t_2) \end{aligned} \tag{2.75}$$

$$\begin{aligned} &F(-\infty,t_2) = F(t_1,-\infty) = F(-\infty,-\infty) = 0 \\ &F(\infty,\infty) = 1 \end{aligned} \tag{2.76}$$

and

$$P(x_1 \le T_1 \le x_2, y_1 \le T_2 \le y_2) = F(x_2,y_2) - F(x_2,y_1) - F(x_1,y_2) + F(x_1,y_1) \ge 0 \tag{2.77}$$

If the partial derivative $\dfrac{\partial^2 F(t_1,t_2)}{\partial t_1 \partial t_2}$ exists everywhere, the bivariate distribution is said to have a density $f(t_1,t_2)$ defined by

$$f(t_1,t_2) = \frac{\partial^2 F(t_1,t_2)}{\partial t_1 \partial t_2} \tag{2.78}$$

The property (2.77) holds if and only if $f(t_1,t_2) \ge 0$. The conditions (2.75), (2.76) and (2.77) are necessary and sufficient conditions that $F(t_1,t_2)$ be a bivariate probability distribution function with marginals $F_1(t_1)$ and $F_2(t_2)$.

Note that the random variables (T_1,T_2) are independent if and only if for all t_1, t_2,

$$F(t_1,t_2) = F_1(t_1)F_2(t_2).$$

The joint probability

$$P(T_1 > t_1, T_2 > t_2) = \overline{F}(t_1, t_2) = 1 - F_1(t_1) - F_2(t_2) + F(t_1, t_2)\,.$$

Even if the conditions (2.75), (2.76) and (2.77) are met, a bivariate probability function is not uniquely determined by its marginals. On the contrary, Fréchet has shown that there is an infinite number of solutions to the problem of determining a distribution from its marginals. Fréchet obtained the condition

$$\max\left(F_1(t_1) + F_2(t_2) - 1, 0\right) \le F(t_1, t_2) \le \min\left(F_1(t_1), F_2(t_2)\right). \tag{2.79}$$

2.13.7 The Bivariate Exponential Distribution

Let the life lengths of two components of bivariate exponential distribution be dependent. Consider that three independent sources of shocks are present in the environment. A shock from source 1 destroys component 1; it occurs at a random time U_1, where $P[U_1 > t] = e^{-\lambda_1 t}$, a shock from source 2 destroys component 2; it occurs at a random time U_2, where $P[U_2 > t] = e^{-\lambda_2 t}$. Finally, a shock from source 3 destroys both components; it occurs at a random time U_{12}, where $P[U_{12} > t] = e^{-\lambda_{12} t}$.

Thus the random life length T_1 of component 1 satisfies

$$T_1 = \min(U_1, U_{12}), \tag{2.80}$$

while the random life length T_2 of component 2 satisfies

$$T_2 = \min(U_2, U_{12}). \tag{2.81}$$

Hence the joint survival probability

$$\overline{F}(t_1,t_2) = P(T_1 > t_1, T_2 > t_2) = e^{-\lambda_1 t_1 - \lambda_2 t_2 - \lambda_{12}\max(t_1,t_2)} \text{ for } t_1 \geq 0,\ t_2 \geq 0 \tag{2.82}$$

is called the bivariate exponential distribution (BVE).

Conversely, if T_1 and T_2 have BVE survival probability, then there exist independent exponential random variables U_1, U_2, U_{12} such that $T_1 = \min(U_1,U_{12})$ and $T_2 = \min(U_2,U_{12})$.

The BVE has exponential marginal distributions with survival probabilities,

$$\overline{F}_1(t_1) = P(T_1 > t_1) = \mathrm{e}^{-(\lambda_1+\lambda_{12})\mathrm{t}_1} \text{ for } \mathrm{t}_1 \geq 0, \tag{2.83}$$

and

$$\overline{F}_2(t_2) = P(T_2 > t_2) = e^{-(\lambda_1+\lambda_{12})t_2} \text{ for } \mathrm{t}_2 \geq 0. \tag{2.84}$$

If T is the life length of a component having exponential life distribution, then

$$P(T > s + t \mid T > s) = P(T > t) \text{ for all } s \geq 0,\ t \geq 0; \tag{2.85}$$

This means that the probability of survival for additional t units of time for a component of age is the same as that of a new component.

The similar property for bivariate exponential distribution takes the following form:

$$P(T_1 > s_1 + t, T_2 > s_2 + t \mid T_1 > t, T_2 > t) = P(T_1 > s_1, T_2 > s_2) \text{ for all } s_1 \geq 0,\ s_2 \geq 0,\ t \geq 0. \tag{2.86}$$

This equation means that the joint survival probability of a pair of components each of age is the same as that of a pair of new components.

Theorem: The BVE is the only bivariate distribution $F(x,y)$ with exponential marginals satisfying following bivariate property:

In terms of the joint survival probability $\overline{F}$, it may be written as:

$$\overline{F}(s_1 + t, s_2 + t) = \overline{F}(s_1, s_2)\overline{F}(t,t) \tag{2.87}$$

for all $s_1 \geq 0$, $s_2 \geq 0$, $t \geq 0$.

Lemma: If (2.87) holds, then

$$\overline{F}(x,y) = \begin{cases} e^{-\theta y}\overline{F}_1(x-y) & \textit{for} \quad x \geq y \\ e^{-\theta x}\overline{F}_2(y-x) & \textit{for} \quad x \leq y \end{cases}$$

and some $\theta > 0$, where $F_1(t) = \mathrm{F}(t,\infty)$ and $F_2(t) = F(\infty,t)$ are the marginal distributions.

CHAPTER THREE
LOCAL DEPENDENCE MEASURES

Global indices were used in probability theory and statistics to measure the dependence between random variables X and Y. Global indices are mentioned in section two. Linear correlation coefficient, Kendall's concordance coefficient, Spearman's coefficient are the famous ones. The linear correlation coefficient is appropriate for normal variables, Kendall's tau and Spearman's rho are independent of the marginal distributions of X and Y.

These indices are defined by the moments of the distribution on the whole plain and although X and Y are not independent, indices can be zero. Therefore, local dependence measure can be used. For example, when X and Y are survival random variables, it is important to identify the time of maximum association. The pairs (X,Y) and $(X^{'},Y^{'})$ can have the same global measure of dependence but may have two different distributions: with a local index it is easy to compare them.

3.1 Global Measures of Dependence

3.1.1 Desired properties of measure of dependence

Mari & Kotz (2001) discussed some axioms that global measures need to satisfy; these axioms are also desired for local measures. The properties are given below:

(i) *Standardization*

The values of an index are between 0 and 1.

(ii) *Independence*

If X and Y are independent, the index should be zero.

(iii) *Functional dependence*

If one variable is a function of the other, the index must be equal to 1.

(iv) *Increasing property*

The index should increase as the dependence increases.

(v) *Invariance*

The index should be invariant with respect to a linear transformation of the variables. A stronger condition would be that the index is marginally free which means that the measure of the dependence is the same as the corresponding measure on the copula.

(vi) *Symmetry*

If the variables are exchangeable, then the index should be symmetric.

(vii) *Relationship with measures for ordinal variables.*

If the index is defined for both ordinal and continuous variables, there should be a close relationship between two measures.

(viii) *Interpretability*

Numerical value of index can be interpreted as qualitative meaningful measure.

3.2 Local Dependence Measures

3.2.1 Local correlation coefficient of Bjerve and Doksum

Bjerve & Doksum (1993), Doksum, Blyth, Bradlow, Meng, & Zhao, H. (1994) and Blyth (1994) introduced and discussed correlation curves as local measures of variance explained by regression. They defined an experiment "heterocorrelations" if the strength of the relationship between a response variable Y and covariate X is different in different regions of the covariate spaces. They introduced a *correlation curve* which measures *heterocorrelacity* in terms of of the variance explained by regression locally at each covariate value. They obtained the squared correlation curve by expressing the linear model *variance explained to total variance* formula in terms of the slope of the conditional mean of $(Y \mid X = x)$.

The regression coefficients measure the relationship between the covariates $(X_1, X_2, ..., X_p)$ and the conditional mean $\mu(x) = E(Y \mid X = x)$ of the response variable *Y*. For the linear model, as it is mentioned before correlation coefficients measure the strength of the relationship between the response variable and the covariates. In fact, correlation coefficients measure how much of the variability of the response variable *Y* can be explained by covariates.

Linear correlation coefficient does not take into account the real association between the random variables (*X,Y*) in some situations. For example, the dependence does not exist on the whole plane or it is not linear or the variance of $(Y \mid X = x)$ is not constant.

Explained variance and strength of relationship depend on the values of covariates, that means they will be local. As an example, let *X* be "a level of symptom" and *Y* be "level of disease", then the regression slope is typically 0 for small $X = x$ and increases as x increases. Doksum, Blyth, Bradlow, Meng, & Zhao (1994) propose to quantify the *heterocorrelaticity* using the local correlation which they call as "correlation curves".

3.2.1.1 Formula for Local Correlation

The relationship between *X* and *Y* is locally linear if $\mu(x) = E(Y \mid X = x)$ is smooth. Using the conditional version of the correlation does not provide a formula for local correlation. For example, in the linear model,

$$Y = \alpha + \beta X + \varepsilon \tag{3.1}$$

then ,

$$\rho^2(X, Y \mid X = x) = \frac{\beta^2 Var(X \mid X = x)}{Var(Y \mid X = x)} = 0 \tag{3.2}$$

whenever $Var(Y \mid X = x) > 0$.

For that reason, it is important to move the measure from linear case to the general case. It is assumed that $\beta(x)$ exists. When the linear case slope β is replaced by the general case slope $\beta(x) = \frac{d}{dx}E(Y \mid X = x)$, the residual variance σ_{ε}^2 is also replaced by the general case residual variance $\sigma^2(x) = Var(Y \mid X = x)$.

So, let $\sigma_1^2 = Var(X)$. Then, for the linear regression model

$$\begin{aligned} \rho^2 &= \frac{\text{variance explained by regression}}{\text{total variance}} \\ &= \frac{Var(\alpha + \beta X)}{Var(\alpha + \beta X + \varepsilon)} \\ &= \frac{\sigma_1^2 \beta^2}{\sigma_1^2 \beta^2 + \sigma_\varepsilon^2}. \end{aligned} \tag{3.3}$$

This formula shows how the squared correlation coefficient measures the strength of the relationship between X and Y in terms of the regression slope and the residual variance, in the linear case. In this general case, the regression slope and the residual variance depend on x. Now, it measures the strength of the relationship between X and Y locally at the given covariate value x. Thus the squared correlation curve at x as

$$\rho^2(x) = (corr.curve)^2 = \frac{\sigma_1^2 \beta^2(x)}{\sigma_1^2 \beta^2(x) + \sigma^2(x)}, \quad x \in S \tag{3.4}$$

where $S = \{x : 0 < F(x) < 1\}$ is the support of the distribution of X and F denotes the distribution function of X.

Note that (3.4) may also be written as

$$\rho^2(x) = \left\{1 + \left[\sigma_1 \beta(x) / \sigma(x)\right]^{-2}\right\}^{-1} . \tag{3.5}$$

It is easily seen that $\rho^2(x)$ is an increasing function of the standardized regression slope $\sigma_1 \beta(x) / \sigma(x)$.

Since the sign of correlation is important, the correlation curve is defined as

$$\rho(x) = \frac{\sigma_1 \beta(x)}{\left[\{\sigma_1 \beta(x)\}^2 + \sigma^2(x)\right]^{1/2}} . \tag{3.6}$$

For the bivariate normal case, $\rho(x)$ is constructed as

$$\rho(x) = \frac{\sigma_1 \beta}{\left[\{\sigma_1 \beta\}^2 + \sigma^2\right]^{1/2}} , \tag{3.7}$$

where β is the slope of the regression line. $\rho(x)$ is not symmetric in *X* and *Y*, and applies only when *X* is a continuous random variable.

3.2.1.2 Properties

(i) $\rho(x)$ is standardized, i.e. $-1 \le \rho(x) \le 1$.

(ii) When (*X*,*Y*) is independent, $\rho(x) \equiv 0 \quad \forall x$.

(iii) When *Y* is a constant, $\beta(x) = \sigma(x) = 0$ and $\rho(x) = 0$ in this case.

(iv) When *Y* is a function of *X*, $\rho(x) = \pm 1$ for almost all *x*.

(v) $\rho(x)$ reduces to the Pearson correlation coefficient in the linear model.

(vi) Generally, $\rho(x)$ is not symmetric, but it is possible to construct a symmetrized $\hat{\rho}$, $\hat{\rho}(x) = sign(\rho(x))\sqrt{\rho(x)\rho(y)}$ if $\rho(x)$ and $\rho(y)$ have the same sign and $\hat{\rho}(x) = 0$ otherwise.

(vii) $\rho(x)$ is scale-free, but $\rho(x)$ is not marginal-free. This means that when the linear transformations of X and Y are $\widetilde{X} = aX + b$ and $\widetilde{Y} = cY + d$, $\rho(x)$ is unchanged but when the transformations $U = F_1(X)$ and $V = F_2(Y)$ are occurred, $\rho(u)$ is different from $\rho(x)$.

3.2.1.3 Estimators of $\rho(x)$

Assume that Y_i (i = *1,2,..,n*) are independently distributed with mean $\mu(x_i) = E(Y \mid x_i)$ and variance $\sigma^2(x_i) = Var(Y \mid x_i)$. Let $w(x)$ be a bounded integrable function with finite support $[-\tau, \tau]$ for some $\tau > 0$. Also it is assumed that

$$\int_{-\tau}^{\tau} w(t)dt = 1 \text{ and } w(\tau) = w(-\tau) = 0.$$

This type of $w(x)$ is often called kernel. Its integral, $W(x) = \int_{-\tau}^{x} w(t)dt$ is called *integrated kernel.*

Let $s_i = (x_i + x_{i+1})/2$, *i=1,...,n.* $s_0 = x_1$ and $s_n = x_n$ and w_j, *j=0,1,2* be three kernels. Kernel estimates for $\mu_1(x) = \mu(x), \mu_2(x) = E(Y^2 \mid x)$ and $\beta(x) = \mu'(x)$ can be constructed as

$$\hat{\mu}_j(x) = -\sum_{i=1}^{n}\left[W_j\left(\frac{x - s_i}{b_{jn}}\right) - W_j\left(\frac{x - s_{i-1}}{b_{jn}}\right)\right]Y_i^j, \qquad j=1,2 \tag{3.8}$$

and

$$\hat{\beta}(x) = -\frac{1}{b_{0n}}\sum_{i=1}^{n}\left[w_0\left(\frac{x-s_i}{b_{0n}}\right) - w_0\left(\frac{x-s_{i-1}}{b_{0n}}\right)\right]Y_i \tag{3.9}$$

where $\{b_{jn}, n \geq 1, j = 0,1,2\}$ are sequences of bandwidths and $W_j(x) = \int_{-\tau}^{x} w_j(x)dx$. Then, the estimate for the correlation curve is given by

$$\hat{\rho}(x) = \frac{\hat{\sigma}_1\hat{\beta}(x)}{\left[\hat{\sigma}_1^2\hat{\beta}^2(x) + \hat{\sigma}^2(x)\right]^{1/2}}, \tag{3.10}$$

where $\hat{\sigma}^2(x) = \hat{\mu}_2(x) - \hat{\mu}_1^2(x)$ and $\hat{\sigma}_1^2$ is the sample variance of the $x_i's$. $\hat{\rho}(x)$ is called the kernel correlation curve (with respect to the kernels w_j and bandwidths $\{b_{jn}, n \geq 1, j = 0,1,2\}$. The estimates used are the Gasser-Müller estimates introduced in Gasser & Müller (1984).

Jones (1996) discussed that the correlation curve does not treat *X* and *Y* on an equal footing. It means that this is really a regression concept and *Y* is a response variable and *X* is a predictor variable. A local dependence function of both *x* and *y* should measure the strength and direction of the association locally, treating variables symmetrically.

Example: Farlie-Gumbel-Morgenstern (FGM) Distribution with Uniform Marginals

Consider random variables (*X,Y*) having FGM distribution with uniform marginals given by the p.d.f.

$$f(x,y) = 1 + \alpha(1-2x)(1-2y), \qquad 0 \le x, y \le 1, \quad -1 \le \alpha \le 1; \tag{3.11}$$

then the correlation curve of (X,Y) is

$$\rho(x) = \frac{\frac{\alpha}{3\sqrt{12}}}{\sqrt{\frac{1}{12} - \frac{\alpha^2}{54}(1-3x)(1-2x)}} \tag{3.12}$$

where

$$\beta(x) = \frac{\alpha}{3} \text{ and } \sigma^2(x) = Var(Y \mid X = x) = \frac{1}{12} - \left(\frac{\alpha(1-2x)}{6}\right)^2.$$

The graphs for the FGM distribution for the different values of ε are given below:

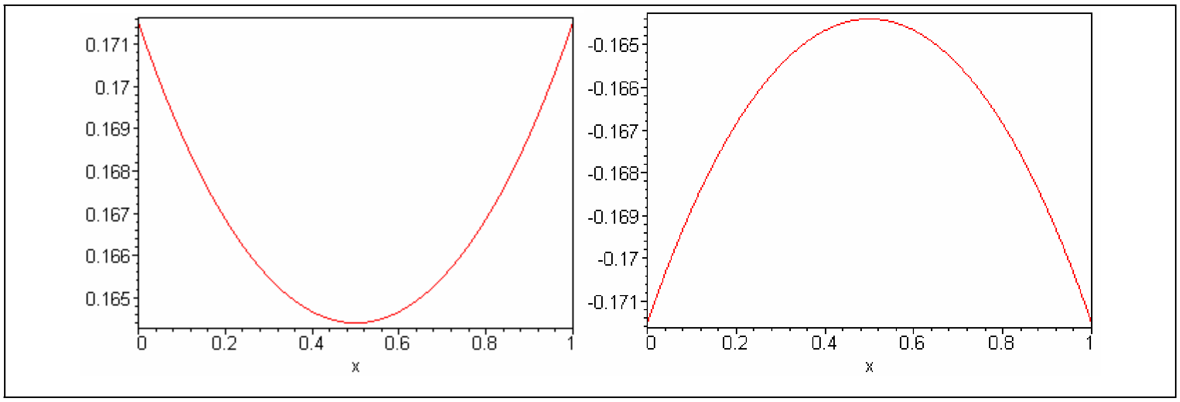


Figure 3.1 Local correlation curve of Bjerve and Doksum of FGM distribution $\alpha = 0.5, -0.5$.

3.2.2 Local Dependence Function of Holland and Wang

Holland & Wang (1987) introduced a local dependence function. Following their work Jones (1996) provided a motivation for the local dependence function by using mixed partial derivative of the log density.

Definition

Let X and Y be random variables with joint density functions $f(x,y)$ and marginal density functions $f_X(x)$ and $f_Y(y)$, then the local dependence function is defined as

$$\gamma(x,y) = \frac{\partial^2 \log f(x,y)}{\partial x \partial y} = \frac{1}{f(x,y)}\left\{ f^{11}(x,y) - \frac{f^{10}(x,y) f^{01}(x,y)}{f(x,y)} \right\} \tag{3.13}$$

where

$$f^{ij}(x,y) = \frac{\partial f(x,y)}{\partial x^i \partial y^j} \tag{3.14}$$

then $\gamma(x,y)$ is a function of the conditional distribution of Y given X, or vice versa and f is a bivariate density function. So, γ may be applied to all bivariate distributions with specified conditionals. But in general, conditionally specified joint distributions are extremely constrained. This function has many of the properties like correlation curve, such as $\gamma(x,y) \equiv \rho$ for all x and y when (X,Y) has the bivariate normal distribution.

Jones (1996) provides alternative motivation of $\gamma(x,y)$ which is based on localization of the Galton correlation coefficient. Also Jones (1998) identifies all other distributions from bivariate normal distribution which also have constant local dependence.

3.2.2.1 Properties

- It is finite everywhere;
- It is zero if and only if X and Y are independent;

- It is constant if f is the bivariate normal density and takes the value $\rho/\left(1-\rho^2\right)$ where ρ is the Pearson correlation coefficient (Jones, 1996) ;
- It is equivariant to monotone marginal transformations;
- It is equivariant to relocation and marginal rescaling;
- The perfect dependence case $Y = g(x)$ is neared, so $\gamma(x, y)$ tends to $\pm\infty$ along $y = g(x)$, depending on whether g is increasing or decreasing.
- The local dependence function has margin-free property; if either marginal is replaced, by multiplying $f(x,y)$ by a new marginal density and dividing by the original marginal, $\gamma(x, y)$ is unaffected.
- Marginal replacement implies that $\gamma(x, y)$ is a function only of the conditional distribution of $(Y|X=x)$ or $(X|Y=y)$.

3.2.2.2 Estimation of γ

Jones (1996) estimated γ by kernel methods (Wand & Jones, 1995) using

$$\hat{\gamma}(x,y) = \frac{\hat{f}_{XY}(x,y) - \hat{f}_1(x,y)^{-1}\hat{f}_X(x,y)\hat{f}_Y(x,y)}{(h_1 h_2 k_2)^2\,\hat{f}_1(x,y)} \tag{3.15}$$

where generically,

$$\hat{f}_Z(x,y) = n^{-1}\sum_{i=1}^{n} Z_i K_{h_1}(X_i - x)K_{h_2}(Y_i - y) \tag{3.16}$$

(i.e. $\hat{f}_{XY}$ has X_iY_i replacing Z_i while $\hat{f}_1$ replaces Z_i by 1 and is the usual kernel estimator of f which will often be written $\hat{f}$) . Here $\{(X_i, Y_i), i = 1,\ldots,n\}$ is the data set h_1 and h_2 are bandwidths controlling how much the data are smoothed in the x and y

directions. K is taken to be a product of symmetric univarite densities, $K_h(.) = h^{-1}K(h^{-1}.)$ and $k_2 = \int u^2 K(u)du$.

