ALMOST SURE CONFIDENCE INTERVALS FOR THE 
CORRELATION COEFFICIENT

by

MARK M. FRIDLINE

Submitted in partial fulfillment of the requirements
For the degree of Doctor of Philosophy

Dissertation Advisor: Dr. Manfred Denker

Department of Statistics
CASE WESTERN RESERVE UNIVERSITY

January 2010
We hereby approve the dissertation of

Mark M. Fridline

candidate for the Doctor of Philosophy degree *

Committee Chair:  
Dr. Manfred Denker  
Dissertation Advisor  
Adjunct Professor  
Department of Statistics

Committee:  
Dr. Wojbor Woyczynski  
Professor  
Department of Statistics

Committee:  
Dr. Patricia Williamson  
Instructor  
Department of Statistics

Committee:  
Dr. David Gurarie  
Professor  
Department of Mathematics

January 2010

*We also certify that written approval has been obtained for any proprietary material contained therein.
Table of Contents

Table of Contents ................................................................. iii
List of Tables ................................................................. v
List of Figures ................................................................. vi
Acknowledgment ............................................................ ix
Abstract ............................................................................. x

1 Introduction ................................................................. 1

2 The Almost Sure Version of Cramér’s Theorem ............................. 12
  2.1 Introduction to ASCLT .................................................. 12
  2.2 Replace the Log Averaging Term in ASCLT ....................... 14
  2.3 Central Limit Theorem .................................................. 16
  2.4 New Notation and Assumptions of the ASCLT ................... 16
  2.5 Cramér’s Theorem ....................................................... 20
  2.6 Auxiliary Lemmas ........................................................ 21
  2.7 The Almost Sure Version of Cramér’s Theorem ............... 28

3 ASCLT for the Correlation Coefficient ................................... 33
  3.1 Introduction to Population Correlation Coefficient .............. 33
  3.2 The Sample Correlation Coefficient ................................ 35
  3.3 The Asymptotic Distribution of Sample Correlation Coefficient .................................................. 36
  3.4 Connecting the ASCLT and the Correlation Coefficient ....... 43
  3.5 Law of Iterated Logarithm for the Correlation Coefficient .. 47

4 ASCLT-based Confidence Interval for the Correlation Coefficient ................................... 53
  4.1 Defined Cumulative Distribution Functions ....................... 53
  4.2 Defined Inverse Cumulative Distribution Functions .......... 54
  4.3 Empirical α-Quantiles ................................................... 56
  4.4 Confidence Intervals for ρ ............................................. 57
  4.5 Confidence Intervals for ρ (Variance Stabilizing Technique) .. 58

5 Numerical Applications .................................................... 61
  5.1 Introduction ................................................................... 61
  5.2 Bivariate Normal Random Deviate Generation ................. 65
  5.3 Bivariate Exponential Deviate Generation ....................... 67
  5.4 ASCLT for the Correlation Coefficient Simulations .......... 69
  5.5 ASCLT for the Correlation Coefficient Simulations – Variance Stabilizing Technique .......................... 81
List of Tables

5.1 Simulation Results for the Bivariate Normal Distribution ............. 115
5.2 Simulation Results for the Bivariate Exponential Distribution ........ 116
5.3 Simulation Results for the Bivariate Poisson Distribution .............. 117
### List of Figures

5.1 Estimated distribution function $H_N(t)$ for simulated samples from bivariate standard normal distributions ............................................................ 72

5.2 Estimated distribution function $\tilde{H}_N(t)$ with $C_N$ for simulated samples from bivariate standard normal distributions ............................................................ 73

5.3 Estimated distribution function $\tilde{H}_N(t)$ with $C_N$ from $n=m(n)$ to $N=max(n)$ 74

5.4 Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from dependent bivariate normal distributions when $\rho=0.3$ ............................................. 75

5.5 Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from dependent bivariate normal distributions when $\rho=0.5$ ............................................. 76

5.6 Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from dependent bivariate normal distributions when $\rho=0.7$ ............................................. 76

5.7 Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from bivariate exponential distributions with parameters $\lambda_1 = \frac{1}{6}, \lambda_2 = \frac{1}{4}, \lambda_{12} = 1 (\rho = 0.7)$. 77

5.8 Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from bivariate exponential distributions with parameters $\lambda_1 = \frac{1}{6}, \lambda_2 = \frac{1}{3}, \lambda_{12} = \frac{1}{2} (\rho = 0.5)$. 78

5.9 Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from bivariate exponential distributions with parameters $\lambda_1 = \frac{1}{2}, \lambda_2 = \frac{1}{2}, \lambda_{12} = \frac{1}{3} (\rho = 0.25)$. 78

