

NONPARAMETRIC STATISTICAL INFERENCE FOR THE
TWO-SAMPLE LOCATION PROBLEM

HAIYAN YANG



Nonparametric Statistical Inference for the Two-Sample
Location Problem

by

©Haiyan Yang

A thesis submitted to the
School of Graduate Studies
in partial fulfilment of the
requirements for degree of
Master of Science

Department of Mathematics and Statistics
Memorial University of Newfoundland

June 2006 Submitted

St. John's

Newfoundland

Canada



Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 978-0-494-30520-1
Our file *Notre référence*
ISBN: 978-0-494-30520-1

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.


Canada

Abstract

The two-sample problem occurs in many scientific fields, with a major frequency in environmental and agricultural research. Its primary goal is either to estimate the difference between certain parameters of the two populations or to test the hypotheses about the difference. The well known nonparametric procedure is Mann-Whitney-Wilcoxon test. The aim of this thesis is to develop a new statistical inference procedure for the problem by extending the Mann-Whitney-Wilcoxon test. In particular, simulated probabilities of the new proposed V statistic for the uniform, normal and exponential distributions are considered. Efficient computation algorithms are proposed to obtain the distribution functions of the V statistic. Power studies via simulation compare the new proposed procedure with Mann-Whitney-Wilcoxon procedure. Also included is a real data set.

Acknowledgement

I would like to express my sincere gratitude to Professor Chu-In Charles Lee for granting me the privilege to write this thesis under his supervision and for his support, both financial and otherwise, through out the course of my program. He has been very generous with his ideas and times. He patiently read through my work several times and offered his very valuable suggestions. This work would have been impossible without his guidance and help.

Many thanks go to my wife, Dr. Yuan Yuan, for making available to me the high performance computing facilities and applied mathematical knowledge.

I would like to thank professors who have shared with me their wealth of statistical knowledge at various points in the course of my program. I would also like to thank the staff of the Mathematics and Statistics for providing me with a friendly atmosphere and necessary facilities to finish my program.

My sincere thanks to my families, friends and classmates whose assistance and encouragement have seen me to the completion of my program.

Contents

1	Introduction	1
2	V Statistic	5
2.1	Introduction	5
2.2	Transformation	6
2.3	The null distribution of V	10
2.4	Table value	15
3	Mean and Variance	31
3.1	Expectations and variances	31
3.2	Sampling from the uniform distribution	50
3.3	Sampling from the normal distribution	59
3.4	Sampling from the exponential distribution	64
3.5	Mean and Variance of V among the tied observations	68
4	Approximations	76
4.1	Normal approximation	76
4.2	Asymptotically distribution-free of V	80
4.3	Simulated power comparison	81

5	Application for Dose-Response Study	84
5.1	Introduction	84
5.2	The proposed testing procedure	85
5.3	An example	87
6	Summary and Further research	90
	Appendix	92
A	Simulation of the Probabilities of Individual V Statistic	92
B	Power Test for LY Test and Mann-Whitney Test	94
	Bibliography	95

List of Tables

2.1	Distributions, Means and Variances of V for the cases $m = 1, n = 2$ and $m = 1, n = 3$	15
2.2	Cumulative Probabilities of Statistic V for Small Values m and n (U: Uniform, N: Normal, E: Exponential, E^* : Negative Exponential)	17
2.3	Critical Values for V (U: Uniform, N: Normal, E: Exponential, E^* : Negative Exponential)	23
2.4	Critical Values for V When $m = n$ (C: Cauchy, D: Double Exponential)	26
3.1	Components of I_1	32
3.2	Components of I_2	33
3.3	Components of I_3	35
3.4	Components of I_4	37
3.5	Components of (I_1, I_2)	40
3.6	Components of (I_1, I_3)	42
3.7	Components of (I_1, I_4)	43
3.8	Components of (I_2, I_3)	44
3.9	Components of (I_2, I_4)	46
3.10	Components of (I_3, I_4)	48
4.1	Normal Approximation Probability (U: Uniform, N: Normal, E: Exponential)	79

4.2	Normal Approximation Probability(C: Cauchy, D: Double Exponential)	80
4.3	Power Test for Different Mean ($\alpha=0.05$)	83
5.1	Reaction Times in Seconds of Mice to Stimuli to their Tails	87

List of Figures

2.1	A plot of distribution of V for different m and n (Exact values when m=1, n=2, Simulated values when n=2, n=3)	27
2.2	A plot of distribution of V (simulated) for different m and n	28
2.3	A plot of distribution of V (simulated) for m = n	29
2.4	A plot of distribution of V (simulated) for m = n	30

Chapter 1

Introduction

In recent years, more and more statisticians have come to appreciate the advantages of non-parametric tests. Not only do nonparametric tests have often surprisingly high efficiency relative to their normal-theory equivalents even under assumption of normality, but they are also less sensitive to the inference of “wild” observations than are the normal-theory equivalents. Moreover, in situations where measurements are costly and/or difficult to obtain but ranking of the potential sample data is relatively easy, the use of statistical methods based on ranked set sampling (RSS) can lead to substantial improvement over analogous methods associated with simple random sampling (SRS) schemes. The approach using ranked set sampling has attracted considerable attention in the recent literature, with principal initial interest being driven by environmental and agricultural issues, where it is clear that pre-sampling judgement ranking can be quite inexpensive relative to the cost of detailed measurement of many quantities of interest. For example, preliminary supporting data can be easily obtained from a contaminated site and analyzed both quickly and inexpensively before final decisions are reached as to where and how to obtain the specific measurement(s) of interest, such as lead or hazardous material content of the site. A similar approach can

be taken when performing standard gasoline octane checks at stations, where preliminary screening at the gasoline pumps can lead to an improved set of ranked set gasoline samples to carry back to the laboratory for more detailed (and costly) analysis. Agricultural settings where such an approach can be used effectively include predicting crop yields or lumber content via preliminary non-destructive measurements (satellite observations or concomitant variables, for example).

Most of the initial research efforts in ranked set sampling have concentrated on parametric and nonparametric estimation and testing procedures for the one- and two-sample settings. See, for example, McIntyre (1952) introduced the concept of ranked set sampling in relation to estimating pasture yield that has promise of application in many other sampling problems. Takahasi & Wakimoto (1968) developed an extensive body of theory that is based upon the assumption of perfect ranking. Halls & Dell (1966) tried the procedure in sampling forage yields. Dell & Clutter (1972) proposed that ranked set sampling employs judgment ordering to obtain an estimate of a population mean. Stokes (1977) showed that improved estimates of the variance of any population can be produced from rank-set sample. Stokes & Sager (1988) provided a characterization of a ranked-set sample that makes the source of additional information intuitively clear and showed that the empirical distribution function of a ranked-set sample has greater precision than that from a random sample. Bohn & Wolfe (1992, 1994) use the empirical distribution function for ranked-set samples to develop nonparametric inference techniques in the two-sample location problem and discussed the asymptotic relative efficiency comparisons between the simple random sample Mann-Whitney-Wilcoxon procedures and their ranked-set analogues, and addressed the issues of imperfect judgment rankings that based on the ranked-set samples analog of the Mann-Whitney-Wilcoxon statistic. Koti & Babu (1996) introduced a sign test for ranked-set sampling. Lately, some researchers have expressed interest in the appropri-

ate allocation of order statistics within a ranked set sample. Kaur, Patil & Taillie (1997) studied the effects of unequal allocation on estimation of the population mean when the underlying distribution is skewed or symmetric. They provided “near” optimality results based on skewness, kurtosis and/or the coefficient of variation. In the parametric set up, Bhoj (1997) constructed an unbiased estimator of the population mean by using a linear combination of the order statistics from the same set in a ranked set sampling environment. Öztürk (1999a, b) and Öztürk & Wolfe (1998, 2000a, b) introduced a design concept for determining the appropriate allocation of the order statistics for the sign and signed rank test statistics. They showed that for all symmetric distributions the best design among all possible allocation procedures is the one that quantifies only the middle observation(s). In this thesis, we extend the same idea to settings where the extension of two-sample ranked set sample Mann-Whitney-Wilcoxon test is appropriate.

In Chapter 2, we will introduce some basic ideas for V statistic. We will present transformation of V that can be seen an extension of Mann-Whitney-Whicoxon test. We list some values of cumulative probabilities when the data distribution is uniform (0,1), normal (0,1), and exponential. If the population distribution is symmetric, then so does the distribution of V . Furthermore, the V statistic is asymptotically normally distributed.

In Chapter 3, we will propose a new test statistic. We study the mean and the variance for different population distribution. We formulate the variance of V and covariances of components of V . We obtain the formula of the mean and variance of V in the presence of tied observations.

In Chapter 4, we will consider the problem of large sample data in two sample studies. Hollander(1967) showed that I_4 has an asymptotically distribution-free procedure under H_0 , we will prove that V has asymptotically distribution-free procedure which trends as a unit normal random variable under H_0 . We will study simulated power comparison for

two test. The evaluation of the simulated power comparison is that the new procedure is significantly more powerful than Mann-Whitney-Wilcoxon's in detecting the difference mean between two-samples when those comparisons are of interest.

In Chapter 5, We consider the problem of identifying the minimum effective dose in dose-response studies. Assessing monotone dose-response relationship is frequently encountered in practice in the context of actively proving a significant monotonous dependence of the response on increasing dose or treatments, but the monotone dose-response assumption is not always satisfied. We will propose a new test(LY test) incorporated into the step-down closed testing scheme under the partial dose-response assumption in which the responses in the dose treatments are larger than that in the control.

Finally, Chapter 6 will provide a summary of the results of this thesis and further areas of research relating to the problem.

Chapter 2

V Statistic

2.1 Introduction

We will discuss in detail only of the two-sample problem. Let X_1, \dots, X_m and Y_1, \dots, Y_n be two independent random samples from continuous distributions with distribution functions F_1 and F_2 , respectively. The excellent properties of the Mann-Whitney-Wilcoxon statistic $\sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)}$ for testing $H_0 : F_1 = F_2 = F$ against translation alternatives $H_1 : F_1(x) = F(x), F_2(x) = F(x - \Delta), \Delta \neq 0$ are well known. Where $I_{(Z)} = 1$ if $Z > 0$ and $= 0$ otherwise. We propose a new statistic V and it is a signed rank sum of all possible pairs (X_i, Y_k) , that is,

$$V = \sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)} \text{rank}(|Y_k - X_i|)$$

We assume $m \leq n$ for convenience; if $m > n$ the symbols x and y can be interchanged. V takes values $0, 1, \dots, mn(mn+1)/2$.

2.2 Transformation

The new test of V statistic is an extension of Mann-Whitney-Wilcoxon statistic. Let $D_{n(i-1)+k} = Y_k - X_i$ and let $D_{(1)} \leq D_{(2)} \leq \dots \leq D_{(mn)}$ be the ordered random variables. It is straight forward that

for $h < r$

$$I_{(D_{(h)}+D_{(r)})} = 1 \iff D_{(r)} > 0 \text{ and } |D_{(h)}| < D_{(r)}.$$

$$\text{when } D_{(r)} > 0, \text{rank}(|D_{(r)}|) = \sum_{h=1}^r I_{(D_{(h)}+D_{(r)})}$$

Therefore

$$\begin{aligned} V &= \sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)} \text{rank}(|Y_k - X_i|) \\ &= \sum_{D_{(r)} > 0} \text{rank}(|D_{(r)}|) = \sum_{r=1}^{mn} I_{(D_{(r)})} \text{rank}(|D_{(r)}|) \\ &= \sum_{r \geq h} \sum I_{(D_{(h)}+D_{(r)})} = \sum_{r \geq h} \sum I_{(D_h + D_r)} \\ &= \sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)} + \frac{1}{2} \sum_{i=1}^m \sum_{k=1}^n \sum_{j=1}^m \sum_{l=1}^n I_{(Y_k - X_i + Y_l - X_j)} \\ &\quad \text{(i,k) \neq (j,l)} \\ &= \sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)} + \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{Y_k + Y_l}{2} - X_i)} \\ &\quad + \sum_{k=1}^n \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m I_{(Y_k - \frac{X_i + X_j}{2})} + 2 \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{Y_k + Y_l}{2} - \frac{X_i + X_j}{2})} \\ &= I_1 + I_2 + I_3 + 2I_4 \end{aligned} \tag{2.1}$$

where

$$I_1 = \sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)}$$

$$I_2 = \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{Y_k + Y_l}{2} - X_i)}$$

$$I_3 = \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m I_{(Y_k - \frac{X_i + X_j}{2})}$$

$$I_4 = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{(\frac{Y_k + Y_l}{2} - \frac{X_i + X_j}{2})}$$

We note that the Wilcoxon signed rank test can be appropriately applied to the paired replicates data, provided the null ($\theta = \theta_0$) distribution of $Y-X$ is symmetric about θ_0 . This assumption is very often inherently satisfied for paired replicates data. The following results are important to our understanding of the V statistic in distribution arguments.

Let $X_1, \dots, X_m, Y_1, \dots, Y_n$ be independent and identically distributed random variables from a continuous distribution F , two theorems are as following.

Theorem 2.1. V is symmetric if distribution function F is symmetric.

Proof: Let $X_1, \dots, X_m, Y_1, \dots, Y_n$ be independent and identically distributed random variables from a continuous distribution that is symmetric about the point μ . It is trivial that

$$\sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)} + \sum_{i=1}^m \sum_{k=1}^n I_{(X_i - Y_k)} = mn.$$

It follows that

$$\sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)} - \frac{mn}{2} = \frac{mn}{2} - \sum_{i=1}^m \sum_{k=1}^n I_{(X_i - Y_k)}.$$

Similarly,

$$\sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{(\frac{Y_k + Y_l}{2} - X_i)} - \frac{mn(n-1)}{4} = \frac{mn(n-1)}{4} - \sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{(X_i - \frac{Y_k + Y_l}{2})}$$

$$\sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m I_{(Y_k - \frac{X_i + X_j}{2})} - \frac{mn(m-1)}{4} = \frac{mn(m-1)}{4} - \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m I_{(\frac{X_i + X_j}{2} - Y_k)}$$

$$\sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{(\frac{Y_k + Y_l}{2} - \frac{X_i + X_j}{2})} - \frac{mn(m-1)(n-1)}{8}$$

$$= \frac{mn(m-1)(n-1)}{8} - \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{\left(\frac{X_i+X_j}{2} - \frac{Y_k+Y_l}{2}\right)} \quad \substack{i < j \\ k < l}$$

implies

$$\begin{aligned} & \sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)} + \sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{\left(\frac{Y_k+Y_l}{2} - X_i\right)} + \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m I_{\left(Y_k - \frac{X_i+X_j}{2}\right)} \\ & + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2}\right)} - \frac{mn(mn+1)}{4} \\ & = \frac{mn(mn+1)}{4} - \left\{ \sum_{i=1}^m \sum_{k=1}^n I_{(X_i - Y_k)} + \sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{\left(X_i - \frac{Y_k+Y_l}{2}\right)} \right. \\ & \left. + \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m I_{\left(\frac{X_i+X_j}{2} - Y_k\right)} + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{\left(\frac{X_i+X_j}{2} - \frac{Y_k+Y_l}{2}\right)} \right\} \end{aligned}$$

Let

$$\begin{aligned} t(X_1, \dots, X_m, Y_1, \dots, Y_n) &= \sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)} + \sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{\left(\frac{Y_k+Y_l}{2} - X_i\right)} \\ & + \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m I_{\left(Y_k - \frac{X_i+X_j}{2}\right)} + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2}\right)} - \frac{mn(mn+1)}{4}. \end{aligned}$$

Then

$$\begin{aligned} t(-X_1, \dots, -X_m, -Y_1, \dots, -Y_n) &= \sum_{i=1}^m \sum_{k=1}^n I_{(X_i - Y_k)} + \sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{\left(X_i - \frac{Y_k+Y_l}{2}\right)} \\ & + \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m I_{\left(\frac{X_i+X_j}{2} - Y_k\right)} + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{\left(\frac{X_i+X_j}{2} - \frac{Y_k+Y_l}{2}\right)} - \frac{mn(mn+1)}{4} \\ & = \frac{mn(mn+1)}{4} - \left\{ \sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)} + \sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{\left(\frac{Y_k+Y_l}{2} - X_i\right)} \right. \\ & \left. + \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m I_{\left(Y_k - \frac{X_i+X_j}{2}\right)} + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2}\right)} \right\} \\ & = -t(X_1, \dots, X_m, Y_1, \dots, Y_n) \end{aligned}$$

and hence $t(\cdot)$ is an odd function. It is trivial that $t(\cdot)$ is translation-invariant, that is

$$t(X_1 + c, \dots, X_m + c, Y_1 + c, \dots, Y_n + c) = t(X_1, \dots, X_m, Y_1, \dots, Y_n)$$

By Corollary 1.3.22 of Randles and Wolfe (1979), $t(\cdot)$ is symmetrically distributed about zero. It follows that V is symmetrically distributed about $mn(mn+1)/4$ for continuous symmetrical distribution F .

Theorem 2.2. V is symmetric if $m = n$ for any continuous distribution F .

Proof: $X_1, \dots, X_n, Y_1, \dots, Y_n$ be independent and identically distributed random variables from a continuous distribution F , then

$$\begin{aligned} & \sum_{i=1}^n \sum_{k=1}^n I_{(Y_k - X_i)} + \sum_{i=1}^n \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{Y_k + Y_l}{2} - X_i)} + \sum_{k=1}^n \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n I_{(Y_k - \frac{X_i + X_j}{2})} \\ & + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{Y_k + Y_l}{2} - \frac{X_i + X_j}{2})} - \frac{n^2(n^2 + 1)}{4} \\ & = \frac{n^2(n^2 + 1)}{4} - \left\{ \sum_{i=1}^n \sum_{k=1}^n I_{(X_i - Y_k)} + \sum_{i=1}^n \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(X_i - \frac{Y_k + Y_l}{2})} \right. \\ & \left. + \sum_{k=1}^n \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n I_{(\frac{X_i + X_j}{2} - Y_k)} + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{X_i + X_j}{2} - \frac{Y_k + Y_l}{2})} \right\} \end{aligned}$$

Let

$$\begin{aligned} W(X_1, \dots, X_n; Y_1, \dots, Y_n) &= \sum_{i=1}^n \sum_{k=1}^n I_{(Y_k - X_i)} + \sum_{i=1}^n \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{Y_k + Y_l}{2} - X_i)} \\ & + \sum_{k=1}^n \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n I_{(Y_k - \frac{X_i + X_j}{2})} + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{Y_k + Y_l}{2} - \frac{X_i + X_j}{2})} - \frac{n^2(n^2 + 1)}{4}. \end{aligned}$$

Then

$$\begin{aligned} W(X_1, \dots, X_n; Y_1, \dots, Y_n) &= \frac{n^2(n^2 + 1)}{4} - \left\{ \sum_{i=1}^n \sum_{k=1}^n I_{(X_i - Y_k)} + \sum_{i=1}^n \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(X_i - \frac{Y_k + Y_l}{2})} \right. \\ & \left. + \sum_{k=1}^n \sum_{i=1}^n \sum_{\substack{j=1 \\ i < j}}^n I_{(\frac{X_i + X_j}{2} - Y_k)} + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{X_i + X_j}{2} - \frac{Y_k + Y_l}{2})} \right\} \\ & = -W(Y_1, \dots, Y_n; X_1, \dots, X_n) \end{aligned}$$

and the statistic $W(\cdot)$ is negative interchangeable. By Corollary 1.3.27 of Randles and Wolfe(1979), $W(\cdot)$ is symmetrically distributed about zero. It follows that V is symmetrically distributed about $n^2(n^2 + 1)/4$ for any continuous distribution F .

2.3 The null distribution of V

Let Z_1, \dots, Z_m be continuous random variables with joint probability density function $f(z_1, \dots, z_m)$.

Then

$$P[Z_1 < Z_2 < \dots < Z_m] = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{z_3} \int_{-\infty}^{z_2} f(z_1, z_2, \dots, z_m) dz_1 dz_2 \dots dz_m.$$

If the marginal density of Z_i 's is $N(0, 1)$, the probabilities $P[Z_1 < Z_2 < \dots < Z_m]$ play a key role in calculation of $P[V = v]$. Making the transformation $U_i = Z_{i+1} - Z_i$ for $i = 1, 2, \dots, m-1$, these probabilities become the orthant probability,

$$P_{n-1} = P[U_1 > 0, U_2 > 0, \dots, U_{n-1} > 0].$$

In which $U = (U_1, U_2, \dots, U_{n-1})$ has a multivariate normal distribution with zero means and correlation coefficient ρ_{ij} . The well known expressions for the first four P_{n-1} are given below, see McFadden(1960), Abrahamson(1964), and Childs(1967):

$$P_1 = \frac{1}{2}$$

$$P_2 = \frac{1}{4} + \frac{1}{2\pi} \sin^{-1} \rho_{12}$$

$$P_3 = \frac{1}{8} + \frac{1}{4\pi} (\sin^{-1} \rho_{12} + \sin^{-1} \rho_{23} + \sin^{-1} \rho_{13})$$

$$P_4 = \frac{1}{16} + \frac{1}{8\pi} \sum_{i < j} \sin^{-1}(\rho_{ij}) + \frac{1}{4\pi^2} \{ \sin^{-1} \rho_{12} \sin^{-1} \rho_{34} + \rho_{12} \rho_{34} \sqrt{(1 - \rho_{12}^2)(1 - \rho_{34}^2)} \times [\frac{\rho_{23}^2}{2!} + \frac{\rho_{23}^4}{4!} (1 + 2\rho_{12}^2)(1 + 2\rho_{34}^2) + \frac{\rho_{23}^6}{6!} (3 + 4\rho_{12}^2 + 8\rho_{12}^4)(3 + 4\rho_{34}^2 + 8\rho_{34}^4) + \dots] \}.$$

The orthant probability that all the U_i 's are simultaneously positive does not have a general closed expression for $n \geq 4$. Although Childs(1967) and Owen(1985) used a

general reduction method for some multivariate integrals, but the integral methods could not give the exactly value for these orthant probability.

For example, let X and Y_1, Y_2, Y_3 be two independent random samples from continuous distributions with same distribution F . The probabilities of $P(V = v)$, $v = 0, 1, \dots, 6$ are:

$$P(V = 0) = P(Y_1 < X, Y_2 < X, Y_3 < X) = 3!P(Y_1 < Y_2 < Y_3 < X) = 1/4$$

$$P(V = 1) = 3P(Y_1 > X, Y_2 < X, Y_3 < X, Y_1 - X < X - Y_2, Y_1 - X < X - Y_3) \\ = 3!P\left(\frac{Y_1+Y_2}{2} < \frac{Y_1+Y_3}{2} < X < Y_1\right)$$

$$P(V = 2) = 3!P(Y_2 < Y_3 < X < Y_1, X - Y_3 < Y_1 - X < X - Y_2) \\ = 3!P\left(Y_3 < \frac{Y_1+Y_2}{2} < X < \frac{Y_1+Y_3}{2}\right) + 3!P\left(\frac{Y_1+Y_2}{2} < Y_3 < X < \frac{Y_1+Y_3}{2}\right)$$

$$P(V = 3) = 3!P(Y_2 < Y_3 < X < Y_1, X - Y_3 < X - Y_2 < Y_1 - X) \\ + 3!P(Y_3 < X < Y_1 < Y_2, Y_1 - X < Y_2 - X < X - Y_3) \\ = 3!P\left(Y_2 < Y_3 < X < \frac{Y_1+Y_2}{2}\right) + 3!P\left(\frac{Y_2+Y_3}{2} < X < Y_1 < Y_2\right)$$

$$P(V = 4) = 3!P(Y_3 < X < Y_1 < Y_2, Y_1 - X < X - Y_3 < Y_2 - X) \\ = 3!P\left(\frac{Y_1+Y_3}{2} < X < \frac{Y_2+Y_3}{2} < Y_1\right) + 3!P\left(\frac{Y_1+Y_3}{2} < X < Y_1 < \frac{Y_2+Y_3}{2}\right)$$

$$P(V = 5) = 3!P(Y_3 < X < Y_1 < Y_2, X - Y_3 < Y_1 - X < Y_2 - X) \\ = 3!P\left(Y_3 < X < \frac{Y_1+Y_3}{2} < \frac{Y_2+Y_3}{2}\right)$$

$$P(V = 6) = 3!P(X < Y_1 < Y_2 < Y_3) = 1/4$$

Therefore, if the distribution is uniform $u(0,1)$, using above formula, then

$$P(V = 0) = 6 \int_0^1 \int_0^x \int_0^{y_3} \int_0^{y_2} dy_1 dy_2 dy_3 dx = \frac{1}{4}$$

$$P(V = 1) = 6 \int_0^1 \int_{y_3}^1 \int_0^{y_3} \int_{\frac{y_1+y_3}{2}}^{y_1} dx dy_2 dy_1 dy_3 = \frac{1}{8}$$

$$P(V = 2) = 6 \int_0^1 \int_0^{y_1} \int_{y_2}^{\frac{y_1+y_2}{2}} \int_{\frac{y_1+y_2}{2}}^{\frac{y_1+y_3}{2}} dx dy_3 dy_2 dy_1 \\ + 6 \int_0^1 \int_0^{y_1} \int_{\frac{y_1+y_2}{2}}^{y_1} \int_{y_3}^{\frac{y_1+y_3}{2}} dx dy_3 dy_2 dy_1 = \frac{1}{16}$$

Since uniform is symmetric distribution, by Theorem 2.1, V statistic is also symmetric.

$$P(V = 3) = \frac{1}{8}$$

$$P(V = 4) = P(V = 2) = \frac{1}{16}$$

$$P(V = 5) = P(V = 1) = \frac{1}{8}$$

$$P(V = 6) = P(V = 0) = \frac{1}{4}$$

Using those probabilities, the mean of V is

$$E(V) = \frac{1}{4} \cdot 0 + \frac{1}{8} \cdot 1 + \frac{1}{16} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{16} \cdot 4 + \frac{1}{8} \cdot 5 + \frac{1}{4} \cdot 6 = 3$$

The variance of V is

$$\begin{aligned} Var(V) &= \frac{1}{4}(0 - 3)^2 + \frac{1}{8}(1 - 3)^2 + \frac{1}{16}(2 - 3)^2 + \frac{1}{8}(3 - 3)^2 + \frac{1}{16}(4 - 3)^2 \\ &\quad + \frac{1}{8}(5 - 3)^2 + \frac{1}{4}(6 - 3)^2 = \frac{45}{8} \end{aligned}$$