$\gamma(x, y)$ is a function only of the conditional distribution of $(Y \mid X = x)$ or $(X \mid Y = y)$. So, $\gamma(x, y)$ is suitable to all bivariate distributions with specified conditionals. But it is known that conditionally specified joint distributions are extremely constrained. (Kotz & Nadarajah, 2003). Hence the local dependence function is specified in terms of the conditional moments $E(X \mid y = y)$ and $E(Y \mid X = x)$, so it is suitable for a wider class of joint distributions. With this motivation, Bairamov et al. (2000) defined a new concept of local dependence function.

Example: Farlie-Gumbel-Morgenstern (FGM) Distribution with Uniform Marginals

Consider random variables (*X,Y*) having FGM distribution with uniform marginals (Bairamov& Kotz, 2002) given by the p.d.f.

$$f(x, y) = 1 + \alpha(1-2x)(1-2y), \qquad 0 \le x, y \le 1, \quad -1 \le \alpha \le 1;$$

then

$$\gamma(x, y) = \frac{1}{1+\alpha(1-2x)(1-2y)}\left[4\alpha - \frac{4\alpha^2(1-2y)(1-2x)}{1+\alpha(1-2x)(1-2y)}\right]. \tag{3.17}$$

The graphs for the FGM distribution for the different values of ε are given below:

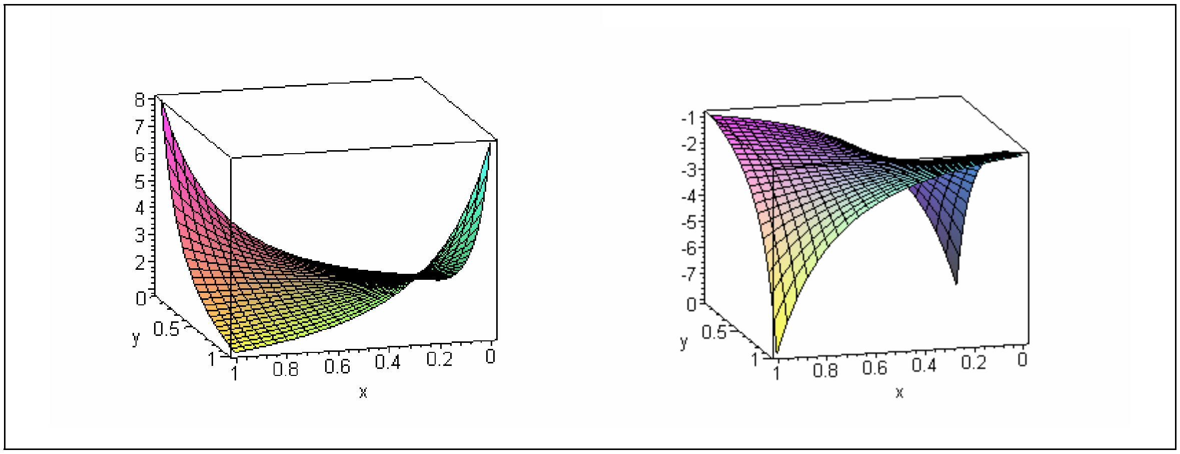


Figure 3.2 Local Dependence function of Holland and Wang of FGM distribution for $\alpha = 0.5,-0.5$.

3.2.3 Local Dependence Function of Bairamov and Kotz

Bairamov & Kotz (2000) have defined a new "local dependence function" $H(x,y)$ that provides a local point of view on dependence at a point (x,y).

Definition

Let X and Y be random variables with marginal distribution functions and densities F_X, f_X and F_Y, f_Y, respectively. Consider the following function of two variables

$$H(x,y) = \frac{E(X - E(X \mid Y = y))(Y - E(Y \mid X = x))}{\sqrt{E(X - E(X \mid Y = y))^2}\sqrt{E(Y - E(Y \mid X = x))^2}}. \tag{3.18}$$

This function is a generalization of Pearson correlation coefficient by replacing mathematical expectation EX and EY by conditional expectations $E(X \mid Y = y)$ and $E(Y \mid X = x)$ and after some algebraic manipulation, one may be rewrite H by

$$H(x,y) = \frac{\rho + \varphi_X(y)\varphi_Y(x)}{\sqrt{(1+\varphi^2{}_X(y))(1+\varphi^2{}_Y(x))}} \tag{3.19}$$

where

$$\varphi_X(y) = \frac{E(X) - E(X|Y=y)}{\sigma_X}, \quad \varphi_Y(x) = \frac{E(Y) - E(Y|X=x)}{\sigma_Y} \tag{3.20}$$

and

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}. \tag{3.21}$$

This local dependence measure has several useful properties. For instance, the new measure is symmetric in X and Y and its expected value is approximately equal the Pearson correlation coefficient. Another property is that, $X^*=0$ and $Y^*=0$ is a saddle point of H and $H(X^*, Y^*) = \rho$ when (X,Y) has bivariate normal distribution with correlation coefficient ρ.

3.2.3.1 Properties

Let (X,Y) have a bivariate distribution with finite second moments, Pearson correlation coefficient ρ, support $N_{X,Y}$, and local dependence function $H = H_{X,Y}$. Then,

1. If X and Y are independent, then $H(x,y) = 0$ for all $(x,y) \in N_{X,Y}$.
2. $|H(x,y)| \leq 1$ for all $(x,y) \in N_{X,Y}$.
3. If $H(x,y) = \pm 1$ for some $(x,y) \in N_{X,Y}$, then $\rho \neq 0$.
4. If $Y = aX+b$, a.s. then $H(X,Y) = sign(a)$, a.s.
5. If $\rho = \pm 1$, then $H(X,Y) = \pm 1$, a.s.
6. If $\widetilde{X} = aX + b$ and $\widetilde{Y} = cY + d$, then
   $H_{\widetilde{X},\widetilde{Y}}(\widetilde{x},\widetilde{y}) = sign(ac) H_{X,Y}(x,y)$

where $\tilde{x} = ax + b$ and $\tilde{y} = cy + d$.

7. If $H(x, y) = 0$ for all $(x,y) \in N_{X,Y}$, then either $E(X) = E(X \mid Y = y)$ or $E(Y) = E(Y \mid X = x)$ for all $(x,y) \in N_{X,Y}$ and $\rho = 0$.
8. The point (x^*, y^*) satisfying $\phi_X(y^*) = \phi_Y(x^*)$ is a saddle point of H and $H(x^*, y^*) = \rho$.

3.2.3.2 Estimation of H(x,y)

The local dependence function $H(x,y)$ is estimated via kernel methods. Bairamov & Kotz (2000) suggested an estimator for $H(x,y)$ by using Nadaraya (1964) and Watson's (1964) estimate for the regression functions $E(X \mid Y = y)$ and $E(Y \mid X = x)$. The estimators are

$$\hat{A}_X(y) = \frac{\sum_{i=1}^{n} X_i K\left(\frac{y - Y_i}{h_n}\right)}{\sum_{i=1}^{n} K\left(\frac{y - Y_i}{h_n}\right)} \quad \text{and} \quad \hat{A}_Y(x) = \frac{\sum_{i=1}^{n} Y_i K\left(\frac{x - X_i}{h_n}\right)}{\sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right)} \tag{3.22}$$

where (X_i, Y_i), i=1,2,...,n are the dataset, K is a kernel function, an integrable function with short tails, and h_n is a width sequence tending to zero at appropriate rates, i.e. $h_n \to 0$ as $n \to \infty$.

A standard estimate for Pearson correlation coefficient ρ is:

$$\hat{\rho} = \frac{n\sum X_i Y_i - \sum_i X_i \sum_j Y_j}{\sqrt{n\sum_i X_i^2 - (\sum_i X_i)^2}\sqrt{n\sum_i Y_i^2 - (\sum_i Y_i)^2}}. \tag{3.23}$$

Therefore, an estimate $\hat{H}(x, y)$ is obtained for $H(x,y)$,

$$\hat{H}(x,y)=\frac{\hat{\rho}+\dfrac{(\overline{X}-\hat{A}_X(y))(\overline{Y}-\hat{A}_Y(x))}{S_X S_y}}{\left(1+\left(\dfrac{\overline{X}-\hat{A}_X(y)}{S_X}\right)^2\right)^{1/2}\left(1+\left(\dfrac{\overline{Y}-\hat{A}_Y(x)}{S_Y}\right)^2\right)^{1/2}} \tag{3.24}$$

where $\hat{\rho}$ is an estimate for Pearson correlation coefficient,

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i,\overline{Y}=\frac{1}{n}\sum_{i=1}^{n}Y_i\ ,$$

and sample variances of X and Y,

$$S_X^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2,S_Y^2=\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\overline{Y})^2\ .$$

Example: Farlie-Gumbel-Morgenstern (FGM) Distribution with Uniform Marginals

Consider random variables (X,Y) having FGM distribution with uniform marginals given by the p.d.f.

$$f(x,y)=1+\alpha(1-2x)(1-2y)\,,\qquad 0\le x,y\le 1,\quad -1\le\alpha\le 1\,.$$

then local dependence function of FGM distribution is

$$H(x,y)=\frac{\alpha+\alpha^2(1-2x)(1-2y)}{\sqrt{3+\alpha^2(1-2x)^2}\sqrt{3+\alpha^2(1-2y)^2}}\,, \tag{3.25}$$

where

$EX = EY = 1/2$ while $Var(X) = Var(Y) = 1/12$,

also,

$E(Y \mid X = x) = 1/2 - \alpha(1-2x)/6$,
$E(X \mid Y = y) = 1/2 - \alpha(1-2y)/6$

Consequently,

$$\phi_Y(x) = \frac{\sqrt{12}\alpha(1-2x)}{6} \text{ and } \phi_Y(y) = \frac{\sqrt{12}\alpha(1-2y)}{6}.$$

In terms of the Pearson correlation coefficient ρ, which equals to $\alpha/3$, the local dependence function (3.25) is,

$$H(x, y) = \frac{\rho + 3\rho^2(1-2x)(1-2y)}{\sqrt{1+3\rho^2(1-2x)^2}\sqrt{1+3\rho^2(1-2y)^2}} \tag{3.26}$$

Here, the symmetry point is $(x, y) = (1/2, 1/2)$, then the local dependence function is equal the correlation curve,

$$H(1/2, 1/2) = \frac{\alpha}{3} = \rho. \tag{3.27}$$

The graphs for the FGM distribution for the different values of ε are given below:

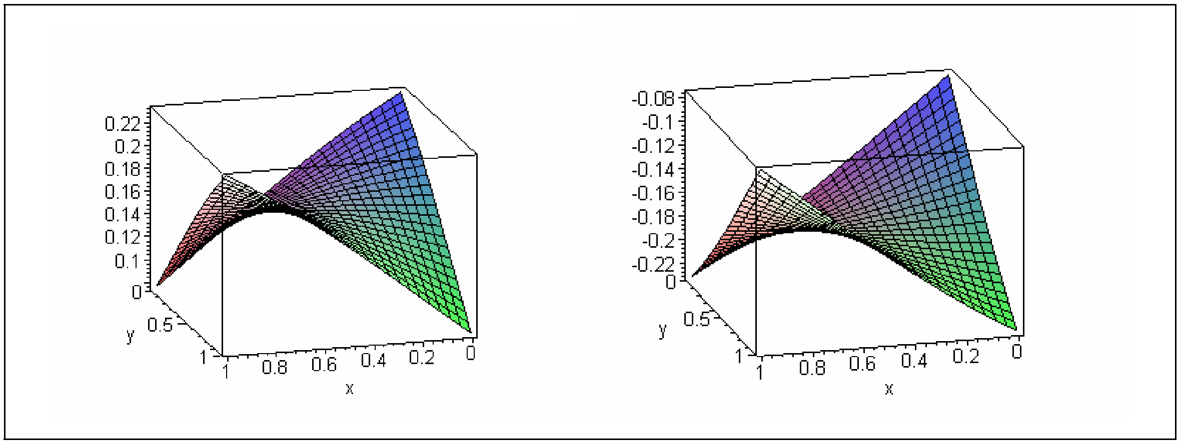


Figure 3.3 Local Dependence function of Bairamov and Kotz of FGM distribution for α = 0.5, -0.5.

CHAPTER FOUR
LOCAL NUMERICAL CHARACTERISTICS
and
LOCAL DEPENDENCE FUNCTION

Moments are important characteristics of random variables. Existence of the first and second moments is a necessary condition of many basic theorems of probability. For instance, important results of probability theory such as the law of large numbers and the central limit theorem require existence of the first and second moments.

For a random variable X, expected value EX may fail to exist, for example a random variable having Cauchy distribution does not have a finite expected value. Hence it is reasonable to consider the expected value of a random variable in some local area from the support set of X. It can be shown that even though the expected value for a random variable fails to exist, the local expected value may exist.

In many practical applications in medicine and biology, one needs to estimate the degree of dependence between two random variables calculated using data taken solely from a localized area. In cancer research, especially, doctors work with the patient scanogram from the healthy tissue, benign tumor and malignant tumor. Cancer cells, unlike normal cells, lack contact inhibition, and patient scanograms providing important information about characteristics of cancer cells are different in infected and noninfected areas. In this case, for example one may need to calculate the degree of dependence of the scanograms of the patients with cancer using the data from the tumor area, which would be different from the estimated values using data taken from the whole body. In an investigation of an association between two characteristics of the scanogram in the healthy and infected areas, two different results would be expected. Then the local dependence functions measuring the dependence in a localized area from the support set of the variables of interest become important. The Pearson correlation coefficient and many other scalar dependence measures, of course, play an important role in understanding the simplest dependence relationships between two random variables. In general, the dependence structure between a pair of random variables is very complex and the scalar dependence measures are inadequate to explain the natural association between them.

The study of the local dependence between random variables has attracted some interest in last years. There are several local dependence measures introduced and studied in statistical literature in the past decade. Bjerve & Doksum (1993) and Blyth (1994) constructed a correlation curve that measures the strength of the association between random variables locally. Holland & Wang (1987) introduced the local dependence function, which is a localized version of the Pearson correlation coefficient and is the second order partial derivative of logarithm of the bivariate density function. Jones (1998) has provided an alternative motivation for studying this function. Bairamov & Kotz (2000) have introduced a new local dependence function based on regression concept which is a radical generalization of the Pearson correlation coefficient.

We introduce a new local dependence function. Its properties and also local dependence function of some bivariate distributions are investigated in this section.

4.1 Local Expected Value and Variance

Let X be a continuous random variable with a distribution function F with support $IR = (-\infty, \infty)$. The function

$$E_{X,\varepsilon}(x) = E(X \mid x - \varepsilon < X < x + \varepsilon) \tag{4.1}$$

is said to be a local expected value of a random variable X in the ε neighbourhood of the point $x \in IR$. Then we have

$$E_{X,\varepsilon}(x) = E(X \mid x-\varepsilon < X < x+\varepsilon) = \frac{E(XI_{(x-\varepsilon,x+\varepsilon)}(X))}{P(x-\varepsilon < X < x+\varepsilon)}$$
$$= \frac{\int_{x-\varepsilon}^{x+\varepsilon} uf(u)du}{F(x+\varepsilon) - F(x-\varepsilon)} \tag{4.2}$$

where

$$I_A(u) = \begin{cases} 1, & u \in A \\ 0, & u \notin A \end{cases} . \tag{4.3}$$

If the support of X is finite interval [a,b], then it is clear that

$$E_{X,\varepsilon}(x) = \frac{\int_{\max(a,x-\varepsilon)}^{\min(b,x+\varepsilon)} uf(u)du}{F(\min(b,x+\varepsilon)) - F(\max(a,x-\varepsilon))} \quad . \tag{4.4}$$

Analogously,

$$\begin{aligned} Var_{X,\varepsilon}(x) &= Var(X \mid \max(a,x-\varepsilon) < X < \min(b,x+\varepsilon)) \\ &= E(X^2 \mid \max(a,x-\varepsilon) < X < \min(b,x+\varepsilon)) \\ &\quad - E^2(X \mid \max(a,x-\varepsilon) < X < \min(b,x+\varepsilon))\big) \end{aligned} \tag{4.5}$$

is a local variance function.

4.1.1 Properties of Local Expected Value Function

Let X has support IR , then

1) $E(c|x-\varepsilon < X < x+\varepsilon) = c, \quad c \neq 0$ where c is a constant.

Proof. $E(c \mid x-\varepsilon < X < x+\varepsilon) = \dfrac{\int_{x-\varepsilon}^{x+\varepsilon} cf(u)du}{F(x+\varepsilon)-F(x-\varepsilon)} = c$

2) **a)** If $\widetilde{X} = aX + b$, then $E_{\widetilde{X},\varepsilon}(x) = aE_{X,\frac{\varepsilon}{a}}\left(\dfrac{x-b}{a}\right) + b$.

Proof. $F_{\widetilde{X}}(t) = P(\widetilde{X} \le t) = P(aX + b \le t) = P\left(X \le \dfrac{t-b}{a}\right) = F_X\left(\dfrac{t-b}{a}\right)$

then

$$f_{\widetilde{X}}(t) = \frac{1}{a} f_X\left(\frac{t-b}{a}\right)$$

So,

$$E_{\widetilde{X},\varepsilon}(x) = \frac{\int_{x-\varepsilon}^{x+\varepsilon} tf_{\widetilde{X}}(t)dt}{\int_{x-\varepsilon}^{x+\varepsilon} f_{\widetilde{X}}(t)dt}$$

by using transformation technique that $\dfrac{t-b}{a} = z$, then

$$E_{\widetilde{X},\varepsilon}(x)=\frac{1}{F_X\left(\frac{x-b+\varepsilon}{a}\right)-F_X\left(\frac{x-b-\varepsilon}{a}\right)}\int_{\frac{x-\varepsilon-b}{a}}^{\frac{x+\varepsilon-b}{a}}(az+b)\frac{1}{a}f_X(z)adz$$

$$=a\left(\frac{1}{F_X\left(\frac{x-b+\varepsilon}{a}\right)-F_X\left(\frac{x-b-\varepsilon}{a}\right)}\int_{\frac{x-\varepsilon-b}{a}}^{\frac{x+\varepsilon-b}{a}}zf_X(z)dz\right)$$

$$+b\left(\frac{1}{F_X\left(\frac{x-b+\varepsilon}{a}\right)-F_X\left(\frac{x-b-\varepsilon}{a}\right)}\int_{\frac{x-\varepsilon-b}{a}}^{\frac{x+\varepsilon-b}{a}}f_X(z)dz\right)$$

$$=aE_{X,\frac{\varepsilon}{a}}\left(\frac{x-b}{a}\right)+b$$

b) If $\widetilde{X}=aX+b$ and $\widetilde{x}=ax+b$, then $E_{\widetilde{X},\varepsilon}(\widetilde{x})=aE_{X,\frac{\varepsilon}{a}}(x)+b$.

Proof. See proof (a) by using $\widetilde{x}=ax+b$

3) $\lim\limits_{\varepsilon\to 0}E_{X,\varepsilon}(x)=x, \quad \forall x\in IR$

Proof. By using L'Hospital Rule and then Leibnitz' rule,

$$\lim_{\varepsilon\to 0}E_{X,\varepsilon}(x)=\lim_{\varepsilon\to 0}\frac{\frac{\partial}{\partial\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon}uf(u)du}{\frac{\partial}{\partial\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon}f(u)du}$$

$$=\lim_{\varepsilon\to 0}\frac{(x+\varepsilon)f(x+\varepsilon)(1)-(x-\varepsilon)f(x-\varepsilon)(-1)}{f(x+\varepsilon)(1)-f(x-\varepsilon)(-1)}$$

$$=x$$

4) $\lim\limits_{\varepsilon\to 0} E(E_{X,\varepsilon}(X)) = EX, \quad \forall x \in IR$

Proof.

$$\lim_{\varepsilon\to 0} E(E_{X,\varepsilon}(X)) = \lim_{\varepsilon\to 0}\int_{-\infty}^{\infty}\frac{\int_{x-\varepsilon}^{x+\varepsilon} uf(u)du}{\int_{x-\varepsilon}^{x+\varepsilon} f(u)du} f(x)dx$$

$$= \int_{-\infty}^{\infty}\lim_{\varepsilon\to 0}\frac{\int_{x-\varepsilon}^{x+\varepsilon} uf(u)du}{\int_{x-\varepsilon}^{x+\varepsilon} f(u)du} f(x)dx$$

$$= \int_{-\infty}^{\infty} xf(x)dx$$

$$= E(X)$$

5) Let $(X \mid x-\varepsilon < X < x+\varepsilon) \equiv Y_{X,\varepsilon}(x)$ be a random variable with expected value $E_{X,\varepsilon}(x) = E(Y_{X,\varepsilon}(x))$. Then for any $\delta > 0$

$$P\left\{\left|Y_{X,\varepsilon}(x) - E_{X,\varepsilon}(x)\right| > \delta\right\} \le \frac{Var\left(Y_{X,\varepsilon}(x)\right)}{\delta^2},$$

Proof .

$$
\begin{aligned}
P\{|Y_{X,\varepsilon}(x)-E_{X,\varepsilon}(x)|>\delta\} &= P\{Y_{X,\varepsilon}(x)\in G=\{y:|y-E_{X,\varepsilon}(x)|>\delta\}\} \\
&= \int_{\{y:|y-E_{X,\varepsilon}(x)|^2>\delta^2\}} dF_Y(y) \\
&= \int_{\left\{y:\frac{|y-E_{X,\varepsilon}(x)|^2}{\delta^2}>1\right\}} dF_Y(y) \\
&\le \frac{1}{\delta^2}\int |y-E_{X,\varepsilon}(x)|^2\, dF_Y(y) \\
&\le \frac{1}{\delta^2} Var(Y_{X,\varepsilon}(x))
\end{aligned}
$$

6) Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of random variables and the distribution function of X_i is F_i , $i = 1, 2, \ldots$; i.e., $Y_{X_1,\varepsilon}(x), Y_{X_2,\varepsilon}(x), \ldots, Y_{X_n,\varepsilon}(x), \ldots$ is a sequence of independent copies of $Y_{X,\varepsilon}(x)$. Consider the sequence $Y_{X_i,\varepsilon}(x), i = 1, 2, \ldots$ It is clear that the distribution function of $Y_{X_i,\varepsilon}(x)$ is

$$
F_{X_i,\varepsilon}(t)=\begin{cases} 0, & t<x-\varepsilon \\ \frac{P\{x-\varepsilon\le X_i\le \min(t,x+\varepsilon)}{F_i(x+\varepsilon)-F_i(x-\varepsilon)}, & x-\varepsilon\le t\le x+\varepsilon \\ 1, & t>x+\varepsilon \end{cases}
$$

Also,

$$E(Y_{X_i,\varepsilon}(x)) = a_i(x)$$

and

$Var(Y_{X_i,\varepsilon}(x)) = D_i(x)$ where $i = 1, 2, \ldots, n, \ldots$

$$\frac{1}{n^2}\sum_{i=1}^{n} D_i(x) \to 0, \ \ \forall x, \ n \to \infty;$$

then for any $x \in IR$

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} Y_{X_i,\varepsilon}(x) - \frac{1}{n}\sum_{i=1}^{n} E_{X_i,\varepsilon}(x)\right| > \delta\right\} \to 0, \qquad n \to \infty$$

Proof. By using Chebyshev's Inequality

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} Y_{X_i,\varepsilon}(x) - E\left(\frac{1}{n}\sum_{i=1}^{n} Y_{X_i,\varepsilon}(x)\right)\right| > \delta\right\} \le \frac{Var\left(\frac{1}{n}\sum_{i=1}^{n} Y_{X_i,\varepsilon}(x)\right)}{\delta^2}$$

$$\le \frac{\frac{1}{n^2}\sum_{i=1}^{n} Var(Y_{X_i,\varepsilon}(x))}{\delta^2}$$

$$\le \frac{1}{n^2\delta^2}\sum_{i=1}^{n} Var(Y_{X_i,\varepsilon}(x))$$

$\lim_{n\to\infty}\frac{1}{n^2\varepsilon^2}\sum_{i=1}^{n} Var(Y_{X_i,\varepsilon}(x)) = 0$, then for any $x \in IR$

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} Y_{X_i,\varepsilon}(x) - \frac{1}{n}\sum_{i=1}^{n} E_{X_i,\varepsilon}(x)\right| > \delta\right\} \to 0, \qquad n \to \infty$$

7) Let $Y_{X_1,\varepsilon}(x), Y_{X_2,\varepsilon}(x), \ldots, Y_{X_n,\varepsilon}(x), \ldots$ be a sequence of independent and identically distributed random variables. Assume $\sigma_\varepsilon^2(x) = Var(Y_{X_1,\varepsilon}(x)) < \infty, \quad \forall x.$ Then

$$\lim_{n\to\infty} P\left\{ \frac{\sum_{i=1}^{n} Y_{X_i,\varepsilon}(x) - nE_{X_1,\varepsilon}(x)}{\sqrt{Var\left(\sum_{i=1}^{n}(Y_{X_i,\varepsilon}(x))\right)}} \le t \right\} = N_{(0,1)}(t)$$

where $Var\left(\sum_{i=1}^{n}(Y_{X_i,\varepsilon}(x))\right) = n\sigma_\varepsilon^2(x)$.