5.10 Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from independent bivariate exponential distributions with parameters $\lambda_i = 3, i=1, 2$ ...... 79

5.11 Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from bivariate poisson distributions with parameters $\lambda_i = 1, i=1, 2$ ............................................. 80

5.12 Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from bivariate poisson distributions with parameters $\lambda_i = 3, i=1, 2$ ............................................. 80
5.13 Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from bivariate poisson distributions with parameters $\lambda_i=10$, $i=1, 2$ ................................. 81

5.14 Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from dependent bivariate normal distributions with $\rho=0.7$ ........................................... 83

5.15 Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from dependent bivariate normal distributions with $\rho=0.5$ ......................................... 84

5.16 Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from dependent bivariate normal distributions with $\rho=0.3$ ......................................... 84

5.17 Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from independent bivariate normal distributions ........................................................... 85

5.18 Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from bivariate exponential distributions with parameters $\lambda_1 = \frac{1}{6}, \lambda_2 = \frac{1}{4}, \lambda_{12} = 1$ ($\rho = 0.7$). 86

5.19 Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from bivariate exponential distributions with parameters $\lambda_1 = \frac{1}{6}, \lambda_2 = \frac{1}{3}, \lambda_{12} = \frac{1}{2}$ ($\rho = 0.5$). 86

5.20 Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from bivariate exponential distributions with parameters $\lambda_1 = \frac{1}{2}, \lambda_2 = \frac{1}{3}, \lambda_{12} = \frac{1}{3}$ ($\rho = 0.25$). 87

5.21 Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from bivariate exponential distributions with parameter $\lambda_i=3$, $i=1, 2$ ......................... 87

5.22 Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from bivariate poisson distributions with parameters $\lambda_i=1$, $i=1, 2$ ................................................. 88

5.23 Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from bivariate poisson distributions with parameters $\lambda_i=3$, $i=1, 2$ ................................................. 89

5.24 Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from bivariate poisson distributions with parameters $\lambda_i=10$, $i=1, 2$ ................................................. 89

5.25 Confidence interval results for $N=10,000$, CL=90%, and $\rho=0.7$ ...... 97

5.26 Confidence interval results for $N=1,000$, CL=90%, and $\rho=0.7$ ...... 98
5.27 Confidence interval results for N=1,000, CL=95%, and ρ=0.5 ........ 98
5.28 Confidence interval results for N=100, CL=90%, and ρ=0.7 ........ 99
5.29 Confidence interval results for N=100, CL=95%, and ρ=0.3 ......... 100
5.30 Confidence interval results for N=100, CL=99%, and ρ=0.0 ......... 100
5.31 Confidence interval results for N=10,000, CL=95%, and ρ=0.7 ...... 102
5.32 Confidence interval results for N=1,000, CL=95%, and ρ=0.7 ....... 103
5.33 Confidence interval results for N=100, CL=95%, and ρ=0.7 ........ 104
5.34 Confidence interval results for N=100, CL=99%, and ρ=0.0 ......... 104
5.35 Confidence interval results for N=10,000, CL=95%, and ρ=0.0 ...... 106
5.36 Confidence interval results for N=1,000, CL=95%, and ρ=0.0 ........ 107
5.37 Confidence interval results for N=100, CL=95%, and ρ=0.0 ......... 107
ACKNOWLEDGMENTS

I would like to express my sincere appreciation to the following people who helped and supported me through this entire dissertation and Ph.D. in Statistics process.

Dr. Patricia Williamson and Dr. David Gurarie, I am appreciative for all your help and effort while participating on my dissertation committee. I am grateful for your outstanding suggestions and advise while I completed my dissertation.

Dr. Wojbor Woyczynski, I extremely appreciative for giving me the encouragement and advice during my initial meeting with you over five years ago. Thank you for understanding me as a student, but also as a person. Your academic and research excellence has been inspiring, and I am blessed that I had the opportunity to work with you during my tenure at CWRU.

Dr. Manfred Denker, I would like to thank you for your incredible help and guidance as my dissertation advisor. Your style and approach while working with me, has re-evaluated how I will work with my students. I am appreciative for your participation in my progress as a student and researcher. I am most grateful to consider you my mentor.

To my wife Sara, who provided continued love, support, and prayers for me during my pursuit of this degree. This accomplishment was a complete team effort, and I would not have accomplished without you by my side. But mostly, I am a better person, husband, and father because of you. To my children Mikaela, Andrew, and Jonathan who gave me the inspiration to show that any dream is possible. To my parents, Jacob and Janice, and my mother-in-law and late father-in-law, Cherie and Rollin, for supporting and believing in me.