If distribution is standard normal, then

$$P(V = 0) = \frac{1}{4}$$

$$\begin{aligned} P(V = 1) &= 6P(Y_3 - Y_2 > 0, X - \frac{Y_1 + Y_3}{2} > 0, Y_1 - X > 0) \\ &= 6P(T_1 > 0, T_2 > 0, T_3 > 0) \end{aligned}$$

where

$$T_1 = Y_3 - Y_2, T_2 = X - \frac{Y_1 + Y_3}{2}, T_3 = Y_1 - X$$

then

$$Var(T_1) = 2, Var(T_2) = \frac{3}{2}, Var(T_3) = 2$$

$$Cov(T_1, T_2) = Cov(Y_3, -\frac{Y_3}{2}) = -\frac{1}{2}$$

$$Cov(T_2, T_3) = Cov(X, -X) + Cov(-\frac{Y_1}{2}, Y_1) = -\frac{3}{2}$$

$$Cov(T_1, T_3) = 0$$

and

$$\rho_{12} = \frac{Cov(T_1, T_2)}{\sqrt{Var(T_1)Var(T_2)}} = -\frac{\sqrt{3}}{6}$$

$$\rho_{23} = \frac{Cov(T_2, T_3)}{\sqrt{Var(T_2)Var(T_3)}} = -\frac{\sqrt{3}}{2}$$

$$\rho_{13} = 0$$

Hence,

$$\begin{aligned}
P(V = 1) &= 6\left[\frac{1}{8} + \frac{1}{4\pi}(\sin^{-1}\rho_{12} + \sin^{-1}\rho_{23} + \sin^{-1}\rho_{13})\right] = 0.1102 \\
P(V = 2) &= 6P\left(\frac{Y_1 + Y_2}{2} - Y_3 > 0, X - \frac{Y_1 + Y_2}{2} > 0, \frac{Y_1 + Y_3}{2} - X > 0\right) \\
&\quad + 6P\left(Y_3 - \frac{Y_1 + Y_2}{2} > 0, X - Y_3 > 0, \frac{Y_1 + Y_3}{2} - X > 0\right) \\
&= 6P(T_1 > 0, T_2 > 0, T_3 > 0) + 6P(T'_1 > 0, T'_2 > 0, Y'_3 > 0)
\end{aligned}$$

where

$$\begin{aligned}
T_1 &= \frac{Y_1 + Y_2}{2} - Y_3, \quad T_2 = X - \frac{Y_1 + Y_2}{2}, \quad T_3 = \frac{Y_1 + Y_3}{2} - X \\
T'_1 &= Y_3 - \frac{Y_1 + Y_2}{2}, \quad T'_2 = X - Y_3, \quad T'_3 = \frac{Y_1 + Y_3}{2} - X
\end{aligned}$$

then

$$\begin{aligned}
\text{Var}(T_1) &= \frac{3}{2}, \quad \text{Var}(T_2) = \frac{3}{2}, \quad \text{Var}(T_3) = \frac{3}{2} \\
\text{Var}(T'_1) &= \frac{3}{2}, \quad \text{Var}(T'_2) = 2, \quad \text{Var}(T'_3) = \frac{3}{2} \\
\text{Cov}(T_1, T_2) &= -\frac{1}{2}, \quad \text{Cov}(T_2, T_3) = -\frac{5}{4}, \quad \text{Cov}(T_1, T_3) = -\frac{1}{4} \\
\text{Cov}(T'_1, T'_2) &= -1, \quad \text{Cov}(T'_2, T'_3) = -\frac{3}{2}, \quad \text{Cov}(T'_1, T'_3) = \frac{1}{4}
\end{aligned}$$

and

$$\begin{aligned}
\rho_{12} &= -\frac{1}{3}, \quad \rho_{23} = -\frac{5}{6}, \quad \rho_{13} = -\frac{1}{6}, \quad \rho'_{12} = -\frac{\sqrt{3}}{3}, \quad \rho'_{23} = -\frac{\sqrt{3}}{2}, \quad \rho'_{13} = \frac{1}{6} \\
P(V = 2) &= 6\left[\frac{1}{8} + \frac{1}{4\pi}(\sin^{-1}(-\frac{1}{3}) + \sin^{-1}(-\frac{5}{6}) + \sin^{-1}(-\frac{1}{6}))\right] \\
&\quad + 6\left[\frac{1}{8} + \frac{1}{4\pi}(\sin^{-1}(-\frac{\sqrt{3}}{3}) + \sin^{-1}(-\frac{\sqrt{3}}{2}) + \sin^{-1}(\frac{1}{6}))\right] = 0.0735
\end{aligned}$$

Similarly,

$$\begin{aligned}
P(V = 3) &= 0.1326 \\
P(V = 4) &= P(V = 2) = 0.0735 \\
P(V = 5) &= P(V = 1) = 0.1102 \\
P(V = 6) &= P(V = 0) = 0.25
\end{aligned}$$

Using those probabilities, the mean of V is

$$E(V) = 0.25 \cdot 0 + 0.1102 \cdot 1 + 0.0735 \cdot 2 + 0.1326 \cdot 3 + 0.0735 \cdot 4 + 0.1102 \cdot 5 + 0.25 \cdot 6 = 3$$

The variance of V is

$$\text{Var}(V) = 0.25(0 - 3)^2 + 0.1102(1 - 3)^2 + 0.0735(2 - 3)^2 + 0.1326(3 - 3)^2$$

$$+0.0735(4-3)^2 + 0.1102(5-3)^2 + 0.25(6-3)^2 = 5.5286$$

If distribution is exponential with unit mean, then

$$\begin{aligned}
P(V=0) &= 6 \int_0^\infty \int_0^x \int_0^{y_3} \int_0^{y_2} \exp(-y_1 - y_2 - y_3 - x) dy_1 dy_2 dy_3 dx = \frac{1}{4} \\
P(V=1) &= 6 \int_0^\infty \int_{y_3}^\infty \int_0^{y_3} \int_{\frac{y_1+y_3}{2}}^{y_1} \exp(-y_1 - y_2 - y_3 - x) dx dy_2 dy_1 dy_3 = \frac{1}{12} \\
P(V=2) &= 6 \int_0^\infty \int_0^{y_1} \int_{y_2}^{\frac{y_1+y_2}{2}} \int_{\frac{y_1+y_3}{2}}^{\frac{y_1+y_2}{2}} \exp(-y_1 - y_2 - y_3 - x) dx dy_3 dy_2 dy_1 \\
&\quad + 6 \int_0^\infty \int_0^{y_1} \int_{\frac{y_1+y_2}{2}}^{y_1} \int_{y_3}^{\frac{y_1+y_3}{2}} \exp(-y_1 - y_2 - y_3 - x) dx dy_3 dy_2 dy_1 = \frac{1}{24} \\
P(V=3) &= 6 \int_0^\infty \int_0^{y_1} \int_{y_2}^{\frac{y_1+y_2}{2}} \int_{y_3}^{\frac{y_1+y_2}{2}} \exp(-y_1 - y_2 - y_3 - x) dx dy_3 dy_2 dy_1 \\
&\quad + 6 \int_0^\infty \int_{y_2}^\infty \int_{y_2}^{\frac{y_1+y_2}{2}} \int_{\frac{y_2+y_3}{2}}^{y_1} \exp(-y_1 - y_2 - y_3 - x) dx dy_1 dy_3 dy_2 = \frac{3}{20} \\
P(V=4) &= 6 \int_0^\infty \int_0^{y_2} \int_{y_2}^{\frac{y_2+y_3}{2}} \int_{\frac{y_2+y_3}{2}}^{\frac{y_1+y_3}{2}} \exp(-y_1 - y_2 - y_3 - x) dx dy_1 dy_3 dy_2 \\
&\quad + 6 \int_0^\infty \int_0^{y_2} \int_{y_3}^{\frac{y_2+y_3}{2}} \int_{\frac{y_1+y_3}{2}}^{y_1} \exp(-y_1 - y_2 - y_3 - x) dx dy_1 dy_3 dy_2 = \frac{3}{40} \\
P(V=5) &= 6 \int_0^\infty \int_0^{y_2} \int_{y_3}^{\frac{y_2+y_3}{2}} \int_{y_3}^{\frac{y_1+y_3}{2}} \exp(-y_1 - y_2 - y_3 - x) dx dy_1 dy_3 dy_2 = \frac{3}{20} \\
P(V=6) &= \frac{1}{4}
\end{aligned}$$

Using those probabilities, the mean of V is

$$E(V) = \frac{1}{4} \cdot 0 + \frac{1}{12} \cdot 1 + \frac{1}{24} \cdot 2 + \frac{3}{20} \cdot 3 + \frac{3}{40} \cdot 4 + \frac{3}{20} \cdot 5 + \frac{1}{4} \cdot 6 = \frac{19}{6}$$

Table 2.1: Distributions, Means and Variances of V for the cases $m = 1, n = 2$ and $m = 1, n = 3$

V	m=1,n=2			m=1,n=3		
	U	N	E	U	N	E
0	1/3	1/3	1/3	1/4	0.25	1/4
1	1/6	1/6	1/9	1/8	0.1102	1/12
2	1/6	1/6	2/9	1/16	0.0735	1/24
3	1/3	1/3	1/3	1/8	0.1326	3/20
4				1/16	0.0735	3/40
5				1/8	0.1102	3/20
6				1/4	0.25	1/4
E(V)	3/2	3/2	14/9	3	3	19/6
Var(V)	19/12	19/12	128/81	45/8	5.5286	497/90

The variance of V is

$$\begin{aligned} Var(V) &= \frac{1}{4}(0 - \frac{19}{6})^2 + \frac{1}{12}(1 - \frac{19}{6})^2 + \frac{1}{24}(2 - \frac{19}{6})^2 + \frac{3}{24}(3 - \frac{19}{6})^2 + \frac{3}{40}(4 - \frac{19}{6})^2 \\ &\quad + \frac{3}{20}(5 - \frac{19}{6})^2 + \frac{1}{4}(6 - \frac{19}{6})^2 = \frac{497}{90}. \end{aligned}$$

In Table 2.1, we list distributions of V, means of V, and variances of V for the cases ($m = 1, n = 2$) and ($m = 1, n = 3$). The V statistic is symmetric if population distribution function F is normal or uniform. When $m \neq n$, the V statistic is not symmetric if population distribution function F is exponential.

2.4 Table value

In this section we give extensive tables of, simulated critical values of V for $m \leq n = 2(1)10$ and $m \leq n = 10(5)25$ and lower $\alpha = 0.001, 0.005, 0.01, 0.025, 0.05, 0.10$. In Table 2.2 we list some values of cumulative probabilities of $P(V = v)$ for small values m and n, and Table

2.3 we list critical values at $\alpha = 0.001, 0.005, 0.01, 0.025, 0.05, 0.10$ when F is uniform (0,1), standard normal and unit exponential. From section 2.3, we know that when $m \neq n$, the V statistic is not symmetric if population distribution F is not symmetric, see Tables 2.1, 2.2 and 2.3. Table 2.4 provides critical values for V for selected values of m and n, $m = n$, and F is cauchy or double exponential. They are the simulation results using 1,000,000 replications.

Tables 2.1 to 2.3 and Figure 2.1 to 2.4 show that if distribution F is uniform or normal, statistic V is symmetric. If distribution F is exponential, the statistic V is not symmetric for small sample size $m < n$. However, statistic V is symmetric for $m = n$ when distribution F is exponential.

Table 2.4 lists critical values v at $\alpha = 0.001, 0.005, 0.01, 0.025, 0.05, 0.10$ when F is cauchy and double exponential. The values are observed from simulation with 1,000,000 replications. These values of individual statistic values v are very close to the values of uniform, normal, and exponential distributions at each single point α .

Table 2.2: Cumulative Probabilities of Statistic V for Small Values m and n (U: Uniform, N: Normal, E: Exponential, E*: Negative Exponential)

v	U	N	E	E*	v	U	N	E	E*	v	U	N	E	E*
m=2, n=3					m=2, n=9					m=2, n=15				
0	0.099897	0.099496	0.099662	0.100434	0	0.018188	0.018333	0.018471	0.018215	0	0.007298	0.007440	0.007169	0.007458
1	0.133446	0.136475	0.133315		1	0.024307	0.024580	0.024494	0.024349	1	0.009836	0.009990	0.009507	0.009941
m=2, n=4					m=2, n=10					m=2, n=15				
0	0.066775	0.066660	0.066837	0.06659	0	0.015137	0.015284	0.015340	0.015227	0	0.011932	0.011555	0.011632	0.011677
1	0.088742	0.090853	0.088971	0.08886	1	0.020201	0.020594	0.020436	0.020257	1	0.018445	0.018404	0.018339	0.018396
2	0.107158	0.108579	0.103010	0.11083	2	0.024328	0.024201	0.023624	0.024798	2	0.011901	0.011712	0.011053	0.012056
m=2, n=5					m=2, n=11					m=2, n=15				
0	0.047408	0.047774	0.047675	0.04760	0	0.032331	0.030486	0.028573	0.033576	0	0.015961	0.014546	0.013416	0.016237
1	0.063264	0.064861	0.063545	0.06354	1	0.037235	0.034377	0.031671	0.038799	1	0.018318	0.016321	0.014906	0.018698
2	0.076526	0.077044	0.073611	0.07866	2	0.040413	0.037705	0.034265	0.042074	2	0.019821	0.017825	0.016198	0.020334
3	0.101657	0.099174	0.089446	0.10953	3	0.044520	0.041550	0.036983	0.046369	3	0.021766	0.019537	0.017565	0.022309
4	0.117091	0.113616	0.099131		4	0.049564	0.045753	0.039966	0.052026					
5	0.127248	0.126090	0.110186		5	0.053558	0.049131	0.042370	0.056418					
m=2, n=6					m=2, n=12					m=2, n=15				
0	0.035725	0.036061	0.035621	0.035579	0	0.027624	0.027889	0.027869	0.027749	0	0.007298	0.007440	0.007169	0.007458
1	0.047619	0.048595	0.047523	0.047286	1	0.036870	0.037620	0.037094	0.036903	1	0.009836	0.009990	0.009507	0.009941
2	0.057760	0.057410	0.054941	0.058412	2	0.044519	0.044368	0.042886	0.045294	2	0.011901	0.011712	0.011053	0.012056
3	0.077146	0.073435	0.066911	0.080641	3	0.059477	0.056448	0.052167	0.062271	3	0.015961	0.014546	0.013416	0.016237
4	0.088607	0.083684	0.074292	0.093614	4	0.068144	0.064098	0.057941	0.072225	4	0.018318	0.016321	0.014906	0.018698
5	0.096107	0.092429	0.080480	0.100886	5	0.074174	0.070703	0.062586	0.078062	5	0.019821	0.017825	0.016198	0.020334
6	0.105512	0.102531	0.088423		6	0.081409	0.078188	0.067649	0.085901	6	0.021766	0.019537	0.017565	0.022309
7	0.117935	0.114007	0.099542		7	0.090604	0.086661	0.073878	0.096656					
8	0.127695	0.123161	0.108290		8	0.098141	0.093535	0.081034	0.104805					
m=2, n=7					m=2, n=13					m=2, n=15				
0	0.027624	0.027889	0.027869	0.027749	0	0.032331	0.030486	0.028573	0.033576	0	0.007298	0.007440	0.007169	0.007458
1	0.036870	0.037620	0.037094	0.036903	1	0.037235	0.034377	0.031671	0.038799	1	0.009836	0.009990	0.009507	0.009941
2	0.044519	0.044368	0.042886	0.045294	2	0.040413	0.037705	0.034265	0.042074	2	0.011901	0.011712	0.011053	0.012056
3	0.059477	0.056448	0.052167	0.062271	3	0.044520	0.041550	0.036983	0.046369	3	0.015961	0.014546	0.013416	0.016237
4	0.068144	0.064098	0.057941	0.072225	4	0.049564	0.045753	0.039966	0.052026	4	0.018318	0.016321	0.014906	0.018698
5	0.074174	0.070703	0.062586	0.078062	5	0.053558	0.049131	0.042370	0.056418	5	0.019821	0.017825	0.016198	0.020334
6	0.081409	0.078188	0.067649	0.085901	6	0.057934	0.052878	0.044881	0.061494	6	0.021766	0.019537	0.017565	0.022309
7	0.090604	0.086661	0.073878	0.096656	7	0.063392	0.057169	0.047854	0.068013					
8	0.098141	0.093535	0.081034	0.104805	8	0.067809	0.061108	0.051347	0.072741					
9	0.106543	0.100928	0.087647		9	0.071946	0.064643	0.054392	0.077330					
10	0.116815	0.109519	0.093392		10	0.075721	0.068134	0.056871	0.081411					
11	0.125077	0.117616	0.099881		11	0.080016	0.072039	0.059830	0.086427					
12	0.132829	0.125213	0.105656		12	0.084080	0.075891	0.062593	0.090861					
m=2, n=8					m=2, n=14					m=2, n=15				
0	0.022276	0.022154	0.022154	0.022238	0	0.032331	0.030486	0.028573	0.033576	0	0.007298	0.007440	0.007169	0.007458
1	0.029764	0.029767	0.029715	0.029703	1	0.037235	0.034377	0.031671	0.038799	1	0.009836	0.009990	0.009507	0.009941
2	0.035975	0.035066	0.034389	0.036486	2	0.040413	0.037705	0.034265	0.042074	2	0.011901	0.011712	0.011053	0.012056
3	0.047846	0.044481	0.041696	0.049900	3	0.044520	0.041550	0.036983	0.046369	3	0.015961	0.014546	0.013416	0.016237
4	0.054986	0.050384	0.046313	0.057710	4	0.049564	0.045753	0.039966	0.052026	4	0.018318	0.016321	0.014906	0.018698
5	0.059756	0.055531	0.050173	0.062459	5	0.053558	0.049131	0.042370	0.056418	5	0.019821	0.017825	0.016198	0.020334
6	0.065507	0.061310	0.054153	0.068574	6	0.057934	0.052878	0.044881	0.061494	6	0.021766	0.019537	0.017565	0.022309
7	0.072927	0.067833	0.058570	0.077232	7	0.063392	0.057169	0.047854	0.068013					
8	0.078731	0.073191	0.062733	0.083699	8	0.067809	0.061108	0.051347	0.072741					
9	0.085389	0.079017	0.068235	0.091322	9	0.071946	0.064643	0.054392	0.077330					
10	0.093837	0.085778	0.073608	0.101558	10	0.075721	0.068134	0.056871	0.081411					
11	0.100312	0.091946	0.077840		11	0.080016	0.072039	0.059830	0.086427					
12	0.106388	0.097665	0.082533		12	0.084080	0.075891	0.062593	0.090861					
13	0.111932	0.103155	0.086722		13	0.087840	0.079419	0.065426	0.095019					
14	0.118445	0.109404	0.091413		14	0.091627	0.082973	0.067980	0.099183					
15	0.124364	0.115842	0.096179		15	0.096509	0.087143	0.070890	0.104921					
16	0.129968	0.121674	0.101077		16	0.100570	0.091022	0.073835						

Table 2.3: Critical Values for V (U: Uniform, N: Normal, E: Exponential, E*: Negative Exponential)

		m=2 n=3	m=2 n=4	m=2 n=5	m=2 n=6	m=2 n=7	m=2 n=8	m=2 n=9	m=2 n=10	m=2 n=15	m=2 n=20	m=2 n=25	m=3 n=3	
.001	U													
	N													
	E													
	E*													
.005	U												2	
	N												2	
	E												2	
	E*												2	
.01	U									1	3	7		
	N									1	4	10		
	E									1	5	12		
	E*									1	3	7		
.025	U						0	1	2	7	15	26		
	N						0	1	2	9	20	36		
	E						0	1	2	11	27	50		
	E*						0	1	2	7	14	25		
.05	U			0	1	2	3	5	7	19	37	61	0	
	N			0	1	2	3	5	8	24	47	79	0	
	E			0	1	2	4	7	10	33	66	109	0	
	E*			0	1	2	3	4	6	18	35	57	0	
.10	U	0	1	2	5	8	10	14	18	45	83	134	3	
	N	0	1	3	5	8	12	16	21	52	98	157	2	
	E	0	1	4	7	11	15	21	27	68	126	202	3	
	E*		1	2	4	7	9	13	17	41	76	123	3	
		m=3 n=4	m=3 n=5	m=3 n=6	m=3 n=7	m=3 n=8	m=3 n=9	m=3 n=10	m=3 n=15	m=3 n=20	m=3 n=25	m=4 n=4	m=4 n=5	m=4 n=6
.001	U								2	7				
	N								2	8				
	E								2	9				
	E*								2	6				
.005	U					0	1	10	24	45				0
	N					0	1	11	29	56				0
	E					0	1	13	38	74				0
	E*					0	1	9	23	43				0
.01	U			0	2	3	5	22	48	83		0	3	
	N			0	2	3	6	26	58	102		0	3	
	E			0	2	4	7	32	77	138		0	3	
	E*			0	2	3	5	20	45	79		0	3	
.025	U	1	3	6	10	14	19	55	107	176	2	7	14	
	N	1	3	6	10	15	21	62	124	207	2	6	13	
	E	1	3	7	12	19	26	79	162	269	2	7	15	
	E*	1	3	5	9	13	18	50	100	165	2	7	13	
.05	U	2	5	10	16	22	30	39	100	188	303	8	16	28
	N	2	5	10	16	23	31	41	108	206	335	7	16	28
	E	2	6	12	19	28	38	50	131	249	404	8	17	29
	E*	2	5	9	14	20	27	35	91	173	279	8	16	26
.10	U	7	14	22	32	43	57	72	172	317	505	18	32	49
	N	7	14	22	32	44	58	73	178	329	523	17	31	49
	E	8	14	24	35	49	65	83	200	370	592	17	32	51
	E*	6	13	20	29	40	52	65	158	291	462	17	30	47

		m = 4 n = 7	m = 4 n = 8	m = 4 n = 9	m = 4 n = 10	m = 4 n = 15	m = 4 n = 20	m = 4 n = 25	m = 5 n = 5	m = 5 n = 6	m = 5 n = 7	m = 5 n = 8	m = 5 n = 9	m = 5 n = 10
.001	U					9	29	59					3	7
	N					10	33	69					3	6
	E					12	40	92					3	7
	E*					9	30	58					3	6
.005	U	2	5	8	13	50	109	193	0	4	10	19	29	42
	N	2	5	8	13	55	125	222	0	4	9	18	29	42
	E	2	5	10	16	70	159	291	0	4	11	20	33	48
	E*	2	5	8	12	47	105	180	0	4	10	18	27	39
.01	U	7	13	20	28	87	175	297	5	12	23	37	52	72
	N	7	13	20	29	94	196	333	5	12	22	34	52	72
	E	8	15	24	34	118	246	422	5	13	24	39	58	80
	E*	7	12	19	26	80	166	277	5	12	22	34	49	65
.025	U	23	34	46	61	164	315	514	17	32	50	73	99	129
	N	22	34	47	63	173	339	554	16	31	49	71	98	130
	E	25	38	54	72	204	399	656	17	32	52	77	107	142
	E*	21	31	43	56	151	292	479	17	30	47	68	93	120
.05	U	42	59	79	101	254	474	762	33	55	81	114	150	192
	N	42	59	80	103	261	493	794	32	53	80	112	150	192
	E	46	65	88	114	293	553	896	32	55	84	119	159	205
	E*	39	55	73	93	235	442	709	32	52	78	108	143	182
.10	U	71	97	126	158	379	695	1106	56	87	124	168	219	277
	N	71	96	126	159	382	704	1120	56	86	124	168	219	277
	E	75	103	134	170	411	759	1210	55	88	128	175	228	288
	E*	67	91	119	150	357	656	1044	55	85	121	164	213	268
		m = 5 n = 15	m = 5 n = 20	m = 5 n = 25	m = 6 n = 6	m = 6 n = 7	m = 6 n = 8	m = 6 n = 9	m = 6 n = 10	m = 6 n = 15	m = 6 n = 20	m = 6 n = 25	m = 7 n = 7	m = 7 n = 8
.001	U	43	105	197		2	7	14	24	108	253	460	8	20
	N	44	117	221		2	7	14	23	114	272	503	8	18
	E	51	151	281		2	7	15	26	137	322	601	8	20
	E*	42	101	191		2	7	13	24	102	237	433	8	19
.005	U	135	279	479	12	25	43	64	88	274	551	936	47	78
	N	144	312	527	12	24	41	61	88	278	587	998	45	74
	E	170	369	642	13	26	44	67	96	321	670	1148	47	78
	E*	126	265	451	13	24	41	60	84	256	522	885	47	74
.01	U	204	406	684	27	47	72	101	137	385	753	1257	79	118
	N	215	438	729	25	45	69	99	134	389	782	1302	75	115
	E	251	508	860	26	47	73	106	146	437	872	1467	77	120
	E*	192	384	639	26	45	68	95	128	357	708	1179	78	114
.025	U	340	646	1059	57	88	127	171	223	580	1101	1799	136	194
	N	349	673	1092	54	86	124	169	220	584	1121	1828	133	190
	E	386	747	1226	56	88	129	177	234	631	1213	1992	134	196
	E*	317	607	995	56	85	121	163	213	550	1046	1710	134	189
.05	U	480	894	1443	89	133	184	243	311	775	1446	2330	195	270
	N	485	914	1465	88	131	182	242	310	778	1456	2346	194	270
	E	523	982	1587	89	133	188	251	324	823	1541	2494	195	275
	E*	454	847	1366	89	130	178	236	302	748	1392	2246	195	269
.10	U	663	1215	1935	134	192	260	338	427	1022	1879	2994	275	371
	N	666	1225	1946	133	192	259	338	427	1028	1887	3010	275	372
	E	699	1287	2054	135	195	266	346	440	1066	1957	3130	276	377
	E*	640	1173	1866	135	191	258	334	422	1007	1839	2930	276	373

		m = 7 n = 9	m = 7 n = 10	m = 7 n = 15	m = 7 n = 20	m = 7 n = 25	m = 8 n = 8	m = 8 n = 9	m = 8 n = 10	m = 8 n = 15	m = 8 n = 20	m = 8 n = 25	m = 9 n = 9	m = 9 n = 10
.001	U	34	56	216	483	840	37	64	96	361	784	1409	105	157
	N	34	50	215	496	895	36	63	95	361	809	1449	103	152
	E	37	60	248	576	1048	38	67	103	406	915	1655	107	155
	E*	34	54	199	452	814	38	65	97	336	730	1339	107	152
.005	U	113	155	465	941	1562	121	174	239	702	1400	2371	254	340
	N	109	150	463	950	1617	117	168	232	698	1430	2429	242	336
	E	117	163	517	1065	1805	119	176	245	758	1553	2644	250	340
	E*	108	146	431	879	1480	119	170	230	672	1348	2265	249	332
.01	U	166	222	619	1213	1984	177	245	326	904	1775	2934	344	455
	N	161	217	619	1230	2044	172	240	322	904	1791	2988	335	449
	E	169	231	671	1345	2250	174	249	336	965	1911	3211	341	456
	E*	159	210	584	1150	1907	173	242	318	876	1708	2829	341	449
.025	U	262	337	878	1678	2725	275	369	479	1242	2369	3853	499	648
	N	258	337	883	1690	2754	271	366	477	1246	2386	3904	494	645
	E	267	352	936	1804	2952	275	375	492	1307	2506	4101	502	656
	E*	255	328	849	1611	2628	275	366	475	1222	2319	3755	502	646
.05	U	358	456	1134	2122	3421	374	496	630	1570	2938	4727	653	834
	N	356	456	1140	2139	3439	372	492	630	1579	2955	4762	652	837
	E	365	471	1190	2240	3620	377	502	646	1630	3062	4942	660	849
	E*	353	449	1113	2072	3340	377	496	630	1557	2898	4650	659	840
.10	U	483	608	1460	2687	4285	503	655	825	1983	3647	5810	850	1073
	N	483	609	1469	2703	4298	503	654	826	1991	3661	5842	852	1079
	E	492	624	1508	2785	4448	508	663	840	2034	3757	5992	861	1089
	E*	482	608	1447	2655	4226	507	658	829	1977	3619	5764	859	1083
		m = 9 n = 15	m = 9 n = 20	m = 9 n = 25	m = 10 n = 10	m = 10 n = 15	m = 10 n = 20	m = 10 n = 25	m = 15 n = 15	m = 15 n = 20	m = 15 n = 25	m = 20 n = 20	m = 20 n = 25	m = 25 n = 25
.001	U	558	1196	2083	233	782	1674	2875	2580	5452	9315	11504	19552	33837
	N	552	1188	2134	217	773	1697	2955	2592	5416	9404	11403	19886	34202
	E	593	1347	2396	223	834	1853	3263	2585	5610	9815	11503	20225	34394
	E*	523	1110	1986	221	757	1603	2755	2609	5387	9224	11624	19628	34596
.005	U	996	1988	3333	459	1346	2663	4465	3860	7648	12804	15310	25726	42971
	N	993	2012	3382	449	1328	2682	4512	3864	7683	12915	15317	25838	43138
	E	1050	2164	3679	460	1400	2837	4816	3868	7882	13300	15506	26160	43430
	E*	970	1916	3208	459	1308	2584	4324	3865	7665	12736	15512	25818	43620
.01	U	1257	2448	4043	601	1660	3220	5336	4537	8858	14635	17346	28819	47662
	N	1247	2468	4084	592	1653	3248	5389	4563	8857	14724	17407	28900	47865
	E	1309	2615	4369	604	1725	3380	5677	4574	9079	15108	17580	29261	48215
	E*	1226	2385	3931	604	1629	3163	5213	4573	8908	14616	17574	28939	48377
.025	U	1673	3179	5192	838	2172	4135	6723	5636	10734	17508	20561	33617	54948
	N	1681	3213	5231	830	2175	4149	6774	5662	10772	17581	20621	33735	55076
	E	1735	3352	5473	846	2243	4282	7041	5691	10946	17923	20818	34103	55429
	E*	1659	3138	5095	847	2157	4079	6635	5702	10820	17554	20785	33785	55553
.05	U	2084	3886	6263	1068	2662	4983	8022	6676	12482	20145	23467	37956	61440
	N	2091	3910	6303	1064	2664	4998	8053	6683	12515	20226	23536	38077	61539
	E	2136	4038	6511	1080	2725	5116	8305	6717	12673	20503	23677	38393	61915
	E*	2073	3861	6188	1081	2656	4951	7967	6724	12555	20188	23686	38185	61928
.10	U	2589	4756	7587	1358	3271	6022	9618	7924	14586	23322	26959	43211	69167
	N	2601	4773	7621	1358	3277	6039	9647	7926	14613	23381	27036	43284	69275
	E	2639	4881	7789	1371	3330	6141	9844	7972	14749	23606	27151	43539	69611
	E*	2585	4742	7537	1372	3270	6003	9573	7976	14672	23357	27153	43401	69585