Proof. Let $S_n(x) = \sum_{i=1}^{n} Y_{X_i}(x)$

$$E(S_n(x)) = E\left(\sum_{i=1}^{n} Y_{X_i,\varepsilon}(x)\right) = \sum_{i=1}^{n} E(Y_{X_i,\varepsilon}(x)) = nE_{X_1,\varepsilon}(x),$$

$$Var(S_n(x)) = Var\left(\sum_{i=1}^{n} Y_{X_i,\varepsilon}(x)\right) = \sum_{i=1}^{n} Var(Y_{X_i,\varepsilon}(x)) = nVar_{X_1,\varepsilon}(x) = n\sigma^2_{X_1,\varepsilon}(x).$$

Assume that $E(Y_{X_i,\varepsilon}(x)) = E_{X_1,\varepsilon}(x)$.

$$\rho_n(x) = \frac{S_n(x) - nE_{X_1,\varepsilon}(x)}{\sigma_{X_1,\varepsilon}(x)\sqrt{n}}.$$

Assume that $E(Y_{X_i,\varepsilon}(x)) = E_{X_1,\varepsilon}(x)$, without loss of generality let $E_{X_1,\varepsilon}(x) = 0$,

Consider $\varphi_{\rho_n}(x) = Ee^{i\frac{\sum_{i=1}^{n} Y_{X_i,\varepsilon}(x)}{\sqrt{n}\sigma(x)}t} = Ee^{i\frac{Y_{X_1,\varepsilon}(x)}{\sqrt{n}\sigma(x)}t} Ee^{i\frac{Y_{X_1,\varepsilon}(x)}{\sqrt{n}\sigma(x)}t} \ldots Ee^{i\frac{Y_{X_1,\varepsilon}(x)}{\sqrt{n}\sigma(x)}t}$.

Since

$$\varphi(t) = \varphi(0) + \varphi^{'}(0)t + \frac{\varphi^{''}(0)}{2!}t^2 + o(t^2)$$

where $\varphi(0)=1, \varphi'(0)=iE_{X,\varepsilon}(x) \quad and \quad \varphi''(0)=-\sigma^2(x)$, then

$$\varphi(t)=\left[1+iE_{X,\varepsilon}(x)t-\frac{\sigma^2(x)t^2}{2}+o(t^2)\right]$$

$$\varphi^n_{\rho_n(x)}\left(\frac{t}{\sqrt{n}\sigma(x)}\right)=\left[1+iE_{X,\varepsilon}(x)\frac{t}{\sqrt{n}\sigma(x)}-\frac{\sigma^2(x)}{2}\frac{t^2}{n\sigma^2(x)}+o\left(\frac{t^2}{n\sigma^2(x)}\right)\right]^n$$

$\lim\limits_{n\to\infty}\ln\varphi_{\rho_n(x)}(t)=-\dfrac{t^2}{2}$ then $\lim\limits_{n\to\infty}\varphi_{Y_{X,\varepsilon}(x)}(t)=e^{-t^2/2}$ where it is the characteristic

function of standard normal distribution.

4.1.2 Properties of Local Variance

1) $Var(c|x-\varepsilon<X<x+\varepsilon)=0,$ where c is a constant.

Proof.

$$\begin{aligned} Var(c\mid x-\varepsilon<X<x+\varepsilon) &= E(c^2\mid x-\varepsilon<X<x+\varepsilon)-\left[E(c\mid x-\varepsilon<X<x+\varepsilon)\right]^2 \\ &= c^2-c^2=0 \end{aligned}$$

2) **a)** If $\widetilde{X}=aX+b$, then $Var_{\widetilde{X},\varepsilon}(x)=a^2 Var_{X,\frac{\varepsilon}{a}}\left(\frac{x-b}{a}\right)$

Proof.

$$
\begin{aligned}
Var_{\tilde{X},\varepsilon}(x) &= E(\tilde{X}^2 \mid x-\varepsilon < \tilde{X} < x+\varepsilon) - \left[E(\tilde{X} \mid x-\varepsilon < \tilde{X} < x+\varepsilon)\right]^2 \\
&= E((aX+b)^2 \mid x-\varepsilon < aX+b < x+\varepsilon) - \left[E((aX+b) \mid x-\varepsilon < aX+b < x+\varepsilon)\right]^2 \\
&= a^2\left[E_{X^2,\frac{\varepsilon}{a}}\left(\frac{x-b}{a}\right) - \left[E_{X,\frac{\varepsilon}{a}}\left(\frac{x-b}{a}\right)\right]^2\right] \\
&= a^2 Var_{X,\frac{\varepsilon}{a}}\left(\frac{x-b}{a}\right)
\end{aligned}
$$

b) If $\tilde{X} = aX+b$ and $\tilde{x} = ax+b$, then $Var_{\tilde{X},\varepsilon}(x) = a^2 Var_{X,\frac{\varepsilon}{a}}(x)$

Proof.

$$
\begin{aligned}
Var_{\tilde{X},\varepsilon}(\tilde{x}) &= E(\tilde{X}^2 \mid \tilde{x}-\varepsilon < \tilde{X} < \tilde{x}+\varepsilon) - \left[E(\tilde{X} \mid \tilde{x}-\varepsilon < \tilde{X} < \tilde{x}+\varepsilon)\right]^2 \\
&= E((aX+b)^2 \mid ax+b-\varepsilon < aX+b < ax+b+\varepsilon) \\
&\quad - \left[E((aX+b) \mid ax+b-\varepsilon < aX+b < ax+b+\varepsilon)\right]^2 \\
&= a^2\left[E_{X^2,\frac{\varepsilon}{a}}(x) - \left[E_{X,\frac{\varepsilon}{a}}(x)\right]^2\right] \\
&= a^2 Var_{X,\frac{\varepsilon}{a}}(x)
\end{aligned}
$$

3) $\lim\limits_{\varepsilon \to 0} Var_{X,\varepsilon}(x) = 0$

Proof.

$$
Var_{X,\varepsilon}(x) = E(X^2 \mid x-\varepsilon < X < x+\varepsilon) - \left(E(X \mid x-\varepsilon < X < x+\varepsilon)\right)^2
$$

By using L'Hospital Rule and then Leibnitz's rule,

$$\lim_{\varepsilon\to 0} E_{X^2,\varepsilon}(x) = \lim_{\varepsilon\to 0}\frac{\frac{\partial}{\partial\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon} u^2 f(u)du}{\frac{\partial}{\partial\varepsilon}\int_{x-\varepsilon}^{x+\varepsilon} f(u)du}$$

$$= \lim_{\varepsilon\to 0}\frac{(x+\varepsilon)^2 f(x+\varepsilon)(1)-(x-\varepsilon)^2 f(x-\varepsilon)(-1)}{f(x+\varepsilon)(1)-f(x-\varepsilon)(-1)}$$

$$= x^2.$$

Also it is proved that $\lim_{\varepsilon\to 0} E_{X,\varepsilon}(x) = x, \quad \forall x \in IR$, then

$$\lim_{\varepsilon\to 0} Var_{X,\varepsilon}(x) = x^2 - x^2 = 0.$$

4) $Var_{X,\varepsilon}(x) = 0$ for $\forall x$ if and only if there exists a constant c such that $P(X = c) = 1$.

Proof.

Let $Y = (X \mid x-\varepsilon < X < x+\varepsilon)$.

$$E(Y - E_{X,\varepsilon}(x))^2 = E(Y^2 - 2YE_{X,\varepsilon}(x) + (E_{X,\varepsilon}(x))^2) \geq 0$$

$$Var(Y) = E(Y - E_{X,\varepsilon}(x))^2 \geq 0$$

$(Y - E_{X,\varepsilon}(x))^2 = 0$ with probability 1, i.e. $Y = E_{X,\varepsilon}(x)$.

4.2 Examples of Some Special Distributions

4.2.1 Uniform Distribution

Let X be a random variable having uniform p.d.f. on $[a,b]$,

$$f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & otherwise. \end{cases}$$

Then the local expected value of X,

$$E_{X,\varepsilon}(x) = \frac{1}{F(\min(b, x+\varepsilon) - F(\max(a, x-\varepsilon))} \int_{\max(a,x-\varepsilon)}^{\min(b,x+\varepsilon)} u \frac{1}{b-a} du$$

$$= \frac{\frac{1}{b-a}\left[\frac{(\min(b, x+\varepsilon))^2 - (\max(a, x-\varepsilon))^2}{2}\right]}{\frac{\min(b, x+\varepsilon) - a}{b-a} - \frac{\max(a, x-\varepsilon) - a}{b-a}}$$

$$= \frac{\min(b, x+\varepsilon) + \max(a, x-\varepsilon)}{2}$$

where

$$\min(b, x+\varepsilon) = \begin{cases} b, & x > b-\varepsilon \\ x+\varepsilon, & x < b-\varepsilon \end{cases}$$

$$\max(a, x-\varepsilon) = \begin{cases} x-\varepsilon, & x > a+\varepsilon \\ a, & x < a+\varepsilon \end{cases}$$

and the variance of X

$$E(X^2 \mid \max(a, x-\varepsilon) < X < \min(b, x+\varepsilon)) =$$
$$\frac{1}{F(\min(b, x+\varepsilon)) - F(\max(a, x-\varepsilon))} \int_{\max(a,x-\varepsilon)}^{\min(b,x+\varepsilon)} u^2 \frac{1}{b-a} du$$

$$= \frac{1}{3}((\min(b, x+\varepsilon))^2 + \min(b, x+\varepsilon)\max(a, x-\varepsilon) + (\max(a, x-\varepsilon))^2)$$

then,

$$Var_{X,\varepsilon}(x) = \frac{1}{3}((\min(b, x+\varepsilon))^2 + \min(b, x+\varepsilon)\max(a, x-\varepsilon) + (\max(a, x-\varepsilon))^2)$$
$$-\left(\frac{\min(b, x+\varepsilon) + \max(a, x-\varepsilon)}{2}\right)^2$$
$$= \frac{(\min(b, x+\varepsilon) - \max(a, x-\varepsilon))^2}{12}.$$

As a result, local expected value and local variance function of uniformly distributed random variable are

$$E_{X,\varepsilon}(x) = \frac{\min(b, x+\varepsilon) + \max(a, x-\varepsilon)}{2}, \tag{4.6}$$

Also, we have

$$Var_{X,\varepsilon}(x) = \frac{(\min(b, x+\varepsilon) - \max(a, x-\varepsilon))^2}{12}. \tag{4.7}$$

In Figure 4.1, we present graphs of local expected value for uniform distribution for different values of ε. It is clear that $\lim_{\varepsilon\to b} E_{X,\varepsilon}(x) = EX$ and $\lim_{\varepsilon\to b} Var_{X,\varepsilon}(x) = Var(X)$. It is seen that if X is a continuous random variable with uniform distribution on $[a,b]$, then $E(E_{X,\varepsilon}(X)) = EX, \forall x$ and ε. Indeed,

$$\begin{aligned} E(E_{X,\varepsilon}(X)) &= \frac{1}{2(b-a)}\left[\int_a^b \min(b, x+\varepsilon)dx + \int_a^b \max(a, x-\varepsilon)dx\right] \\ &= \frac{1}{2(b-a)}\left[\int_a^{b-\varepsilon}(x+\varepsilon)dx + \int_{b-\varepsilon}^{b} b\,dx + \int_a^{a+\varepsilon} a\,dx + \int_{a+\varepsilon}^{b}(x-\varepsilon)dx\right] \\ &= \frac{1}{2(b-a)}(b^2-a^2) = \frac{a+b}{2} \end{aligned} \tag{4.8}$$

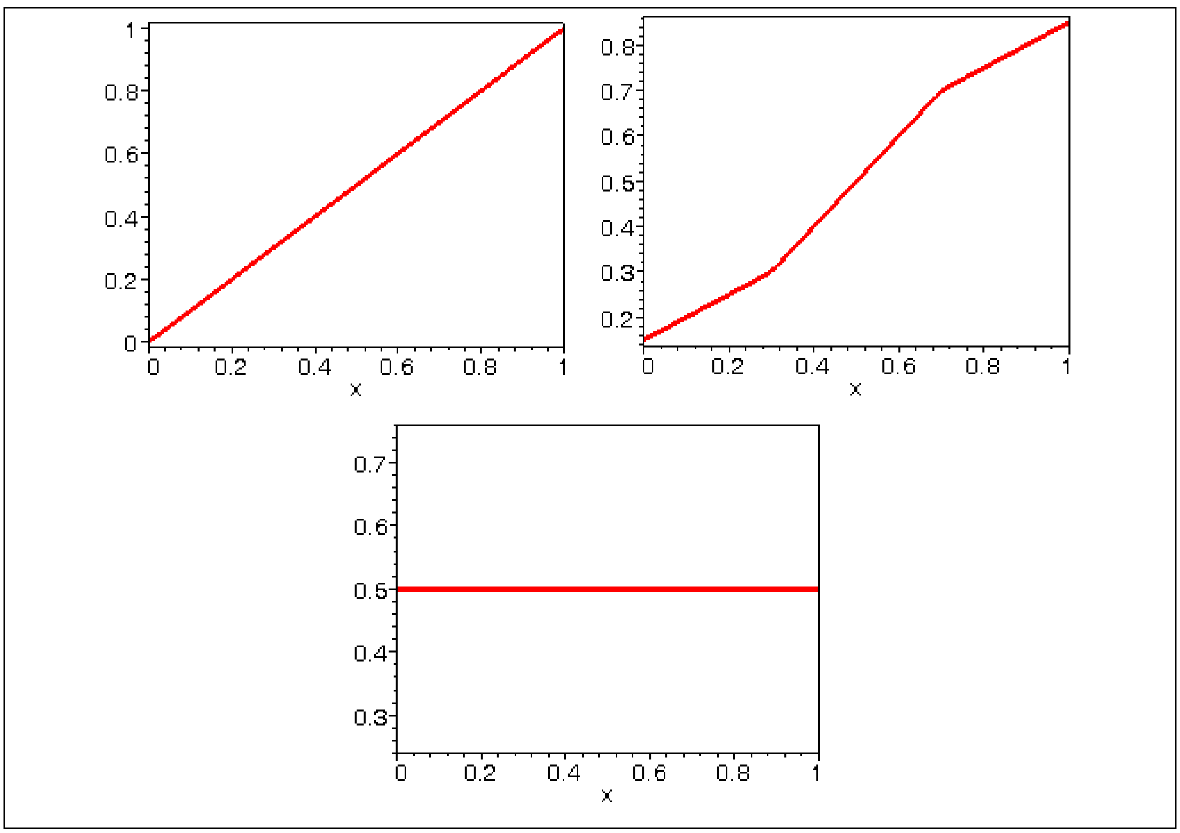


Figure 4.1 Graphs for Local expected value function of Uniform Distribution for the different values of $\varepsilon = 0.01, 0.3, 1, \quad 0 < x < 1$

4.2.2 Exponential Distribution

Let X be a random variable having exponential p.d.f.

$$f(x)=\begin{cases}\lambda e^{-\lambda x}, & x>0\\ 0, & otherwise\end{cases}.$$

The local expected value can be obtained as

$$E_{X,\varepsilon}(x)=\frac{1}{F(x+\varepsilon)-F(\max(0,x-\varepsilon))}\int_{\max(0,x-\varepsilon)}^{x+\varepsilon} u\lambda e^{-\lambda u}\,du$$

$$=\frac{-(x+\varepsilon)e^{-\lambda(x+\varepsilon)}+\max(0,x-\varepsilon)e^{-\lambda\max(0,x-\varepsilon)}-\dfrac{e^{-\lambda(x+\varepsilon)}}{\lambda}+\dfrac{e^{-\lambda\max(0,x-\varepsilon)}}{\lambda}}{e^{-\lambda\max(0,x-\varepsilon)}-e^{-\lambda(x+\varepsilon)}}.$$

After some algebraic manipulations, the local expected value is

$$E_{X,\varepsilon}(x)=\frac{e^{-\lambda\max(0,x-\varepsilon)}(\max(0,x-\varepsilon)+\frac{1}{\lambda})-e^{-\lambda(x+\varepsilon)}(x+\varepsilon+\frac{1}{\lambda})}{e^{-\lambda\max(0,x-\varepsilon)}-e^{-\lambda(x+\varepsilon)}}. \tag{4.9}$$

We present the graphs of the local expected value of exponentially distributed random variable for different values of ε in Figure 4.2.