Most importantly, I would like to thank my Lord and Savior Jesus Christ for answering all my prayers while helping me obtain my dream. I can do everything through Him who gives me strength (Philippians 4:13). I give Him all the glory for this accomplishment.
Almost Sure Confidence Intervals for the Correlation Coefficient

Abstract

by

Mark M. Fridline

The dissertation develops a new estimation technique for the correlation coefficient. Although this method seems to be similar to a bootstrap method, it is nearly based on sequential sampling and sampling without replacement. This paper will emphasize the features, advantages, and applications of this new procedure. It also will explain the theoretical background and explain the necessary theory to apply this method successfully.
Chapter 1

Introduction

The subject of probability theory is the foundation upon which all statistics was created, providing a procedure for making decisions for populations when the probabilistic model is unknown. Through these models, statisticians are provided the tools to draw inferences about populations, while examining only a portion of the population. In this dissertation, we will consider the probability space $(\Omega, \mathcal{B}, \mathcal{P})$, where $\Omega$ is the sample space, $\mathcal{B}$ is the Sigma Field, and $\mathcal{P}$ is a family of probabilities on $\mathcal{B}$ where all inferential decisions will be made. For a more complete set-up and explanation of probability theory and how it relates to statistical hypothesis, please refer to Lehmann and Romano (2005).

In the situation of statistical inference, as in confidence intervals or hypothesis testing, these conclusions are made based on the quantiles of an unknown distribution. Therefore, a highly important topic in statistical inference is the estimation of quantiles from an unknown distribution. This dissertation will address methods of estimating quantiles using a theorem called the Almost Sure Central Limit Theorem.

The Almost Sure Central Limit Theorem (ASCLT) was first developed independently by the researchers Fisher (1987), Brosamler (1988), and Schatte (1988) under different degrees of generality. In the past decade, there have been several authors that have investigated the ASCLT and related logarithmic limit theorems for partial sums of
independent random variables. We refer to Atlagh (1993), Atlagh and Weber (1992), and
Berkes and Dehling (1993) for surveys of this field. The simplest form of the ASCLT
states that if $X_1, X_2, X_3, \ldots, X_n$ are independently and identically distributed random
variables with mean 0, variance 1, we have for any fixed $t \in \mathbb{R}$:

$$\frac{1}{\log N} \sum_{n=1}^{N} d_n I\left\{ \frac{S_n}{\sqrt{n}} \leq t\right\} \xrightarrow{a.s.} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} \, dx$$

(1.1)

or

$$\frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} I\left\{ \frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}} \leq t\right\} \xrightarrow{a.s.} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} \, dx$$

(1.2)

where $d_n = \frac{1}{\sqrt{n}}$, the partial sums $S_n = \sum_{n=1}^{N} X_n$, and $I_{\{A\}}$ is the indicator function of the
set $A$. By the ASCLT, the above averages converge almost surely to $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} \, dx$,
the standard normal cumulative distribution function, $N(0, 1)$. By this ASCLT, the
averages on the left hand side of (1.1) and (1.2) will converge almost surely to $N(0, 1)$ if
$d_n = \frac{1}{\sqrt{n}}$ (logarithmic averages), but will diverge to the ordinary average if $d_n = 1$
(Hörmann, 2007). It should be noted that the result in (1.1) was first presented by Lévy
(1937, p. 270) but he did not specify the conditions and gave no proof. For a complete
and detailed proof of the ASCLT, please refer to Brosamler (1988).

It should be noted that any results utilizing the ASCLT are asymptotic in nature and are
derived from logarithmic averages. Therefore, the rates of convergence for (1.1) and
(1.2) will be very slow. Due to this issue, general data analysis applications that use the
ASCLT are nearly impossible when dealing with small sample sizes. Later in this

2
dissertation we will address this small sample issue and propose an approximation with asymptotic results.

As mentioned earlier, a concern when working with the ASCLT is the rate of convergence to a normal distribution. Even for very large sample sizes, the rate of convergence is very slow. In chapter 2, this dissertation will address this ASCLT rate of convergence by using a proposal from Thangavelu (2005) that we replace the “Log N” in (1.1) or (1.2) with the averaging term \( \sum_{n=1}^{N} \frac{1}{n} \) to create the following,

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left( \frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}} \right) \leq t \quad (1.3)
\]

or

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \left( \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sqrt{n}} \right) \leq t \quad (1.4)
\]

It will be proven that (1.3) and (1.4) are cumulative distribution functions.