Table 2.4: Critical Values for V When $m = n$ (C: Cauchy, D: Double Exponential)

		m = 3 n = 3	m = 4 n = 4	m = 5 n = 5	m = 6 n = 6	m = 7 n = 7	m = 8 n = 8	m = 9 n = 9	m = 10 n = 10	m = 15 n = 15
.001	C					7	34	98	208	2545
	D					8	34	101	211	2567
.005	C			0	11	42	113	237	448	3855
	D			0	11	44	113	239	447	3841
.01	C			4	24	72	167	331	590	4580
	D			4	25	74	168	332	589	4567
.025	C		2	15	52	130	268	495	843	5733
	D		2	16	54	131	270	497	833	5687
.05	C		7	31	86	193	373	659	1083	6767
	D		7	31	87	192	371	654	1069	6711
.10	C	2	16	54	133	276	508	862	1378	8025
	D	2	17	55	134	274	504	855	1364	7959

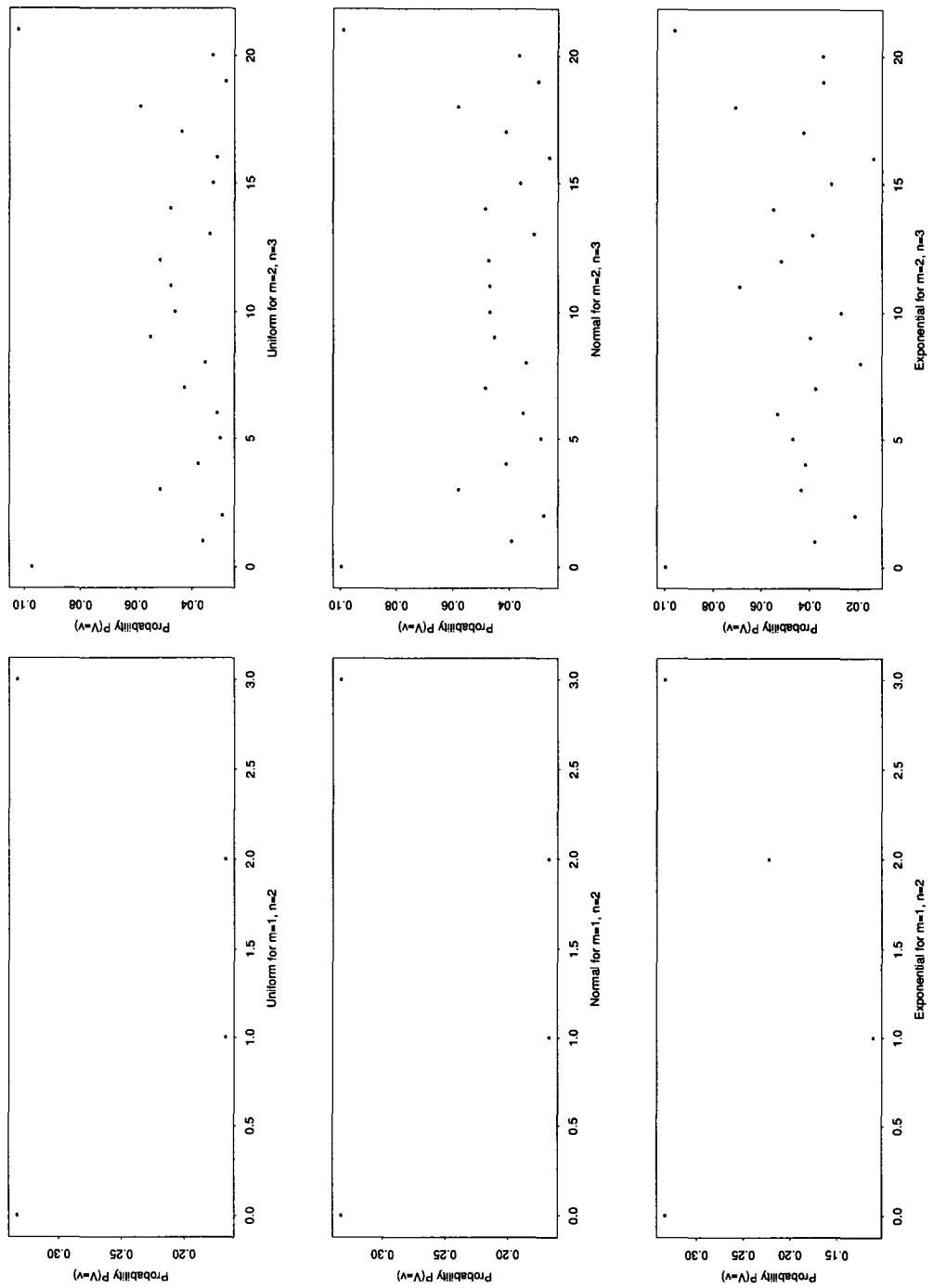


Figure 2.1: A plot of distribution of V for different m and n (Exact values when $m=1, n=2$, Simulated values when $n=2, n=3$)

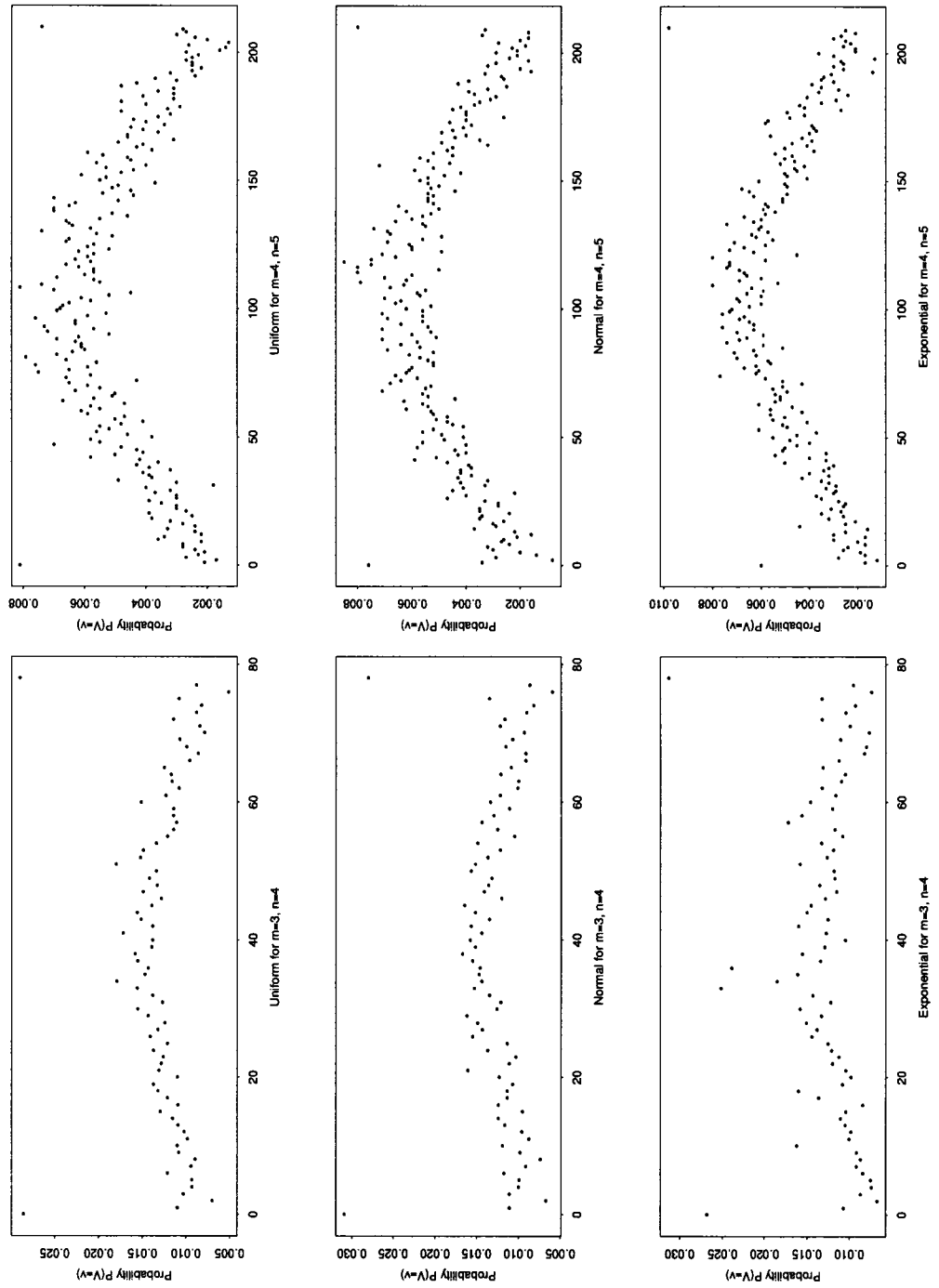


Figure 2.2: A plot of distribution of V (simulated) for different m and n

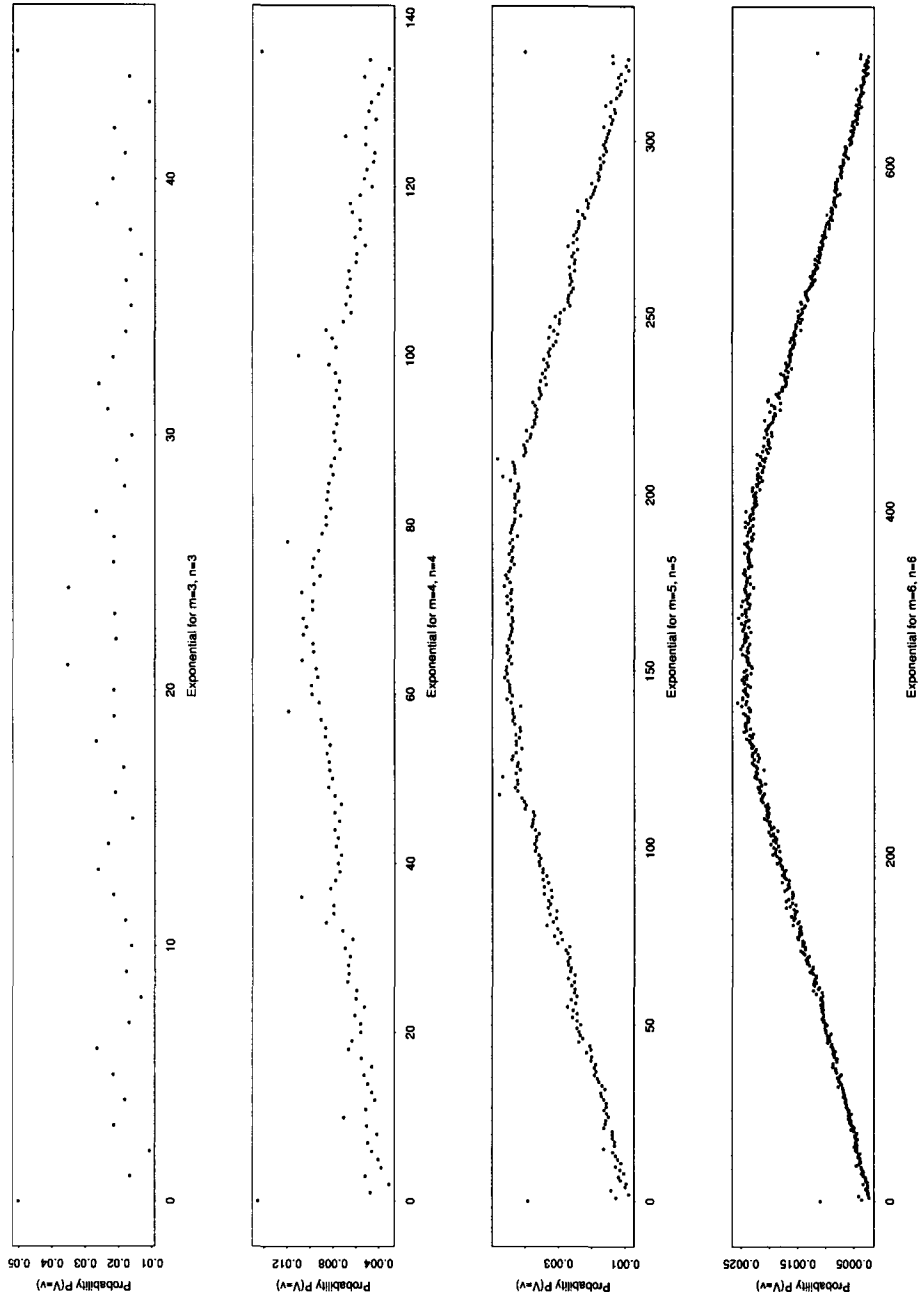


Figure 2.3: A plot of distribution of V (simulated) for $m = n$

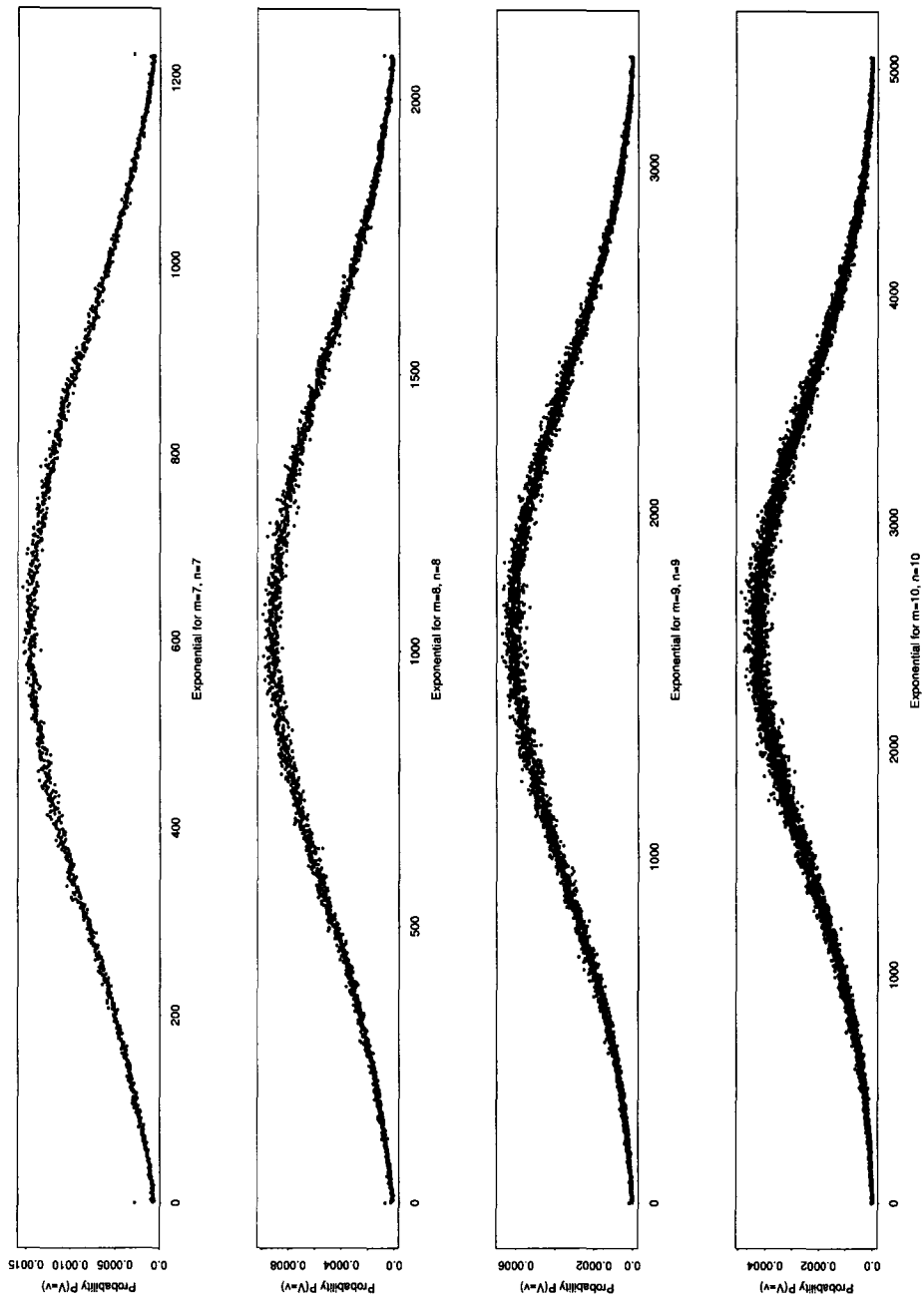


Figure 2.4: A plot of distribution of V (simulated) for $m = n$

Chapter 3

Mean and Variance

3.1 Expectations and variances

It has been defined in that

$$\begin{aligned}
 I_1 &= \sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)} \\
 I_2 &= \sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{(\frac{Y_k + Y_l}{2} - X_i)} \\
 I_3 &= \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m I_{(Y_k - \frac{X_i + X_j}{2})} \\
 I_4 &= \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{(\frac{Y_k + Y_l}{2} - \frac{X_i + X_j}{2})}
 \end{aligned}$$

Then, the mean of I_1 is

$$EI_1 = \sum_{i=1}^m \sum_{k=1}^n EI_{(Y_k - X_i)} = \sum_{i=1}^m \sum_{k=1}^n P(X_i < Y_k) \quad (3.1)$$

In order to get the variance of I_1 , we list components of I_1 as four separate parts in Table 3.1. The covariances of $I_{(Y_k - X_i)}$ with each individual component parts have the same pattern. Clearly, the general case of the variances of I_1 is:

Table 3.1: Components of I_1

$I_{(Y_k - X_i)}$	$I_{(Y_l - X_i)}$ $l \neq k$
$I_{(Y_k - X_j)}$ $j \neq i$	$I_{(Y_l - X_j)}$ $j \neq i, l \neq k$

$$\begin{aligned}
Var I_1 &= Var\left(\sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)}\right) \\
&= \sum_{i=1}^m \sum_{k=1}^n Var I_{(Y_k - X_i)} + \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k \neq l}}^n Cov(I_{(Y_k - X_i)}, I_{(Y_l - X_i)}) \\
&\quad + \sum_{i=1}^m \sum_{\substack{j=1 \\ i \neq j}}^m \sum_{k=1}^n Cov(I_{(Y_k - X_i)}, I_{(Y_k - X_j)}) + \sum_{i=1}^m \sum_{\substack{j=1 \\ i \neq j}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k \neq l}}^n Cov(I_{(Y_k - X_i)}, I_{(Y_l - X_j)}) \\
&= \sum_{i=1}^m \sum_{k=1}^n [P(X_i < Y_k) - P(X_i < Y_k)^2] \\
&\quad + \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k \neq l}}^n [P(X_i < Y_k, X_i < Y_l) - P(X_i < Y_k)P(X_i < Y_l)] \\
&\quad + \sum_{i=1}^m \sum_{\substack{j=1 \\ i \neq j}}^m \sum_{k=1}^n [P(X_i < Y_k, X_j < Y_k) - P(X_i < Y_k)P(X_j < Y_k)] \\
&\quad + \sum_{i=1}^m \sum_{\substack{j=1 \\ i \neq j}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k \neq l}}^n [P(X_i < Y_k, X_j < Y_l) - P(X_i < Y_k)P(X_j < Y_l)] \quad (3.2)
\end{aligned}$$

For I_2 , we have that

$$EI_2 = \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n EI_{(Y_{\frac{k+l}{2}} - X_i)} = \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n P(X_i < \frac{Y_k + Y_l}{2}) \quad (3.3)$$

Similarly, the components of I_2 are listed in Table 3.2. The covariances of $I_{(Y_{\frac{k+l}{2}} - X_i)}$, $k < l$, with each individual component of I_2 have the same pattern.

Table 3.2: Components of I_2

$I_{(\frac{Y_k+Y_l}{2}-X_i)}$	$I_{(\frac{Y_k+Y_{l'}}{2}-X_i)}$ $l' \neq (k, l)$	$I_{(\frac{Y_{k'}+Y_l}{2}-X_i)}$ $k' \neq (k, l)$	$I_{(\frac{Y_{k'}+Y_{l'}}{2}-X_i)}$ $(k', l') \neq (k, l), k' < l'$
$I_{(\frac{Y_k+Y_l}{2}-X_j)}$ $j \neq i$	$I_{(\frac{Y_k+Y_{l'}}{2}-X_j)}$ $l' \neq (k, l), j \neq i$	$I_{(\frac{Y_{k'}+Y_l}{2}-X_j)}$ $k' \neq (k, l), j \neq i$	$I_{(\frac{Y_{k'}+Y_{l'}}{2}-X_j)}$ $(k', l') \neq (k, l), j \neq i$

the variance of I_2 is

$$\begin{aligned}
 Var I_2 &= Var\left(\sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{Y_k+Y_l}{2}-X_i)}\right) \\
 &= \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n Var I_{(\frac{Y_k+Y_l}{2}-X_i)} + \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{l'=1 \\ l' \neq (k, l)}}^n Cov[I_{(\frac{Y_k+Y_l}{2}-X_i)}, I_{(\frac{Y_k+Y_{l'}}{2}-X_i)}] \\
 &\quad + \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{k'=1 \\ k' \neq (k, l)}}^n Cov[I_{(\frac{Y_k+Y_l}{2}-X_i)}, I_{(\frac{Y_{k'}+Y_l}{2}-X_i)}] \\
 &\quad + \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{k'=1 \\ k' < l'}}^n \sum_{\substack{l'=1 \\ (k', l') \neq (k, l)}}^n Cov[I_{(\frac{Y_k+Y_l}{2}-X_i)}, I_{(\frac{Y_{k'}+Y_{l'}}{2}-X_i)}] \\
 &\quad + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ j \neq i, k < l}}^n Cov[I_{(\frac{Y_k+Y_l}{2}-X_i)}, I_{(\frac{Y_k+Y_l}{2}-X_j)}] \\
 &\quad + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ j \neq i, k < l}}^n \sum_{\substack{l'=1 \\ l' \neq (k, l)}}^n Cov[I_{(\frac{Y_k+Y_l}{2}-X_i)}, I_{(\frac{Y_k+Y_{l'}}{2}-X_{j'})}] \\
 &\quad + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ j \neq i, k < l}}^n \sum_{\substack{k'=1 \\ k' \neq (k, l)}}^n Cov[I_{(\frac{Y_k+Y_l}{2}-X_i)}, I_{(\frac{Y_{k'}+Y_l}{2}-X_{j'})}] \\
 &\quad + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ j \neq i, k < l}}^n \sum_{\substack{k'=1 \\ k' < l'}}^n \sum_{\substack{l'=1 \\ (k', l') \neq (k, l)}}^n Cov[I_{(\frac{Y_k+Y_l}{2}-X_i)}, I_{(\frac{Y_{k'}+Y_{l'}}{2}-X_{j'})}] \\
 &= \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n [P(X_i < \frac{Y_k+Y_l}{2}) - P(X_i < \frac{Y_k+Y_l}{2})^2]
 \end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{\substack{l'=1 \\ k < l \quad l' \neq (k,l)}}^n [P(X_i < \frac{Y_k+Y_l}{2}, X_i < \frac{Y_k+Y_{l'}}{2}) - P(X_i < \frac{Y_k+Y_l}{2})P(X_i < \frac{Y_k+Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{\substack{k'=1 \\ k < l \quad k' \neq (k,l)}}^n [P(X_i < \frac{Y_k+Y_l}{2}, X_i < \frac{Y_{k'}+Y_l}{2}) - P(X_i < \frac{Y_k+Y_l}{2})P(X_i < \frac{Y_{k'}+Y_l}{2})] \\
& + \sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{\substack{k'=1 \quad l'=1 \\ k < l \quad k' < l' \\ (k',l') \neq (k,l)}}^n [P(X_i < \frac{Y_k+Y_l}{2}, X_i < \frac{Y_{k'}+Y_{l'}}{2}) - P(X_i < \frac{Y_k+Y_l}{2})P(X_i < \frac{Y_{k'}+Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ j \neq i \quad k < l}}^n [P(X_i < \frac{Y_k+Y_l}{2}, X_j < \frac{Y_k+Y_l}{2}) - P(X_i < \frac{Y_k+Y_l}{2})P(X_j < \frac{Y_k+Y_l}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \quad l'=1 \\ j \neq i \quad k < l \quad l' \neq (k,l)}}^n [P(X_i < \frac{Y_k+Y_l}{2}, X_j < \frac{Y_k+Y_{l'}}{2}) - P(X_i < \frac{Y_k+Y_l}{2})P(X_j < \frac{Y_k+Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \quad k'=1 \\ j \neq i \quad k < l \quad k' \neq (k,l)}}^n [P(X_i < \frac{Y_k+Y_l}{2}, X_j < \frac{Y_{k'}+Y_l}{2}) - P(X_i < \frac{Y_k+Y_l}{2})P(X_j < \frac{Y_{k'}+Y_l}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \quad k'=1 \quad l'=1 \\ j \neq i \quad k < l \quad k' < l' \\ (k',l') \neq (k,l)}}^n [P(X_i < \frac{Y_k+Y_l}{2}, X_j < \frac{Y_{k'}+Y_{l'}}{2}) \\
& \quad - P(X_i < \frac{Y_k+Y_l}{2})P(X_j < \frac{Y_{k'}+Y_{l'}}{2})] \tag{3.4}
\end{aligned}$$

For I_3 , we have that

$$EI_3 = \sum_{k=1}^n \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m EI_{(Y_k - \frac{X_i+X_j}{2})} = \sum_{k=1}^n \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m P(\frac{X_i + X_j}{2} < Y_k) \tag{3.5}$$

The components of I_3 are listed in Table 3.3. The covariances of $I_{(Y_k - \frac{X_i+X_j}{2})}$, $i < j$, with each individual component of the seven parts of I_3 have the same pattern.

the variance of I_3 is

$$\begin{aligned}
Var I_3 & = Var(\sum_{k=1}^n \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m I_{(Y_k - \frac{X_i+X_j}{2})}) \\
& = \sum_{k=1}^n \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m Var I_{(Y_k - \frac{X_i+X_j}{2})} + \sum_{k=1}^n \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m \sum_{\substack{j'=1 \\ j' \neq (i,j)}}^m Cov[I_{(Y_k - \frac{X_i+X_j}{2})}, I_{(Y_k - \frac{X_i+X_{j'}}{2})}]
\end{aligned}$$