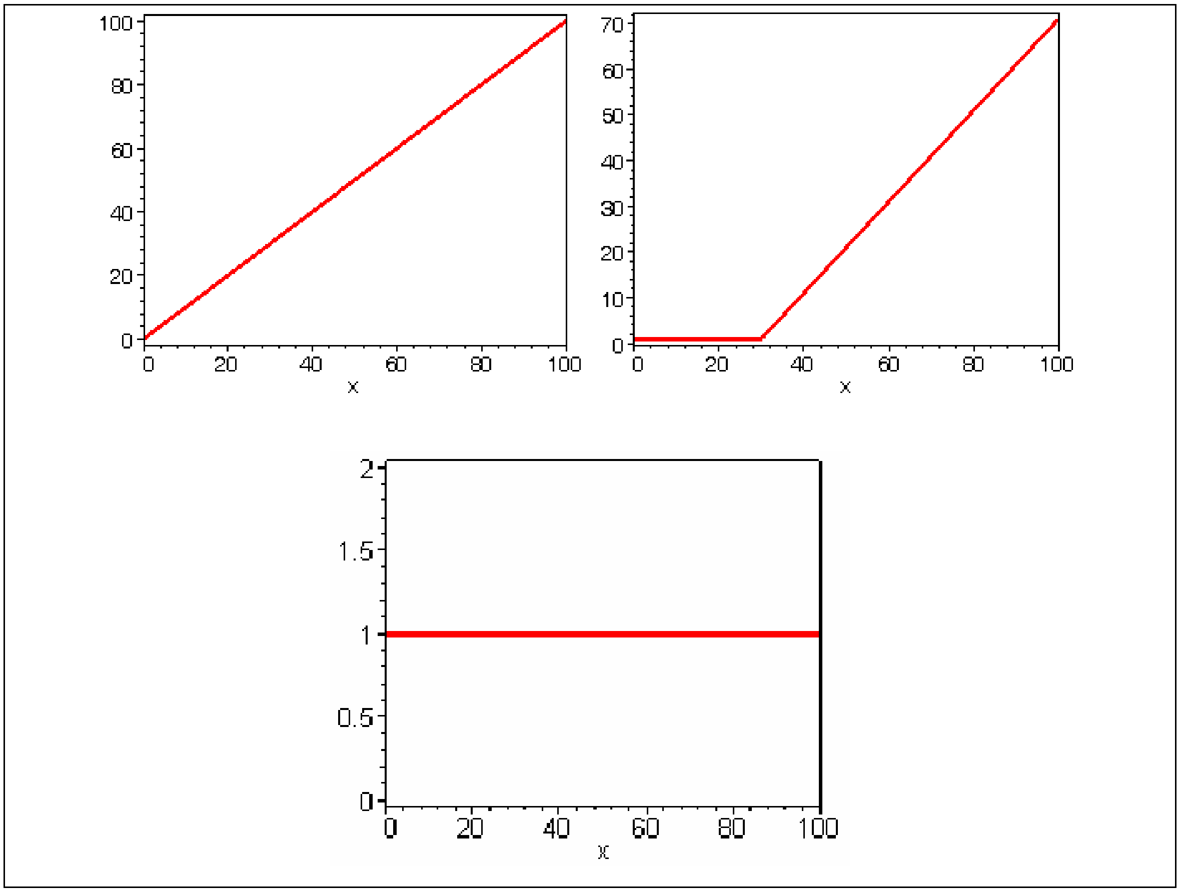


Figure 4.2 Graphs for Local expected value function of Exponential Distribution for the different values of $\varepsilon = 0.1, 30, 100$ and $\lambda = 0.1$, $0 < x < 100$

If X is a continuous random variable having an exponential distribution on $[0,\infty)$, then for all x,

$$\lim_{\varepsilon \to 0} E(E_{X,\varepsilon}(x)) = E(X) = \frac{1}{\lambda} \tag{4.10}$$

$E(\,E_{X,\varepsilon}(X)\,)=$

$$\int_0^{\varepsilon}\frac{-(x+\varepsilon)e^{-\lambda(x+\varepsilon)}\lambda e^{-\lambda x}}{1-e^{-\lambda(x+\varepsilon)}}dx-\int_{\varepsilon}^{\infty}\frac{(x+\varepsilon)e^{-\lambda(x+\varepsilon)}\lambda e^{-\lambda x}}{e^{-\lambda(x-\varepsilon)}-e^{-\lambda(x+\varepsilon)}}dx+\int_{\varepsilon}^{\infty}\frac{(x-\varepsilon)e^{-\lambda(x-\varepsilon)}\lambda e^{-\lambda x}}{e^{-\lambda(x-\varepsilon)}-e^{-\lambda(x+\varepsilon)}}dx$$

$$-\int_0^{\varepsilon}\frac{e^{-\lambda(x+\varepsilon)}e^{-\lambda x}}{1-e^{-\lambda(x+\varepsilon)}}dx-\int_{\varepsilon}^{\infty}\frac{e^{-\lambda(x+\varepsilon)}e^{-\lambda x}}{e^{-\lambda(x-\varepsilon)}-e^{-\lambda(x+\varepsilon)}}dx+\int_0^{\varepsilon}\frac{e^{-\lambda x}}{1-e^{-\lambda(x+\varepsilon)}}dx+\int_{\varepsilon}^{\infty}\frac{e^{-\lambda(x-\varepsilon)}e^{-\lambda x}}{e^{-\lambda(x-\varepsilon)}-e^{-\lambda(x+\varepsilon)}}dx$$

$$=\frac{e^{\lambda\varepsilon}}{\lambda}(-di\log(1-u),u=1-e^{-\lambda\varepsilon}\ldots 1-e^{-2\lambda\varepsilon})-\frac{2\varepsilon}{e^{\lambda\varepsilon}-e^{-\lambda\varepsilon}}+(\varepsilon+\frac{2}{\lambda})$$

$$=\frac{e^{\lambda\varepsilon}}{\lambda}(e^{-2\lambda\varepsilon}-e^{-\lambda\varepsilon})-\frac{2\varepsilon}{e^{\lambda\varepsilon}-e^{-\lambda\varepsilon}}+\varepsilon+\frac{2}{\lambda}$$

$$=\frac{e^{-\lambda\varepsilon}}{\lambda}+\frac{1}{\lambda}+\varepsilon-\frac{2\varepsilon}{e^{\lambda\varepsilon}-e^{-\lambda\varepsilon}}=\frac{1-e^{-2\lambda\varepsilon}-2\lambda\varepsilon}{e^{\lambda\varepsilon}(1-e^{-2\lambda\varepsilon})}+\frac{1}{\lambda}+\varepsilon$$

$$\cong\frac{1}{\lambda}+\varepsilon$$

where

$$-\,di\log(1-u),u=1-e^{-\lambda\varepsilon}\ldots 1-e^{2\lambda\varepsilon}=\int_{1-e^{-\lambda\varepsilon}}^{1-e^{-2\lambda\varepsilon}}\frac{\ln(1-u)}{u}du\cong\frac{e^{\lambda\varepsilon}}{\lambda}(e^{-2\lambda\varepsilon}-e^{-\lambda\varepsilon})$$

4.2.3 Cauchy Distribution

Consider a random variable having Cauchy distribution, given by the p.d.f.

$$f(x)=\begin{cases}\dfrac{1}{\pi\left(1+(x-\theta)^2\right)}, & -\infty<x<\infty \quad -\infty<\theta<\infty,\\ 0, & \textit{otherwise}\end{cases}$$

The local expected value can be obtained as

$$E_{X,\varepsilon}(x)=\frac{1}{F(x+\varepsilon)-F(x-\varepsilon)}\int_{x-\varepsilon}^{x+\varepsilon} u\,\frac{1}{\pi}\,\frac{1}{1+(u-\theta)^2}\,du$$

$$=\frac{\ln\left(\dfrac{1+(x+\varepsilon-\theta)^2}{1+(x-\varepsilon-\theta)^2}\right)^{1/2}}{\arctan(x+\varepsilon-\theta)-\arctan(x+\varepsilon-\theta)}+\theta$$

The local expected value is

$$E_{X,\varepsilon}(x)=\frac{\ln(\dfrac{1+(x+\varepsilon-\theta)^2}{1+(x-\varepsilon-\theta)^2})^{1/2}}{\arctan(x+\varepsilon-\theta)-\arctan(x-\varepsilon-\theta)}+\theta \qquad (4.11)$$

Let $\theta = 0$, then

$$E_{X,\varepsilon}(x) = \frac{\ln\left(\frac{1+(x+\varepsilon)^2}{1+(x-\varepsilon)^2}\right)^{1/2}}{\arctan(x+\varepsilon)-\arctan(x+\varepsilon)} = \frac{\ln\left(\frac{1+(x+\varepsilon)^2}{1+(x-\varepsilon)^2}\right)^{1/2}}{\arctan\left(\frac{\varepsilon}{1+x^2-\varepsilon^2}\right)} = \frac{\frac{1}{2}\ln\left(1+\frac{4x\varepsilon}{1+(x-\varepsilon)^2}\right)}{\arctan\left(\frac{2\varepsilon}{1+x^2-\varepsilon^2}\right)}$$

that is,

$$E_{X,\varepsilon}(x) = \frac{\frac{1}{2}\ln(1+\frac{4x\varepsilon}{1+(x-\varepsilon)^2})}{\arctan(\frac{2\varepsilon}{1+x^2-\varepsilon^2})} \tag{4.12}$$

In Figure 4.3, we give graphs of local expected value for Cauchy distribution for different values of ε. It can be observed when ε goes to infinity, $E_{X,\varepsilon}(x)$ also goes to infinity as it is expected since the distribution doesn't have finite expected value.

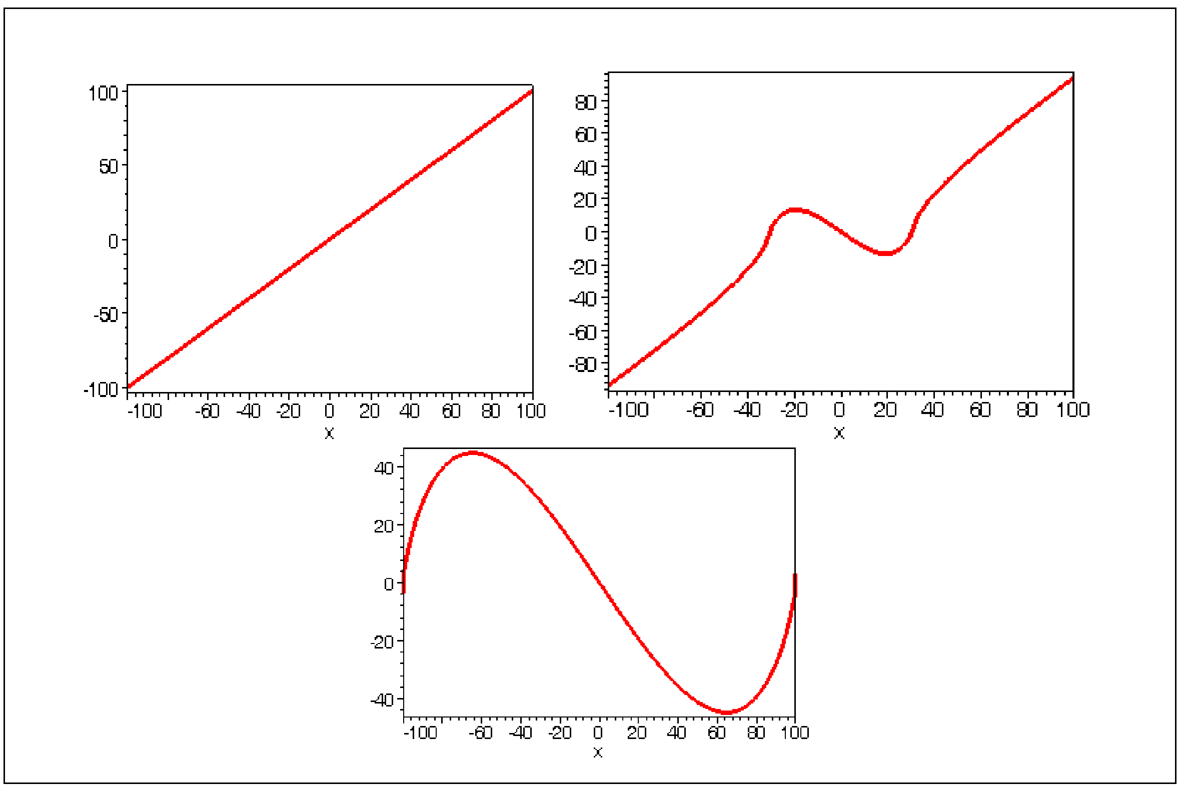


Figure 4.3 Graphs for Local expected value function of Cauchy Distribution for the different values of ε = 0.1, 30, 100,   -100 < x < 100

4.2.4 Two-Sided Power Distribution

Van Dorp & Kotz (2002b) introduced a two sided power distribution with the probability density function

$$f(x) = \begin{cases} n\left(\dfrac{x}{\theta}\right)^{n-1}, & if \quad 0 < x \le \theta \\ n\left(\dfrac{1-x}{1-\theta}\right)^{n-1}, & if \quad \theta < x \le 1 \end{cases}$$

where $n > 0$ and $0 \le \theta \le 1$.

For $0 < x \le \theta$

$$E_{X,\varepsilon}(x) = \frac{\displaystyle\int_{\max(0,x-\varepsilon)}^{\min(\theta,x+\varepsilon)} un\left(\frac{u}{\theta}\right)^{n-1} du}{F(\min(\theta, x+\varepsilon)) - F(\max(0, x-\varepsilon))}$$

$$= \frac{\dfrac{n}{\theta^{n-1}} \displaystyle\int_{\max(0,x-\varepsilon)}^{\min(\theta,x+\varepsilon)} u^n du}{\dfrac{(\min(\theta, x+\varepsilon))^n}{\theta^{n-1}} - \dfrac{(\max(0, x-\varepsilon))^n}{\theta^{n-1}}}$$

$$= \frac{\dfrac{n}{n+1}\left[(\min(\theta, x+\varepsilon))^{n+1} - (\max(0, x-\varepsilon))^{n+1}\right]}{(\min(\theta, x+\varepsilon))^n - (\max(0, x-\varepsilon))^n}$$

For $\theta < x \le 1$

$$E_{X,\varepsilon}(x) = \frac{\displaystyle\int_{\max(\theta,x-\varepsilon)}^{\min(1,x+\varepsilon)} un\left(\frac{1-u}{1-\theta}\right)^{n-1} du}{F(\min(1,x+\varepsilon)) - F(\max(\theta,x-\varepsilon))}$$

$$= \frac{\dfrac{n}{(1-\theta)^{n-1}} \displaystyle\int_{\max(\theta,x-\varepsilon)}^{\min(1,x+\varepsilon)} u(1-u)^{n-1} du}{(1-\theta)\left[\left(\dfrac{1-\max(\theta,x-\varepsilon)}{1-\theta}\right)^n - \left(\dfrac{1-\min(1,x+\varepsilon)}{1-\theta}\right)^n\right]}$$

$$= \frac{\max(\theta,x-\varepsilon)(1-\max(\theta,x-\varepsilon))^n - \min(1,x+\varepsilon)(1-\min(1,x+\varepsilon))^n}{(1-\max(\theta,x-\varepsilon))^n - (1-\min(1,x+\varepsilon))^n}$$

$$+ \frac{1}{n+1}\frac{(1-\max(\theta,x-\varepsilon))^{n+1} - (1-\min(1,x+\varepsilon))^{n+1}}{(1-\max(\theta,x-\varepsilon))^n - (1-\min(1,x+\varepsilon))^n}$$

$$E_{X,\varepsilon}(x) = \begin{cases} \dfrac{n}{n+1}\dfrac{\left[(\min(\theta,x+\varepsilon))^{n+1} - (\max(0,x-\varepsilon))^{n+1}\right]}{\left[(\min(\theta,x+\varepsilon))^{n} - (\max(0,x-\varepsilon))^{n}\right]}, & if \quad 0 < x \le \theta \\ \dfrac{\max(\theta,x-\varepsilon)(1-\max(\theta,x-\varepsilon))^n - \min(1,x+\varepsilon)(1-\min(1,x+\varepsilon))^n}{(1-\max(\theta,x-\varepsilon))^n - (1-\min(1,x+\varepsilon))^n} & if \quad \theta < x \le 1 \\ \quad + \dfrac{1}{n+1}\dfrac{(1-\max(\theta,x-\varepsilon))^{n+1} - (1-\min(1,x+\varepsilon))^{n+1}}{(1-\max(\theta,x-\varepsilon))^n - (1-\min(1,x+\varepsilon))^n} & \end{cases} \tag{4.13}$$

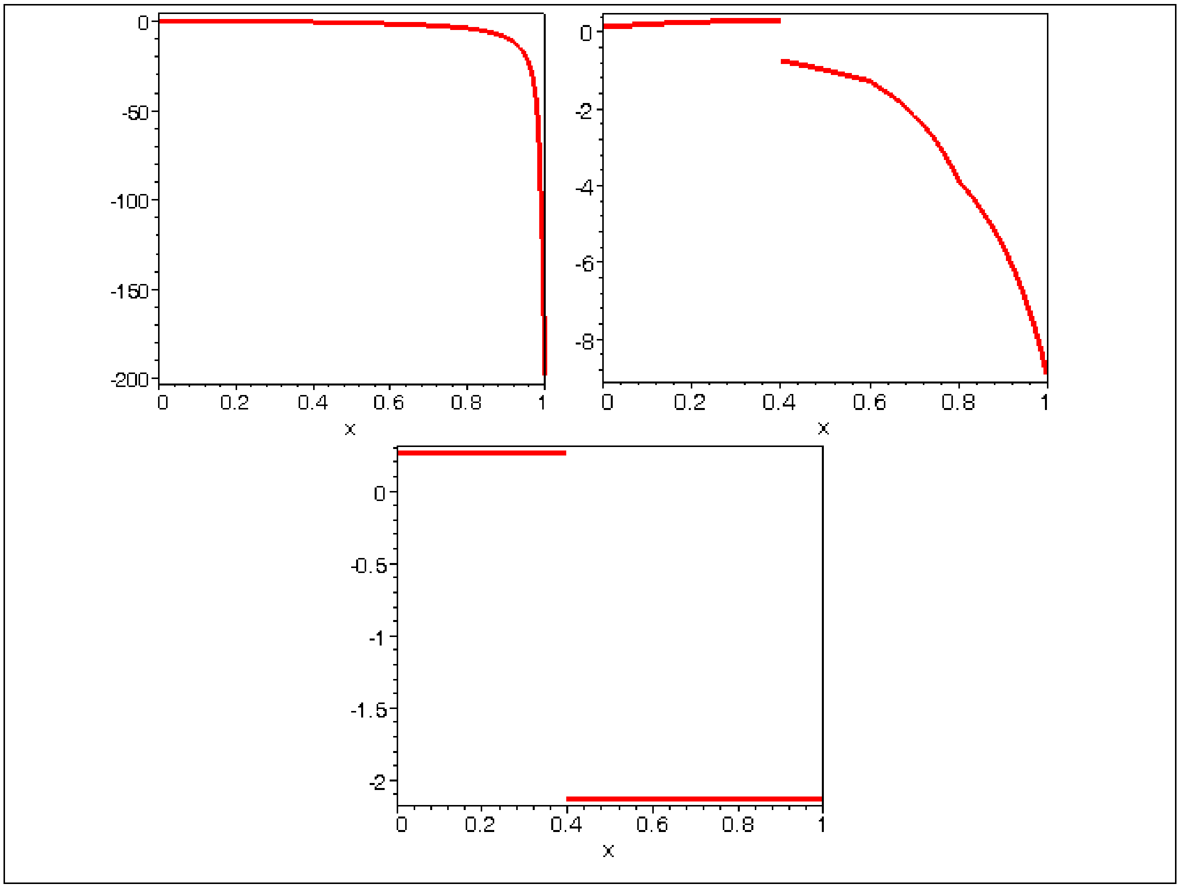


Figure 4.4 Graphs for Local expected value function of Two-sided Power Distribution for the different values of ε =0.01, 0.2, 1, θ=0.4

4.2.5 Pareto Distribution

Consider the random variable having Pareto distribution whose p.d.f is

$$f(x)=\begin{cases}\dfrac{1}{x^2}, & x\geq 1\\ 0, & otherwise.\end{cases}$$

$$E_{X,\varepsilon}(x) = \frac{\int\limits_{\max(1,x-\varepsilon)}^{x+\varepsilon} u \ \frac{1}{u^2} du}{\int\limits_{\max(1,x-\varepsilon)}^{x+\varepsilon} \frac{1}{u^2} dy}$$

$$= \frac{\ln(x+\varepsilon) - \ln(\max(1,x-\varepsilon))}{\frac{1}{\max(1,x-\varepsilon)} - \frac{1}{x+\varepsilon}}$$

$$= \frac{(x+\varepsilon)\max(1,x-\varepsilon)\ln\left(\frac{x+\varepsilon}{\max(1,x-\varepsilon)}\right)}{(x+\varepsilon) - \max(1,x-\varepsilon)}$$

The local expected value of this random variable is

$$E_{X,\varepsilon}(x) = \frac{\ln\left(\frac{x+\varepsilon}{\max(1,x-\varepsilon)}\right)^{(x+\varepsilon)\max(1,x-\varepsilon)}}{(x+\varepsilon) - \max(1,x-\varepsilon)} \tag{4.14}$$

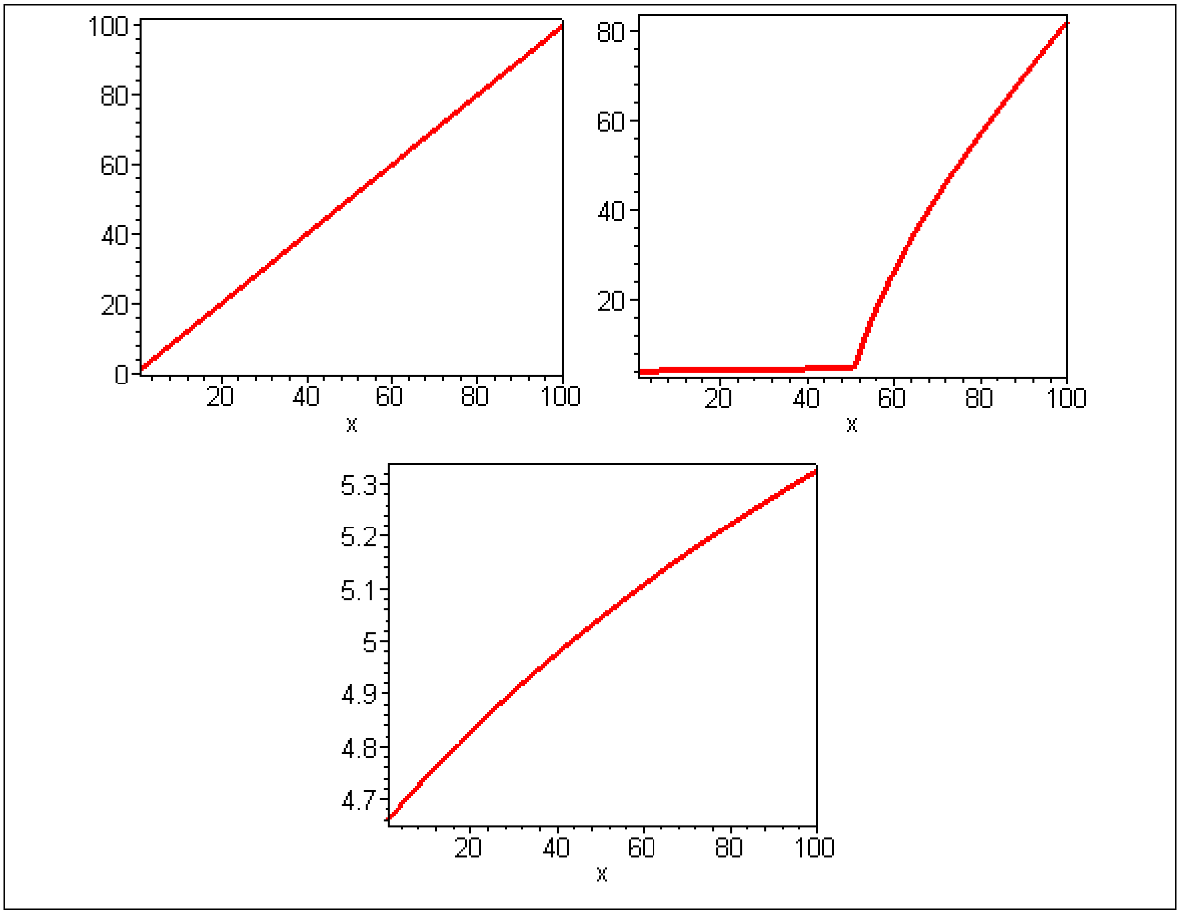


Figure 4.5 Graphs for Local expected value function of Pareto Distribution for the different values of $\varepsilon = 0.1, 50{,}100$.