Table 3.3: Components of I_3

$I_{(Y_k - \frac{x_i + x_j}{2})}$	$I_{(Y_k - \frac{x_i + x_{j'}}{2})}$ $j' \neq (i, j)$	$I_{(Y_k - \frac{x_{i'} + x_j}{2})}$ $i' \neq (i, j)$	$I_{(Y_k - \frac{x_{i'} + x_{j'}}{2})}$ $(i', j') \neq (i, j), i' < j'$
$I_{(Y_l - \frac{x_i + x_j}{2})}$ $l \neq k$	$I_{(Y_l - \frac{x_i + x_{j'}}{2})}$ $j' \neq (i, j), l \neq k$	$I_{(Y_l - \frac{x_{i'} + x_j}{2})}$ $i' \neq (i, j), l \neq k$	$I_{(Y_l - \frac{x_{i'} + x_{j'}}{2})}$ $(i', j') \neq (i, j), i' < j', l \neq k$

$$\begin{aligned}
& + \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \text{Cov}[I_{(Y_k - \frac{x_i + x_j}{2})}, I_{(Y_k - \frac{x_{i'} + x_j}{2})}] \\
& \quad \quad \quad i < j \quad i' \neq (i, j) \\
& + \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \text{Cov}[I_{(Y_k - \frac{x_i + x_j}{2})}, I_{(Y_k - \frac{x_{i'} + x_{j'}}{2})}] \\
& \quad \quad \quad i < j \quad i' < j' \\
& \quad \quad \quad (i', j') \neq (i, j) \\
& + \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{j=1}^m \text{Cov}[I_{(Y_k - \frac{x_i + x_j}{2})}, I_{(Y_l - \frac{x_i + x_j}{2})}] \\
& \quad \quad \quad l \neq k \quad i < j \\
& + \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \text{Cov}[I_{(Y_k - \frac{x_i + x_j}{2})}, I_{(Y_l - \frac{x_i + x_{j'}}{2})}] \\
& \quad \quad \quad l \neq k \quad i < j \quad j' \neq (i, j) \\
& + \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \text{Cov}[I_{(Y_k - \frac{x_i + x_j}{2})}, I_{(Y_l - \frac{x_{i'} + x_j}{2})}] \\
& \quad \quad \quad l \neq k \quad i < j \quad i' \neq (i, j) \\
& + \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \text{Cov}[I_{(Y_k - \frac{x_i + x_j}{2})}, I_{(Y_l - \frac{x_{i'} + x_{j'}}{2})}] \\
& \quad \quad \quad l \neq k \quad i < j \quad i' < j' \\
& \quad \quad \quad (i', j') \neq (i, j) \\
& = \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m [P(\frac{x_i + x_j}{2} < Y_k) - P(\frac{x_i + x_j}{2} < Y_k)^2] \\
& \quad \quad \quad i < j \\
& + \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m [P(\frac{x_i + x_j}{2} < Y_k, \frac{x_i + x_{j'}}{2} < Y_k) - P(\frac{x_i + x_j}{2} < Y_k)P(\frac{x_i + x_{j'}}{2} < Y_k)] \\
& \quad \quad \quad i < j \quad j' \neq (i, j) \\
& + \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m [P(\frac{x_i + x_j}{2} < Y_k, \frac{x_{i'} + x_j}{2} < Y_k) - P(\frac{x_i + x_j}{2} < Y_k)P(\frac{x_{i'} + x_j}{2} < Y_k)] \\
& \quad \quad \quad i < j \quad i' \neq (i, j)
\end{aligned}$$

$$\begin{aligned}
& + \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m [P(\frac{X_i+X_j}{2} < Y_k, \frac{X_{i'}+X_{j'}}{2} < Y_k) - P(\frac{X_i+X_j}{2} < Y_k)P(\frac{X_{i'}+X_{j'}}{2} < Y_k)] \\
& \quad \quad \quad \begin{matrix} i < j & i' < j' \\ (i', j') \neq (i, j) \end{matrix} \\
& + \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{j=1}^m [P(\frac{X_i+X_j}{2} < Y_k, \frac{X_i+X_j}{2} < Y_l) - P(\frac{X_i+X_j}{2} < Y_k)P(\frac{X_i+X_j}{2} < Y_l)] \\
& \quad \quad \quad \begin{matrix} l \neq k & i < j \end{matrix} \\
& + \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m [P(\frac{X_i+X_j}{2} < Y_k, \frac{X_i+X_{j'}}{2} < Y_l) - P(\frac{X_i+X_j}{2} < Y_k)P(\frac{X_i+X_{j'}}{2} < Y_l)] \\
& \quad \quad \quad \begin{matrix} l \neq k & i < j & j' \neq (i, j) \end{matrix} \\
& + \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m [P(\frac{X_i+X_j}{2} < Y_k, \frac{X_{i'}+X_j}{2} < Y_l) - P(\frac{X_i+X_j}{2} < Y_k)P(\frac{X_{i'}+X_j}{2} < Y_l)] \\
& \quad \quad \quad \begin{matrix} l \neq k & i < j & i' \neq (i, j) \end{matrix} \\
& + \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m [P(\frac{X_i+X_j}{2} < Y_k, \frac{X_{i'}+X_{j'}}{2} < Y_l) - P(\frac{X_i+X_j}{2} < Y_k)P(\frac{X_{i'}+X_{j'}}{2} < Y_l)] \\
& \quad \quad \quad \begin{matrix} l \neq k & i < j & i' < j' \\ (i', j') \neq (i, j) \end{matrix}
\end{aligned} \tag{3.6}$$

For I_4 , we have that

$$\begin{aligned}
EI_4 & = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n EI_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})} \\
& \quad \quad \quad \begin{matrix} i < j & k < l \end{matrix} \\
& = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n P(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2}) \\
& \quad \quad \quad \begin{matrix} i < j & k < l \end{matrix}
\end{aligned} \tag{3.7}$$

The components of I_4 are listed in Table 3.4. The covariances of $I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})}$, $k < l$, $i < j$, with each individual component of I_4 have the same pattern. The variances of I_4 is

$$\begin{aligned}
\text{Var}(I_4) & = \text{Var} \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})} \\
& \quad \quad \quad \begin{matrix} i < j & k < l \end{matrix} \\
& = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \text{Var} I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})} \\
& \quad \quad \quad \begin{matrix} i < j & k < l \end{matrix} \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \text{Cov}(I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})}, I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_{j'}}{2})}) \\
& \quad \quad \quad \begin{matrix} i < j & j' \neq (i, j) & k < l \end{matrix} \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{l=1}^n \text{Cov}(I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})}, I_{(\frac{Y_k+Y_l}{2} - \frac{X_{i'}+X_j}{2})}) \\
& \quad \quad \quad \begin{matrix} i < j & i' \neq (i, j) & k < l \end{matrix}
\end{aligned}$$

Table 3.4: Components of I_4

$I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2}\right)}$	$I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_i+X_{j'}}{2}\right)}$ $j' \neq (i, j)$	$I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_{i'}+X_j}{2}\right)}$ $i' \neq (i, j)$	$I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}$ $(i', j') \neq (i, j), i' < j'$
$I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_i+X_j}{2}\right)}$ $l' \neq (k, l)$	$I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}$ $j' \neq (i, j)$ $l' \neq (k, l)$	$I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_{i'}+X_j}{2}\right)}$ $i' \neq (i, j)$ $l' \neq (k, l)$	$I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}$ $(i', j') \neq (i, j), i' < j'$ $l' \neq (k, l)$
$I_{\left(\frac{Y_{k'}+Y_l}{2} - \frac{X_i+X_j}{2}\right)}$ $k' \neq (k, l)$	$I_{\left(\frac{Y_{k'}+Y_l}{2} - \frac{X_i+X_{j'}}{2}\right)}$ $j' \neq (i, j)$ $k' \neq (k, l)$	$I_{\left(\frac{Y_{k'}+Y_l}{2} - \frac{X_{i'}+X_j}{2}\right)}$ $i' \neq (i, j)$ $k' \neq (k, l)$	$I_{\left(\frac{Y_{k'}+Y_l}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}$ $(i', j') \neq (i, j), i' < j'$ $k' \neq (k, l)$
$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_i+X_j}{2}\right)}$ $(k', l') \neq (k, l), k' < l'$	$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}$ $j' \neq (i, j)$ $(k', l') \neq (k, l), k' < l'$	$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_j}{2}\right)}$ $i' \neq (i, j)$ $(k', l') \neq (k, l), k' < l'$	$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}$ $(i', j') \neq (i, j), i' < j'$ $(k', l') \neq (k, l), k' < l'$

$$\begin{aligned}
 & + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{\substack{i < j \\ i' < j' \\ (i', j') \neq (i, j)}} \text{Cov}\left(I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2}\right)}, I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}\right) \\
 & + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l'=1}^n \sum_{\substack{i < j \\ k < l \\ l' \neq (k, l)}} \text{Cov}\left(I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2}\right)}, I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_i+X_j}{2}\right)}\right) \\
 & + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{\substack{i < j \\ k < l \\ k' \neq (k, l)}} \text{Cov}\left(I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2}\right)}, I_{\left(\frac{Y_{k'}+Y_l}{2} - \frac{X_i+X_j}{2}\right)}\right) \\
 & + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \sum_{\substack{i < j \\ k < l \\ k' < l' \\ (k', l') \neq (k, l)}} \text{Cov}\left(I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2}\right)}, I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_i+X_j}{2}\right)}\right) \\
 & + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l'=1}^n \sum_{\substack{i < j \\ j' \neq (i, j) \\ k < l \\ l' \neq (k, l)}} \text{Cov}\left(I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2}\right)}, I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}\right)
 \end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \text{Cov}(I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})}, I_{(\frac{Y_{k'}+Y_l}{2} - \frac{X_i+X_{j'}}{2})}) \\
& \quad \quad \quad i < j \quad j' \neq (i,j) \quad k < l \quad k' \neq (k,l) \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l'=1}^n \text{Cov}(I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})}, I_{(\frac{Y_k+Y_{l'}}{2} - \frac{X_{i'}+X_j}{2})}) \\
& \quad \quad \quad i < j \quad i' \neq (i,j) \quad k < l \quad l' \neq (k,l) \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \text{Cov}(I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})}, I_{(\frac{Y_{k'}+Y_l}{2} - \frac{X_{i'}+X_j}{2})}) \\
& \quad \quad \quad i < j \quad i' \neq (i,j) \quad k < l \quad k' \neq (k,l) \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \text{Cov}(I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})}, I_{(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2})}) \\
& \quad \quad \quad i < j \quad j' \neq (i,j) \quad k < l \quad k' < l' \\
& \quad \quad \quad (k',l') \neq (k,l) \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \text{Cov}(I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})}, I_{(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2})}) \\
& \quad \quad \quad i < j \quad i' \neq (i,j) \quad k < l \quad k' < l' \\
& \quad \quad \quad (k',l') \neq (k,l) \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l'=1}^n \text{Cov}(I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})}, I_{(\frac{Y_k+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2})}) \\
& \quad \quad \quad i < j \quad i' < j' \quad k < l \quad l' \neq (k,l) \\
& \quad \quad \quad (i',j') \neq (i,j) \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \text{Cov}(I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})}, I_{(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2})}) \\
& \quad \quad \quad i < j \quad i' < j' \quad k < l \quad k' \neq (k,l) \\
& \quad \quad \quad (i',j') \neq (i,j) \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \text{Cov}(I_{(\frac{Y_k+Y_l}{2} - \frac{X_i+X_j}{2})}, I_{(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2})}) \\
& \quad \quad \quad i < j \quad i' < j' \quad k < l \quad k' < l' \\
& \quad \quad \quad (i',j') \neq (i,j) \quad (k',l') \neq (k,l) \\
& = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n [P(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2}) - P(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2})^2] \\
& \quad \quad \quad i < j \quad k < l \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n [P(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2}, \frac{X_i+X_{j'}}{2} < \frac{Y_k+Y_l}{2}) \\
& \quad \quad \quad i < j \quad j' \neq (i,j) \quad k < l \\
& \quad \quad \quad - P(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2})P(\frac{X_i+X_{j'}}{2} < \frac{Y_k+Y_l}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{l=1}^n [P(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2}, \frac{X_{i'}+X_j}{2} < \frac{Y_k+Y_l}{2}) \\
& \quad \quad \quad i < j \quad i' \neq (i,j) \quad k < l \\
& \quad \quad \quad - P(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2})P(\frac{X_{i'}+X_j}{2} < \frac{Y_k+Y_l}{2})]
\end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n [P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}, \frac{X_{i'} + X_{j'}}{2} < \frac{Y_k + Y_l}{2}) \\
& \quad \substack{i < j \\ i' < j' \\ (i', j') \neq (i, j)} \\
& \quad - P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}) P(\frac{X_{i'} + X_{j'}}{2} < \frac{Y_k + Y_l}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l'=1}^n [P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}, \frac{X_i + X_j}{2} < \frac{Y_k + Y_{l'}}{2}) \\
& \quad \substack{i < j \\ k < l \\ l' \neq (k, l)} \\
& \quad - P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}) P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n [P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}, \frac{X_i + X_j}{2} < \frac{Y_{k'} + Y_l}{2}) \\
& \quad \substack{i < j \\ k < l \\ k' \neq (k, l)} \\
& \quad - P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}) P(\frac{X_i + X_j}{2} < \frac{Y_{k'} + Y_l}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n [P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}, \frac{X_i + X_j}{2} < \frac{Y_{k'} + Y_{l'}}{2}) \\
& \quad \substack{i < j \\ k < l \\ k' < l' \\ (k', l') \neq (k, l)} \\
& \quad - P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}) P(\frac{X_i + X_j}{2} < \frac{Y_{k'} + Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l'=1}^n [P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}, \frac{X_i + X_{j'}}{2} < \frac{Y_k + Y_{l'}}{2}) \\
& \quad \substack{i < j \\ j' \neq (i, j) \\ k < l \\ l' \neq (k, l)} \\
& \quad - P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}) P(\frac{X_i + X_{j'}}{2} < \frac{Y_k + Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n [P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}, \frac{X_i + X_{j'}}{2} < \frac{Y_{k'} + Y_l}{2}) \\
& \quad \substack{i < j \\ j' \neq (i, j) \\ k < l \\ k' \neq (k, l)} \\
& \quad - P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}) P(\frac{X_i + X_{j'}}{2} < \frac{Y_{k'} + Y_l}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l'=1}^n [P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}, \frac{X_{i'} + X_j}{2} < \frac{Y_k + Y_{l'}}{2}) \\
& \quad \substack{i < j \\ i' \neq (i, j) \\ k < l \\ l' \neq (k, l)} \\
& \quad - P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}) P(\frac{X_{i'} + X_j}{2} < \frac{Y_k + Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n [P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}, \frac{X_{i'} + X_j}{2} < \frac{Y_{k'} + Y_l}{2}) \\
& \quad \substack{i < j \\ i' \neq (i, j) \\ k < l \\ k' \neq (k, l)} \\
& \quad - P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}) P(\frac{X_{i'} + X_j}{2} < \frac{Y_{k'} + Y_l}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n [P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}, \frac{X_i + X_{j'}}{2} < \frac{Y_{k'} + Y_{l'}}{2}) \\
& \quad \substack{i < j \\ j' \neq (i, j) \\ k < l \\ k' < l' \\ (k', l') \neq (k, l)} \\
& \quad - P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}) P(\frac{X_i + X_{j'}}{2} < \frac{Y_{k'} + Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n [P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}, \frac{X_i + X_{j'}}{2} < \frac{Y_{k'} + Y_{l'}}{2}) \\
& \quad \substack{i < j \\ j' \neq (i, j) \\ k < l \\ k' < l' \\ (k', l') \neq (k, l)} \\
& \quad - P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_l}{2}) P(\frac{X_i + X_{j'}}{2} < \frac{Y_{k'} + Y_{l'}}{2})]
\end{aligned}$$

Table 3.5: Components of (I_1, I_2)

$I_{\left(\frac{Y_k+Y_{l'}}{2}-X_i\right)}$ $l' \neq k$	$I_{\left(\frac{Y_{k'}+Y_{l'}}{2}-X_i\right)}$ $(k', l') \neq k, k' < l'$
$I_{\left(\frac{Y_k+Y_{l'}}{2}-X_{i'}\right)}$ $l' \neq k, i' \neq i$	$I_{\left(\frac{Y_{k'}+Y_{l'}}{2}-X_{i'}\right)}$ $(k', l') \neq k, k' < l', i' \neq i$

$$\begin{aligned}
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n [P\left(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2}, \frac{X_{i'}+X_j}{2} < \frac{Y_{k'}+Y_{l'}}{2}\right) \\
& \quad - P\left(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2}\right) P\left(\frac{X_{i'}+X_j}{2} < \frac{Y_{k'}+Y_{l'}}{2}\right)] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l'=1}^n [P\left(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2}, \frac{X_{i'}+X_{j'}}{2} < \frac{Y_k+Y_{l'}}{2}\right) \\
& \quad - P\left(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2}\right) P\left(\frac{X_{i'}+X_{j'}}{2} < \frac{Y_k+Y_{l'}}{2}\right)] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n [P\left(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2}, \frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2}\right) \\
& \quad - P\left(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2}\right) P\left(\frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2}\right)] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n [P\left(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2}, \frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2}\right) \\
& \quad - P\left(\frac{X_i+X_j}{2} < \frac{Y_k+Y_l}{2}\right) P\left(\frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2}\right)] \\
\end{aligned} \tag{3.8}$$

The components of (I_1, I_2) are listed in Table 3.5. The covariances of $(I_{(Y_k-X_i)}, I_{\left(\frac{Y_{k'}+Y_{l'}}{2}-X_i\right)})$, $k' < l'$ with each individual component of the four parts of (I_1, I_2) have the same pattern.

The covariance of (I_1, I_2) are

$$Cov(I_1, I_2) = Cov\left[\sum_{i=1}^m \sum_{k=1}^n I_{(Y_k-X_i)}, \sum_{i'=1}^m \sum_{k'=1}^n \sum_{l'=1}^n I_{\left(\frac{Y_{k'}+Y_{l'}}{2}-X_{i'}\right)}\right]$$

$$\begin{aligned}
&= \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k \neq l'}}^n Cov(I_{(Y_k - X_i)}, I_{(\frac{Y_k + Y_{l'}}{2} - X_i)}) + \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k' < l' \\ (k', l') \neq k}}^n Cov(I_{(Y_k - X_i)}, I_{(\frac{Y_{k'} + Y_{l'}}{2} - X_i)}) \\
&+ \sum_{i=1}^m \sum_{\substack{i'=1 \\ i' \neq i}}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ l' \neq k}}^n Cov(I_{(Y_k - X_i)}, I_{(\frac{Y_k + Y_{l'}}{2} - X_{i'})}) \\
&+ \sum_{i=1}^m \sum_{\substack{i'=1 \\ i' \neq i}}^m \sum_{k=1}^n \sum_{\substack{k'=1 \\ k' < l' \\ (k', l') \neq k}}^n \sum_{l'=1}^n Cov(I_{(Y_k - X_i)}, I_{(\frac{Y_{k'} + Y_{l'}}{2} - X_{i'})}) \\
&= \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k \neq l'}}^n [P(X_i < Y_k, X_i < \frac{Y_k + Y_{l'}}{2}) - P(X_i < Y_k)P(X_i < \frac{Y_k + Y_{l'}}{2})] \\
&+ \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{k'=1 \\ k' < l' \\ (k', l') \neq k}}^n \sum_{l'=1}^n [P(X_i < Y_k, X_i < \frac{Y_{k'} + Y_{l'}}{2}) - P(X_i < Y_k)P(X_i < \frac{Y_{k'} + Y_{l'}}{2})] \\
&+ \sum_{i=1}^m \sum_{\substack{i'=1 \\ i' \neq i}}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ l' \neq k}}^n [P(X_i < Y_k, X_{i'} < \frac{Y_k + Y_{l'}}{2}) - P(X_i < Y_k)P(X_{i'} < \frac{Y_k + Y_{l'}}{2})] \\
&+ \sum_{i=1}^m \sum_{\substack{i'=1 \\ i' \neq i}}^m \sum_{k=1}^n \sum_{\substack{k'=1 \\ k' < l' \\ (k', l') \neq k}}^n \sum_{l'=1}^n [P(X_i < Y_k, X_{i'} < \frac{Y_{k'} + Y_{l'}}{2}) - P(X_i < Y_k)P(X_{i'} < \frac{Y_{k'} + Y_{l'}}{2})]
\end{aligned} \tag{3.9}$$

It is noted that the pair of indices (k, l') does not require that $k < l'$.

The components of (I_1, I_3) are listed in Table 3.6. The covariances of $(I_{(Y_k - X_i)}, I_{(\frac{Y_{k'} - X_{i'} + X_{j'}}{2} - X_i)})$, $i' < j'$, with each individual component of the four parts of (I_1, I_3) have the same pattern.

Clearly, the covariances of (I_1, I_3) are the follow.

$$\begin{aligned}
Cov(I_1, I_3) &= Cov[\sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)}, \sum_{k'=1}^n \sum_{\substack{i'=1 \\ i' < j'}}^m \sum_{j'=1}^m I_{(\frac{Y_{k'} - X_{i'} + X_{j'}}{2} - X_i)}] \\
&= \sum_{i=1}^m \sum_{\substack{j'=1 \\ j' \neq i}}^m \sum_{k=1}^n Cov(I_{(Y_k - X_i)}, I_{(\frac{Y_k - X_i + X_{j'}}{2} - X_i)}) + \sum_{i=1}^m \sum_{\substack{i'=1 \\ i' < j'}}^m \sum_{j'=1}^m \sum_{k=1}^n Cov(I_{(Y_k - X_i)}, I_{(\frac{Y_{k'} - X_{i'} + X_{j'}}{2} - X_i)})
\end{aligned}$$

Table 3.6: Components of (I_1, I_3)

$I_{(Y_{k'} - \frac{x_i + x_{j'}}{2})}$ $j' \neq i$	$I_{(Y_{k'} - \frac{x_{i'} + x_{j'}}{2})}$ $(i', j') \neq i, i' < j'$
$I_{(Y_{k'} - \frac{x_i + x_{j'}}{2})}$ $k' \neq k, j' \neq i$	$I_{(Y_{k'} - \frac{x_{i'} + x_{j'}}{2})}$ $(i', j') \neq i, i' < j', k' \neq k$

$$\begin{aligned}
& + \sum_{i=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{\substack{j' \neq i \\ k' \neq k}} Cov(I_{(Y_k - X_i)}, I_{(Y_{k'} - \frac{x_i + x_{j'}}{2})}) \\
& + \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{\substack{i' < j' \\ k' \neq k \\ (i', j') \neq i}} Cov(I_{(Y_k - X_i)}, I_{(Y_{k'} - \frac{x_{i'} + x_{j'}}{2})}) \\
& = \sum_{i=1}^m \sum_{j'=1}^m \sum_{k=1}^n [P(X_i < Y_k, \frac{X_i + X_{j'}}{2} < Y_k) - P(X_i < Y_k)P(\frac{X_i + X_{j'}}{2} < Y_k)] \\
& + \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{\substack{i' < j' \\ (i', j') \neq i}} [P(X_i < Y_k, \frac{X_{i'} + X_{j'}}{2} < Y_k) - P(X_i < Y_k)P(\frac{X_{i'} + X_{j'}}{2} < Y_k)] \\
& + \sum_{i=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{\substack{j' \neq i \\ k' \neq k \\ (i', j') \neq i}} [P(X_i < Y_k, \frac{X_i + X_{j'}}{2} < Y_{k'}) - P(X_i < Y_k)P(\frac{X_i + X_{j'}}{2} < Y_{k'})] \\
& + \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{\substack{i' < j' \\ k' \neq k \\ (i', j') \neq i}} [P(X_i < Y_k, \frac{X_{i'} + X_{j'}}{2} < Y_{k'}) - P(X_i < Y_k)P(\frac{X_{i'} + X_{j'}}{2} < Y_{k'})] \\
\end{aligned} \tag{3.10}$$

It is noted that the pair of indices (i, j') does not require that $i < j'$.

The components of (I_1, I_4) are listed in Table 3.7. The covariances of $(I_{(Y_k - X_i)}, I_{(\frac{Y_{k'} + Y_{l'}}{2} - \frac{x_{i'} + x_{j'}}{2})})$, $k' < l', i' < j'$, with each individual component of the four parts (I_1, I_4) have the same pattern.

The covariances of (I_1, I_4) are

Table 3.7: Components of (I_1, I_4)

$I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}$ $l' \neq k, j' \neq i$	$I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}$ $(i', j') \neq i, i' < j', l' \neq k$
$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}$ $(k', l') \neq k, k' < l'$ $j' \neq i$	$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}$ $(k', l') \neq k, k' < l'$ $(i', j') \neq i, i' < j'$

$$\begin{aligned}
Cov(I_1, I_4) &= Cov\left[\sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)}, \sum_{i'=1}^m \sum_{j'=1}^n \sum_{k'=1}^n \sum_{l'=1}^n I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}\right] \\
&= \sum_{i=1}^m \sum_{\substack{j'=1 \\ i \neq j'}}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k \neq l'}}^n Cov(I_{(Y_k - X_i)}, I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}) \\
&\quad + \sum_{i=1}^m \sum_{\substack{i'=1 \\ i' < j'}}^m \sum_{j'=1}^n \sum_{k=1}^n \sum_{\substack{l'=1 \\ k \neq l'}}^n Cov(I_{(Y_k - X_i)}, I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}) \\
&\quad + \sum_{i=1}^m \sum_{\substack{j'=1 \\ j' \neq i}}^m \sum_{k=1}^n \sum_{\substack{k'=1 \\ (k', l') \neq k}}^n \sum_{l'=1}^n Cov(I_{(Y_k - X_i)}, I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}) \\
&\quad + \sum_{i=1}^m \sum_{\substack{i'=1 \\ (i', j') \neq i}}^m \sum_{j'=1}^n \sum_{k=1}^n \sum_{\substack{k'=1 \\ (k', l') \neq k}}^n \sum_{l'=1}^n Cov(I_{(Y_k - X_i)}, I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}) \\
&= \sum_{i=1}^m \sum_{\substack{j'=1 \\ i \neq j'}}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k \neq l'}}^n [P(X_i < Y_k, \frac{X_i + X_{j'}}{2} < \frac{Y_k + Y_{l'}}{2}) - P(X_i < Y_k)P(\frac{X_i + X_{j'}}{2} < \frac{Y_k + Y_{l'}}{2})] \\
&\quad + \sum_{i=1}^m \sum_{\substack{i'=1 \\ i' < j'}}^m \sum_{j'=1}^n \sum_{k=1}^n \sum_{\substack{l'=1 \\ k \neq l'}}^n [P(X_i < Y_k, \frac{X_{i'} + X_{j'}}{2} < \frac{Y_k + Y_{l'}}{2}) - P(X_i < Y_k)P(\frac{X_{i'} + X_{j'}}{2} < \frac{Y_k + Y_{l'}}{2})] \\
&\quad + \sum_{i=1}^m \sum_{\substack{j'=1 \\ j' \neq i}}^m \sum_{k=1}^n \sum_{\substack{k'=1 \\ (k', l') \neq k}}^n \sum_{l'=1}^n [P(X_i < Y_k, \frac{X_i + X_{j'}}{2} < \frac{Y_{k'} + Y_{l'}}{2}) - P(X_i < Y_k)P(\frac{X_i + X_{j'}}{2} < \frac{Y_{k'} + Y_{l'}}{2})]
\end{aligned}$$

Table 3.8: Components of (I_2, I_3)

$I_{(Y_k - \frac{x_i + x_{j'}}{2})}$ $j' \neq i$	$I_{(Y_k - \frac{x_{i'} + x_{j'}}{2})}$ $(i', j') \neq i, i' < j'$
$I_{(Y_{k'} - \frac{x_i + x_{j'}}{2})}$ $k' \neq (k, l), j \neq i$	$I_{(Y_{k'} - \frac{x_{i'} + x_{j'}}{2})}$ $(i', j') \neq i, i' < j', k' \neq (k, l)$

$$+ \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n [P(X_i < Y_k, \frac{X_{i'} + X_{j'}}{2} < \frac{Y_{k'} + Y_{l'}}{2}) - P(X_i < Y_k)P(\frac{X_{i'} + X_{j'}}{2} < \frac{Y_{k'} + Y_{l'}}{2})]$$

$(i', j') \neq i$ $(k', l') \neq k$

(3.11)

It is noted that the two pairs of indices (i, j') and (k, l') do not require that $i < j'$ and $k < l'$.