In Figure 4.5, we give graphs of local expected value for Pareto distribution for different values of ε. As in the Cauchy distribution, Pareto distribution also does not have finite expected value; we see again when ε goes to infinity, $E_{X,\varepsilon}(x)$ also goes to infinity .

4.3 Local Dependence Function

Let X and Y be continuous random variables with distribution function $F_{X,Y}(x,y)$ and marginals $F_X(x)$ and $F_Y(y)$, respectively with support $N_{X,Y}$. The function $\rho_{X,Y,\varepsilon_1,\varepsilon_2}(x,y)$ is a local dependence function of X and Y at the $(\varepsilon_1,\varepsilon_2)$ neighbourhood of the point $(x,y) \in N_{X,Y.}$ The localized version of Hoeffding's formula (Block & Fang, 1988) can be written as

$$\rho_{X,Y,\varepsilon_1,\varepsilon_2}(x,y) = \frac{\int_{y-\varepsilon_2}^{y+\varepsilon_2}\int_{x-\varepsilon_1}^{x+\varepsilon_1}[F_{X,Y}(u,v) - F_X(u)F_Y(v)]dudv}{\sqrt{Var(X)}\sqrt{Var(Y)}} \tag{4.15}$$

We call this function the $\varepsilon-$ local dependence function.

If the support of X and Y is a finite interval $[a,b],$ then it is clear that

$$\rho_{X,Y,\varepsilon_1,\varepsilon_2}(x,y) = \frac{\int_{\max(a,y-\varepsilon_2)}^{\min(b,y+\varepsilon_2)}\int_{\max(a,x-\varepsilon_1)}^{\min(b,x+\varepsilon_1)}[F_{X,Y}(u,v) - F_X(u)F_Y(v)]dudv}{\sqrt{Var(X)}\sqrt{Var(Y)}} \tag{4.16}$$

4.3.1 Properties of Local Dependence Function

1) If X and Y are independent, then $\rho_{X,Y,\varepsilon_1,\varepsilon_2}(x,y) = 0$ for all $(x,y) \in N_{X,Y}$.

Proof. If X and Y are independent then the bivariate distribution function is $F_{XY}(u,v) = F_X(u)F_Y(v)$.

So,

$$\rho_{X,Y,\varepsilon_1,\varepsilon_2}(x,y) = \frac{\int_{y-\varepsilon_2}^{y+\varepsilon_2}\int_{x-\varepsilon_1}^{x+\varepsilon_1}[F_X(u)F_Y(v) - F_X(u)F_Y(v)]dudv}{\sqrt{Var(X)}\sqrt{Var(Y)}} = 0$$

2) $\left|\rho_{X,Y,\varepsilon_1,\varepsilon_2}(x,y)\right| \le 1$ for all $(x,y) \in N_{X,Y}$.

Proof. If the support of X and Y is a finite interval $[a,b]$,

$$\left|\rho(x,y)\right| = \left|\frac{\int_a^b\int_a^b[F_{XY}(u,v) - F_X(u)F_Y(v)]dudv}{\sqrt{Var(X)}\sqrt{Var(Y)}}\right| \le 1 \text{ so,}$$

$$\left|\rho_{X,Y,\varepsilon_1,\varepsilon_2}(x,y)\right| = \left|\frac{\int_{y-\varepsilon_2}^{y+\varepsilon_2}\int_{x-\varepsilon_1}^{x+\varepsilon_1}[F_{XY}(u,v) - F_X(u)F_Y(v)]dudv}{\sqrt{Var(X)}\sqrt{Var(Y)}}\right| \le 1.$$

3) $\rho_{X,Y,\varepsilon_1,\varepsilon_2}(x,y) = \rho_{X,Y,\varepsilon_1,\varepsilon_2}(y,x)$ under the condition that $F_{XY}(x,y)$ is symmetric.

Proof. $$\rho_{X,Y,\varepsilon_1,\varepsilon_2}(x,y) = \frac{\int_{y-\varepsilon_2}^{y+\varepsilon_2}\int_{x-\varepsilon_1}^{x+\varepsilon_1}[F_{XY}(u,v) - F_X(u)F_Y(v)]dudv}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

and

$$\rho_{X,Y,\varepsilon_1,\varepsilon_2}(y,x) = \frac{\int\limits_{x-\varepsilon_1}^{x+\varepsilon_1}\int\limits_{y-\varepsilon_2}^{y+\varepsilon_2}[F_{XY}(v,u) - F_X(v)F_Y(u)]dvdu}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

then $\rho_{X,Y,\varepsilon_1,\varepsilon_2}(x,y) = \rho_{X,Y,\varepsilon_1,\varepsilon_2}(y,x)$ under the condition that $F_{XY}(x,y)$ is symmetric.

4) If $\widetilde{X} = aX + b$ and $\widetilde{Y} = cY + d$, then $Cov_{\widetilde{X},\widetilde{Y},\varepsilon_1,\varepsilon_2}(\widetilde{x},\widetilde{y}) = acCov_{X,Y,\frac{\varepsilon_1}{a},\frac{\varepsilon_2}{c}}(x,y)$. So the local dependence function,

$$\rho_{\widetilde{X},\widetilde{Y},\varepsilon_1,\varepsilon_2}(\widetilde{x},\widetilde{y}) = \frac{Cov_{X,Y,\frac{\varepsilon_1}{a},\frac{\varepsilon_2}{c}}(x,y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

where $\widetilde{x} = ax + b$ and $\widetilde{y} = cy + d$.

Proof.

$$\begin{aligned} F_{\widetilde{X}\widetilde{Y}}(u,v) &= P\{\widetilde{X} \le u, \widetilde{Y} \le v\} = P\{aX + b \le u, cY + d \le v\} \\ &= P\left\{X \le \frac{u-b}{a}, Y \le \frac{v-d}{c}\right\} \\ &= F_{X,Y}\left(\frac{u-b}{a}, \frac{v-d}{c}\right) \end{aligned}$$

$$\begin{aligned} F_{\widetilde{X}}(u) &= P\{\widetilde{X} \le u\} = P\{aX + b \le u\} \\ &= P\left\{X \le \frac{u-b}{a}\right\} \\ &= F_X\left(\frac{u-b}{a}\right) \end{aligned}$$

$$F_{\tilde{Y}}(v) = P\{\tilde{Y} \le v\} = P\{cY + d \le v\}$$
$$= P\left\{Y \le \frac{v-d}{c}\right\}$$
$$= F_Y\left(\frac{v-d}{c}\right)$$

$$Cov_{\tilde{X},\tilde{Y},\varepsilon_1,\varepsilon_2}(x,y) = \int_{y-\varepsilon_2}^{y+\varepsilon_2}\int_{x-\varepsilon_1}^{x+\varepsilon_1}\left[F_{XY}\left(\frac{u-b}{a},\frac{v-d}{c}\right) - F_X\left(\frac{u-b}{a}\right)F_Y\left(\frac{v-d}{c}\right)\right]dudv$$

By using transformation technique,

$\dfrac{u-b}{a} = z$ and $\dfrac{v-d}{c} = t$

we have

$$Cov_{\tilde{X},\tilde{Y},\varepsilon_1,\varepsilon_2}(x,y) = ac\int_{\frac{y-\varepsilon_2-d}{c}}^{\frac{y+\varepsilon_2-d}{c}}\int_{\frac{x-\varepsilon_1-b}{a}}^{\frac{x+\varepsilon_1-b}{a}}\left[F_{XY}(z,t) - F_X(z)F_Y(t)\right]dzdt$$
$$= acCov_{X,Y,\frac{\varepsilon_1}{a},\frac{\varepsilon_2}{c}}\left(\frac{x-b}{a},\frac{y-d}{c}\right)$$

and

$$Cov_{\tilde{X},\tilde{Y},\varepsilon_1,\varepsilon_2}(\tilde{x},\tilde{y}) = acCov_{X,Y,\frac{\varepsilon_1}{a},\frac{\varepsilon_2}{c}}(x,y)$$

so,

$$\rho_{\widetilde{X},\widetilde{Y},\varepsilon_1,\varepsilon_2}(\widetilde{x},\widetilde{y}) = \frac{acCov_{X,Y,\frac{\varepsilon_1}{a},\frac{\varepsilon_2}{c}}(x,y)}{ac\sqrt{Var(X)}\sqrt{Var(Y)}}$$
$$= \rho_{X,Y,\frac{\varepsilon_1}{a},\frac{\varepsilon_2}{c}}(x,y).$$

5) When $\widetilde{Y} = aX + b$ then,

$$Cov_{X,\widetilde{Y},\varepsilon_1,\varepsilon_2}(x,\widetilde{y}) = aCov_{X,X,\varepsilon_1,\frac{\varepsilon_2}{a}}(x,x)$$

and

$$\rho_{X,\widetilde{Y},\varepsilon_1,\varepsilon_2}(x,\widetilde{y}) = \frac{Cov_{X,X,\varepsilon_1,\frac{\varepsilon_2}{a}}(x,x)}{Var(X)}$$

where $\widetilde{y} = ax + b$.

Proof

$$F_{X\widetilde{Y}}(u,v) = P\{X \le u, \widetilde{Y} \le v\} = P\{X \le u, aX + b \le v\}$$
$$= P\left\{X \le u, X \le \frac{v-b}{a}\right\}$$
$$= F_{XX}\left(u, \frac{v-b}{a}\right)$$

$$F_{\widetilde{Y}}(v) = P\{\widetilde{Y} \le v\} = P\{aX + b \le v\}$$
$$= P\left\{X \le \frac{v-b}{a}\right\}$$
$$= F_X\left(\frac{v-b}{a}\right)$$

$$Cov_{X,\widetilde{Y},\varepsilon_1,\varepsilon_2}(x,y) = \int_{y-\varepsilon_2}^{y+\varepsilon_2}\int_{x-\varepsilon_1}^{x+\varepsilon_1}\left[F_{XX}\left(u,\frac{v-b}{a}\right) - F_X(u)F_X\left(\frac{v-b}{a}\right)\right]dudv$$

By using transformation technique,

$$\frac{v-b}{a} = z$$

$$Cov_{X,\widetilde{Y},\varepsilon_1,\varepsilon_2}(x,y) = a\int_{\frac{y-\varepsilon_2-b}{a}}^{\frac{y+\varepsilon_2-b}{a}}\int_{x-\varepsilon_1}^{x+\varepsilon_1}\left[F_{XX}(u,z) - F_X(u)F_X(z)\right]dudz$$
$$= aCov_{X,X,\varepsilon_1,\frac{\varepsilon_2}{a}}\left(x,\frac{y-b}{a}\right)$$

and

$$Cov_{X,\widetilde{Y},\varepsilon_1,\varepsilon_2}(\widetilde{x},\widetilde{y}) = aCov_{X,X,\varepsilon_1,\frac{\varepsilon_2}{a}}(x,x)$$

so,

$$\rho_{X,\widetilde{Y},\varepsilon_1,\varepsilon_2}(\widetilde{x},\widetilde{y}) = \frac{aCov_{X,X,\varepsilon_1,\frac{\varepsilon_2}{a}}(x,x)}{aVar(X)}$$
$$= \rho_{X,X,\varepsilon_1,\frac{\varepsilon_2}{a}}(x,x)$$

4.3.2 Examples of Some Special Bivariate Distributions

4.3.2.1 Farlie-Gumbel-Morgenstern Copula

Consider Farlie-Gumbel-Morgenstern distributions with uniform marginals (Bairamov & Kotz, 2002). The p.d.f. is

$$f_\alpha(x,y) = 1 + \alpha(1-2x)(1-2y), \qquad 0 \le x \le 1,\ 0 \le y \le 1, \qquad -1 \le \alpha \le 1$$

Let $\varepsilon_1 = \varepsilon_2 = \varepsilon,$ then the local covariance function is

$$\begin{aligned} Cov_{X,Y,\varepsilon,\varepsilon}(x,y) &= \int_{\max(0,y-\varepsilon)}^{\min(1,y+\varepsilon)} \int_{\max(0,x-\varepsilon)}^{\min(1,x+\varepsilon)} \alpha uv(1-u)(1-v)\,dudv \\ &= \alpha\left[\left(\frac{c_2^2}{2}-\frac{c_2^3}{3}\right)-\left(\frac{c_1^2}{2}-\frac{c_1^3}{3}\right)\right]\left[\left(\frac{k_2^2}{2}-\frac{k_2^3}{3}\right)-\left(\frac{k_1^2}{2}-\frac{k_1^3}{3}\right)\right] \end{aligned} \tag{4.17}$$

where

$$c_2 = \min(1, x+\varepsilon) = \begin{cases} 1, & x > 1-\varepsilon \\ x+\varepsilon, & x < 1-\varepsilon \end{cases}$$

$$k_2 = \min(1, y+\varepsilon) = \begin{cases} 1, & y > 1-\varepsilon \\ y+\varepsilon, & y < 1-\varepsilon \end{cases}$$

$$c_1 = \max(0, x-\varepsilon) = \begin{cases} x-\varepsilon, & x > \varepsilon \\ 0, & x < \varepsilon \end{cases}$$

$$k_1 = \max(0, y-\varepsilon) = \begin{cases} y-\varepsilon, & y > \varepsilon \\ 0, & y < \varepsilon \end{cases}$$

$Cov_{X,Y,\varepsilon,\varepsilon}(x,y) =$

$$
\begin{cases}
\alpha\left[\left(\frac{(x+\varepsilon)^2}{2}-\frac{(x+\varepsilon)^3}{3}\right)\right]\left[\left(\frac{(y+\varepsilon)^2}{2}-\frac{(y+\varepsilon)^3}{3}\right)\right], & 0<x<\varepsilon \quad 0<y<\varepsilon \\
\alpha\left[\left(\frac{(x+\varepsilon)^2}{2}-\frac{(x+\varepsilon)^3}{3}\right)-\left(\frac{(x-\varepsilon)^2}{2}-\frac{(x-\varepsilon)^3}{3}\right)\right]\left[\left(\frac{(y+\varepsilon)^2}{2}-\frac{(y+\varepsilon)^3}{3}\right)\right], & \varepsilon<x<1-\varepsilon \quad 0<y<\varepsilon \\
\alpha\left[\frac{1}{6}-\left(\frac{(x-\varepsilon)^2}{2}-\frac{(x-\varepsilon)^3}{3}\right)\right]\left[\left(\frac{(y+\varepsilon)^2}{2}-\frac{(y+\varepsilon)^3}{3}\right)\right], & 1-\varepsilon<x<1 \quad 0<y<\varepsilon \\
\alpha\left[\left(\frac{(x+\varepsilon)^2}{2}-\frac{(x+\varepsilon)^3}{3}\right)\right]\left[\left(\frac{(y+\varepsilon)^2}{2}-\frac{(y+\varepsilon)^3}{3}\right)-\left(\frac{(y-\varepsilon)^2}{2}-\frac{(y-\varepsilon)^3}{3}\right)\right], & 0<x<\varepsilon \quad \varepsilon<y<1-\varepsilon \\
\alpha\left[\left(\frac{(x+\varepsilon)^2}{2}-\frac{(x+\varepsilon)^3}{3}\right)-\left(\frac{(x-\varepsilon)^2}{2}-\frac{(x-\varepsilon)^3}{3}\right)\right]\left[\left(\frac{(y+\varepsilon)^2}{2}-\frac{(y+\varepsilon)^3}{3}\right)-\left(\frac{(y-\varepsilon)^2}{2}-\frac{(y-\varepsilon)^3}{3}\right)\right], & \varepsilon<x<1-\varepsilon \quad \varepsilon<y<1-\varepsilon \\
\alpha\left[\frac{1}{6}-\left(\frac{(x-\varepsilon)^2}{2}-\frac{(x-\varepsilon)^3}{3}\right)\right]\left[\left(\frac{(y+\varepsilon)^2}{2}-\frac{(y+\varepsilon)^3}{3}\right)-\left(\frac{(y-\varepsilon)^2}{2}-\frac{(y-\varepsilon)^3}{3}\right)\right], & 1-\varepsilon<x<1 \quad \varepsilon<y<1-\varepsilon \\
\alpha\left[\left(\frac{(x+\varepsilon)^2}{2}-\frac{(x+\varepsilon)^3}{3}\right)\right]\left[\frac{1}{6}-\left(\frac{(y-\varepsilon)^2}{2}-\frac{(y-\varepsilon)^3}{3}\right)\right], & 0<x<\varepsilon \quad 1-\varepsilon<y<1 \\
\alpha\left[\left(\frac{(x+\varepsilon)^2}{2}-\frac{(x+\varepsilon)^3}{3}\right)-\left(\frac{(x-\varepsilon)^2}{2}-\frac{(x-\varepsilon)^3}{3}\right)\right]\left[\frac{1}{6}-\left(\frac{(y-\varepsilon)^2}{2}-\frac{(y-\varepsilon)^3}{3}\right)\right], & \varepsilon<x<1-\varepsilon \quad 1-\varepsilon<y<1 \\
\alpha\left[\frac{1}{6}-\left(\frac{(x-\varepsilon)^2}{2}-\frac{(x-\varepsilon)^3}{3}\right)\right]\left[\frac{1}{6}-\left(\frac{(y-\varepsilon)^2}{2}-\frac{(y-\varepsilon)^3}{3}\right)\right], & 1-\varepsilon<x<1 \quad 1-\varepsilon<y<1
\end{cases}
$$

In Figure 4.6, we have presented graphs of local dependence function for FGM distributions with different values of ε. It can be observed that $\lim_{\varepsilon\to 1} Cov_{X,Y,\varepsilon,\varepsilon}(x,y) = \frac{\alpha}{3}$, which is the correlation coefficient of FGM distribution.

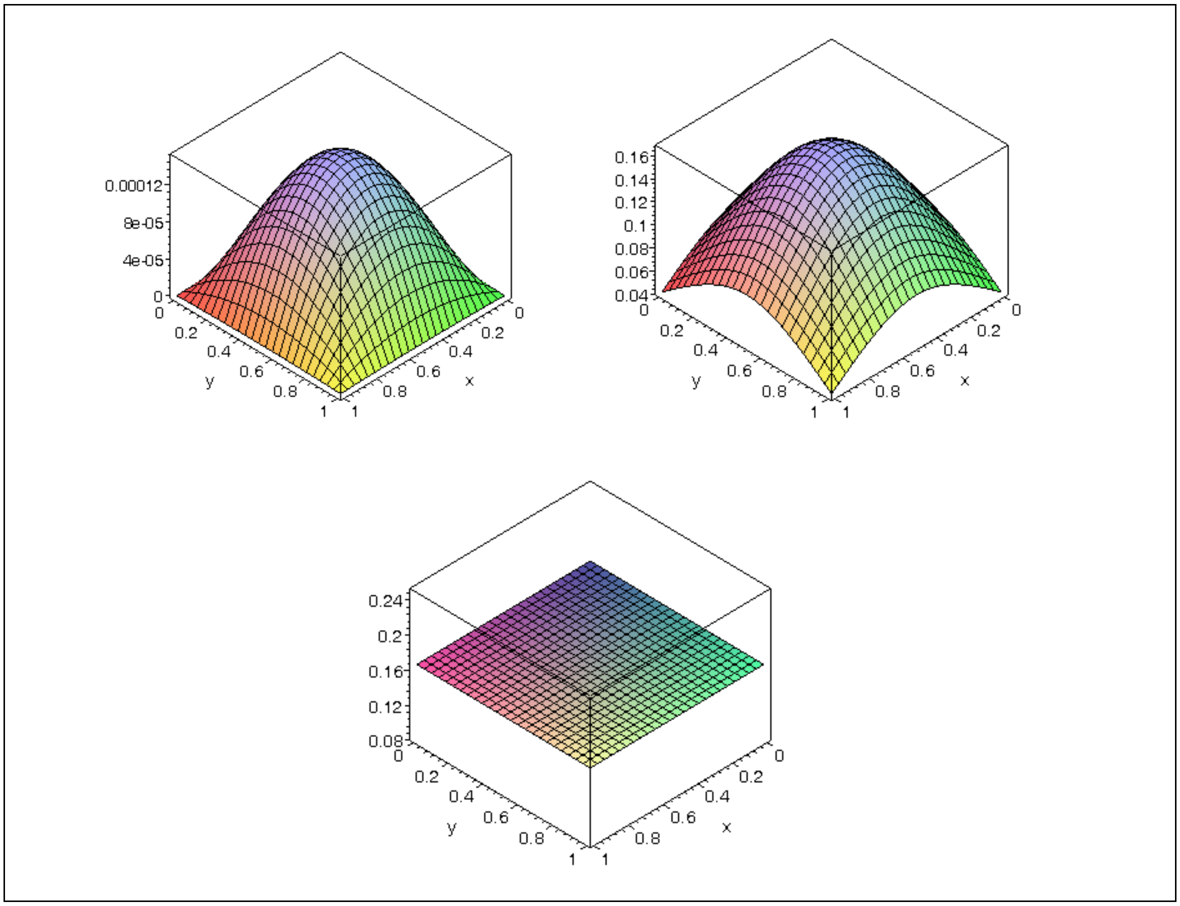


Figure 4.6 Graphs for Local dependence function of FGM Distribution for the different values of $\varepsilon = 0.01, 0.5, 1$, $\alpha = 0.5$, Var$X$=Var$Y$= $\frac{1}{12}$

4.3.2.2 Sarmanov-Lee Copula

Consider the Sarmanov-Lee bivariate density with uniform marginals introduced by Lee (1996). The p.d.f is

$$f_\alpha(x,y) = 1 + \alpha\left(x - \frac{1}{2}\right)\left(y - \frac{1}{2}\right) \qquad 0 \le x \le 1,\ 0 \le y \le 1, \qquad -4 \le \alpha \le 4$$

Let $\varepsilon_1 = \varepsilon_2 = \varepsilon,$ then the local covariance function is

$$
\begin{aligned}
Cov_{X,Y,\varepsilon,\varepsilon}(x,y) &= \int_{\max(0,y-\varepsilon)}^{\min(1,y+\varepsilon)} \int_{\max(0,x-\varepsilon)}^{\min(1,x+\varepsilon)} \alpha\left(\frac{u^2}{2}-\frac{u}{2}\right)\left(\frac{v^2}{2}-\frac{v}{2}\right)dudv \\
&= \alpha\left[\left(\frac{c_2^3}{6}-\frac{c_2^2}{4}\right)-\left(\frac{c_1^3}{6}-\frac{c_1^2}{4}\right)\right]\left[\left(\frac{k_2^3}{6}-\frac{k_2^2}{4}\right)-\left(\frac{k_1^3}{6}-\frac{k_1^2}{4}\right)\right]
\end{aligned} \tag{4.18}
$$

$$Cov_{X,Y,\varepsilon,\varepsilon}(x,y) =$$

$$
\begin{cases}
\alpha\left[\left(\frac{(x+\varepsilon)^3}{6}-\frac{(x+\varepsilon)^2}{4}\right)\right]\left[\left(\frac{(y+\varepsilon)^3}{6}-\frac{(y+\varepsilon)^2}{4}\right)\right], & 0<x<\varepsilon \quad 0<y<\varepsilon \\
\alpha\left[\left(\frac{(x+\varepsilon)^3}{6}-\frac{(x+\varepsilon)^2}{4}\right)-\left(\frac{(x-\varepsilon)^3}{6}-\frac{(x-\varepsilon)^2}{4}\right)\right]\left[\left(\frac{(y+\varepsilon)^3}{6}-\frac{(y+\varepsilon)^2}{4}\right)\right], & \varepsilon<x<1-\varepsilon \quad 0<y<\varepsilon \\
\alpha\left[-\frac{1}{12}-\left(\frac{(x-\varepsilon)^3}{6}-\frac{(x-\varepsilon)^2}{4}\right)\right]\left[\left(\frac{(y+\varepsilon)^3}{6}-\frac{(y+\varepsilon)^2}{4}\right)\right], & 1-\varepsilon<x<1 \quad 0<y<\varepsilon \\
\alpha\left[\left(\frac{(x+\varepsilon)^3}{6}-\frac{(x+\varepsilon)^2}{4}\right)\right]\left[\left(\frac{(y+\varepsilon)^3}{6}-\frac{(y+\varepsilon)^2}{4}\right)-\left(\frac{(y-\varepsilon)^3}{6}-\frac{(y-\varepsilon)^2}{4}\right)\right], & 0<x<\varepsilon \quad \varepsilon<y<1-\varepsilon \\
\alpha\left[\left(\frac{(x+\varepsilon)^3}{6}-\frac{(x+\varepsilon)^2}{4}\right)-\left(\frac{(x-\varepsilon)^3}{6}-\frac{(x-\varepsilon)^2}{4}\right)\right]\left[\left(\frac{(y+\varepsilon)^3}{6}-\frac{(y+\varepsilon)^2}{4}\right)-\left(\frac{(y-\varepsilon)^3}{6}-\frac{(y-\varepsilon)^2}{4}\right)\right], & \varepsilon<x<1-\varepsilon \quad \varepsilon<y<1-\varepsilon \\
\alpha\left[-\frac{1}{12}-\left(\frac{(x-\varepsilon)^3}{6}-\frac{(x-\varepsilon)^2}{4}\right)\right]\left[\left(\frac{(y+\varepsilon)^3}{6}-\frac{(y+\varepsilon)^2}{4}\right)-\left(\frac{(y-\varepsilon)^3}{6}-\frac{(y-\varepsilon)^2}{4}\right)\right], & 1-\varepsilon<x<1 \quad \varepsilon<y<1-\varepsilon \\
\alpha\left[\left(\frac{(x+\varepsilon)^3}{6}-\frac{(x+\varepsilon)^2}{4}\right)\right]\left[-\frac{1}{12}-\left(\frac{(y-\varepsilon)^3}{6}-\frac{(y-\varepsilon)^2}{4}\right)\right], & 0<x<\varepsilon \quad 1-\varepsilon<y<1 \\
\alpha\left[\left(\frac{(x+\varepsilon)^3}{6}-\frac{(x+\varepsilon)^2}{4}\right)-\left(\frac{(x-\varepsilon)^3}{6}-\frac{(x-\varepsilon)^2}{4}\right)\right]\left[-\frac{1}{12}-\left(\frac{(y-\varepsilon)^3}{6}-\frac{(y-\varepsilon)^2}{4}\right)\right], & \varepsilon<x<1-\varepsilon \quad 1-\varepsilon<y<1 \\
\alpha\left[-\frac{1}{12}-\left(\frac{(x-\varepsilon)^3}{6}-\frac{(x-\varepsilon)^2}{4}\right)\right]\left[-\frac{1}{12}-\left(\frac{(y-\varepsilon)^3}{6}-\frac{(y-\varepsilon)^2}{4}\right)\right], & 1-\varepsilon<x<1 \quad 1-\varepsilon<y<1
\end{cases}
$$

In Figure 4.7, for different values of ε, we presented graphs of Sarmanov-Lee distribution with uniform marginals. As in the local dependence function for FGM distribution; it is observed that $\lim_{\varepsilon \to 1} Cov_{X,Y,\varepsilon,\varepsilon}(x,y) = \frac{\alpha}{12}$, which is the correlation coefficient of Sarmanov-Lee distribution.

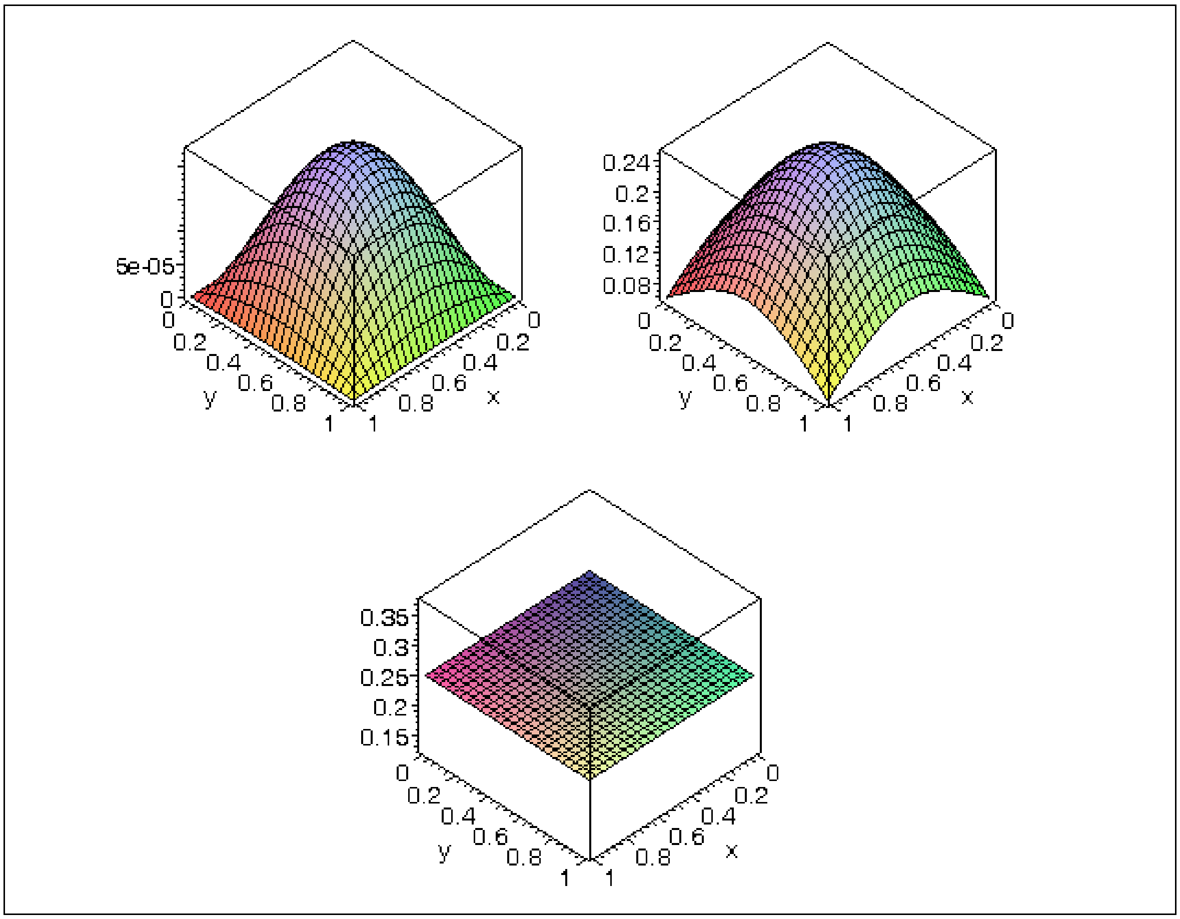


Figure 4.7 Graphs for Local dependence function of Sarmanov-Lee Distribution for the different values of $\varepsilon = 0.01, 0.5, 1$, $\alpha = 0.5$.

4.3.2.2 Gumbel's Bivariate Exponential Distribution

Let X and Y be random variables having Gumbel's Bivariate Exponential density function (Gumbel, 1960) given as

$$f(x,y)=e^{-x-y-\theta xy}\{(1+\theta x)(1+\theta y)-\theta\}, \quad x,y>0, \quad 0\le\theta\le 1$$

Let $\varepsilon_1=\varepsilon_2=\varepsilon,$ then the local covariance function is

$$\begin{aligned}
Cov_{X,Y,\varepsilon,\varepsilon}(x,y) &= \int_{\max(0,y-\varepsilon)}^{y+\varepsilon}\int_{\max(0,x-\varepsilon)}^{x+\varepsilon} e^{-(u+v)}(e^{-\theta uv}-1)dudv \\
&= \tfrac{e^{-(x+\varepsilon)}e^{1/\theta}}{\theta}\Big[e^{(y+\varepsilon)}\,\mathrm{Ei}(1,\frac{1}{\theta}(\theta(y+\varepsilon)+1)(\theta(x+\varepsilon)+1)) \\
&\quad - e^{-\max(0,y-\varepsilon)}\,\mathrm{Ei}(1,\frac{1}{\theta}(\theta\max(0,y-\varepsilon)+1)(\theta(x+\varepsilon)+1))\Big] \\
&\quad + \tfrac{e^{-\max(0,x-\varepsilon)}e^{1/\theta}}{\theta}\Big[-e^{(y+\varepsilon)}\,\mathrm{Ei}(1,\frac{1}{\theta}(\theta(y+\varepsilon)+1)(\theta\max(0,x-\varepsilon)+1)) \\
&\quad + e^{\max(0,y-\varepsilon)}e^{(y+\varepsilon)}\,\mathrm{Ei}(1,\frac{1}{\theta}(\theta\max(0,x-\varepsilon)+1)(\theta\max(0,y-\varepsilon)+1))\Big] \\
&\quad +(e^{-(y+\varepsilon)}-e^{-\max(0,y-\varepsilon)})(e^{-\max(0,x-\varepsilon)}-e^{-(x+\varepsilon)})
\end{aligned} \tag{4.19}$$

where $\mathrm{Ei}(u)=\int_u^\infty v^{-1}e^{-v}dv$ and Pearson correlation coefficient is $\rho=\theta^{-1}e^{1/\theta}\,\mathrm{Ei}(\theta^{-1})$.

$Cov_{X,Y,\varepsilon,\varepsilon}(x,y) =$

$$
\begin{cases}
\begin{aligned}
&\frac{e^{-(x+\varepsilon)}e^{1/\theta}}{\theta}\left[e^{(y+\varepsilon)}Ei(1,\frac{1}{\theta}(\theta(y+\varepsilon)+1)(\theta(x+\varepsilon)+1)-Ei(1,\frac{1}{\theta}(\theta(x+\varepsilon)+1)\right] \\
&+\frac{e^{1/\theta}}{\theta}\left[e^{-(y+\varepsilon)}Ei(1,\frac{1}{\theta}(\theta(y+\varepsilon)+1))+e^{(y+\varepsilon)}Ei(1,\frac{1}{\theta})\right] \\
&+(e^{-(y+\varepsilon)}-1)(1-e^{-(x+\varepsilon)}) \\
&x<\varepsilon \quad y<\varepsilon
\end{aligned} \\[2ex]
\begin{aligned}
&\frac{e^{-(x+\varepsilon)}e^{1/\theta}}{\theta}\left[e^{(y+\varepsilon)}Ei(1,\frac{1}{\theta}(\theta(y+\varepsilon)+1)(\theta(x+\varepsilon)+1)\right] \\
&+\frac{e^{1/\theta}}{\theta}\left[e^{-(y+\varepsilon)}Ei(1,\frac{1}{\theta}(\theta(y+\varepsilon)+1))+e^{(y-\varepsilon)}Ei(1,\frac{1}{\theta}(\theta(y-\varepsilon)+1))\right] \\
&+(e^{-(y+\varepsilon)}-e^{-(y-\varepsilon)})(1-e^{-(x+\varepsilon)}) \\
&x<\varepsilon \quad y>\varepsilon
\end{aligned} \\[2ex]
\begin{aligned}
&\frac{e^{-(x+\varepsilon)}e^{1/\theta}}{\theta}\left[e^{(y+\varepsilon)}Ei(1,\frac{1}{\theta}(\theta(y+\varepsilon)+1)(\theta(x+\varepsilon)+1))-Ei(1,\frac{1}{\theta}(\theta(x+\varepsilon)+1)\right] \\
&+\frac{e^{-(x-\varepsilon)}e^{1/\theta}}{\theta}\left[e^{-(y+\varepsilon)}Ei(1,\frac{1}{\theta}(\theta(y+\varepsilon)+1)(\theta(x-\varepsilon)+1))+Ei(1,\frac{1}{\theta}(\theta(x-\varepsilon)+1))\right] \\
&+(e^{-(x-\varepsilon)}-e^{-(x+\varepsilon)})(e^{-(y+\varepsilon)}-1) \\
&x>\varepsilon \quad y<\varepsilon
\end{aligned} \\[2ex]
\begin{aligned}
&\frac{e^{-(x+\varepsilon)}e^{1/\theta}}{\theta}\left[e^{(y+\varepsilon)}Ei(1,\frac{1}{\theta}(\theta(y+\varepsilon)+1)(\theta(x+\varepsilon)+1))-e^{-(y-\varepsilon)}Ei(1,\frac{1}{\theta}(\theta(y-\varepsilon)+1)(\theta(x+\varepsilon)+1))\right] \\
&+\frac{e^{-(x-\varepsilon)}e^{1/\theta}}{\theta}\left[e^{(y-\varepsilon)}Ei(1,\frac{1}{\theta}(\theta(y+\varepsilon)+1)(\theta(x-\varepsilon)+1))+e^{(y-\varepsilon)}Ei(1,\frac{1}{\theta}(\theta(y-\varepsilon)+1)(\theta(x-\varepsilon)+1))\right] \\
&+(e^{-(x-\varepsilon)}-e^{-(x+\varepsilon)})(e^{-(y+\varepsilon)}-e^{-(y-\varepsilon)}) \\
&x>\varepsilon \quad y>\varepsilon
\end{aligned}
\end{cases}
$$

In Figure 4.8, we give graphs of local dependence function for different values of ε. It is observed that $\lim_{\varepsilon\to\infty} Cov_{X,Y,\varepsilon,\varepsilon}(x,y) = \theta^{-1}e^{1/\theta}\,\mathrm{Ei}(\theta^{-1})$ which is the correlation coefficient of Gumbel Bivariate Exponential distribution.

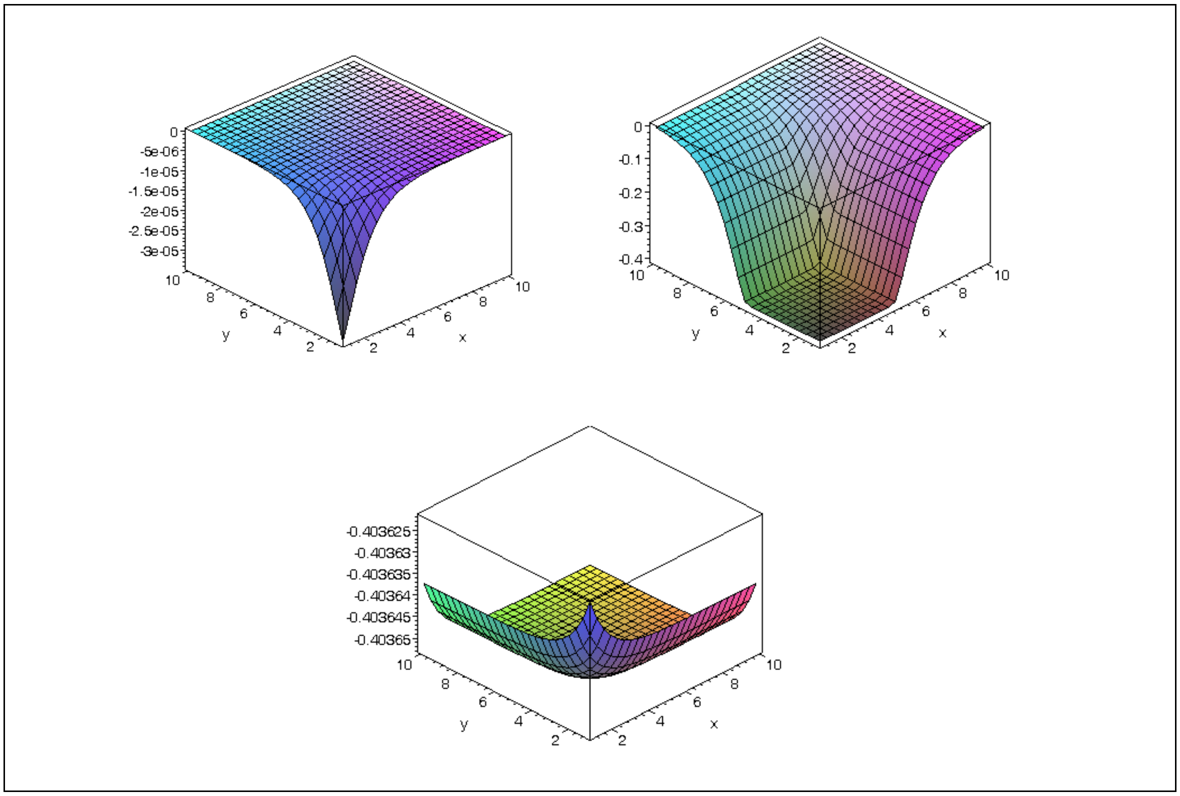


Figure 4.8 Graphs of Local dependence function of Gumbel's Bivariate Exponential Distribution for different values of ε=0.01,5,10, 0<x<10, 0<y<10, θ=1

CHAPTER FIVE
DEPENDENCE MAP AND PERMUTATION TEST

Because of the difficulty in interpreting the estimated local dependence function, Jones & Koch (2003) use the dependence maps via local permutation testing by using Holland & Wang's (1987) local dependence function.

Using the local dependence functions one can construct the dependence maps that make it possible to interpret the full dependence structure of the data set. Dependence maps help us to determine the dependence structure between random variables visually. Local dependence function becomes a more interpretable tool through dependence maps. A dependence map is generally separated into three regions: positive (significant), negative (significant) and zero (nonsignificant) regions. So, by the help of dependence maps we can easily interpret the dependence structure of the data. We construct the dependence map of the given data set by using the natural estimators of local dependence functions. Permutation test algorithm is applied for approximately 500 times to obtain accurate results. Several examples, concerning inflation rate with stock index rate of return with percentage change in the US-Dollar for Turkey and also simulated bivariate normal data are provided.

5.1 Estimation of $\rho_{X,Y,\varepsilon,\varepsilon}(x,y)$

We estimate the local dependence function for the data available such as;

$$\hat{\rho}_{X,Y,\varepsilon,\varepsilon}(x,y) = \frac{\frac{1}{n}\sum_{i=1}^{n}\left[\max(X_i, x+\varepsilon) - \max(X_i, x-\varepsilon)\right]\left[\max(Y_i, y+\varepsilon) - \max(Y_i, y-\varepsilon)\right]}{S_X S_Y} - \frac{\frac{1}{n}\sum_{i=1}^{n}\left[\max(X_i, x+\varepsilon) - \max(X_i, x-\varepsilon)\right]\frac{1}{n}\sum_{I=1}^{N}\left[\max(Y_i, y+\varepsilon) - \max(Y_i, y-\varepsilon)\right]}{S_X S_Y} \tag{5.1}$$

$$F^*_{XY}(u,v) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le u, Y_i \le v\}}\ ,\ F^*_X(u) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le u\}}\ \text{ and }\ F^*_Y(v) = \frac{1}{n}\sum_{i=1}^{n} I_{\{Y_i \le v\}} \tag{5.2}$$

where I is an indicator function such as

$$I_A(u) = \begin{cases} 1, & u \in A \\ 0, & u \notin A \end{cases}$$

where *n* is the sample size and S_X and S_Y are the sample standard deviations of *X* and *Y*. By using this formula, we can compute the sample local dependence function.