The components of (I_2, I_3) are listed in Table 3.8. The covariances of $(I_{(Y_{\frac{k+Y_l}{2}} - X_i)}, I_{(Y_{k'} - \frac{x_{i'} + x_{j'}}{2})})$, $k < l, i' < j'$, with each individual component of the six parts of (I_2, I_3) have the same pattern. The covariances of (I_2, I_3) are

$$Cov(I_2, I_3) = Cov[\sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n I_{(Y_{\frac{k+Y_l}{2}} - X_i)}, \sum_{k'=1}^n \sum_{i'=1}^m \sum_{j'=1}^m I_{(Y_{k'} - \frac{x_{i'} + x_{j'}}{2})}]$$

$$= \sum_{i=1}^m \sum_{j'=1}^m \sum_{k'=1}^n \sum_{l=1}^n Cov(I_{(Y_{\frac{k'+Y_l}{2}} - X_i)}, I_{(Y_{k'} - \frac{x_i + x_{j'}}{2})})$$

$i \neq j' \quad k' \neq l$

$$+ \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k'=1}^n \sum_{l=1}^n Cov(I_{(Y_{\frac{k'+Y_l}{2}} - X_i)}, I_{(Y_{k'} - \frac{x_{i'} + x_{j'}}{2})})$$

$i' < j' \quad k' \neq l$

$(i', j') \neq i$

$$+ \sum_{i=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n Cov(I_{(Y_{\frac{k+Y_l}{2}} - X_i)}, I_{(Y_{k'} - \frac{x_i + x_{j'}}{2})})$$

$i \neq j' \quad k < l \quad k' \neq (k, l)$

$$+ \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n Cov(I_{(Y_{\frac{k+Y_l}{2}} - X_i)}, I_{(Y_{k'} - \frac{x_{i'} + x_{j'}}{2})})$$

$i' < j' \quad k < l \quad k' \neq (k, l)$

$(i', j') \neq i$

$$\begin{aligned}
&= \sum_{i=1}^m \sum_{\substack{j'=1 \\ i \neq j'}}^m \sum_{k'=1}^n \sum_{\substack{l=1 \\ k' \neq l}}^n [P(X_i < \frac{Y_{k'} + Y_l}{2}, \frac{X_i + X_{j'}}{2} < Y_{k'}) - P(X_i < \frac{Y_{k'} + Y_l}{2})P(\frac{X_i + X_{j'}}{2} < Y_{k'})] \\
&+ \sum_{i=1}^m \sum_{\substack{i'=1 \\ i' < j'}}^m \sum_{j'=1}^m \sum_{\substack{k'=1 \\ k' \neq l}}^n \sum_{l=1}^n P(X_i < \frac{Y_{k'} + Y_l}{2}, \frac{X_{i'} + X_{j'}}{2} < Y_{k'}) - P(X_i < \frac{Y_{k'} + Y_l}{2})P(\frac{X_{i'} + X_{j'}}{2} < Y_{k'}) \\
&+ \sum_{i=1}^m \sum_{\substack{j'=1 \\ i \neq j'}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{k'=1 \\ k' \neq (k,l)}}^n [P(X_i < \frac{Y_k + Y_l}{2}, \frac{X_i + X_{j'}}{2} < Y_{k'}) - P(X_i < \frac{Y_k + Y_l}{2})P(\frac{X_i + X_{j'}}{2} < Y_{k'})] \\
&+ \sum_{i=1}^m \sum_{\substack{i'=1 \\ i' < j'}}^m \sum_{j'=1}^m \sum_{\substack{k=1 \\ k < l}}^n \sum_{\substack{l=1 \\ k' \neq (k,l)}}^n \sum_{k'=1}^n [P(X_i < \frac{Y_k + Y_l}{2}, \frac{X_{i'} + X_{j'}}{2} < Y_{k'}) \\
&\quad - P(X_i < \frac{Y_k + Y_l}{2})P(\frac{X_{i'} + X_{j'}}{2} < Y_{k'})] \tag{3.12}
\end{aligned}$$

It is noted that the two pairs of indices (i, j') , (k', l) do not require that $i < j'$ and $k' < l$.

The components of (I_2, I_4) are listed in Table 3.9. The covariances of $(I_{(\frac{Y_k + Y_l}{2} - X_i)}, I_{(\frac{Y_{k'} + Y_{l'}}{2} - \frac{X_{i'} + X_{j'}}{2})})$, $k < l, k' < j', i' < j'$, with each individual component of the eight parts of (I_2, I_4) have the same pattern.

The covariance of (I_2, I_4) are

$$\begin{aligned}
Cov(I_2, I_4) &= Cov[\sum_{i=1}^m \sum_{\substack{k=1 \\ k < l}}^n \sum_{l=1}^n I_{(\frac{Y_k + Y_l}{2} - X_i)}, \sum_{\substack{i'=1 \\ i' < j'}}^m \sum_{j'=1}^m \sum_{\substack{k'=1 \\ k' < l'}}^n \sum_{l'=1}^n I_{(\frac{Y_{k'} + Y_{l'}}{2} - \frac{X_{i'} + X_{j'}}{2})}] \\
&= \sum_{i=1}^m \sum_{\substack{j'=1 \\ i \neq j'}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n Cov(I_{(\frac{Y_k + Y_l}{2} - X_i)}, I_{(\frac{Y_k + Y_l}{2} - \frac{X_i + X_{j'}}{2})}) \\
&+ \sum_{i=1}^m \sum_{\substack{i'=1 \\ i' < j'}}^m \sum_{j'=1}^m \sum_{\substack{k=1 \\ k < l}}^n \sum_{l=1}^n Cov(I_{(\frac{Y_k + Y_l}{2} - X_i)}, I_{(\frac{Y_k + Y_l}{2} - \frac{X_{i'} + X_{j'}}{2})}) \\
&+ \sum_{i=1}^m \sum_{\substack{j'=1 \\ i \neq j'}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{l'=1 \\ l' \neq (k,l)}}^n Cov(I_{(\frac{Y_k + Y_l}{2} - X_i)}, I_{(\frac{Y_k + Y_{l'}}{2} - \frac{X_i + X_{j'}}{2})})
\end{aligned}$$

Table 3.9: Components of (I_2, I_4)

$I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_i+X_{j'}}{2}\right)}$ $j' \neq i$	$I_{\left(\frac{Y_k+Y_l}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}$ $(i', j') \neq i, i' < j'$
$I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}$ $l' \neq (k, l), j' \neq i$	$I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}$ $(i', j') \neq i, i' < j', l' \neq (k, l)$
$I_{\left(\frac{Y_{k'}+Y_l}{2} - \frac{X_i+X_{j'}}{2}\right)}$ $k' \neq (k, l), j' \neq i$	$I_{\left(\frac{Y_{k'}+Y_l}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}$ $(i', j') \neq i, i' < j', k' \neq (k, l)$
$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}$ $j' \neq i$ $(k', l') \neq (k, l), k' < l'$	$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}$ $(i', j') \neq i, i' < j'$ $(k', l') \neq (k, l), k' < l'$

$$\begin{aligned}
& + \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l'=1}^n \sum_{\substack{i' < j' \\ (i', j') \neq i \\ k < l \\ l' \neq (k, l)}} \text{Cov}\left(I_{\left(\frac{Y_k+Y_l}{2} - X_i\right)}, I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}\right) \\
& + \sum_{i=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{\substack{i \neq j' \\ k < l \\ k' \neq (k, l)}} \text{Cov}\left(I_{\left(\frac{Y_k+Y_l}{2} - X_i\right)}, I_{\left(\frac{Y_{k'}+Y_l}{2} - \frac{X_i+X_{j'}}{2}\right)}\right) \\
& + \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{\substack{i' < j' \\ k < l \\ k' \neq (k, l)}} \text{Cov}\left(I_{\left(\frac{Y_k+Y_l}{2} - X_{i'}\right)}, I_{\left(\frac{Y_{k'}+Y_l}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}\right) \\
& + \sum_{i=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \sum_{\substack{i \neq j' \\ k < l \\ k' < l' \\ (k', l') \neq (k, l)}} \text{Cov}\left(I_{\left(\frac{Y_k+Y_l}{2} - X_i\right)}, I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}\right) \\
& + \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \sum_{\substack{i' < j' \\ k < l \\ k' < l' \\ (i', j') \neq i \\ (k', l') \neq (k, l)}} \text{Cov}\left(I_{\left(\frac{Y_k+Y_l}{2} - X_{i'}\right)}, I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}\right) \\
& = \sum_{i=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{i \neq j'} \sum_{k < l} \left[P\left(X_i < \frac{Y_k+Y_l}{2}, \frac{X_i+X_{j'}}{2} < \frac{Y_k+Y_l}{2}\right) - P\left(X_i < \frac{Y_k+Y_l}{2}\right) P\left(\frac{X_i+X_{j'}}{2} < \frac{Y_k+Y_l}{2}\right) \right]
\end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n [P(X_i < \frac{Y_k+Y_l}{2}, \frac{X_{i'}+X_{j'}}{2} < \frac{Y_k+Y_l}{2}) \\
& \quad \substack{i' < j' \\ (i', j') \neq i} \\
& \quad \quad \quad - P(X_i < \frac{Y_k+Y_l}{2}) P(\frac{X_{i'}+X_{j'}}{2} < \frac{Y_k+Y_l}{2})] \\
& + \sum_{i=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l'=1}^n [P(X_i < \frac{Y_k+Y_l}{2}, \frac{X_i+X_{j'}}{2} < \frac{Y_k+Y_{l'}}{2}) \\
& \quad \substack{i \neq j' \\ k < l \quad l' \neq (k, l)} \\
& \quad \quad \quad - P(X_i < \frac{Y_k+Y_l}{2}) P(\frac{X_i+X_{j'}}{2} < \frac{Y_k+Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l'=1}^n [P(X_i < \frac{Y_k+Y_l}{2}, \frac{X_{i'}+X_{j'}}{2} < \frac{Y_k+Y_{l'}}{2}) \\
& \quad \substack{i' < j' \\ (i', j') \neq i} \\
& \quad \quad \quad - P(X_i < \frac{Y_k+Y_l}{2}) P(\frac{X_{i'}+X_{j'}}{2} < \frac{Y_k+Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n [P(X_i < \frac{Y_k+Y_l}{2}, \frac{X_i+X_{j'}}{2} < \frac{Y_{k'}+Y_l}{2}) \\
& \quad \substack{i \neq j' \\ k < l \quad k' \neq (k, l)} \\
& \quad \quad \quad - P(X_i < \frac{Y_k+Y_l}{2}) P(\frac{X_i+X_{j'}}{2} < \frac{Y_{k'}+Y_l}{2})] \\
& + \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n [P(X_i < \frac{Y_k+Y_l}{2}, \frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_l}{2}) \\
& \quad \substack{i' < j' \\ (i', j') \neq i} \\
& \quad \quad \quad - P(X_i < \frac{Y_k+Y_l}{2}) P(\frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_l}{2})] \\
& + \sum_{i=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n [P(X_i < \frac{Y_k+Y_l}{2}, \frac{X_i+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2}) \\
& \quad \substack{i \neq j' \\ k < l \quad k' < l' \\ (k', l') \neq (k, l)} \\
& \quad \quad \quad - P(X_i < \frac{Y_k+Y_l}{2}) P(\frac{X_i+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n [P(X_i < \frac{Y_k+Y_l}{2}, \frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2}) \\
& \quad \substack{i' < j' \\ (i', j') \neq i} \\
& \quad \quad \quad \substack{k < l \quad k' < l' \\ (k', l') \neq (k, l)} \\
& \quad \quad \quad - P(X_i < \frac{Y_k+Y_l}{2}) P(\frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2})]
\end{aligned} \tag{3.13}$$

It is noted that the three pairs of indices (i, j') , (k, l') , and (k', l) do not require that $i < j'$, $k < l'$ and $k' < l$.

The components (I_3, I_4) are listed in Table 3.10. The covariances of $(I_{(Y_k - \frac{X_i+X_{j'}}{2})})$,

Table 3.10: Components of (I_3, I_4)

$I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_i+X_j}{2}\right)}$ $l' \neq k$	$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_i+X_j}{2}\right)}$ $(k', l') \neq k, k' < l'$
$I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}$ $j' \neq (i, j), l' \neq k$	$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}$ $(k', l') \neq k, k' < l', j' \neq (i, j)$
$I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_{i'}+X_j}{2}\right)}$ $i' \neq (i, j), l' \neq k$	$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_j}{2}\right)}$ $(k', l') \neq k, k' < l', i' \neq (i, j)$
$I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}$ $(i', j') \neq (i, j), i' < j'$ $l' \neq k$	$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}$ $(i', j') \neq (i, j), i' < j'$ $(k', l') \neq k, k' < l'$

$I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}, i < j, k' < l', i' < j'$, with each individual component of the eight parts

of (I_3, I_4) have the same pattern. The covariance of (I_3, I_4) are

$$\begin{aligned}
 Cov(I_3, I_4) &= Cov\left(\sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m I_{\left(Y_k - \frac{X_i+X_j}{2}\right)}, \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k'=1}^n \sum_{l'=1}^n I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_{i'}+X_{j'}}{2}\right)}\right) \\
 &= \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l'=1}^n Cov\left(I_{\left(Y_k - \frac{X_i+X_j}{2}\right)}, I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_i+X_j}{2}\right)}\right) \\
 &\quad + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n Cov\left(I_{\left(Y_k - \frac{X_i+X_j}{2}\right)}, I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_i+X_j}{2}\right)}\right) \\
 &\quad + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^n \sum_{k=1}^n \sum_{l'=1}^n Cov\left(I_{\left(Y_k - \frac{X_i+X_j}{2}\right)}, I_{\left(\frac{Y_k+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}\right) \\
 &\quad + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^n \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n Cov\left(I_{\left(Y_k - \frac{X_i+X_j}{2}\right)}, I_{\left(\frac{Y_{k'}+Y_{l'}}{2} - \frac{X_i+X_{j'}}{2}\right)}\right)
 \end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{l'=1}^n \text{Cov}(I_{(Y_k - \frac{X_i + X_j}{2})}, I_{(\frac{Y_k + Y_{l'}}{2} - \frac{X_{i'} + X_j}{2})}) \\
& \quad \substack{i < j \\ i' \neq (i, j) \\ k \neq l'} \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \text{Cov}(I_{(Y_k - \frac{X_i + X_j}{2})}, I_{(\frac{Y_{k'} + Y_{l'}}{2} - \frac{X_{i'} + X_j}{2})}) \\
& \quad \substack{i < j \\ i' \neq (i, j) \\ (k', l') \neq k \\ k' < l'} \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^n \sum_{k=1}^n \sum_{l'=1}^n \text{Cov}(I_{(Y_k - \frac{X_i + X_j}{2})}, I_{(\frac{Y_k + Y_{l'}}{2} - \frac{X_{i'} + X_{j'}}{2})}) \\
& \quad \substack{i < j \\ i' < j' \\ k \neq l'} \\
& \quad \substack{(i', j') \neq (i, j)} \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^n \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \text{Cov}(I_{(Y_k - \frac{X_i + X_j}{2})}, I_{(\frac{Y_{k'} + Y_{l'}}{2} - \frac{X_{i'} + X_{j'}}{2})}) \\
& \quad \substack{i < j \\ i' < j' \\ k' < l'} \\
& \quad \substack{(i', j') \neq (i, j) \\ (k', l') \neq k} \\
= & \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l'=1}^n [P(\frac{X_i + X_j}{2} < Y_k, \frac{X_i + X_j}{2} < \frac{Y_k + Y_{l'}}{2}) \\
& \quad \substack{i < j \\ k \neq l'} \\
& \quad - P(\frac{X_i + X_j}{2} < Y_k) P(\frac{X_i + X_j}{2} < \frac{Y_k + Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n [P(\frac{X_i + X_j}{2} < Y_k, \frac{X_i + X_j}{2} < \frac{Y_{k'} + Y_{l'}}{2}) \\
& \quad \substack{i < j \\ k' < l'} \\
& \quad \substack{(k', l') \neq k} \\
& \quad - P(\frac{X_i + X_j}{2} < Y_k) P(\frac{X_i + X_j}{2} < \frac{Y_{k'} + Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^n \sum_{k=1}^n \sum_{l'=1}^n [P(\frac{X_i + X_j}{2} < Y_k, \frac{X_i + X_{j'}}{2} < \frac{Y_k + Y_{l'}}{2}) \\
& \quad \substack{i < j \\ j' \neq (i, j) \\ k \neq l'} \\
& \quad - P(\frac{X_i + X_j}{2} < Y_k) P(\frac{X_i + X_{j'}}{2} < \frac{Y_k + Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^n \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n [P(\frac{X_i + X_j}{2} < Y_k, \frac{X_i + X_{j'}}{2} < \frac{Y_{k'} + Y_{l'}}{2}) \\
& \quad \substack{i < j \\ j' \neq (i, j) \\ k' < l'} \\
& \quad \substack{(k', l') \neq k} \\
& \quad - P(\frac{X_i + X_j}{2} < Y_k) P(\frac{X_i + X_{j'}}{2} < \frac{Y_{k'} + Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{l'=1}^n [P(\frac{X_i + X_j}{2} < Y_k, \frac{X_{i'} + X_j}{2} < \frac{Y_k + Y_{l'}}{2}) \\
& \quad \substack{i < j \\ i' \neq (i, j) \\ k \neq l'} \\
& \quad - P(\frac{X_i + X_j}{2} < Y_k) P(\frac{X_{i'} + X_j}{2} < \frac{Y_k + Y_{l'}}{2})] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n [P(\frac{X_i + X_j}{2} < Y_k, \frac{X_{i'} + X_j}{2} < \frac{Y_{k'} + Y_{l'}}{2}) \\
& \quad \substack{i < j \\ i' \neq (i, j) \\ k' < l'} \\
& \quad \substack{(k', l') \neq k} \\
& \quad - P(\frac{X_i + X_j}{2} < Y_k) P(\frac{X_{i'} + X_j}{2} < \frac{Y_{k'} + Y_{l'}}{2})]
\end{aligned}$$

$$\begin{aligned}
& -P\left(\frac{X_i+X_j}{2} < Y_k\right)P\left(\frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2}\right)] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l'=1}^n [P\left(\frac{X_i+X_j}{2} < Y_k, \frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2}\right) \\
& \quad \substack{i < j & i' < j' & k \neq l' \\ (i', j') \neq (i, j)} \\
& -P\left(\frac{X_i+X_j}{2} < Y_k\right)P\left(\frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2}\right)] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n [P\left(\frac{X_i+X_j}{2} < Y_k, \frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2}\right) \\
& \quad \substack{i < j & i' < j' & k' < l' \\ (i', j') \neq (i, j) & (k', l') \neq k} \\
& -P\left(\frac{X_i+X_j}{2} < Y_k\right)P\left(\frac{X_{i'}+X_{j'}}{2} < \frac{Y_{k'}+Y_{l'}}{2}\right)] \\
& \hspace{15em} (3.14)
\end{aligned}$$

It is noted that the pairs of indices (i, j') , (i', j) , (k, l') do not require that $i < j'$, $i' < j$ and $k < l'$.

3.2 Sampling from the uniform distribution

For uniform case, substituting in formula (3.1) to (3.14) with the probability values from integral methods, we have the mean of I_1

$$EI_1 = \sum_{i=1}^m \sum_{k=1}^n \frac{1}{2} = \frac{mn}{2} \quad (3.15)$$

The variance of I_1 is

$$\begin{aligned}
Var I_1 &= \sum_{i=1}^m \sum_{k=1}^n \left[\frac{1}{2} - \left(\frac{1}{2}\right)^2\right] + \sum_{i=1}^m \sum_{k=1}^n \sum_{l=1}^n \left[\frac{1}{3} - \left(\frac{1}{2}\right)^2\right] \\
& \quad + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \left[\frac{1}{3} - \left(\frac{1}{2}\right)^2\right] + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \left[\frac{1}{4} - \left(\frac{1}{2}\right)^2\right] \\
& \quad \substack{i \neq j & i \neq j & k \neq l} \\
&= mn\left(\frac{1}{2} - \frac{1}{4}\right) + mn(n-1)\left(\frac{1}{3} - \frac{1}{4}\right) \\
& \quad + m(m-1)n\left(\frac{1}{3} - \frac{1}{4}\right) + m(m-1)n(n-1)\left(\frac{1}{4} - \frac{1}{4}\right) \\
&= \frac{mn(m+n+1)}{12} \hspace{15em} (3.16)
\end{aligned}$$

The mean of I_2 is

$$EI_2 = \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \frac{1}{2} = \frac{m \binom{n}{2}}{2} \quad (3.17)$$

The variance of I_2 is

$$\begin{aligned} & \text{Var}(I_2) \\ &= \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \left[\frac{1}{2} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{l'=1 \\ l' \neq (k,l)}}^n \left[\frac{5}{12} - \left(\frac{1}{2}\right)^2 \right] \\ &+ \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{l'=1 \\ k' \neq (k,l)}}^n \left[\frac{5}{12} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{l'=1 \\ k' < l' \\ (k',l') \neq (k,l)}}^n \left[\frac{23}{60} - \left(\frac{1}{2}\right)^2 \right] \\ &+ \sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq i}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \left[\frac{7}{24} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq i}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{l'=1 \\ l' \neq (k,l)}}^n \left[\frac{13}{18} - \left(\frac{1}{2}\right)^2 \right] \\ &+ \sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq i}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{l'=1 \\ k' \neq (k,l)}}^n \left[\frac{13}{18} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{\substack{j=1 \\ j \neq i}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{l'=1 \\ k' < l' \\ (k',l') \neq (k,l)}}^n \left[\frac{1}{4} - \left(\frac{1}{2}\right)^2 \right] \\ &= m \binom{n}{2} \left[\frac{1}{2} - \frac{1}{4} \right] + m \binom{n}{2} (n-2) \left[\frac{5}{12} - \frac{1}{4} \right] \\ &+ m \binom{n}{2} (n-2) \left[\frac{5}{12} - \frac{1}{4} \right] + m \binom{n}{2} \binom{n-2}{2} \left[\frac{23}{60} - \frac{1}{4} \right] \\ &+ m \binom{n}{2} (m-1) \left[\frac{7}{24} - \frac{1}{4} \right] + m \binom{n}{2} (m-1)(n-2) \left[\frac{13}{48} - \frac{1}{4} \right] \\ &+ m \binom{n}{2} (m-1)(n-2) \left[\frac{13}{48} - \frac{1}{4} \right] + m \binom{n}{2} (m-1) \binom{n-2}{2} \left[\frac{1}{4} - \frac{1}{4} \right] \\ &= mn \left(\frac{n}{30} - \frac{1}{80} - \frac{13n^2}{240} + \frac{n^3}{30} - \frac{mn}{24} + \frac{mn^2}{48} + \frac{m}{48} \right) \quad (3.18) \end{aligned}$$

The mean of I_3 is

$$EI_3 = \sum_{k=1}^n \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m \frac{1}{2} = \frac{\binom{m}{2} n}{2} \quad (3.19)$$

The variance of I_3 is

$$\begin{aligned}
& \text{Var}(I_3) \\
&= \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i < j} \left[\frac{1}{2} - \left(\frac{1}{2}\right)^2 \right] + \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{i < j} \sum_{i' \neq (i,j)} \left[\frac{5}{12} - \left(\frac{1}{2}\right)^2 \right] \\
&+ \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{i < j} \sum_{i' < j'} \sum_{(i',j') \neq (i,j)} \left[\frac{5}{12} - \left(\frac{1}{2}\right)^2 \right] + \sum_{k=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{i < j} \sum_{i' < j'} \sum_{(i',j') \neq (i,j)} \left[\frac{23}{60} - \left(\frac{1}{2}\right)^2 \right] \\
&+ \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{l \neq k} \sum_{i < j} \left[\frac{7}{24} - \left(\frac{1}{2}\right)^2 \right] + \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \sum_{l \neq k} \sum_{i < j} \sum_{j' \neq (i,j)} \left[\frac{13}{48} - \left(\frac{1}{2}\right)^2 \right] \\
&+ \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{l \neq k} \sum_{i < j} \sum_{i' \neq (i,j)} \left[\frac{13}{48} - \left(\frac{1}{2}\right)^2 \right] + \sum_{k=1}^n \sum_{l=1}^n \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{l \neq k} \sum_{i < j} \sum_{i' < j'} \sum_{(i',j') \neq (i,j)} \left[\frac{1}{4} - \left(\frac{1}{2}\right)^2 \right] \\
&= n \binom{m}{2} \left[\frac{1}{2} - \frac{1}{4} \right] + n \binom{m}{2} (m-2) \left[\frac{5}{12} - \frac{1}{4} \right] \\
&+ n \binom{m}{2} (m-2) \left[\frac{5}{12} - \frac{1}{4} \right] + n \binom{m}{2} \binom{m-2}{2} \left[\frac{23}{60} - \frac{1}{4} \right] \\
&+ n \binom{m}{2} (n-1) \left[\frac{7}{24} - \frac{1}{4} \right] + n \binom{m}{2} (n-1)(m-2) \left[\frac{13}{48} - \frac{1}{4} \right] \\
&+ n \binom{m}{2} (n-1)(m-2) \left[\frac{13}{48} - \frac{1}{4} \right] + n \binom{m}{2} (n-1) \binom{m-2}{2} \left[\frac{1}{4} - \frac{1}{4} \right] \\
&= mn \left(\frac{m}{30} - \frac{1}{80} - \frac{13m^2}{240} + \frac{m^3}{30} - \frac{mn}{24} + \frac{nm^2}{48} + \frac{n}{48} \right) \tag{3.20}
\end{aligned}$$

The mean of I_4 is

$$EI_4 = \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{i < j} \sum_{k < l} \frac{1}{2} = \frac{\binom{m}{2} \binom{n}{2}}{2} \tag{3.21}$$

The variance of I_4 is

$$\begin{aligned}
Var I_4 = & \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \left[\frac{1}{2} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{j=1}^m \sum_{j' \neq (i,j)}^m \sum_{k=1}^n \sum_{l=1}^n \left[\frac{23}{60} - \left(\frac{1}{2}\right)^2 \right] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i' \neq (i,j)}^m \sum_{k=1}^n \sum_{l=1}^n \left[\frac{23}{60} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{j=1}^m \sum_{i' < j'}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \left[\frac{1}{3} - \left(\frac{1}{2}\right)^2 \right] \\
& \quad \quad \quad (i', j') \neq (i, j) \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l' \neq (k,l)}^n \left[\frac{23}{60} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k' \neq (k,l)}^n \left[\frac{23}{60} - \left(\frac{1}{2}\right)^2 \right] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \left[\frac{1}{3} - \left(\frac{1}{2}\right)^2 \right] \\
& \quad \quad \quad (k', l') \neq (k, l) \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j' \neq (i,j)}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l' \neq (k,l)}^n \left[\frac{1}{3} - \left(\frac{1}{2}\right)^2 \right] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j' \neq (i,j)}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k' \neq (k,l)}^n \left[\frac{1}{3} - \left(\frac{1}{2}\right)^2 \right] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i' \neq (i,j)}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l' \neq (k,l)}^n \left[\frac{1}{3} - \left(\frac{1}{2}\right)^2 \right] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i' \neq (i,j)}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k' \neq (k,l)}^n \left[\frac{1}{3} - \left(\frac{1}{2}\right)^2 \right] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{j' \neq (i,j)}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \left[\frac{17371}{59640} - \left(\frac{1}{2}\right)^2 \right] \\
& \quad \quad \quad (k', l') \neq (k, l) \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i' \neq (i,j)}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \left[\frac{17371}{59640} - \left(\frac{1}{2}\right)^2 \right] \\
& \quad \quad \quad (k', l') \neq (k, l) \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{l' \neq (k,l)}^n \left[\frac{17371}{59640} - \left(\frac{1}{2}\right)^2 \right] \\
& \quad \quad \quad (i', j') \neq (i, j)
\end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{\substack{i < j \\ i' < j' \\ k < l \\ k' \neq (k,l) \\ (i',j') \neq (i,j)}} \left[\frac{17371}{59640} - \left(\frac{1}{2}\right)^2 \right] \\
& + \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \sum_{\substack{i < j \\ i' < j' \\ k < l \\ k' < l' \\ (i',j') \neq (i,j) \\ (k',l') \neq (k,l)}} \left[\frac{1}{4} - \left(\frac{1}{2}\right)^2 \right] \\
& = \binom{m}{2} \binom{n}{2} \left(\frac{1}{2} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} (m-2) \left(\frac{23}{60} - \frac{1}{4}\right) \\
& + \binom{m}{2} \binom{n}{2} (m-2) \left(\frac{23}{60} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} \binom{m-2}{2} \left(\frac{1}{3} - \frac{1}{4}\right) \\
& + \binom{m}{2} \binom{n}{2} (n-2) \left(\frac{23}{60} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} (n-2) \left(\frac{23}{60} - \frac{1}{4}\right) \\
& + \binom{m}{2} \binom{n}{2} \binom{n-2}{2} \left(\frac{1}{3} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) \\
& + \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) \\
& + \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} \binom{n-2}{2} (m-2) \left(\frac{17371}{59640} - \frac{1}{4}\right) \\
& + \binom{m}{2} \binom{n}{2} \binom{n-2}{2} (m-2) \left(\frac{17371}{59640} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} \binom{m-2}{2} (n-2) \left(\frac{17371}{59640} - \frac{1}{4}\right) \\
& + \binom{m}{2} \binom{n}{2} \binom{m-2}{2} (n-2) \left(\frac{17371}{59640} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} \binom{n-2}{2} \binom{m-2}{2} \left(\frac{1}{4} - \frac{1}{4}\right) \\
& = \binom{m}{2} \binom{n}{2} \left(-\frac{473mn}{5964} - \frac{2437n^2}{59640} + \frac{2461mn^2}{59640} + \frac{2461nm^2}{59640} \right. \\
& \quad \left. + \frac{619m}{11928} + \frac{619n}{11928} - \frac{2437m^2}{59640} + \frac{157}{5964} \right) \tag{3.22}
\end{aligned}$$

Substituting in formula (3.9) the probability values from integral methods, we have covari-

ance of (I_1, I_2)

$$\begin{aligned}
Cov(I_1, I_2) &= \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k \neq l'}}^n \left[\frac{5}{12} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k' < l'}}^n \sum_{\substack{(k', l') \neq k}} \left[\frac{17}{48} - \left(\frac{1}{2}\right)^2 \right] \\
&\quad + \sum_{i=1}^m \sum_{\substack{i'=1 \\ i \neq i'}}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k \neq l'}}^n \left[\frac{14}{48} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{\substack{i'=1 \\ i \neq i'}}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k' < l'}}^n \sum_{\substack{(k', l') \neq k}} \left[\frac{1}{4} - \left(\frac{1}{2}\right)^2 \right] \\
&= mn(n-1) \left[\frac{5}{12} - \frac{1}{4} \right] + mn \binom{n-1}{2} \left[\frac{17}{48} - \frac{1}{4} \right] \\
&\quad + mn(m-1)(n-1) \left[\frac{14}{48} - \frac{1}{4} \right] + mn(m-1) \binom{n-1}{2} \left[\frac{1}{4} - \frac{1}{4} \right] \\
&= mn \left(-\frac{n}{32} - \frac{1}{48} + \frac{5n^2}{96} + \frac{mn}{24} - \frac{m}{24} \right) \tag{3.23}
\end{aligned}$$

The covariance of (I_1, I_3) is

$$\begin{aligned}
Cov(I_1, I_3) &= \sum_{i=1}^m \sum_{\substack{j'=1 \\ j' \neq i}}^m \sum_{k=1}^n \left[\frac{5}{12} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{\substack{i'=1 \\ (i', j') \neq i}}^m \sum_{j'=1}^m \sum_{k=1}^n \left[\frac{17}{48} - \left(\frac{1}{2}\right)^2 \right] \\
&\quad + \sum_{i=1}^m \sum_{\substack{j'=1 \\ j' \neq i}}^m \sum_{k=1}^n \sum_{\substack{k'=1 \\ k' \neq k}}^n \left[\frac{14}{48} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{\substack{i'=1 \\ (i', j') \neq i}}^m \sum_{j'=1}^m \sum_{\substack{k=1 \\ k' \neq k}}^n \sum_{\substack{k'=1}}^n \left[\frac{1}{4} - \left(\frac{1}{2}\right)^2 \right] \\
&= mn(m-1) \left[\frac{5}{12} - \frac{1}{4} \right] + mn \binom{m-1}{2} \left[\frac{17}{48} - \frac{1}{4} \right] \\
&\quad + mn(m-1)(n-1) \left[\frac{14}{48} - \frac{1}{4} \right] + mn(n-1) \binom{m-1}{2} \left[\frac{1}{4} - \frac{1}{4} \right] \\
&= mn \left(-\frac{m}{32} - \frac{1}{48} + \frac{5m^2}{96} + \frac{mn}{24} - \frac{n}{24} \right) \tag{3.24}
\end{aligned}$$

The covariance of (I_1, I_4) is

$$\begin{aligned}
Cov(I_1, I_4) &= \sum_{i=1}^m \sum_{\substack{j'=1 \\ i \neq j'}}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k \neq l'}}^n \left[\frac{3}{8} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{i'=1}^m \sum_{\substack{j'=1 \\ i' < j'}}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k \neq l'}}^n \left[\frac{37}{120} - \left(\frac{1}{2}\right)^2 \right] \\
&\quad + \sum_{i=1}^m \sum_{\substack{j'=1 \\ j' \neq i}}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k' < l'}}^n \sum_{\substack{l=1 \\ (k', l') \neq k}}^n \left[\frac{37}{120} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{i'=1}^m \sum_{\substack{j'=1 \\ i' < j'}}^m \sum_{k=1}^n \sum_{\substack{l'=1 \\ k' < l'}}^n \sum_{\substack{l=1 \\ (k', l') \neq k}}^n \left[\frac{1}{4} - \left(\frac{1}{2}\right)^2 \right] \\
&= mn(m-1)(n-1) \left[\frac{3}{8} - \frac{1}{4} \right] + mn \binom{m-1}{2} (n-1) \left[\frac{37}{120} - \frac{1}{4} \right] \\
&\quad + mn \binom{n-1}{2} (m-1) \left[\frac{37}{120} - \frac{1}{4} \right] + mn \binom{m-1}{2} \binom{n-1}{2} \left[\frac{1}{4} - \frac{1}{4} \right] \\
&= mn \left(-\frac{mn}{20} + \frac{m}{48} + \frac{n}{48} + \frac{1}{120} + \frac{7m^2n}{240} - \frac{7m^2}{240} + \frac{7mn^2}{240} - \frac{7n^2}{240} \right) \quad (3.25)
\end{aligned}$$

The covariance of (I_2, I_3) is

$$\begin{aligned}
Cov(I_2, I_3) &= \sum_{i=1}^m \sum_{\substack{j'=1 \\ i \neq j'}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k \neq l}}^n \left[\frac{13}{36} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{i'=1}^m \sum_{\substack{j'=1 \\ i' < j'}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k' \neq l}}^n \left[\frac{29}{96} - \left(\frac{1}{2}\right)^2 \right] \\
&\quad + \sum_{i=1}^m \sum_{\substack{j'=1 \\ i \neq j'}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{l'=1 \\ k' \neq (k, l)}}^n \left[\frac{29}{96} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{i'=1}^m \sum_{\substack{j'=1 \\ i' < j'}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \sum_{\substack{l'=1 \\ k' \neq (k, l)}}^n \left[\frac{1}{4} - \left(\frac{1}{2}\right)^2 \right] \\
&= 2m \binom{n}{2} (m-1) \left[\frac{13}{36} - \frac{1}{4} \right] + 2m \binom{n}{2} \binom{m-1}{2} \left[\frac{29}{96} - \frac{1}{4} \right] \\
&\quad + m \binom{n}{2} (n-2)(m-1) \left[\frac{29}{96} - \frac{1}{4} \right] + m \binom{n}{2} \binom{m-1}{2} (n-2) \left[\frac{1}{4} - \frac{1}{4} \right] \\
&= m \binom{n}{2} \left(-\frac{11m}{288} - \frac{1}{72} + \frac{5m^2}{96} + \frac{5mn}{96} - \frac{5n}{96} \right) \quad (3.26)
\end{aligned}$$

The covariance of (I_3, I_4) is

$$\begin{aligned}
Cov(I_3, I_4) &= \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{l'=1}^n \sum_{\substack{i < j \\ k \neq l'}} \left[\frac{19}{48} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \sum_{\substack{i < j \\ k' < l' \\ (k', l') \neq k}} \left[\frac{37}{120} - \left(\frac{1}{2}\right)^2 \right] \\
&+ \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l'=1}^n \sum_{\substack{i < j \\ j' \neq (i, j) \\ k \neq l'}} \left[\frac{57}{160} - \left(\frac{1}{2}\right)^2 \right] + \sum_{i=1}^m \sum_{j=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \sum_{\substack{i < j \\ j' \neq (i, j) \\ (k', l') \neq k \\ k' < l'}} \left[\frac{57}{160} - \left(\frac{1}{2}\right)^2 \right] \\
&+ \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{l'=1}^n \sum_{\substack{i < j \\ i' \neq (i, j) \\ k \neq l'}} \left[\frac{67}{240} - \left(\frac{1}{2}\right)^2 \right] \\
&+ \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \sum_{\substack{i < j \\ i' \neq (i, j) \\ k' < l' \\ (k', l') \neq k}} \left[\frac{67}{240} - \left(\frac{1}{2}\right)^2 \right] \\
&+ \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{l'=1}^n \sum_{\substack{i < j \\ i' < j' \\ k \neq l' \\ (i', j') \neq (i, j)}} \left[\frac{931}{2880} - \left(\frac{1}{2}\right)^2 \right] \\
&+ \sum_{i=1}^m \sum_{j=1}^m \sum_{i'=1}^m \sum_{j'=1}^m \sum_{k=1}^n \sum_{k'=1}^n \sum_{l'=1}^n \sum_{\substack{i < j \\ i' < j' \\ k' < l' \\ (i', j') \neq (i, j) \\ (k', l') \neq k}} \left[\frac{1}{4} - \left(\frac{1}{2}\right)^2 \right] \\
&= n \binom{m}{2} (n-1) \left[\frac{19}{48} - \frac{1}{4} \right] + n \binom{m}{2} \binom{n-1}{2} \left[\frac{37}{120} - \frac{1}{4} \right] \\
&+ 2n \binom{m}{2} (n-1)(m-2) \left[\frac{57}{160} - \frac{1}{4} \right] + 2n \binom{m}{2} \binom{n-1}{2} (m-2) \left[\frac{67}{240} - \frac{1}{4} \right] \\
&+ n \binom{m}{2} \binom{m-2}{2} (n-1) \left[\frac{931}{2880} - \frac{1}{4} \right] + n \binom{m}{2} \binom{n-1}{2} \binom{m-2}{2} \left[\frac{1}{4} - \frac{1}{4} \right] \\
&= n \binom{m}{2} \left(\frac{9n}{320} + \frac{1}{960} - \frac{7n^2}{240} - \frac{67mn}{1152} + \frac{167m}{5760} + \frac{7n^2m}{240} + \frac{211nm^2}{5760} - \frac{211m^2}{5760} \right)
\end{aligned} \tag{3.28}$$

Substituting from formula (3.15), (3.17), (3.19), and (3.21), equation (3.29) is the mean of

V.

$$E(V) = EI_1 + EI_2 + EI_3 + 2EI_4 = \frac{mn(mn+1)}{4} \tag{3.29}$$

Substituting from formula (3.16), (3.18), (3.20), and (3.22)~(3.28), the variance of V is

$$\begin{aligned}
 Var(V) = & \frac{15881}{357840}mn + \frac{1333}{1431360}m^4n - \frac{523}{23856}m^3n + \frac{1153}{11928}m^3n^2 + \frac{12697}{477120}nm^2 \\
 & - \frac{13297}{715680}m^2n^2 + \frac{1153}{11928}m^2n^3 + \frac{1333}{1431360}mn^4 - \frac{523}{23856}mn^3 + \frac{12697}{477120}mn^2 \\
 & - \frac{2537}{286272}m^4n^2 + \frac{2461}{59640}m^4n^3 - \frac{449}{9940}m^3n^3 + \frac{2461}{59640}m^3n^4 - \frac{2537}{286272}m^2n^4
 \end{aligned} \tag{3.30}$$

Using formula (3.29) and (3.30), the mean and the variance of V are $\frac{3}{2}$, $\frac{19}{12}$, and 3 , $\frac{45}{8}$, when $m=1$, $n=2$, and $m=1$, $n=3$, respectively. Those results are exactly the same as the ones when we use the exact value of the distribution of V to calculate the mean and the variance of V in Table 2.1.

3.3 Sampling from the normal distribution

For normal case, substituting in formula (3.1) to (3.14) and using the orthant probability values, we have the mean of I_1

$$EI_1 = \sum_{i=1}^m \sum_{k=1}^n \frac{1}{2} = \frac{mn}{2} \tag{3.31}$$

The variance of I_1 is

$$\begin{aligned}
 VarI_1 = & mn\left(\frac{1}{2} - \frac{1}{4}\right) + mn(n-1)\left(\frac{1}{3} - \frac{1}{4}\right) \\
 & + m(m-1)n\left(\frac{1}{3} - \frac{1}{4}\right) + m(m-1)n(n-1)\left(\frac{1}{4} - \frac{1}{4}\right) \\
 = & \frac{mn(m+n+1)}{12}
 \end{aligned} \tag{3.32}$$

The mean of I_2 is

$$EI_2 = \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \frac{1}{2} = \frac{m \binom{n}{2}}{2} \tag{3.33}$$

The variance of I_2 is

$$\begin{aligned}
Var(I_2) &= m \binom{n}{2} \left[\frac{1}{2} - \frac{1}{4} \right] + m \binom{n}{2} (n-2) \left[0.4068 - \frac{1}{4} \right] \\
&\quad + m \binom{n}{2} (n-2) \left[0.4068 - \frac{1}{4} \right] + m \binom{n}{2} \binom{n-2}{2} \left[0.36614 - \frac{1}{4} \right] \\
&\quad + m \binom{n}{2} (m-1) \left[0.304087 - \frac{1}{4} \right] + m \binom{n}{2} (m-1)(n-2) \left[0.27665 - \frac{1}{4} \right] \\
&\quad + m \binom{n}{2} (m-1)(n-2) \left[0.27665 - \frac{1}{4} \right] + m \binom{n}{2} (m-1) \binom{n-2}{2} \left[\frac{1}{4} - \frac{1}{4} \right] \\
&= mn(0.0268915n - 0.0118665 - 0.04406n^2 + 0.029035n^3 - 0.0529065mn \\
&\quad + 0.02665mn^2 + 0.0262565m) \tag{3.34}
\end{aligned}$$

The mean of I_3 is

$$EI_3 = \sum_{k=1}^n \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m \frac{1}{2} = \frac{\binom{m}{2} n}{2} \tag{3.35}$$

The variance of I_3 is

$$\begin{aligned}
Var(I_3) &= n \binom{m}{2} \left[\frac{1}{2} - \frac{1}{4} \right] + n \binom{m}{2} (m-2) \left[0.4068 - \frac{1}{4} \right] \\
&\quad + n \binom{m}{2} (m-2) \left[0.4068 - \frac{1}{4} \right] + n \binom{m}{2} \binom{m-2}{2} \left[0.36614 - \frac{1}{4} \right] \\
&\quad + n \binom{m}{2} (n-1) \left[0.304087 - \frac{1}{4} \right] + n \binom{m}{2} (n-1)(m-2) \left[0.27665 - \frac{1}{4} \right] \\
&\quad + n \binom{m}{2} (n-1)(m-2) \left[0.27665 - \frac{1}{4} \right] + n \binom{m}{2} (n-1) \binom{m-2}{2} \left[\frac{1}{4} - \frac{1}{4} \right] \\
&= mn(0.0268915m - 0.0118665 - 0.04406m^2 + 0.029035m^3 - 0.0529065mn \\
&\quad + 0.02665nm^2 + 0.0262565n) \tag{3.36}
\end{aligned}$$

The mean of I_4 is

$$EI_4 = \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \frac{1}{2} = \frac{\binom{m}{2} \binom{n}{2}}{2} \tag{3.37}$$

The variance of I_4 is

$$\begin{aligned}
& \text{Var } I_4 \\
&= \binom{m}{2} \binom{n}{2} \left(\frac{1}{2} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} (m-2) \left(0.385 - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} (m-2) \left(0.385 - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} \binom{m-2}{2} \left(\frac{1}{3} - \frac{1}{4}\right) \\
&- \frac{1}{4} + \binom{m}{2} \binom{n}{2} (n-2) \left(0.385 - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} (n-2) \left(0.385 - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} \binom{n-2}{2} \left(\frac{1}{3} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} \binom{n-2}{2} (m-2) \left(0.2902 - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} \binom{n-2}{2} (m-2) \left(0.2902 - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} \binom{m-2}{2} (n-2) \left(0.2902 - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} \binom{m-2}{2} (n-2) \left(0.2902 - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} \binom{n-2}{2} \binom{m-2}{2} \left(\frac{1}{4} - \frac{1}{4}\right) \\
&= \binom{m}{2} \binom{n}{2} \left(-0.068667mn - 0.0387333n^2 + 0.0402mn^2 + 0.0402nm^2 + 0.0382m \right. \\
&\quad \left. + 0.0382n - 0.0387333m^2 + 0.0385333\right) \tag{3.38}
\end{aligned}$$

And, the covariance of (I_1, I_2) is

$$\begin{aligned}
\text{Cov}(I_1, I_2) &= mn(n-1) \left[\frac{5}{12} - \frac{1}{4}\right] + mn \binom{n-1}{2} \left[0.3480 - \frac{1}{4}\right] \\
&+ mn(m-1)(n-1) \left[0.2966 - \frac{1}{4}\right] + mn(m-1) \binom{n-1}{2} \left[\frac{1}{4} - \frac{1}{4}\right] \\
&= mn(-0.026933n - 0.0220667 + 0.049n^2 + 0.0466mn - 0.0466m) \tag{3.39}
\end{aligned}$$

The covariance of (I_1, I_3) is

$$\begin{aligned}
Cov(I_1, I_3) &= mn(m-1)\left[\frac{5}{12} - \frac{1}{4}\right] + mn \binom{m-1}{2} \left[0.3480 - \frac{1}{4}\right] \\
&\quad + mn(m-1)(n-1)\left[0.2966 - \frac{1}{4}\right] + mn(n-1) \binom{m-1}{2} \left[\frac{1}{4} - \frac{1}{4}\right] \\
&= mn(-0.026933m - .0220667 + .049m^2 + .0466mn - .0466n)
\end{aligned} \tag{3.40}$$

The covariance of (I_1, I_4) is

$$\begin{aligned}
Cov(I_1, I_4) &= mn(m-1)(n-1)\left[\frac{3}{8} - \frac{1}{4}\right] + mn \binom{m-1}{2} (n-1)\left[0.3075 - \frac{1}{4}\right] \\
&\quad + mn \binom{n-1}{2} (m-1)\left[0.3075 - \frac{1}{4}\right] + mn \binom{m-1}{2} \binom{n-1}{2} \left[\frac{1}{4} - \frac{1}{4}\right] \\
&= mn(-0.0475mn + 0.01875m + 0.01875n + .01 + 0.02875m^2n \\
&\quad - 0.02875m^2 + 0.02875mn^2 - 0.02875n^2)
\end{aligned} \tag{3.41}$$

The covariance of (I_2, I_3) is

$$\begin{aligned}
Cov(I_2, I_3) &= 2m \binom{n}{2} (m-1)\left[0.3661 - \frac{1}{4}\right] + 2m \binom{n}{2} \binom{m-1}{2} \left[0.304087 - \frac{1}{4}\right] \\
&\quad + m \binom{n}{2} (n-2)(m-1)\left[0.304087 - \frac{1}{4}\right] \\
&\quad + m \binom{n}{2} \binom{m-1}{2} (n-2)\left[\frac{1}{4} - \frac{1}{4}\right] \\
&= m \binom{n}{2} (-.038235m - .015852 + .054087m^2 + .054087mn - .054087n)
\end{aligned} \tag{3.42}$$

The covariance of (I_2, I_4) is

$$\begin{aligned}
Cov(I_2, I_4) &= m \binom{n}{2} (m-1) [0.40204 - \frac{1}{4}] + m \binom{n}{2} \binom{m-1}{2} [0.31693 - \frac{1}{4}] \\
&\quad + 2m \binom{n}{2} (m-1)(n-2) [0.3549 - \frac{1}{4}] \\
&\quad + 2m \binom{n}{2} \binom{m-1}{2} (n-2) [0.28272 - \frac{1}{4}] \\
&\quad + m \binom{n}{2} \binom{n-2}{2} (m-1) [0.31693 - \frac{1}{4}] \\
&\quad + m \binom{n}{2} \binom{m-1}{2} \binom{n-2}{2} [\frac{1}{4} - \frac{1}{4}] \\
&= m \binom{n}{2} (.029155m + .00282 - .031975m^2 - .055685mn + .022965n \\
&\quad + .03272m^2n + .033465mn^2 - .033465n^2) \tag{3.43}
\end{aligned}$$

The covariance of (I_3, I_4) is

$$\begin{aligned}
Cov(I_3, I_4) &= n \binom{m}{2} (n-1) [0.40204 - \frac{1}{4}] + n \binom{m}{2} \binom{n-1}{2} [0.3169 - \frac{1}{4}] \\
&\quad + 2n \binom{m}{2} (n-1)(m-2) [0.3549 - \frac{1}{4}] \\
&\quad + 2n \binom{m}{2} \binom{n-1}{2} (m-2) [0.28272 - \frac{1}{4}] \\
&\quad + n \binom{m}{2} \binom{m-2}{2} (n-1) [0.31693 - \frac{1}{4}] \\
&\quad + n \binom{m}{2} \binom{n-1}{2} \binom{m-2}{2} [\frac{1}{4} - \frac{1}{4}] \\
&= n \binom{m}{2} (.029155n + .00282 - .031975n^2 - .055685mn + .022965m \\
&\quad + .03272n^2m + .033465nm^2 - .033465m^2) \tag{3.44}
\end{aligned}$$

Substituting from formula (3.31), (3.33), (3.35), and (3.37), the equation (3.45) is the mean of V ,

$$E(V) = EI_1 + EI_2 + EI_3 + 2EI_4 = \frac{mn(mn+1)}{4} \tag{3.45}$$

Substituting from formula (3.32), (3.34), (3.36), and (3.38)~(3.44), the variance of V is as follow.

$$\begin{aligned}
Var(V) = & .0544388mn + .0008383m^4n - .0152703m^3n + .0738473m^3n^2 + .003717nm^2 \\
& +.0310913m^2n^2 + .0738473m^2n^3 + .0008383mn^4 - .0152703mn^3 + .003717mn^2 \\
& - .0120033m^4n^2 + .0402m^4n^3 - .018187m^3n^3 + .0402m^3n^4 - .0120033m^2n^4
\end{aligned} \tag{3.46}$$

Using formula (3.45), the mean of V are $\frac{3}{2}$ and 3, when m=1, n=2, and m=1, n=3, respectively. This results are exactly the same as the ones when we use the exact value of the distribution of V to calculate the mean of V in Table 2.1.

Using formula (3.46), the variance of V are 1.583337 and 5.5228808, when m=1, n=2, and m=1, n=3, respectively. This results are very close to the ones when we use the exact value of the distribution of V to calculate the variance of V in Table 2.1.

3.4 Sampling from the exponential distribution

For exponential case, substituting in formula (3.1) to (3.14) with the probability values from integral methods, we have the mean of I_1

$$EI_1 = \sum_{i=1}^m \sum_{k=1}^n \frac{1}{2} = \frac{mn}{2} \tag{3.47}$$

Substituting in formula (3.2) the probability values from integral methods, we have the variance of I_1 .

$$\begin{aligned}
Var I_1 = & mn\left(\frac{1}{2} - \frac{1}{4}\right) + mn(n-1)\left(\frac{1}{3} - \frac{1}{4}\right) \\
& + m(m-1)n\left(\frac{1}{3} - \frac{1}{4}\right) + m(m-1)n(n-1)\left(\frac{1}{4} - \frac{1}{4}\right) \\
= & \frac{mn(m+n+1)}{12}
\end{aligned} \tag{3.48}$$

$$EI_2 = \sum_{i=1}^m \sum_{\substack{k=1 \\ k < l}}^n \sum_{l=1}^n \frac{5}{9} = \frac{5m \binom{n}{2}}{9} \quad (3.49)$$

$$\begin{aligned} \text{Var}(I_2) &= m \binom{n}{2} \left[\frac{5}{9} - \left(\frac{5}{9}\right)^2 \right] + m \binom{n}{2} (n-2) \left[\frac{7}{15} - \left(\frac{5}{9}\right)^2 \right] \\ &\quad + m \binom{n}{2} (n-2) \left[\frac{7}{15} - \left(\frac{5}{9}\right)^2 \right] + m \binom{n}{2} \binom{n-2}{2} \left[\frac{53}{125} - \left(\frac{5}{9}\right)^2 \right] \\ &\quad + m \binom{n}{2} (m-1) \left[\frac{13}{36} - \left(\frac{5}{9}\right)^2 \right] + m \binom{n}{2} (m-1)(n-2) \left[\frac{1}{3} - \left(\frac{5}{9}\right)^2 \right] \\ &\quad + m \binom{n}{2} (m-1)(n-2) \left[\frac{1}{3} - \left(\frac{5}{9}\right)^2 \right] + m \binom{n}{2} (m-1) \binom{n-2}{2} \left[\frac{25}{81} - \left(\frac{5}{9}\right)^2 \right] \\ &= mn \left(\frac{1171n}{81000} - \frac{97}{27000} - \frac{134n^2}{3375} + \frac{292n^3}{10125} - \frac{31mn}{648} + \frac{2mn^2}{81} + \frac{5m}{216} \right) \quad (3.50) \end{aligned}$$

$$EI_3 = \sum_{k=1}^n \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m \frac{4}{9} = \frac{4 \binom{m}{2} n}{9} \quad (3.51)$$

Substituting in formula (3.6) the probability values from integral methods, we have the variance of I_3 .

$$\begin{aligned} \text{Var}(I_3) &= n \binom{m}{2} \left[\frac{4}{9} - \left(\frac{4}{9}\right)^2 \right] + n \binom{m}{2} (m-2) \left[\frac{16}{45} - \left(\frac{4}{9}\right)^2 \right] \\ &\quad + n \binom{m}{2} (m-2) \left[\frac{16}{45} - \left(\frac{4}{9}\right)^2 \right] + n \binom{m}{2} \binom{m-2}{2} \left[\frac{352}{1125} - \left(\frac{4}{9}\right)^2 \right] \\ &\quad + n \binom{m}{2} (n-1) \left[\frac{1}{4} - \left(\frac{4}{9}\right)^2 \right] + n \binom{m}{2} (n-1)(m-2) \left[\frac{2}{9} - \left(\frac{4}{9}\right)^2 \right] \\ &\quad + n \binom{m}{2} (n-1)(m-2) \left[\frac{2}{9} - \left(\frac{4}{9}\right)^2 \right] + n \binom{m}{2} (n-1) \binom{m-2}{2} \left[\frac{16}{81} - \left(\frac{4}{9}\right)^2 \right] \\ &= mn \left(\frac{1171m}{81000} - \frac{97}{27000} - \frac{134m^2}{3375} + \frac{292m^3}{10125} - \frac{31mn}{648} + \frac{2nm^2}{81} + \frac{5n}{216} \right) \quad (3.52) \end{aligned}$$

$$EI_4 = \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n \frac{1}{2} = \frac{\binom{m}{2} \binom{n}{2}}{2} \quad (3.53)$$

$Var I_4$

$$\begin{aligned}
&= \binom{m}{2} \binom{n}{2} \left(\frac{1}{2} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} (m-2) \left(\frac{7}{18} - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} (m-2) \left(\frac{7}{18} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} \binom{m-2}{2} \left(\frac{1}{3} - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} (n-2) \left(\frac{7}{18} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} (n-2) \left(\frac{7}{18} - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} \binom{n-2}{2} \left(\frac{1}{3} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} \binom{n-2}{2} (m-2) \left(\frac{125}{432} - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} \binom{n-2}{2} (m-2) \left(\frac{125}{432} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} \binom{m-2}{2} (n-2) \left(\frac{125}{432} - \frac{1}{4}\right) \\
&+ \binom{m}{2} \binom{n}{2} \binom{m-2}{2} (n-2) \left(\frac{125}{432} - \frac{1}{4}\right) + \binom{m}{2} \binom{n}{2} \binom{n-2}{2} \binom{m-2}{2} \left(\frac{1}{4} - \frac{1}{4}\right) \\
&= \binom{m}{2} \binom{n}{2} \left(-\frac{13mn}{216} + \frac{17m^2n}{432} - \frac{n^2}{27} + \frac{17mn^2}{432} + \frac{7m}{216} + \frac{7n}{216} - \frac{m^2}{27} + \frac{1}{36}\right) \quad (3.54)
\end{aligned}$$

$Cov(I_1, I_2)$

$$\begin{aligned}
&= mn(n-1) \left[\frac{4}{9} - \left(\frac{1}{2}\right) \left(\frac{5}{9}\right)\right] + mn \binom{n-1}{2} \left[\frac{3}{8} - \left(\frac{1}{2}\right) \left(\frac{5}{9}\right)\right] \\
&+ mn(m-1)(n-1) \left[\frac{29}{90} - \left(\frac{1}{2}\right) \left(\frac{5}{9}\right)\right] + mn(m-1) \binom{n-1}{2} \left[\frac{5}{18} - \left(\frac{1}{2}\right) \left(\frac{5}{9}\right)\right] \\
&= mn \left(-\frac{17n}{720} - \frac{1}{40} + \frac{7n^2}{144} + \frac{2mn}{45} - \frac{2m}{45}\right) \quad (3.55)
\end{aligned}$$