Proof. Since,

$$Cov_{X,Y,\varepsilon_1,\varepsilon_2}(x,y) = \int_{y-\varepsilon_2}^{y+\varepsilon_2}\int_{x-\varepsilon_1}^{x+\varepsilon_1}\left[F_{XY}(u,v) - F_X(u)F_Y(v)\right]dudv$$

then the estimator for the local covariance function is

$$
\begin{aligned}
Cov^{*}{}_{X,Y,\varepsilon_1,\varepsilon_2}(x,y) &= \int_{y-\varepsilon_2}^{y+\varepsilon_2}\int_{x-\varepsilon_1}^{x+\varepsilon_1}\left[F_{XY}^{*}(u,v)-F_{X}^{*}(u)F_{Y}^{*}(v)\right]dudv \\
&= \int_{y-\varepsilon_2}^{y+\varepsilon_2}\int_{x-\varepsilon_1}^{x+\varepsilon_1}\left[\frac{1}{n}\sum_{i=1}^{n} I_{\{X_i\le u, Y_i\le v\}} - \left(\frac{1}{n}\sum_{i=1}^{n} I_{\{X_i\le u\}}\right)\left(\frac{1}{n}\sum_{i=1}^{n} I_{\{Y_i\le v\}}\right)\right]dudv \\
&= \frac{1}{n}\sum_{i=1}^{n}\int_{\max(Y_i,y-\varepsilon_2)}^{\max(Y_i,y+\varepsilon_2)}\int_{\max(X_i,x-\varepsilon_1)}^{\max(X_i,x+\varepsilon_1)} 1dudv - \left(\frac{1}{n}\sum_{i=1}^{n}\int_{\max(X_i,x-\varepsilon_1)}^{\max(X_i,x+\varepsilon_1)} 1dudv\right)\left(\frac{1}{n}\sum_{i=1}^{n}\int_{\max(Y_i,y-\varepsilon_1)}^{\max(Y_i,y+\varepsilon_1)} 1dudv\right) \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\max(X_i,x+\varepsilon_1)-\max(X_i,x-\varepsilon_1)\right]\left[\max(Y_i,y+\varepsilon_2)-\max(Y_i,y-\varepsilon_2)\right] \\
&\quad -\frac{1}{n}\sum_{i=1}^{n}\left[\max(X_i,x+\varepsilon_1)-\max(X_i,x-\varepsilon_1)\right]\frac{1}{n}\sum_{i=1}^{n}\left[\max(Y_i,y+\varepsilon_2)-\max(Y_i,y-\varepsilon_2)\right]
\end{aligned}
$$

then

$$
\rho^{*}{}_{X,Y,\varepsilon_1,\varepsilon_2}(x,y) = \frac{Cov^{*}{}_{X,Y,\varepsilon_1,\varepsilon_2}(x,y)}{S_X S_Y} \tag{5.3}
$$

5.2 Local Permutation Test

Permutation test interests the variables being correlated can be classified on certain attributes, primarily. This test is a type of statistical significance test and based on permuting the observed data points across all possible outcomes. Permutation test can also be called a randomization test or re-randomization test. The permutation test is designed to determine whether the observed difference between sample means is large enough to reject the null hypothesis which means that two groups have identical probability distribution. This test is a kind of non-parametric test. It also has only one assumption that it is possible that all of the treatments are equivalent and also every member of them is the same before sampling began. Permutation test exists in many situations where parametric tests are not valid, also for any test statistic it exists regardless of the distribution. An assumption behind this test is to be able to exchange the observations under the null hypothesis; therefore tests of difference in location require equal variance. An alternative of this test is to use bootstrap-based test.

In local permutation testing, for each (x_i, y_i), we firstly compute $\hat{\rho}_{X,Y,\varepsilon,\varepsilon}(x,y)$ for an appropriate ε-value firstly. It is considered making local permutation test by comparing these original estimated values. It is randomly permuted to obtain new samples satisfying the independence hypothesis, which is $\hat{\rho}_{X,Y,\varepsilon,\varepsilon}(x,y) = 0$

This permutation operation is repeated P times, so we obtain P samples which are satisfying the null hypothesis. $\hat{\rho}_P$ is computed for each permuted data set, $p=1,\ldots,P$. For each permuted data set, the significant estimated local correlations at the ε-neighbourhood are taken to the list by comparing the original estimated local correlations. The simulated local correlations in this list are ranked. At α significance level, we decide the dependence map value, such that if the observed value of $\hat{\rho}_P$ is in the highest *(α/2)%* of simulated $\hat{\rho}_P$'s then the dependence map value is +1, if the observed value of $\hat{\rho}_P$ is in the lowest *(α/2)%* of simulated $\hat{\rho}_P$'s then the dependence map value is -1, otherwise it equals to 0. The construction of dependence map does not change for the value greater than P = 500. Also α = 0.05 level is preferred to test the significance. An application-code is developed with Excel Visual Basic which is the application of the local permutation test.

5.3 Examples

5.3.1 Standard Bivariate Normal Data

A standard-normal distributed data set is generated for n = 100 and the dependence map is plotted for different values of *ε*. For this data set, the Pearson correlation coefficient which measures the linear global correlation coefficient is 0.87. Dependence map is constructed for standard bivariate normal distributed data set. In Figure 5.1, this dependence map shows how *X* increases with *Y* by moderate values of *X* and *Y*. There are also areas of no local dependence between X and Y. In spite of small ε, wider positive region occurs since dependence between X and Y is strong; but also zero region is existing. For moderate values of X and Y, it can be said that there is positive dependence.

It is easily said that as ε increases it approximates the Pearson correlation coefficient. For $\varepsilon = 1$, dependence between X and Y is generally positive as expected; but also for large values of X and Y and small values of X and Y, it is seen that there is independence.

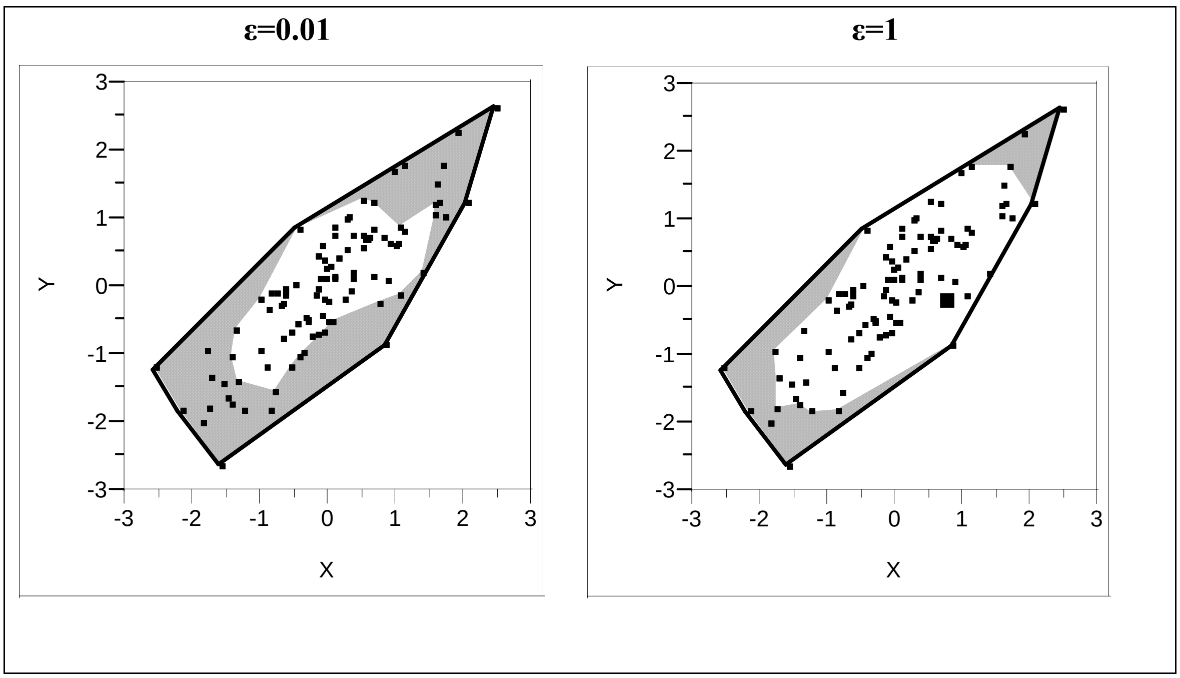


Figure 5.1 Dependence map for Standard Bivariate Normal Data; zero local dependence light grey, positive local dependence white.

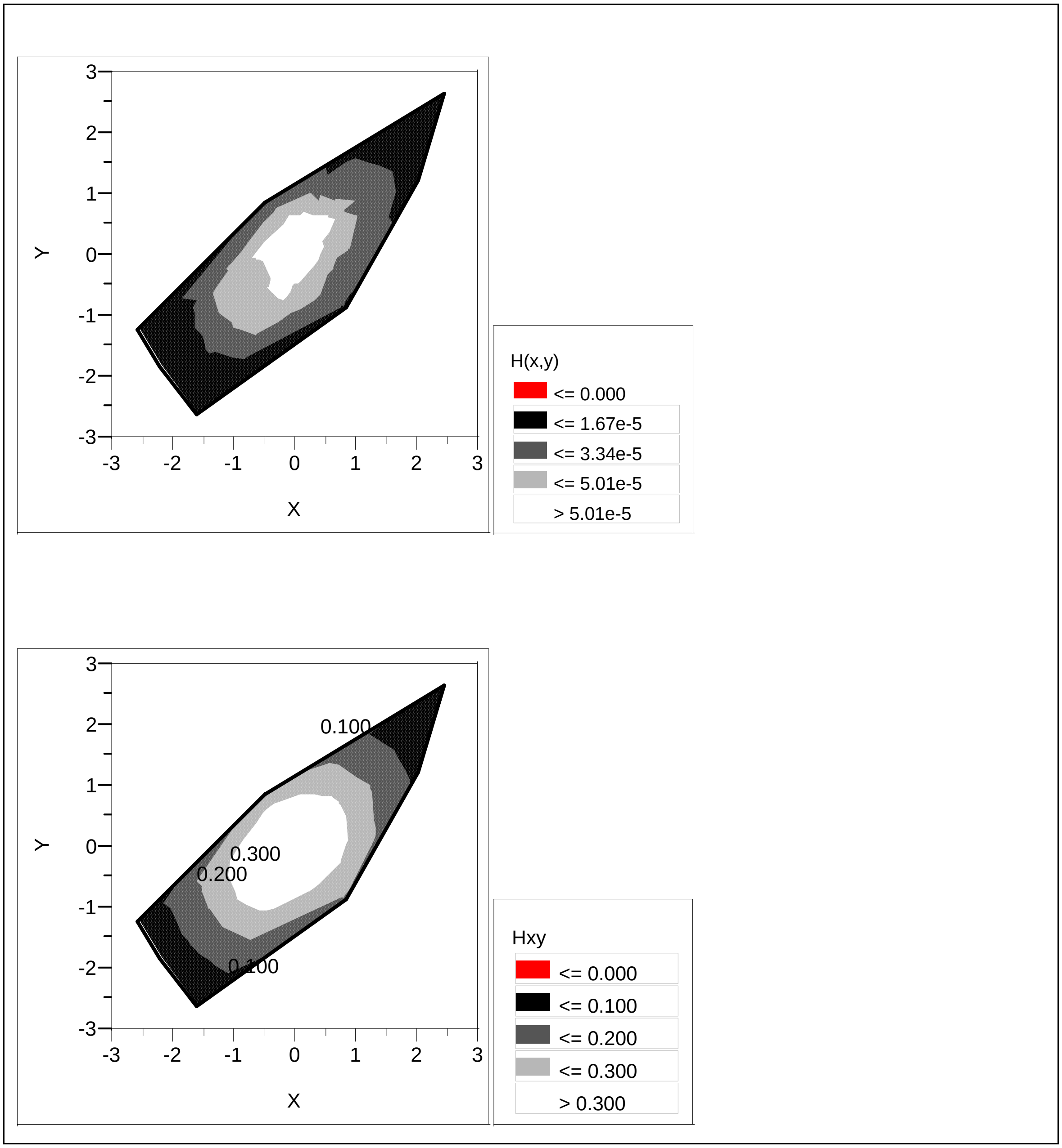


Figure 5.2 Contour Plot of Local Dependence Function for Standard Normally Distributed Data with $\rho = 0.87$

In the above figures, it is seen that the higher dependence value occurs in the middle of the graphs and it gets smaller through the edge of the graphs. There is higher dependence in the second graph as it is expected since it approximates the correlation coefficient when ε gets larger.

5.3.2 Percentage Change in CPI - Percentage Change in US Dollar

In this example, it is concerned that the dependence between monthly percentage change in consumer price index and monthly percentage change in US dollar between the years 1995-2005 for Turkey. They are filtered in case of autocorrelation. There is not strong linear correlation $r = 0.314$. The overall positive correlation has latent dependence structure at the local level. In Figure 5.3, a dependence map for consumer price index monthly change rate and dollar monthly change rate is constructed. In this figure, it is seen that dollar percentage changes around zero is positively dependent with inflation for $\varepsilon = 2.5$. Also when $\varepsilon = 9.5$, there is generally positive dependence except from the lowest and highest values of percentage change in dollar, because there is independence at the lowest and highest values of change in dollar.

In Figure 5.4, contour plot of local dependence function for consumer price index monthly change rate and dollar monthly change rate for $\varepsilon = 2.5$ and $\varepsilon = 9.5$ is constructed. Higher dependence occurs around 0 and the value of index gets smaller through the edge of the plot.

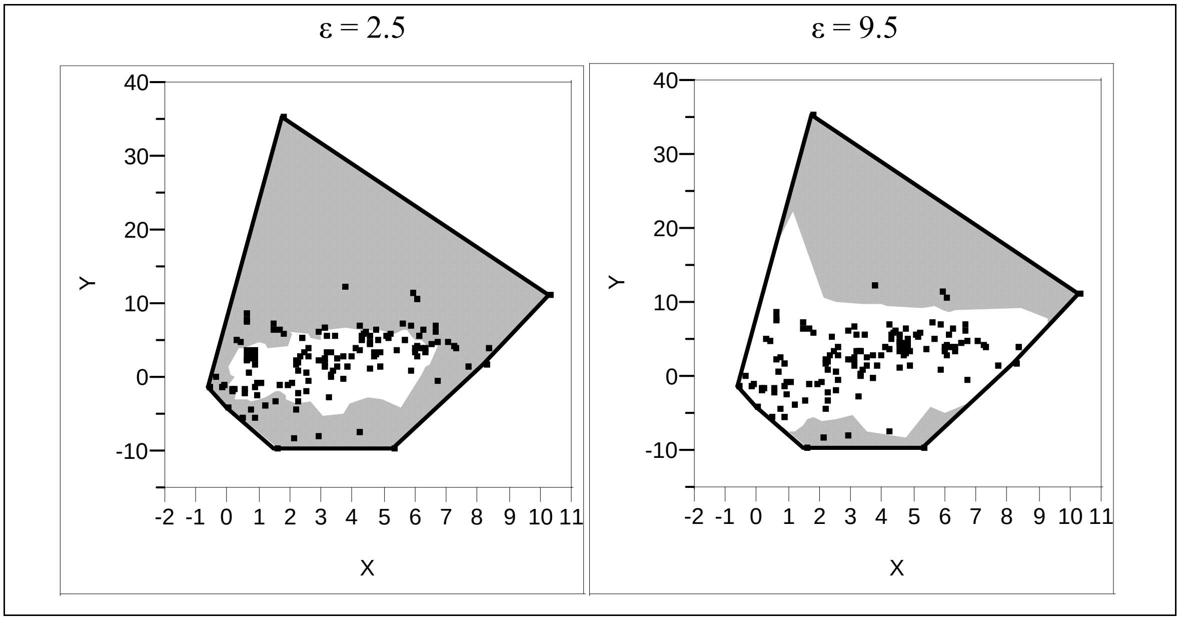


Figure 5.3 Dependence map for Consumer Price Index monthly change rate and dollar monthly change rate; zero local dependence light grey, positive local dependence white. (X: inflation, Y: dollar)for ε = 2.5 and ε = 9.5

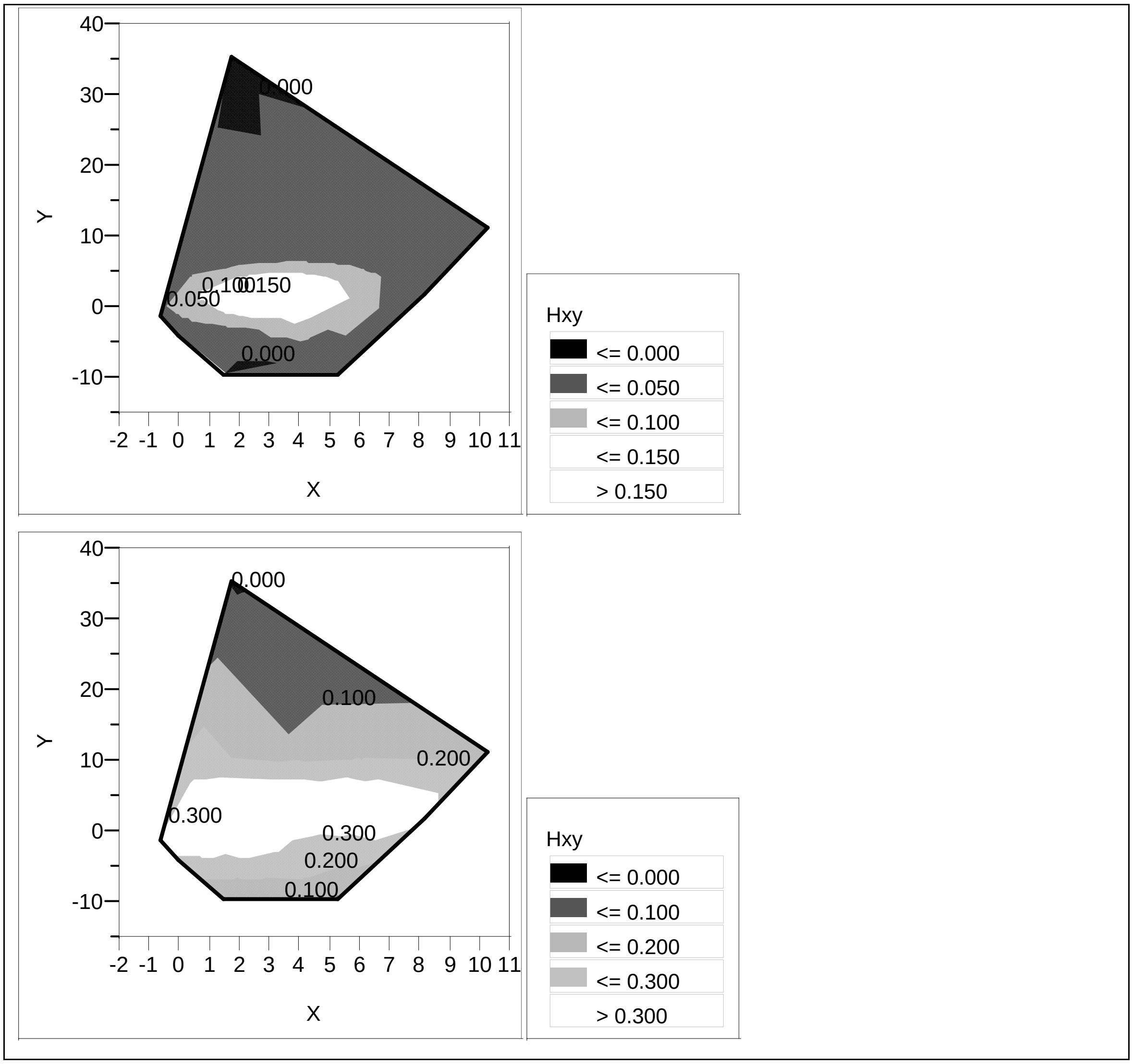


Figure 5.4 Contour Plot of Local Dependence Function for for Consumer Price Index monthly change rate and dollar monthly change rate for ε =2.5 and ε = 9.5

5.3.3 Percentage Change in CPI – Stock Index Rate of Return

In this example, the dependence structure between the variables, monthly percentage change in consumer price index and stock index rate of return is constructed. They are also filtered in case of autocorrelation. A dependence map between these variables is constructed in Figure 5.5. In the case of ε =2.5, there is small positive region that is for the stock index rate of return around 10, but there is generally independence. When ε gets larger, it does not affect the dependence structure very much since there is weak dependence between X and Y.

Contour plot of local dependence function for CPI and return is constructed as shown in Figure 5.6. Higher dependence values occur around 10.