$Cov(I_1, I_3)$

$$\begin{aligned}
&= mn(m-1) \left[\frac{7}{18} - \left(\frac{1}{2}\right) \left(\frac{4}{9}\right)\right] + mn \binom{m-1}{2} \left[\frac{23}{72} - \left(\frac{1}{2}\right) \left(\frac{4}{9}\right)\right] \\
&+ mn(m-1)(n-1) \left[\frac{4}{15} - \left(\frac{1}{2}\right) \left(\frac{4}{9}\right)\right] + mn(n-1) \binom{m-1}{2} \left[\frac{2}{9} - \left(\frac{1}{2}\right) \left(\frac{4}{9}\right)\right] \\
&= mn \left(-\frac{17m}{720} - \frac{1}{40} + \frac{7m^2}{144} + \frac{2mn}{45} - \frac{2n}{45}\right) \quad (3.56)
\end{aligned}$$

$$\text{Cov}(I_1, I_4)$$

$$\begin{aligned} &= mn(m-1)(n-1)\left[\frac{3}{8} - \frac{1}{4}\right] + mn \binom{m-1}{2} (n-1)\left[\frac{11}{36} - \frac{1}{4}\right] \\ &\quad + mn \binom{n-1}{2} (m-1)\left[\frac{11}{36} - \frac{1}{4}\right] + mn \binom{m-1}{2} \binom{n-1}{2} \left[\frac{1}{4} - \frac{1}{4}\right] \\ &= mn\left(-\frac{mn}{24} + \frac{m}{72} + \frac{n}{72} + \frac{1}{72} + \frac{m^2n}{36} - \frac{m^2}{36} + \frac{mn^2}{36} - \frac{n^2}{36}\right) \end{aligned} \quad (3.57)$$

$$\text{Cov}(I_2, I_3)$$

$$\begin{aligned} &= 2m \binom{n}{2} (m-1)\left[\frac{13}{36} - \left(\frac{5}{9}\right)\left(\frac{4}{9}\right)\right] + 2m \binom{n}{2} \binom{m-1}{2} \left[\frac{44}{147} - \left(\frac{5}{9}\right)\left(\frac{4}{9}\right)\right] \\ &\quad + m \binom{n}{2} (n-2)(m-1)\left[\frac{44}{147} - \left(\frac{5}{9}\right)\left(\frac{4}{9}\right)\right] + m \binom{n}{2} \binom{m-1}{2} (n-2)\left[\frac{20}{81} - \left(\frac{5}{9}\right)\left(\frac{4}{9}\right)\right] \\ &= m \binom{n}{2} \left(-\frac{89m}{2646} - \frac{149}{7938} + \frac{208m^2}{3969} + \frac{208mn}{3969} - \frac{208n}{3969}\right) \end{aligned} \quad (3.58)$$

$$\text{Cov}(I_2, I_4)$$

$$\begin{aligned} &= m \binom{n}{2} (m-1)\left[\frac{31}{72} - \left(\frac{5}{9}\right)\left(\frac{1}{2}\right)\right] + m \binom{n}{2} \binom{m-1}{2} \left[\frac{773}{2250} - \left(\frac{5}{9}\right)\left(\frac{1}{2}\right)\right] \\ &\quad + 2m \binom{n}{2} (m-1)(n-2)\left[\frac{55}{144} - \left(\frac{5}{9}\right)\left(\frac{1}{2}\right)\right] + 2m \binom{n}{2} \binom{m-1}{2} (n-2)\left[\frac{139}{450} - \left(\frac{5}{9}\right)\left(\frac{1}{2}\right)\right] \\ &\quad + m \binom{n}{2} \binom{n-2}{2} (m-1)\left[\frac{11}{32} - \left(\frac{5}{9}\right)\left(\frac{1}{2}\right)\right] + m \binom{n}{2} \binom{m-1}{2} \binom{n-2}{2} \left[\frac{5}{18} - \left(\frac{5}{9}\right)\left(\frac{1}{2}\right)\right] \\ &= m \binom{n}{2} \left(\frac{793m}{36000} + \frac{263}{36000} - \frac{11m^2}{375} - \frac{719mn}{14400} + \frac{271n}{14400} + \frac{7m^2n}{225} + \frac{19mn^2}{576} - \frac{19n^2}{576}\right) \end{aligned} \quad (3.59)$$

$$\text{Cov}(I_3, I_4)$$

$$\begin{aligned} &= n \binom{m}{2} (n-1)\left[\frac{3}{8} - \left(\frac{4}{9}\right)\left(\frac{1}{2}\right)\right] + n \binom{m}{2} \binom{n-1}{2} \left[\frac{36}{125} - \left(\frac{4}{9}\right)\left(\frac{1}{2}\right)\right] \\ &\quad + 2n \binom{m}{2} (n-1)(m-2)\left[\frac{47}{144} - \left(\frac{4}{9}\right)\left(\frac{1}{2}\right)\right] + 2n \binom{m}{2} \binom{n-1}{2} (m-2)\left[\frac{19}{75} - \left(\frac{4}{9}\right)\left(\frac{1}{2}\right)\right] \\ &\quad + n \binom{m}{2} \binom{m-2}{2} (n-1)\left[\frac{83}{288} - \left(\frac{4}{9}\right)\left(\frac{1}{2}\right)\right] + n \binom{m}{2} \binom{n-1}{2} \binom{m-2}{2} \left[\frac{2}{9} - \left(\frac{4}{9}\right)\left(\frac{1}{2}\right)\right] \\ &= n \binom{m}{2} \left(\frac{793n}{36000} + \frac{263}{36000} - \frac{11n^2}{375} - \frac{719mn}{14400} + \frac{271m}{14400} + \frac{7n^2m}{225} + \frac{19nm^2}{576} - \frac{19m^2}{576}\right) \end{aligned} \quad (3.60)$$

Substituting from (3.47), (3.49), (3.51), and (3.53), equation (3.61) is the mean of V.

$$\begin{aligned} E(V) = EI_1 + EI_2 + EI_3 + 2EI_4 &= \frac{mn(9mn - m + n + 9)}{36} \\ &= \frac{mn(mn + 1)}{4} + \frac{mn(n - m)}{36} \end{aligned} \quad (3.61)$$

Substituting from (3.48), (3.50), (3.52), and (3.54)~(3.60), equation (3.62) is the variance of V.

$$\begin{aligned} Var(V) = & \frac{194599}{3969000}mn - \frac{31}{324000}m^4n - \frac{104507}{7938000}m^3n + \frac{93229}{1323000}m^3n^2 + \frac{183601}{15876000}nm^2 \\ & + \frac{158771}{7938000}m^2n^2 + \frac{93229}{1323000}m^2n^3 - \frac{31}{324000}mn^4 - \frac{104507}{7938000}mn^3 + \frac{183601}{15876000}mn^2 \\ & - \frac{1}{96}m^4n^2 + \frac{17}{432}m^4n^3 - \frac{13}{900}m^3n^3 + \frac{17}{432}m^3n^4 - \frac{1}{96}m^2n^4 \end{aligned} \quad (3.62)$$

Using formula (3.61) and (3.62), the mean and the variance of V are $\frac{14}{9}$, $\frac{128}{81}$, and $\frac{19}{6}$, $\frac{497}{90}$, when $m=1$, $n=2$, and $m=1$, $n=3$, respectively. Those results are exactly same as the results when we use the exact value of the distribution of V to calculate the mean and the variance of V in Table 2.1.

3.5 Mean and Variance of V among the tied observations

Let x_1, \dots, x_m and y_1, \dots, y_n be two sets of measurements taken from distribution F. Let $\mathbf{Z} = \{z_1, \dots, z_N\} = \{x_1, \dots, x_m, y_1, \dots, y_n\}$, $N=m+n$. Every partition of the z 's into two sets of m and n observations has equal probability.

If distribution F is continuous, the probability of ties among the z_k is zero. If there are ties among the z_k , the distribution of V depends on these specific ties. Ties may occur within groups or across them. Ties within groups have no effect on the test statistic, but those across groups do. When we use the large-sample approximation, we may adjust the formula for the test statistic.

Let t_1 be the number of observations of the smallest observed rank and t_2 that of the next smallest rank of \mathbf{Z} , etc, so that $t_1 + t_2 + \dots = N$. Let p be the probability that if two ranks are selected at random from among z_1, \dots, z_N , they are not equal. Thus

$$p = 1 - \frac{\sum t(t-1)}{N(N-1)}.$$

Similarly, if three ranks are selected at random from among z_1, \dots, z_N , the three ranks are not all equal, the probability is as follow:

$$p = 1 - \frac{\sum t(t-1)(t-2)}{N(N-1)(N-2)},$$

Let F be normal distribution. Then the adjust formulas are as follow:

$$\begin{aligned} & \text{var} I_1 \\ &= mn\left(\frac{1}{2} - \frac{1}{4}\right)\left(1 - \frac{\sum t(t-1)}{N(N-1)}\right) + mn(n-1)\left(\frac{1}{3} - \frac{1}{4}\right)\left(1 - \frac{\sum t(t-1)(t-2)}{N(N-1)(N-2)}\right) \\ & \quad + m(m-1)n\left(\frac{1}{3} - \frac{1}{4}\right)\left(1 - \frac{\sum t(t-1)(t-2)}{N(N-1)(N-2)}\right) \\ &= \frac{mn(N+1)}{12}\left(1 - \frac{\sum (t^3 - t)}{N^3 - N}\right) \end{aligned} \quad (3.63)$$

similarly, the variance of I_2 is

$Var I_2 =$

$$\begin{aligned}
& m \binom{n}{2} \left[\frac{1}{2} - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)}{N(N-1)(N-2)} \right] \\
& + m \binom{n}{2} (n-2) \left[0.4068 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right] \\
& + m \binom{n}{2} (n-2) \left[0.4068 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right] \\
& + m \binom{n}{2} \binom{n-2}{2} \left[0.36614 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right] \\
& + m \binom{n}{2} (m-1) \left[0.304087 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right] \\
& + m \binom{n}{2} (m-1)(n-2) \left[0.27665 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right] \\
& + m \binom{n}{2} (m-1)(n-2) \left[0.27665 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right]
\end{aligned} \tag{3.64}$$

If more than four ranks are selected at random from among z_1, \dots, z_N , when we use the large-sample approximation, the probabilities that they are equal have negligible effect unless a large proportion of observations are tied or there are ties of considerable extent. Therefore, the variance of I_2 is approximately as follow:

$$\begin{aligned}
& Var I_2 \\
& \approx mn(0.0268915n - 0.0118665 - 0.04406n^2 + 0.029035n^3 - 0.0529065mn \\
& + 0.02665mn^2 + 0.0262565m) - m \binom{n}{2} \left[\frac{\sum t(t-1)(t-2)}{4N(N-1)(N-2)} \right. \\
& \left. + (0.3136(n-2) + 0.054087(m-1)) \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right]
\end{aligned} \tag{3.65}$$

Similarly, the variance of I_3 , $Var I_3$, is as follow:

$$\begin{aligned}
& n \binom{m}{2} \left[\frac{1}{2} - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)}{N(N-1)(N-2)} \right] \\
& + n \binom{m}{2} (m-2) \left[0.4068 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right] \\
& + n \binom{m}{2} (m-2) \left[0.4068 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right] \\
& + n \binom{m-2}{2} \binom{m}{2} \left[0.36614 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right] \\
& + n \binom{m}{2} (n-1) \left[0.304087 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right] \\
& + n \binom{m}{2} (n-1)(m-2) \left[0.27665 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right] \\
& + n \binom{m}{2} (n-1)(m-2) \left[0.27665 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right] \\
& \approx mn(0.0268915m - 0.0118665 - 0.04406m^2 + 0.029035m^3 - 0.0529065mn \\
& + 0.02665nm^2 + 0.0262565n) - n \binom{m}{2} \left[\frac{\sum t(t-1)(t-2)}{4N(N-1)(N-2)} \right. \\
& \left. + (0.3136(m-2) + 0.054087(n-1)) \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right]
\end{aligned} \tag{3.66}$$

the variance of I_4 , $Var I_4$, equals

$$\begin{aligned}
& \binom{m}{2} \binom{n}{2} \left(\frac{1}{2} - \frac{1}{4} \right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right) \\
& + \binom{m}{2} \binom{n}{2} (m-2) \left(0.385 - \frac{1}{4} \right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right) \\
& + \binom{m}{2} \binom{n}{2} (m-2) \left(0.385 - \frac{1}{4} \right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right) \\
& + \binom{m}{2} \binom{n}{2} \binom{m-2}{2} \left(\frac{1}{3} - \frac{1}{4} \right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)}{N(N-1)(N-2)(N-3)(N-4)(N-5)} \right) \\
& + \binom{m}{2} \binom{n}{2} (n-2) \left(0.385 - \frac{1}{4} \right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right) \\
& + \binom{m}{2} \binom{n}{2} (n-2) \left(0.385 - \frac{1}{4} \right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right) \\
& + \binom{m}{2} \binom{n}{2} \binom{n-2}{2} \left(\frac{1}{3} - \frac{1}{4} \right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)}{N(N-1)(N-2)(N-3)(N-4)(N-5)} \right)
\end{aligned}$$

$$\begin{aligned}
& + \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)}{N(N-1)(N-2)(N-3)(N-4)(N-5)}\right) \\
& + \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)}{N(N-1)(N-2)(N-3)(N-4)(N-5)}\right) \\
& + \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)}{N(N-1)(N-2)(N-3)(N-4)(N-5)}\right) \\
& + \binom{m}{2} \binom{n}{2} (n-2)(m-2) \left(\frac{1}{3} - \frac{1}{4}\right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)}{N(N-1)(N-2)(N-3)(N-4)(N-5)}\right) \\
& + \binom{m}{2} \binom{n}{2} \binom{n-2}{2} (m-2) \left(0.2902 - \frac{1}{4}\right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)(t-6)}{N(N-1)(N-2)(N-3)(N-4)(N-5)(N-6)}\right) \\
& + \binom{m}{2} \binom{n}{2} \binom{n-2}{2} (m-2) \left(0.2902 - \frac{1}{4}\right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)(t-6)}{N(N-1)(N-2)(N-3)(N-4)(N-5)(N-6)}\right) \\
& + \binom{m}{2} \binom{n}{2} \binom{m-2}{2} (n-2) \left(0.2902 - \frac{1}{4}\right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)(t-6)}{N(N-1)(N-2)(N-3)(N-4)(N-5)(N-6)}\right) \\
& + \binom{m}{2} \binom{n}{2} \binom{m-2}{2} (n-2) \left(0.2902 - \frac{1}{4}\right) \left(1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)(t-6)}{N(N-1)(N-2)(N-3)(N-4)(N-5)(N-6)}\right) \\
& \approx \binom{m}{2} \binom{n}{2} (-0.068667mn - 0.0387333n^2 + 0.0402mn^2 + 0.0402nm^2 + 0.0382m \\
& \quad + 0.0382n - 0.0387333m^2 + 0.0385333) - \binom{m}{2} \binom{n}{2} \frac{\sum t(t-1)(t-2)(t-3)}{4N(N-1)(N-2)(N-3)}
\end{aligned} \tag{3.67}$$

Using the orthant probability values, the covariance of (I_1, I_2) is as follow:

$$\begin{aligned}
& Cov(I_1, I_2) \\
& = mn(n-1) \left[\frac{5}{12} - \frac{1}{4}\right] \left[1 - \frac{\sum t(t-1)(t-2)}{N(N-1)(N-2)}\right] \\
& \quad + mn \binom{n-1}{2} \left[0.3480 - \frac{1}{4}\right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)}\right] \\
& \quad + mn(m-1)(n-1) \left[0.2966 - \frac{1}{4}\right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)}\right] \\
& \approx mn(-0.026933n - 0.0220667 + 0.049n^2 + 0.0466mn - 0.0466m) - mn(n-1) \\
& \quad \left[\frac{\sum t(t-1)(t-2)}{6N(N-1)(N-2)} + (0.049(n-2) + 0.0466(m-1)) \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)}\right]
\end{aligned} \tag{3.68}$$

The covariance of (I_1, I_3) is as follow:

$Cov(I_1, I_3)$

$$\begin{aligned}
&= mn(m-1)\left[\frac{5}{12} - \frac{1}{4}\right]\left[1 - \frac{\sum t(t-1)(t-2)}{N(N-1)(N-2)}\right] \\
&\quad + mn\binom{m-1}{2}\left[0.3480 - \frac{1}{4}\right]\left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)}\right] \\
&\quad + mn(m-1)(n-1)\left[0.2966 - \frac{1}{4}\right]\left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)}\right] \\
&\approx mn(-0.026933m - .0220667 + .049m^2 + .0466mn - .0466n) - mn(m-1) \\
&\quad \left[\frac{\sum t(t-1)(t-2)}{6N(N-1)(N-2)} + (0.049(m-2) + 0.0466(n-1))\frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)}\right]
\end{aligned} \tag{3.69}$$

the covariance of I_4 under H_{00} is as follow:

$Cov(I_1, I_4)$

$$\begin{aligned}
&= mn(m-1)(n-1)\left[\frac{3}{8} - \frac{1}{4}\right]\left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)}\right] \\
&\quad + mn\binom{m-1}{2}(n-1)\left[0.3075 - \frac{1}{4}\right]\left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)}\right] \\
&\quad + mn\binom{n-1}{2}(m-1)\left[0.3075 - \frac{1}{4}\right]\left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)}\right] \\
&\approx mn(-0.0475mn + 0.01875m + 0.01875n + .01 + 0.02875m^2n - 0.02875m^2 \\
&\quad + 0.02875mn^2 - 0.02875n^2) - \binom{m}{2}\binom{n}{2}\frac{\sum t(t-1)(t-2)(t-3)}{2N(N-1)(N-2)(N-3)}
\end{aligned} \tag{3.70}$$

$$\begin{aligned}
& Cov(I_2, I_3) \\
&= m \binom{n}{2} (m-1) \left[0.3661 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right] \\
&\quad + m \binom{n}{2} (m-1) \left[0.3661 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right] \\
&\quad + m \binom{n}{2} \binom{m-1}{2} \left[0.304087 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right] \\
&\quad + m \binom{n}{2} \binom{m-1}{2} \left[0.304087 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right] \\
&\quad + m \binom{n}{2} (n-2)(m-1) \left[0.304087 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right] \\
&= m \binom{n}{2} (-.038235m - .015852 + .054087m^2 + .054087mn - .054087n) \\
&\quad - 0.4644 \binom{m}{2} \binom{n}{2} \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)}
\end{aligned}$$

the covariance of (I_2, I_4) under H_{00} is as follow:

$$\begin{aligned}
& Cov(I_2, I_4) \\
&= m \binom{n}{2} (m-1) \left[0.40204 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right] \\
&\quad + m \binom{n}{2} \binom{m-1}{2} \left[0.31693 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right] \\
&\quad + m \binom{n}{2} (m-1)(n-2) \left[0.3549 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right] \\
&\quad + m \binom{n}{2} (m-1)(n-2) \left[0.3549 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)} \right] \\
&\quad + m \binom{n}{2} \binom{m-1}{2} (n-2) \left[0.28272 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)}{N(N-1)(N-2)(N-3)(N-4)(N-5)} \right] \\
&\quad + m \binom{n}{2} \binom{m-1}{2} (n-2) \left[0.28272 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)}{N(N-1)(N-2)(N-3)(N-4)(N-5)} \right] \\
&\quad + m \binom{n}{2} \binom{n-2}{2} (m-1) \left[0.31693 - \frac{1}{4} \right] \left[1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)}{N(N-1)(N-2)(N-3)(N-4)(N-5)} \right] \\
&\approx m \binom{n}{2} (.029155m + .00282 - .031975m^2 - .055685mn + .022965n + .03272m^2n \\
&\quad + .033465mn^2 - .033465n^2) - 0.30408 \binom{m}{2} \binom{n}{2} \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)}
\end{aligned}$$

(3.71)

the covariance of (I_3, I_4) is as follow:

$$\begin{aligned}
& Cov(I_3, I_4) \\
&= n \binom{m}{2} (n-1) [0.40204 - \frac{1}{4}] [1 - \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)}] \\
&\quad + n \binom{m}{2} \binom{n-1}{2} [0.3169 - \frac{1}{4}] [1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)}] \\
&\quad + n \binom{m}{2} (n-1)(m-2) [0.3549 - \frac{1}{4}] [1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)}] \\
&\quad + n \binom{m}{2} (n-1)(m-2) [0.3549 - \frac{1}{4}] [1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)}{N(N-1)(N-2)(N-3)(N-4)}] \\
&\quad + n \binom{m}{2} \binom{n-1}{2} (m-2) [0.28272 - \frac{1}{4}] [1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)}{N(N-1)(N-2)(N-3)(N-4)(N-5)}] \\
&\quad + n \binom{m}{2} \binom{n-1}{2} (m-2) [0.28272 - \frac{1}{4}] [1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)}{N(N-1)(N-2)(N-3)(N-4)(N-5)}] \\
&\quad + n \binom{m}{2} \binom{m-2}{2} (n-1) [0.31693 - \frac{1}{4}] [1 - \frac{\sum t(t-1)(t-2)(t-3)(t-4)(t-5)}{N(N-1)(N-2)(N-3)(N-4)(N-5)}] \\
&\approx n \binom{m}{2} (.029155n + .00282 - .031975n^2 - .055685mn + .022965m + .03272n^2m \\
&\quad + .033465nm^2 - .033465m^2) - 0.30408 \binom{m}{2} \binom{n}{2} \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)}
\end{aligned} \tag{3.72}$$

Therefore, the variance of V is obtained.

$$\begin{aligned}
& Var(V) \approx \\
&.0544388mn + .0008383m^4n - .0152703m^3n + .0738473m^3n^2 + .003717nm^2 + .0310913m^2n^2 \\
&+ .0738473m^2n^3 + .0008383mn^4 - .0152703mn^3 + .003717mn^2 - .0120033m^4n^2 + .0402m^4n^3 \\
&- .018187m^3n^3 + .0402m^3n^4 - .0120033m^2n^4 - \left\{ \frac{mn \sum t(t-1)(13t-20)}{24N(N-1)} \right. \\
&\quad \left. + [1.5288m \binom{n}{3} + 1.5288n \binom{m}{3} + 6.5778 \binom{m}{2} \binom{n}{2}] \frac{\sum t(t-1)(t-2)(t-3)}{N(N-1)(N-2)(N-3)} \right\}
\end{aligned} \tag{3.73}$$

Chapter 4

Approximations

4.1 Normal approximation

In Tables 4.1, 4.2 we examine the accuracy of normal approximation at the two significance levels $\alpha = 0.01, 0.05$. In each cell we calculate the standard normal probabilities. The procedure is as follows. Using a continuity correction of $1/2$, the normal probabilities in the table correspond to the unit normal is approximated by

$$Pr(\mathbf{V} \leq v) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt,$$

where

$$x = \frac{v - E(V) + \frac{1}{2}}{\sigma(V)}.$$

The value v comes from Tables 2.3 and 2.4, the values of $E(V)$ and $\sigma(V)$ for unit uniform, standard normal, and exponential with unit mean distribution see Chapter 3, the values of $E(V)$ and $\sigma(V)$ for Cauchy, and double exponential with unit mean distribution are the simulation results with 1,000,000 replications.

For example, let $m = n = 5$, the critical values v from Table 2.3 are 5, 5, 5 and 33, 32, 32 for unit uniform, standard normal, and exponential with unit mean distribution, the critical

values v from Table 2.4 are 5, 5 and 31, 31 for Cauchy, and double exponential with unit mean distribution at the two significance levels $\alpha = 0.01, 0.05$. Using formula (3.29) and (3.30), $E(V) = 162.5, \sigma(V) = 77.7$ for unit uniform distribution. Using formula (3.45) and (3.46), $E(V) = 162.5, \sigma(V) = 78.0$ for standard normal distribution. Using formula (3.61) and (3.62), $E(V) = 162.5, \sigma(V) = 77.7$ for exponential with unit mean distribution. The simulation ones of mean of V and standard deviation of V are 162.48, 162.63 and 78.43, 78.26 for Cauchy, and double exponential with unit mean distribution, respectively. Therefore, using the above formula, the normal approximation probabilities was found to be .0217, .0221, .0217, .0220, and .0217 at the significance level $\alpha = 0.01$, the probabilities was found to be .0484, .0478, .0472, .0475, and .0469 at the significance levels $\alpha = 0.05$ for unit uniform, standard normal, exponential, cauchy, and double exponential with unit mean distribution, respectively.

In Tables 4.1 and 4.2, it was found that for smaller samples the normal probability of the test based on standard normal can be much larger than the true nominal level. For example, when the true nominal level is .01, the normal probabilities was found to be .0217, .0221, .0217, .0220, and .0217 when $m = n = 5$ for unit uniform, standard normal, exponential, cauchy, and double exponential with unit mean distribution, respectively. Thus the use of standard normal may not be advisable for these smaller sample sizes. However, for the larger sample sizes considered, the normal probabilities tabulated lend some credence to the large-sample comparison.

One consideration in the use of standard normal is whether the actual significance levels obtained with standard normal are the same as the significance levels for the exact distributions. Comparison of the standard normal with the exact distributions in Tables 4.1 and 4.2 indicates that there is good agreement of all the time for the .05 tabular value even the values of m and n are both small. Cochran (1952) recommended that the standard normal

is acceptable if the normal probability falls within the rang .04-.06 for the .05 tabular value and within .007-.015 for the .01 tabular value. Tables 4.1 and 4.2 show clearly that the standard normal works rather well when the values of m and n are both greater and equal than 10 for the three different data distributions.

Table 4.1: Normal Approximation Probability (U: Uniform, N: Normal, E: Exponential)

		m = 2 n = 5	m = 2 n = 6	m = 2 n = 15	m = 2 n = 20	m = 2 n = 25	m = 3 n = 3	m = 3 n = 4	m = 3 n = 25	m = 4 n = 4	m = 4 n = 5	m = 4 n = 25
.01	U			.0393	.0385	.0382			.0230		.0244	.0177
	N			.0363	.0354	.0350			.0226		.0249	.0180
	E			.0299	.0288	.0284			.0213		.0239	.0188
.05	U	.0515	.0529	.0527	.0524	.0524	.0463	.0481	.0479	.0478	.0480	.0481
	N	.0512	.0518	.0531	.0523	.0535	.0478	.0485	.0495	.0475	.0476	.0489
	E	.0460	.0464	.0510	.0522	.0523	.0474	.0475	.0523	.0480	.0475	.0522
		m = 5 n = 5	m = 5 n = 25	m = 6 n = 6	m = 6 n = 15	m = 6 n = 20	m = 6 n = 25	m = 7 n = 7	m = 7 n = 10	m = 7 n = 15	m = 7 n = 20	m = 7 n = 25
.01	U	.0217	.0157	.0188	.0157	.0150	.0148	.0173	.0160	.0150	.0145	.0139
	N	.0221	.0159	.0185	.0154	.0151	.0148	.0167	.0154	.0146	.0143	.0140
	E	.0217	.0170	.0182	.0161	.0158	.0159	.0166	.0156	.0151	.0151	.0151
.05	U	.0484	.0494	.0492	.0497	.0497	.0499	.0491	.0492	.0492	.0495	.0497
	N	.0478	.0490	.0483	.0491	.0491	.0490	.0487	.0489	.0489	.0492	.0489
	E	.0472	.0515	.0482	.0510	.0509	.0514	.0482	.0498	.0505	.0511	.0514
		m = 8 n = 8	m = 8 n = 9	m = 8 n = 10	m = 8 n = 15	m = 8 n = 20	m = 8 n = 25	m = 9 n = 9	m = 9 n = 10	m = 9 n = 15	m = 9 n = 20	m = 9 n = 25
.01	U	.0162	.0157	.0154	.0144	.0140	.0137	.0155	.0151	.0143	.0137	.0133
	N	.0157	.0152	.0149	.0141	.0137	.0136	.0148	.0146	.0137	.0134	.0131
	E	.0154	.0152	.0150	.0146	.0143	.0145	.0146	.0143	.0140	.0141	.0141
.05	U	.0494	.0498	.0492	.0494	.0496	.0495	.0494	.0494	.0499	.0495	.0497
	N	.0487	.0487	.0488	.0493	.0492	.0492	.0489	.0494	.0495	.0493	.0493
	E	.0489	.0492	.0497	.0503	.0506	.0506	.0491	.0495	.0498	.0510	.0508
		m = 10 n = 10	m = 10 n = 15	m = 10 n = 20	m = 10 n = 25	m = 15 n = 15	m = 15 n = 20	m = 15 n = 25	m = 20 n = 20	m = 20 n = 25	m = 25 n = 25	
.01	U	.0147	.0140	.0133	.0131	.0131	.0126	.0123	.0122	.0121	.0118	
	N	.0141	.0135	.0131	.0129	.0129	.0121	.0121	.0119	.0117	.0114	
	E	.0141	.0139	.0133	.0136	.0124	.0124	.0124	.0118	.0117	.0113	
.05	U	.0496	.0498	.0499	.0498	.0504	.0498	.0498	.0500	.0501	.0502	
	N	.0487	.0490	.0493	.0491	.0496	.0492	.0494	.0494	.0496	.0492	
	E	.0493	.0500	.0500	.0510	.0491	.0497	.0499	.0494	.0498	.0494	

Table 4.2: Normal Approximation Probability(C: Cauchy, D: Double Exponential)

		m = 3	m = 4	m = 5	m = 6	m = 7	m = 8	m = 9	m = 10	m = 15
		n = 3	n = 4	n = 5	n = 6	n = 7	n = 8	n = 9	n = 10	n = 15
.01	C			.0220	.0182	.0159	.0148	.0140	.0132	.0119
	D			.0217	.0184	.0166	.0153	.0144	.0139	.0126
.05	C		.0474	.0475	.0470	.0478	.0482	.0489	.0490	.0491
	D		.0472	.0469	.0479	.0481	.0486	.0490	.0490	.0495

4.2 Asymptotically distribution-free of V

Some U-statistics are distribution-free; others are not. A well-known distribution-free test is the rank sum W -test of Wilcoxon (1945) which introduced a ranking method for determining the significance of the difference between two treatments. Since in such methods ranks 1, 2, 3 ... n are substituted for the numerical data, there is certain sacrifice of information, so that Wilcoxon regarded his method as giving a "rapid approximation idea of the significance of the differences". Another distribution-free test is the rank sum test of Siegel & Tukey (1960). This test statistic has the same probability distribution as W . In some instance, statistic is not distribution-free but it is asymptotically distribution-free.