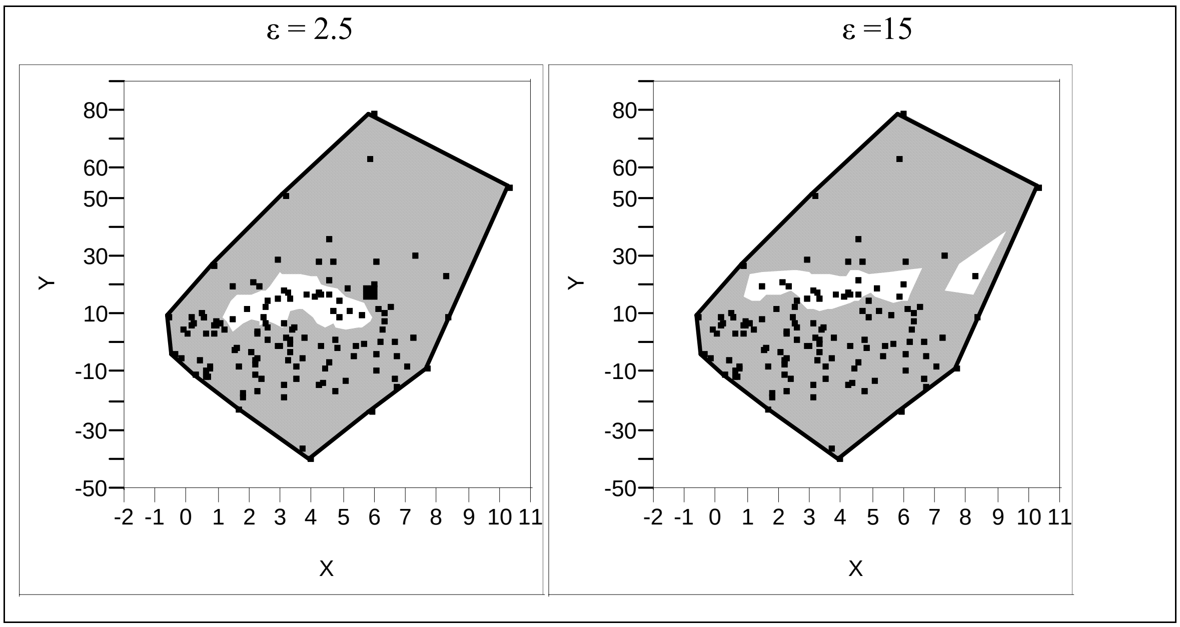


Figure 5.5 Dependence map for consumer price index monthly change rate and stock index rate of return; zero local dependence light grey, positive local dependence white. (X: inflation Y: stock return) ε =2.5 and ε = 15

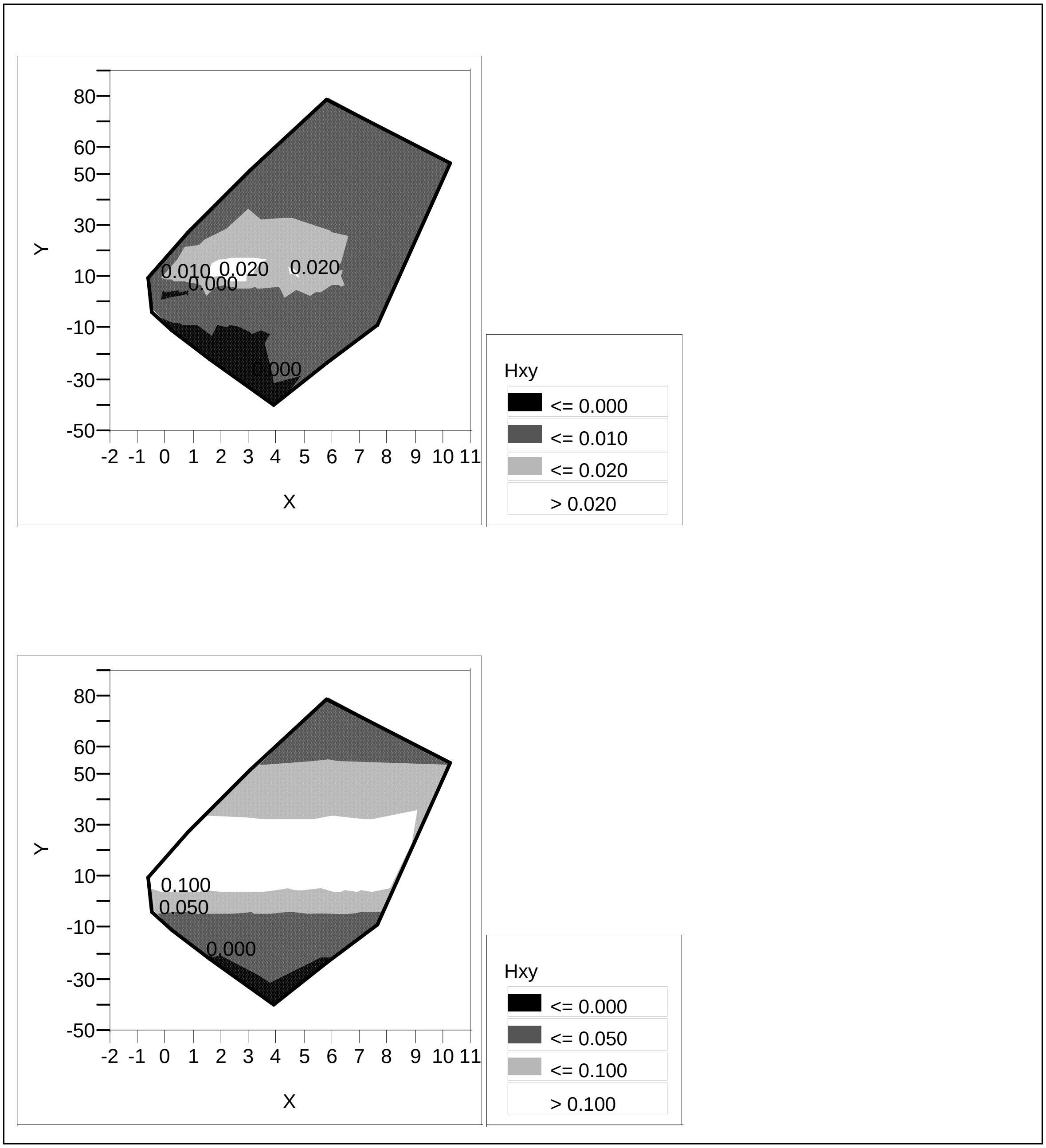


Figure 5.6 Contour Plot of Local Dependence Function for consumer price index monthly change rate and stock index rate of return for ε =2.5 and ε =15.

CHAPTER SIX
CONCLUSIONS

The concept of dependence among random variables is important in statistical theory and applications. Unless specific assumptions are made about the dependence, no meaningful statistical model can be constructed. The Pearson correlation coefficient and many other scalar dependence measures, of course, play an important role in understanding the simplest dependence relationships between two random variables. In general, the dependence structure between a pair of random variables is very complex and the scalar dependence measures can not be adequate to explain the natural association between them.

In this thesis, we introduce a new local dependence function characterizing the dependence structure between a pair of random variables in an ε-neighbourhood of a particular point (x,y) from the domain of underlying bivariate distribution. This local dependence function is essentially different from all local dependence functions introduced earlier, because it depends on the point itself as well as the ε-value which depends on the nature of the considered model. The local dependence function is constructed by the natural way considering the localized version of the Hoeffding's formula and possesses many properties that must satisfy any dependence measure. These properties are investigated. Local dependence functions of Farlie-Gumbel-Morgensten distribution, Sarmanov-Lee distribution and Gumbel's bivariate exponential distribution are also provided.

Dependence maps are constructed for real data examples to simplify the interpretation of the values of the local dependence function. Dependence maps help us to determine the dependence structure between random variables visually. Local dependence function becomes a more interpretable tool through dependence maps. Dependence maps display just three categories; positive dependence, negative dependence and zero (independence). Dependence map is constructed by using permutation test algorithm. Permutation test is a kind of statistical significance test. In this test, correlated variables are classified on certain attributes: positive, negative and zero correlated.

Local dependence maps carry out the information about the estimated local dependence between two random variables better for exploratory use. Dependence is a scale-dependent concept; therefore data may exhibit different types of dependence at different scales.

Moreover, local numerical characteristics of random variables are introduced. When one observes data, and either inadvertently or by design, allows a subset of the data to be observable, the statistics calculated from the data will be the localized statistics as discussed in this thesis. Local moments, especially local expected value and local variances, may be very important in many cases. Local expected value and local variances of some distributions, such as Uniform distribution, Exponential distribution, Cauchy distribution, Two-sided power distribution and Pareto distribution, are obtained. Also many of the properties of the local expected value and the local variances are investigated.

REFERENCES

Arnold C. B., & Strauss D. (1988). Bivariate distributions with exponential conditionals. *JASA, 83* (402), 522-527.

Bairamov, I., & Kotz, S. (2000). On local dependence function for multivariate distributions. *New Trends in Probability and Statistics, 5*, 27-44.

Bairamov, I., & Kotz, S. (2002). Dependence structure and symmetry of Huang-Kotz FGM distributions and their extensions. *Metrika, 56*, 55-72.

Bairamov, I., Kotz, S., & Kozubowski, T. (2003). A new measure of linear local dependence. *Statistics, 37* (3), 243-258.

Barlow, R. E., & Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing.* USA: Holt, Rinehart and Winston, Inc.

Bayes, T. (1763). An essay toward solving a problem in the doctrine of chance. *Phil. Trans. Roy. Soc. 53,* 370-418.

Bjerve, S., & Doksum, K. (1993). Correlation Curves: measures of association as functions of covariate values. *Annals of Statistics, 21*, 890-902.

Block, W. H., & Fang, Z. (1988). A multivariate extension of Hoeffding's lemma. *The Annals of Probability, 16* (4), 1803-1820.

Blyth S., (1994). Measuring local association: an introduction to the correlation curve. *Sociol Meth. 24,* 171-197.

Bracquemond, C., Gaudoin, O., Roy, D., & Xie, M. (2001). In Yu. Hayakava, T. Irony & M. Xie, (Ed.), *In Sysytem and Bayesian Reliability, Essays in Honor of Professor Richard E. Barlow in His 70 th Birthday.* 185-197. Singapore: World Scientific Publishing.

Doksum, K., Blyth S., Bradlow, E., Meng, X., & Zhao, H. (1994). Correlation curve as local measures of variance explained by regression, *JASA, 89* (426), 571-582.

Esary, J.D., Proschan, D., & Walkup, D. W. (1967). Association of random variables, with applications. *The Annals of Mathematical Statistics, 38* (5), 1466-1474.

Esary, J. D., & Proschan, F. (1972). Relationships among some concepts of bivariate dependence. *The Annals of Mathematical Statistics, 43* (2), 651-655.

Gasser, T. & Müller, H.G. (1984). Estimating regressşon functions and their derivatives with kernel methods. *Scandinavian Journal of Statistics,11,* 171-185.

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

Hoeffding, W. (1940). Maastabinvariante Korelations-theorie. *Schr. Math. Inst. Univ. Berlin, 5,* 181-233.

Holland, P.W., & Wang, Y.J. (1987). Dependence function for continuous bivariate densities. *Communucations in Statistics-Theory and Methods, 16*, 863-876.

Joe, H. (1997). *Multivariate Models and Depndence Concepts.* London: Chapman & Hall.

Jones, M.C. (1996). The local dependence function. *Biometrika, 83*, 899-904.

Jones, M.C. (1998). Constant local dependence. *Journal of Multivariate Analysis, 64*, 148-155.

Jones, M.C., & Koch, I. (2003). Dependence Maps:Local Dependence in Practice. *Statistics and Computing, 13*, 241-255.

Kotz, S., & Nadarajah, S. (2003). Local dependence functions for the elliptically symmetric distributions. *Sankhya, 65,* 207-223.

Lai, C.D., Xie, M., & Bairamov, I. G. (2001). Dependence and Ageing Properties of Bivariate Lomax Distribution. In Yu. Hayakava, T. Irony & M. Xie, (Ed.), *In Sysytem and Bayesian Reliability, Essays in Honor of Professor Richard E. Barlow in His 70 th Birthday.* 243-255. Singapore: World Scientific Publishing.

Lancaster, H.O. (1963). Correlation and complete dependence of random variables. *The Annals of Mathematical Statistics, 34* (4), 1315-1321.

Laplace, P. S. (1810). Memoires sur les integrales definies et leurs applications aux probabilities. *Mem. Inst. Imp.,* 279-347.

Lehmann, E. L. (1966). Some concepts of dependence. *The Annals of Mathematical Statistics, 37* (5), 1137-1153.

Lee, M.L., (1996). Properties and applications of Sarmanov family of bivariate distributions. *Communications and Statistics Theory and Methods, 25* (6), 1207-1222.

Mari, D., & Kotz, S. (2001). *Correlation and Dependence.* London: Imperial College Press.

Matteis, R. (2001). Fitting Copulas to Data, *Diploma Thesis Institute of Mathematics of the University Zurich.*

Nadarajah, S., Mitov, K., & Kotz, S. (2003). Local dependence functions for extreme value distributions. *Journal of Applied Statistics, 30* (10), 1081-1100.

Nadaraja, E. A. (1964). On estimating regression. *Theory of Probability and Its Applications, 9,* 141-142.

Nelsen, R.B. (1999). *An Introduction to Copulas.* New York: Spring-Verlag.

Renyi, A. (1959). On measures of dependence. *Acta Mathematica Academiae Scientarium Hungaricae, 10*, 441-451.

Schweizer, B., & Wolff, E. F. (1981). On parametric measures of dependence for random variables. *Annals of Statistics, 9*, 879-885.

Sklar, A. (1959). Fonctions de répartation à dimensions et leurs marges. *Publ. Inst. Univ. Paris 8*, 229-231.

Van Dorp, J.R, & Kotz, S. (2002b). The standard two-sided power distribution and its properties: with application in financial engineering. *The American Statistician, 56,* 90-99.

Wand, M.P, & Jones, M.C. (1995). *Kernel Smoothing.* London: Chapman and Hall.

Watson, G.S. (1964). Smooth regression analysis. *Sankhya A, 26,* 359-372.

Yule, U. (1897). On the theory of correlation. *Journal of Royal Statistical Society, 60* (4), 812-854.

APPENDICES

APPENDIX I

Visual Basic Script for Constructing Dependence Map

```
Option Explicit

Dim MinX, MaxX, MinY, MaxY
Function StDevX() As Double
  StDevX = Cells(12, 2)
End Function

Function StDevY() As Double
  StDevY = Cells(13, 2)
End Function

Function SampleN() As Integer
   SampleN = Cells(4, 2)
End Function

Function DataN() As Integer
   DataN = Cells(2, 2)
End Function

Function Eps() As Double
   Eps = Cells(3, 2)
End Function

Function Sqrt(ByVal Sayi As Double) As Double
   Sqrt = Sayi ^ 0.5
 End Function

Function XI(ByVal I As Long) As Double
  XI = Cells(I + 1, 4)
End Function

Function YI(ByVal I As Long, ByVal Sampling As Boolean) As
Double
  If Sampling Then
    YI = Cells(I + 1, 7)
  Else
    YI = Cells(I + 1, 5)
  End If
```

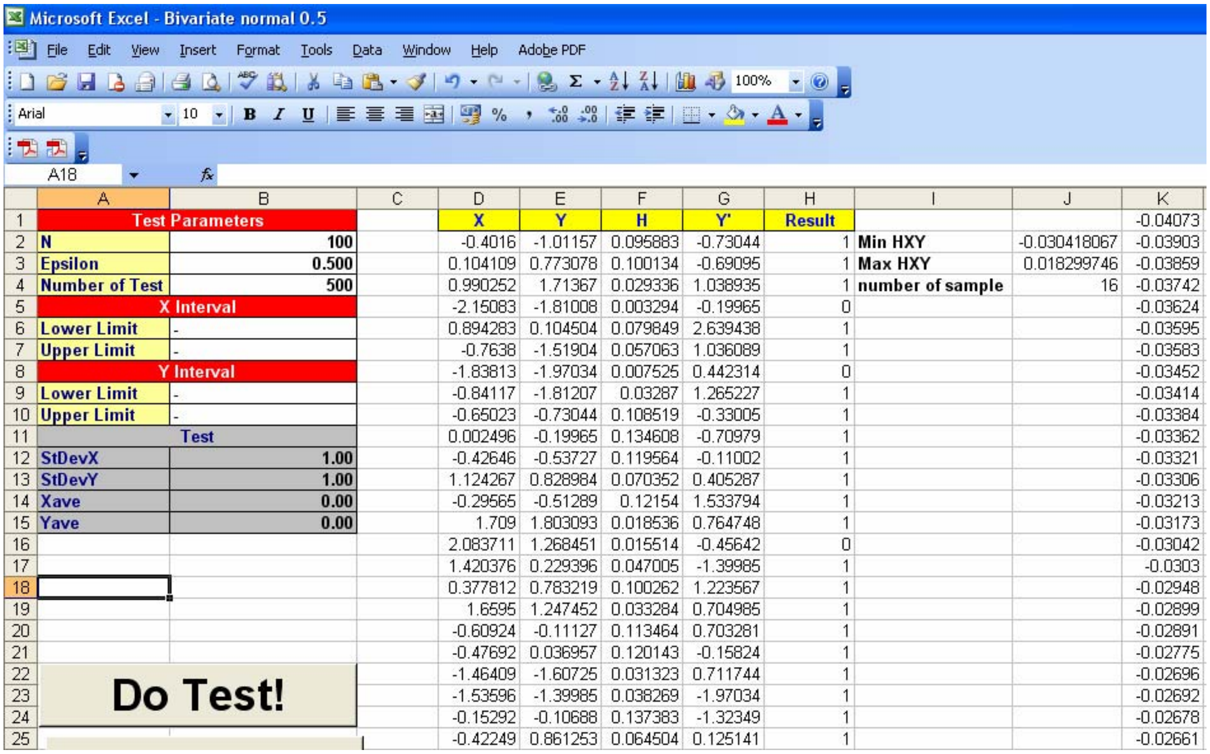
```
End Function

Function OrgHXY(ByVal I As Long) As Double
  OrgHXY = Cells(I + 1, 6)
End Function


Function GetMax(ByVal P1 As Double, ByVal P2 As Double) As
Double
  If P1 > P2 Then
    GetMax = P1
  Else
    GetMax = P2
  End If
End Function

Function SXY(ByVal PX As Double, ByVal PY As Double, ByVal
Sampling As Boolean) As Double
  Dim I As Long
  Dim ASum As Double

  ASum = 0
  For I = 1 To DataN
    ASum = ASum + ((GetMax(XI(I), PX + Eps) - GetMax(XI(I), PX -
Eps)) * (GetMax(YI(I, Sampling), PY + Eps) - GetMax(YI(I,
Sampling), PY - Eps)))
  Next I
  SXY = (ASum / DataN)
End Function

Function SX(ByVal PX As Double, ByVal Sampling As Boolean) As
Double
  Dim I As Long
  Dim ASum As Double

  ASum = 0
  For I = 1 To DataN
    ASum = ASum + (GetMax(XI(I), PX + Eps) - GetMax(XI(I), PX -
Eps))
  Next I
  SX = (ASum / DataN)
End Function

Function SY(ByVal PY As Double, ByVal Sampling As Boolean) As
Double
  Dim I As Long
  Dim ASum As Double

  ASum = 0
```

```
  For I = 1 To DataN
    ASum = ASum + (GetMax(YI(I, Sampling), PY + Eps) -
GetMax(YI(I, Sampling), PY - Eps))
  Next I
  SY = (ASum / DataN)
End Function

Function HXY(ByVal PX As Double, ByVal PY As Double, ByVal
Sampling As Boolean) As Double
  HXY = (SXY(PX, PY, Sampling) - SX(PX, Sampling) * SY(PY,
Sampling)) / (StDevX * StDevY)
End Function

Sub DoSample(SourceColIndex, TargetColIndex)
   Dim SourceList As New Collection
   Dim TargetList As New Collection
   Dim I As Integer
   Dim Index As Integer

   For I = 2 To DataN + 1
     SourceList.Add (Cells(I, SourceColIndex))
   Next

   Randomize
   For I = 1 To DataN
     Index = Int(((DataN - I + 1) * Rnd) + 1)
     TargetList.Add (SourceList(Index))
     SourceList.Remove (Index)
   Next

   For I = 1 To DataN
     Cells(I + 1, TargetColIndex) = TargetList(I)
   Next I
 End Sub

 Sub DoTest()
   Dim InitTime
   Dim I, J As Integer
   Dim MinHXY, MaxHXY, TempHXY As Double
   Dim Index, HXYNumber As Integer
   Dim SumHXY As New Collection
   Dim MinHXY As New Collection


   InitTime = Now

   For I = 1 To DataN
     Cells(I + 1, 6) = HXY(XI(I), YI(I, False), False)
   Next I
```

```
    Randomize

    For I = 1 To SampleN
      DoSample 5, 7
      Application.StatusBar = "Sampling " & I & " " &
CStr(CDate(Now - InitTime))
      For J = 1 To DataN
        TempHXY = HXY(XI(J), YI(J, True), True)
        If Abs(TempHXY) >= Abs(OrgHXY(J)) Then
          ToplamHXY.Add (TempHXY)
        End If
      Next J
      DoEvents
    Next I

    Columns(11).Clear
    For I = 1 To SumHXY.Count
      Cells(I, 11) = SumHXY.Item(I)
    Next
    Columns(11).Select

    HXYNumber = Int(SumHXY.Count * 0.025)
    If HXYSayisi = 0 Then
      HXYSayisi = 1
    End If

    EnKucukHXY = Cells(HXYSayisi, 11)
    EnBuyukHXY = Cells(SumHXY.Count - HXYNumber + 1, 11)

    Cells(2, 9) = "Min HXY"
    Cells(2, 10) = MinHXY
    Cells(3, 9) = "Max HXY"
    Cells(3, 10) = MaxHXY
    Cells(4, 9) = "sample number"
    Cells(4, 10) = HXYNumber

    For I = 1 To DataN
     Cells(I + 1, 8) = 0
     If Cells(I + 1, 6) > MaxHXY Then
       Cells(I + 1, 8) = 1
     End If
     If Cells(I + 1, 6) < MinHXY Then
       Cells(I + 1, 8) = -1
     End If
    Next I

    Application.StatusBar = ""
    Cells(4, 10).Select
```

```
   MsgBox "Test completed" & Chr(13) & Chr(10) & "("
&CStr(CDate(Now - InitTime)) & ")", vbOKOnly
End Sub


Sub x()
  DoSample 5, 7
End Sub
```

APPENDIX II



Microsoft Excel - Bivariate normal 0.5

File Edit View Insert Format Tools Data Window Help Adobe PDF

	A	B	C	D	E	F	G	H	I	J	K
1	Test Parameters			X	Y	H	Y'	Result			-0.04073
2	N	100		-0.4016	-1.01157	0.095883	-0.73044	1	Min HXY	-0.030418067	-0.03903
3	Epsilon	0.500		0.104109	0.773078	0.100134	-0.69095	1	Max HXY	0.018299746	-0.03859
4	Number of Test	500		0.990252	1.71367	0.029336	1.038935	1	number of sample	16	-0.03742
5	X Interval			-2.15083	-1.81008	0.003294	-0.19965	0			-0.03624
6	Lower Limit	-		0.894283	0.104504	0.079849	2.639438	1			-0.03595
7	Upper Limit	-		-0.7638	-1.51904	0.057063	1.036089	1			-0.03583
8	Y Interval			-1.83813	-1.97034	0.007525	0.442314	0			-0.03452
9	Lower Limit	-		-0.84117	-1.81207	0.03287	1.265227	1			-0.03414
10	Upper Limit	-		-0.65023	-0.73044	0.108519	-0.33005	1			-0.03384
11	Test			0.002496	-0.19965	0.134608	-0.70979	1			-0.03362
12	StDevX	1.00		-0.42646	-0.53727	0.119564	-0.11002	1			-0.03321
13	StDevY	1.00		1.124267	0.828984	0.070352	0.405287	1			-0.03306
14	Xave	0.00		-0.29565	-0.51289	0.12154	1.533794	1			-0.03213
15	Yave	0.00		1.709	1.803093	0.018536	0.764748	1			-0.03173
16				2.083711	1.268451	0.015514	-0.45642	0			-0.03042
17				1.420376	0.229396	0.047005	-1.39985	1			-0.0303
18				0.377812	0.783219	0.100262	1.223567	1			-0.02948
19				1.6595	1.247452	0.033284	0.704985	1			-0.02899
20				-0.60924	-0.11127	0.113464	0.703281	1			-0.02891
21				-0.47692	0.036957	0.120143	-0.15824	1			-0.02775
22	Do Test!			-1.46409	-1.60725	0.031323	0.711744	1			-0.02696
23				-1.53596	-1.39985	0.038269	-1.97034	1			-0.02692
24				-0.15292	-0.10688	0.137383	-1.32349	1			-0.02678
25				-0.42249	0.861253	0.064504	0.125141	1			-0.02661