Theorem 4.1. Statistic V is asymptotically distribution-free.

Proof: From chapter 3, we have obtained different values of the mean and the variance of V for various F and thus the distribution of V will depend on F. Statistic V is not distribution-free.

In order to proof statistic V is asymptotically distribution-free,

let

$$W = \frac{\sum_{\substack{i=1 \\ i < j}}^m \sum_{j=1}^m \sum_{\substack{k=1 \\ k < l}}^n \sum_{l=1}^n I_{\left(\frac{Y_k + Y_l}{2} - \frac{X_i + X_j}{2}\right)}}{2 \binom{m}{2} \binom{n}{2}}$$

then

$$\begin{aligned}
\lim_{\substack{m \rightarrow \infty \\ n \rightarrow \infty}} \left(\frac{V}{2 \binom{m}{2} \binom{n}{2}} \right) &= \lim_{\substack{m \rightarrow \infty \\ n \rightarrow \infty}} \left[\sum_{i=1}^m \sum_{k=1}^n I_{(Y_k - X_i)} + \sum_{i=1}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{Y_k + Y_l}{2} - X_i)} \right. \\
&+ \sum_{k=1}^n \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m I_{(Y_k - \frac{X_i + X_j}{2})} + 2 \sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{Y_k + Y_l}{2} - \frac{X_i + X_j}{2})} \Big] / [2 \binom{m}{2} \binom{n}{2}] \\
&= \frac{\sum_{i=1}^m \sum_{\substack{j=1 \\ i < j}}^m \sum_{k=1}^n \sum_{\substack{l=1 \\ k < l}}^n I_{(\frac{Y_k + Y_l}{2} - \frac{X_i + X_j}{2})}}{2 \binom{m}{2} \binom{n}{2}} = W
\end{aligned}$$

So, V and W are asymptotically equivalent test statistic, and Hollander (1967) has shown that W is asymptotically distribution-free. Therefore, statistic V is asymptotically distribution-free.

4.3 Simulated power comparison

The powers of the new test (LY-test), Mann-Whitney-Wilcoxon (U-test) are investigated in this section. Our simulation study used for x_i and y_j are the uniform (U), normal (N) and exponential (E) distributions. At each occasion, 10,000 observations of three different distribution random vectors are generated. Based on the simulation we analyzed the behavior of the measures V_i for sample sizes under the null hypothesis $F_x = F_y$ and under the alternatives $F_y(\cdot) = F_x(\cdot - \theta)$ with varying θ .

Let $m = n = 5, 10, 15, 20, 25$ be the sizes of the two samples. The uniform distributions under consideration are over the range (0, 1) for F_x and the designated location alternatives F_y correspond to values of range (0, $2\theta + 1$); the normal distributions have the mean zero and $\sigma^2 = 5$ for F_x and the same variance (five) with different mean θ for alternatives F_y ; the exponential distributions have scale parameter 1 for F_x and various scale parameters $\theta + 1$ for F_y . The results are provided in the Table 4.3. It is found that power of the new

procedure is slightly larger than that of U procedure. As a consequence, the new procedure is consistently more powerful than Mann-Whitney-Wilcoxon test in detecting the difference mean between two-samples when those comparisons are of interest, the new procedure is recommended. On the other hand, our new procedure as a test statistic is more powerful than U test.

Table 4.3: Power Test for Different Mean ($\alpha=0.05$)

sample size		mean difference									
		0		1.5		3		4.5		6	
		LY-test	U-test	LY-test	U-test	LY-test	U-test	LY-test	U-test	LY-test	U-test
m=5,n=5	U	.0496	.0485	.8036	.6995	.9281	.8457	.9639	.8993	.9773	.9279
	N	.0483	.0464	.2424	.2360	.5939	.5728	.8865	.8695	.9852	.9784
	E	.0500	.0498	.3180	.2915	.5476	.5052	.6832	.6239	.7646	.7030
m=10,n=10		0		1		2		3		4	
	U	.0506	.0438	.9457	.8611	.9946	.9682	.9990	.9912	.9996	.9947
	N	.0453	.0412	.2446	.2190	.6064	.5637	.8871	.8572	.9869	.9783
m=15,n=15		0		1		1.5		2.5		3	
	U	.0498	.0487	.9928	.9636	.9985	.9902	1	.9993	1	.9993
	N	.0485	.0476	.3369	.3056	.5621	.5230	.9171	.8965	.9773	.9679
m=20,n=20		0		0.5		1.5		2		3	
	U	.0489	.0451	.9459	.8825	.9999	.9989	1	.9998	1	.9999
	N	.0495	.0477	.1673	.1632	.6519	.6311	.8665	.8498	.9931	.9914
m=25,n=25		0		0.5		1.5		2		2.5	
	U	.0507	.0486	.9766	.9354	1	.9997	1	1	1	1
	N	.0469	.0440	.1911	.1782	.7569	.7302	.9284	.9136	.9908	.9864
	E	.0506	.0486	.3757	.3256	.9046	.8508	.9724	.9413	.9971	.9888

Chapter 5

Application for Dose-Response Study

5.1 Introduction

In drug development studies, several increasing dose levels of a substance are usually compared with the zero-dose control to investigate the effect of the substance. For this propose, a dose-response experiment is often conducted in a one-way layout in which the doses of the substance under consideration are administered to separate groups of subjects. The primary interest is identifying the lowest dose level producing a desirable effect over that of the zero-dose control, which is commonly referred as the minimum effective dose (MED; Ruberg, 1989). This paper mainly discusses a test-based approach to identify the MED in drug development studies. Note that, in these drug studies, increasing dose levels are frequently expected to produce stronger or at least equal treatment effects. However, it also happens often that, due to the toxic effects at high doses, an ordering in the treatment effects is anticipated that it is monotonically increasing up to a point, follows by a monotonic decrease. Since this corresponds to an up-down ordering of the treatment effects, they are said to follow an umbrella patten.

The problem of identifying the MED has been investigated by several authors for normally distributed responses with a common variance. For example, Williams (1971) considered a closed testing procedure based on the isotonic regression of the sample means for a monotonic dose-response relationship. Ruberg (1989) suggested tests based on different contrasts of sample means to identify the MED. Tamhane, Hochberg, and Dunnett (1996) further proposed contrast-based closed testing procedures for identifying the MED. In dose-response studies, however, it occurs frequently that the normal assumption is not tenable or the observations are too few to rely on the central limit theorem for normality. In this cases, nonparametric procedures providing practical alternatives for identifying the MED are then needed. Williams (1971) considered a nonparametric test for contrasting increasing dose levels and Williams (1986) further suggested a modification of Shirley's (1977) test. Chen (1999) utilized the Mann-Whitney statistic for identifying the MED of modification of Chen and Wolfe (1993) test. We consider in this paper employing the step-down closed testing (SDT) scheme suggested by Tamhane et al. (1996), but we utilize the V statistic for identifying the MED.

5.2 The proposed testing procedure

For the i th sample ($i = 0, 1, \dots, k; k \geq 4$), let Y_{i1}, \dots, Y_{in_i} be independent and identically distributed random variables, each with a continuous distribution function F_i . Suppose that the zero population ($i = 0$) is the zero-dose control and the other k populations correspond to increasing dose treatments. Furthermore, assume that the $k + 1$ samples are independent of each other. We consider estimation of the MED, which is the smallest i so that the response in the i th population is stochastically larger than that in the control, in other words, the i th treatment is the smallest number of treatments that their mean are larger than that in

the control, when the dose-response relationship is either monotonic (ordered; $F_0 \geq F_1 \geq \dots \geq F_k$) or nonmonotonic with a down turn (umbrella patterned; $F_0 \geq F_1 \geq \dots \geq F_l \leq \dots \leq F_k$) for some $l, 1 < l < k$.

The two-sample V statistic comparing the i th dose group with the combined group of all the lower dose levels (including the control) is

$$V_i = \sum_{j=0}^{i-1} \sum_{u=1}^{n_i} \sum_{v=1}^{n_j} I_{(Y_{iu} - Y_{jv})} \text{rank}(|Y_{iu} - Y_{jv}|), \quad i = 1, 2, \dots, k,$$

Let

$$V_i^* = [V_i - \mu(V_i)] / \sqrt{\sigma^2(V_i)}, \quad i = 1, 2, \dots, k,$$

where

$$\mu(V_i) = \frac{N_{i-1}n_i(N_{i-1}n_i + 1)}{4}$$

$$\begin{aligned} \sigma^2(V_i) = & .0544388N_{i-1}n_i + .0008383N_{i-1}^4n_i - .0152703N_{i-1}^3n_i + .0738473N_{i-1}^3n_i^2 \\ & + .003717n_iN_{i-1}^2 + .0310913N_{i-1}^2n_i^2 + .0738473N_{i-1}^2n_i^3 + .0008383N_{i-1}n_i^4 \\ & - .0152703N_{i-1}n_i^3 + .003717N_{i-1}n_i^2 - .0120033N_{i-1}^4n_i^2 + .0402N_{i-1}^4n_i^3 \\ & - .018187N_{i-1}^3n_i^3 + .0402N_{i-1}^3n_i^4 - .0120033N_{i-1}^2n_i^4 \end{aligned}$$

where $\mu(V_i)$ and $\sigma^2(V_i)$ with $N_i = \sum_{j=0}^i n_j$, are mean and variance of V_i under H_{0i} , respectively. Then the test based on V_i^* is appropriate for testing against the alternative hypothesis $H_{1i} : (F_0 = F_1 = \dots = F_{i-1} > F_i)$, $i=1,2,\dots,k$. When ties occur in the N_i observations, the values of $\mu(V_i)$ are used and a modification of variance of V_i is obtained by reducing the expression

$$\begin{aligned} & \frac{N_{i-1}n_i \sum t(t-1)(13t-20)}{24N_i(N_i-1)} + [1.5288N_{i-1} \binom{n_i}{3} + 1.5288n_i \binom{N_{i-1}}{3} \\ & + 6.5778 \binom{N_{i-1}}{3} \binom{n_i}{3}] \frac{\sum t(t-1)(t-2)(t-3)}{N_i(N_i-1)(N_i-2)(N_i-3)} \end{aligned}$$

Table 5.1: Reaction Times in Seconds of Mice to Stimuli to their Tails

	Reaction time									
Group 0	2.4	3.0	3.0	2.2	2.2	2.2	2.2	2.8	2.0	3.0
Group 1	2.8	2.2	3.8	9.4	8.4	3.0	3.2	4.4	3.2	7.4
Group 2	9.8	3.2	5.8	7.8	2.6	2.2	6.2	9.4	7.8	3.4
Group 3	7.0	9.8	9.4	8.8	8.8	3.4	9.0	8.4	2.4	7.8

As noted in Tamhane et al. (1996), the family of null hypotheses $H = H_{0i}$, where $H_{0i} : (F_0 = F_1 = \dots = F_{i-1} = F_i)$ for $i = 1, 2, \dots, k$, is closed under intersection in the sense that $H_{0i} \in H$ and $H_{0j} \in H$ imply $H_{0i} \cap H_{0j} \in H$. Hence, a level- α closed procedure that includes separate level- α tests of individual H_{0i} applied in a step-down manner can be employed in finding the MED. To estimate the MED, we first let $k_1 = k$ and find $V_{(k_1)}^*$, where $V_{(k_1)}^*$ is the maximum of $V_1^*, V_2^*, \dots, V_{k_1}^*$. Since the results in Terpstra (1952) imply that the statistics $V_1^*, V_2^*, \dots, V_{k_1}^*$ are stochastically independent and we have showed that they are standard normal approach under the null hypothesis H_{0k_1} . Therefore, we observe that $P\{V_{k_1}^* \leq z(\alpha) | H_{0k_1}\} = [P\{V_1^* \leq z(\alpha) | H_{0k_1}\}]^{k_1} \approx (1 - \alpha)^{k_1}$, where $z(\alpha)$ is the upper α th percentile of the standard normal distribution. let $\alpha(k_1) = 1 - (1 - \alpha)^{1/k_1}$. Define $d(k_1)$ to be antirank of $V_{(k_1)}^*$ that $V_{(k_1)}^* = V_{d(k_1)}^*$. Then, if $V_{(k_1)}^* \geq z(\alpha(k_1))$, reject $H_{0j}, j = d(k_1), \dots, k_1$, and go to the second step with $k_2 = d(k_1) - 1$; otherwise, stop testing and accept all hypothesis.

5.3 An example

Consider the data set in Table 5.1 analyzed in Shirley (1977) which contains three dose levels and a zero-dose control, there are 10 observations in each group.

- (1) **LY test:** the V statistics, their corresponding means and standard deviation are

obtained in the following: $V_1 = 4829, V_2 = 16274, V_3 = 39633, \mu(V_1) = 2525, \mu(V_2) = 10050, \mu(V_3) = 22575, \sigma(V_1) = 881.32, \sigma(V_2) = 3058, \sigma(V_3) = 6490, V_1^* = 2.614, V_2^* = 2.035, V_3^* = 2.628$. Note that the largest statistic among the three V_i^* 's is V_3^* , so $d(3)=3$. Since at the level $\alpha = 0.05, V_3^* = 2.628 > z(1 - (.95)^{1/3}) = 2.1212$, we go to the second step with $k_2 = 2$. We observe that $d(1)=1$ and $V_1^* = 2.614 > z(1 - (.95)) = 1.6448$. Therefore, we estimate that, at the 5% significant level, the MED is the first dose level.

(2) **Chen test:** the statistic is:

$$T_i = \sum_{j=0}^{i-1} \sum_{u=1}^{n_i} \sum_{v=0}^{n_j} I_{(Y_{iu} - Y_{jv})}, \quad i = 1, 2, \dots, k,$$

let

$$T_i^* = \frac{T_i - \mu(T_i)}{\sqrt{\sigma^2(T_i)}}, \quad i = 1, 2, \dots, k,$$

where $\mu(T_i) = n_i N_{i-1} / 2$ and $\sigma^2(T_i) = n_i N_{i-1} (N_i + 1) / 12$, with $N_i = \sum_{k=1}^i n_j$, are the mean and variance of T_i under H_{0i} . If there are ties among the N_i observations, a modification of T_i^* is obtained by replacing the $N_i + 1$ in $\sigma^2(T_i)$ with

$$N_i + 1 - \sum t(t^2 - 1) / [N_i(N_i - 1)],$$

their corresponding means and variances are obtained in the following: $T_1 = 84, T_2 = 145, T_3 = 235, \mu(T_1) = 50, \mu(T_2) = 100, \mu(T_3) = 150, \sigma^2(T_1) = 170.79, \sigma^2(T_2) = 510.69, \sigma^2(T_3) = 1018.94, T_1^* = 2.60, T_2^* = 1.99, T_3^* = 2.66$. We estimate that, at the 5% significant level, the MED is the first dose level.

(3) **Williams test:** the statistic is:

$$t_i = \left[\max_{1 \leq \mu \leq i} \left(\sum_{j=\mu}^i n_j \bar{R}_{ij} / \sum_{j=\mu}^i n_j \right) - \bar{R}_{i0} \right] / \left[\sigma_i^2 \left(\frac{1}{n_i} + \frac{1}{n_0} \right) \right]^{1/2}, \quad i = 1, 2, \dots, k,$$

where $\sigma_i^2 = \frac{N_i(N_i+1)}{12}$, $i = 1, 2, \dots, k$. When ties occur in the rankings, average ranks are used and corrections for ties in the rankings are made by reducing the expression $N_i(N_i + 1) / 12$ by $\sum t(t^2 - 1) / 12(N_i - 1)$.

The mean of the ranks for the three groups are:

$$\begin{aligned}\bar{R}_{30} &= 8.25 & \bar{R}_{31} &= 20.70 & \bar{R}_{32} &= 23.65 & \bar{R}_{33} &= 29.40 \\ \bar{R}_{20} &= 7.8 & \bar{R}_{21} &= 18.3 & \bar{R}_{22} &= 20.4 & & \\ \bar{R}_{10} &= 6.7 & \bar{R}_{11} &= 14.3 & & & & \end{aligned}$$

The variance σ_i^2 for the three treatment groups are: $\sigma_1^2 = 34.16$, $\sigma_2^2 = 76.60$, $\sigma_3^2 = 135.86$.

The statistics t_i for the three treatment groups are $t_1 = 2.91$, $t_2 = 3.22$, $t_3 = 4.06$. The

MED is the first dose level.

Chapter 6

Summary and Further research

The application of this new test we have described does not require any assumption that the distribution(s) from which the samples were drawn is normal or any other specific shape. The test has the additional advantage of being directly applicable to non-numerical ordinal data. Observations which are ordered but not measured on any scale arise frequently, especially in research in the behavioral sciences.

Exact inference has important place in statistical inference of discrete data, in particular for sparse contingency table problems for which large-sample chi-squared statistics are often unreliable. However, approximate results are sometimes more useful than exact results, because of the inherent conservativeness of exact methods. Mann and Whitney (1947) proved that the limit distribution of U is normal if m and n go to infinity in any arbitrary manner, and they suggest that this approximation may be used to obtain critical values of U when both m and n is greater than 20. Statistic V is normal approximation, and is asymptotically normal for m and n larger than 10.

A well-known distribution-free test is the rank sum W -test of Wilcoxon (1945) which introduced a ranking method for determining the significance of the difference between

two treatments. Unfortunately not all statistics are distribution-free. In this thesis, we show that V statistic is asymptotically distribution-free.

We are investigated the relative power of the new test with respect to the U test. We applied the new procedure to independent two-sample problems, and found that the test is efficient. The power comparison show that the new procedure is consistently more powerful than Mann-Whitney-Wilcoxon's in detecting mean difference between two-samples. The new procedure is efficient and powerful nonparametric test, it may be a reasonable method and test for k -sample problems.

Appendix A

Simulation of the Probabilities of Individual V Statistic

```
# simulation of the probabilities of individual V statistic

k<-0;m<-value;n<-value;N<-m*n*(m*n+1)/2+1
dy<-matrix(0,m,n);udy<-matrix(0,m,n);edy<-matrix(0,m,n)
c215<-rep(0,N);uc215<-rep(0,N); ec215<-rep(0,N)
while(k<1000000) {
  k<-k+1; tr<-0;utr<-0;etr<-0
  y0<-rnorm(m,0,1);uy0<-runif(m,0,1);ey0<-rexp(m,1)
  y1<-rnorm(n,0,1);uy1<-runif(n,0,1);ey1<-rexp(n,1)
  for (i in 1:m) {
    for (j in 1:n) {dy[i,k]<-y1[j]-y0[i]
                    udy[i,k]<-uy1[j]-uy0[i]
                    edy[i,k]<-ey1[j]-ey0[i]}
  }
}
```

```

        }
    }
    dr<-matrix(rank(abs(dy)),m,n)
    udr<-matrix(rank(abs(udy)),m,n)
    edr<-matrix(rank(abs(edy)),m,n)
    for(i in 1:m) { for( j in 1:n)
        {if(dy[i,k]>0) tr<-tr+dr[i,k]
        if(udy[i,k]>0) utr<-utr+udr[i,k]
        if(edy[i,k]>0) etr<-etr+edr[i,k]
        }
    }
    c215[tr+1]<-c215[tr+1]+1
    uc215[utr+1]<-uc215[utr+1]+1
    ec215[etr+1]<-ec215[etr+1]+1
}
sc<-0;suc<-0;sec<-0
for(i in 1:N) { suc<-suc+uc215[i]
    sc<-sc+c215[i]
    sec<-sec+ec215[i]
    cat(i-1," ", suc/1000000, " ",
        sc/1000000, " ",
        sec/1000000, "\n",file="filename",append=TRUE)
}

```

Appendix B

Power Test for LY Test and Mann-Whitney Test

```
# Power test for LY test and Mann-Whitney test

mus<-0
sd<-sqrt(variance)
nsim<-10000; k<-0
n<-# of variables
cs1<-0;cs0<-0
ccs1<-0;ccs0<-0
while(k<nsim) {k<-k+1
  ys0<-rnorm(n,mus,sd)
  ys1<-rnorm(n,mus+different mean,sd)
  s1<-rep(0,n^2)
  g<-0
```

```

for( i in 1:n) { for(j in 1: n)
                    { g<-g+1; s1[g]<-ys1[i]-ys0[j]
                      }
                    }
rk<-rank(abs(s1)); sumr<-0
for(i in 1:length(s1)) { if (s1[i]<0)
                        sumr<-sumr+rk[i]
                        }
if (sumr < data)  cs1<-cs1+1
                else cs0<-cs0+1
s2<-rank(c(ys0,ys1))
if ((sum(s2[1:n])-n*(n+1)/2) < 227)
    ccs1<-ccs1+1    else ccs0<-ccs0+1
}
cat("sample size m=n=  ",n,
    "two sample difference mean: value"," \n",
    " power for LY test is  ",cs1, " \n",
    " power for Mann-Whitney test is  ", ccs1, " \n")

```

Bibliography

- Abrahamson, I. G. (1964). Orthant Probabilities for the Quadrivariate Normal Distribution. *Annals of Mathematical Statistics*. **35** 1685-1703.
- Bhoj, D. S. (1997). New Parametric Ranked Set Sampling. *Journal of applied statistical Science*. **6** 275-289.
- Bohn, L. L. and Wolfe, D. A. (1992). Nonparametric Two-Sample Procedures for Ranked-Set Samples Data. *Journal of the American Statistical Association*. **87** 552-561.
- Bohn, L. L. and Wolfe, D. A. (1994). The Effect of Imperfect Judgment on Properties of Procedures Based on the Ranked-Set Samples Analog of the Mann-Whitney-Wilcoxon Statistic. *Journal of the American Statistical Association*. **89** 168-176.
- Chen, Y. I. (1999). Nonparametric Identification of the Minimum Effective Dose. *Biometrics*. **55** 1236-1240.
- Chen, Y. I. and Wolfe, D. A. (1993). Nonparametric Procedures for Comparing Umbrella Pattern Treatment Effects with a Control in a One-Way Layout. *Biometrics*. **49** 455-465.
- Childs, D. R. (1967). Reduction of the Multivariate Normal Integral to Characteristic Form. *Biometrika* . **54** 293-300.

- Cochran, W. G. (1952). The χ^2 Test of Goodness of Fit. *Annals of Mathematical Statistics*. **23** 315-345.
- Dell, T. R. and Clutter, J. L. (1972). Ranked-Set Sampling Theory with Order Statistics Background. *Biometrics*. **28** 545-555.
- Halls, L. K. and Dell, T. R. (1966). Trial of Ranked-Set Sampling for Forage Yields. *Forest Science*. **12** 22-26.
- Hollander, M. (1967). Asymptotic Efficiency of Two Nonparametric Competitors of Wilcoxon's Two Sample Test. *Journal of the American Statistical Association*. **62** 939-949.
- Kaur, A., Patil, G. P. and Taillie, C. (1997). Unequal Allocation Models for Ranked Set Sampling with Skew Distributions. *Biometrics*. **53** 123-130.
- Koti, M. K. and Babu, G. J. (1996). Sign Test for Ranked-Set Sampling. *Communications in Statistics - Theory and Methods*. **25** 1617-1630.
- McFadden, J. A. (1960). Two Expansions for the Quadrivariate Normal Integral. *Biometrika*. **47** 325-333.
- McIntyre, G. A. (1952). A Method of Unbiased Selective Sampling. *Australian Journal of Agricultural Research*. **3** 385-390.
- Owen, D. B. (1985). Orthant Probabilities. *Encyclopedia of statistical sciences*. **6** 521-523.
- Öztürk, Ö. (1999a). One and Two Sample Sign Tests for Ranked Set Sampling with Selective Designs. *Communications in Statistics - Theory and Methods*. **28** 1231-1245.

- Öztürk, Ö. (1999b). Two Sample Inference Based On One Sample Ranked Set Sample Sign Statistic. *Journal of Nonparametric Statistics*. **10** 197-212.
- Öztürk, Ö. and Wolfe, D. A. (1998). Optimal Ranked Set Sampling Protocol for the Signed Rank Test. Technical Report No. 630, Department of Statistics, The Ohio State University, Columbus, OH.
- Öztürk, Ö. and Wolfe, D. A. (2000a). Optimal Allocation Procedure in Ranked Set Sampling for Unimodal and Multi-Modal Distributions. *Environmental and Ecological Statistics*. **7** 343-356.
- Öztürk, Ö. and Wolfe, D. A. (2000b). Alternative Ranked Set Sampling Protocols for the Sign Test. *Statistics and Probability Letters*. **47** 15-23.
- Randles, R. H. and Wolfe, D. A. (1979). Introduction to The theory of Nonparametric Statistics. New York: Wiley.
- Ruberg, S. J. (1989). Contrast for Identifying the Minimum Effective Dose. *Journal of the American Statistical Associations*. **84** 816-822.
- Shirley, E. (1977). A Nonparametric Equivalent of William's Test for Contrasting Increasing Dose Levels of a treatment. *Biometrics*. **33** 386-389.
- Siegel, S. and Tukey, J. W. (1960). A Nonparametric Sum of Ranks Procedure for Relative Spread in Unpaired Samples. *Journal of the American Statistical Associations*. **55** 429-445.
- Stokes, S. L. (1977). Ranked Set Sampling with Concomitant Variables. *Communications in Statistics - Theory and Methods* **6** 1207-1212.

- Stokes, S. L. and Sager, T. W. (1988), Characterization of a Ranked-Set Sample with Application to Estimating Distribution Functions. *Journal of the American Statistical Association*. **83** 374-381.
- Tamhane, A. C., Hochberg, Y., and Dunnett, C. (1996). Multiple Test Procedures for Dose Finding. *Biometrics*. **52** 21-37.
- Takahasi, K. and Wakimoto, K. (1968). On Unbiased Estimates of the Population Mean Based On the Sample Stratified by Means of Ordering. *Annals of the Institute of Statistical mathematics*. **20** 1-31.
- Terpstra, T. J. (1952). The Asymptotic Normality and Consistency of Kendall's Test Against Trend When Ties Are Present in One Ranking. *Indagationes Mathematica*. **14** 327-333.
- Williams, D. A. (1971). A Test for Differences Between Treatment Means When Several Dose Levels Are Compared with a Zero Dose Control. *Biometrics*. **27** 103-117.
- Williams, D. A. (1986). A Note on Shirley's Nonparametric Test for Comparing Several Dose Levels with a Zero-Dose Control. *Biometrics*. **42**. 183-186.