This dissertation will address this rate of convergence by applying Cramér’s Theorem (see Ferguson, 2002). Cramér’s Theorem for smooth functions of sample moments is one of the basic results in statistics used to diminish variation and allow the construction of smaller, more precise confidence intervals. This theorem states that if

\[
\sqrt{n}(X_n - \mu) \xrightarrow{L} N(0, \Sigma) \quad \text{where } \Sigma \text{ is a } d \times d \text{ covariance matrix, then}
\]

\[
\sqrt{n}(g(X_n) - g(\mu)) \xrightarrow{L} N(0, \hat{g}(\mu)\Sigma \hat{g}(\mu)^T). \quad \text{In this chapter, we will use a version of Cramér’s Theorem to extend the ASCLT to show that the result converges almost surely}
\]
towards a limiting distribution. Observe the following version of the ASCLT that includes the mean,

$$\frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \sqrt{n} (X_n - \mu) \leq t \right\} \overset{a.s.}{\longrightarrow} N(0, \Sigma) \quad (1.5)$$

The following almost sure version of Cramér’s Theorem, which is a new result, will be stated and proved in this chapter,

$$\frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \sqrt{n} (g(X_n) - g(\mu)) \leq t \right\} \overset{a.s.}{\longrightarrow} N(0, \tilde{g}(\mu) \Sigma \tilde{g}(\mu)^T) \quad (1.6)$$

Later in this dissertation we will be connecting the above ideas and connect them with the correlation coefficient.

The correlation coefficient has a history that started over a century ago. Rogers and Nicewander (1988) gave the following brief history of the development of the correlation coefficient. Sir Francis Galton first discussed the term bivariate correlation in 1885. In that year, he published an article that included the first bivariate scatterplot showing the idea of correlation. However, it was not until a decade later, in 1895, did Karl Pearson develop an index that we use to measure the association between two random variables. We will refer to this measure as the correlation coefficient; however it is frequently referred to as the Pearson's product-moment correlation coefficient, which is interesting since it was first introduced by Galton.

The basic idea of correlation was considered before 1885. In the article by Pearson in 1920, “Notes on the History of Correlation,” he credited Carl Friedrich Gauss for developing the normal surface of $n$ correlated variates in 1823. Gauss was not
particularly interested in the correlation as a statistic, but was interested in it as a parameter in his distributional equations. Pearson also credited Auguste Bravais in a previously published paper in 1895, with referring to one parameter of the bivariate normal distribution as “une correlation.” But like Gauss, Bravais did not see the importance of correlation as a stand alone statistic. Charles Darwin who was Dalton’s cousin, used the concept of correlation in 1868 by discussing that “all the parts of the organisation are to a certain extent connected or correlated together.”

Inferences based on the correlation coefficient of two random variables have been discussed by many authors. Fisher (1915) and Hotelling (1953) have derived various forms of the distribution function for the sample correlation coefficient. These distribution functions can be used to develop confidence intervals for correlation coefficient. Fisher derived the density for the correlation statistic \( r \) and presented it in two forms: one is expressed in terms of an infinite sum, and the other is expressed in terms of derivatives. Due to the complexity of calculations in obtaining the confidence intervals from each form, Fisher (1921) introduces the extremely useful Fisher’s z-transformation (see section 5.9) that simplified the confidence interval calculations. This method is currently still the most common method for the calculation of the confidence intervals for the correlation coefficient. Whereas Hotelling recommended the density of \( r \) to be expressed in a form that included the hypergeometric distribution. He proved that his derived distribution function converges rapidly even for small samples. Rubin (1966) also suggested a simpler method for approximating the confidence interval for the correlation coefficient, where his calculation presented a simple approximate
normalization for the correlation coefficient in samples that compared favorably to the classic large-sample normalizing transformation by Fisher. Sun and Wong (2007) suggested a likelihood-based higher-order asymptotic method to obtain confidence intervals for the correlation coefficient.

Also, in the literature, results can be found that finds the distribution of the correlation coefficient, and subsequently derives the confidence interval. Some of these techniques are very simple and require little effort in calculating the confidence intervals (see Samiuddin, 1970), and others are more complex (see Mudholkar and Chaubey, 1978). In an article written by Boomsma (1977), he compares confidence interval results from Fisher (1921), Samiuddin (1970), Mudholkar and Chaubey (1978), and Rubin (1966). In this article, he shows that it is possible to calculate improved approximations of $\rho$ using Ruben’s technique in principle, however in practice Fisher’s technique is most frequently used. Later in this dissertation we will be presenting two new results that are ASCLT-based confidence intervals for $\rho$.

In chapter 3 we will introduce the population correlation coefficient (see Ferguson, 2002) and prove that the asymptotic distribution of $\sqrt{n}(r_n - \rho)$ converges to the normal distribution with a mean of 0 and a variance of $\gamma^2$, 

$$
\gamma^2 = \frac{1}{4} \rho^2 \left[ b_{33} \frac{b_{33}}{\sigma_x^4} + 2 b_{34} \frac{b_{34}}{\sigma_x^2 \sigma_y^2} + b_{44} \frac{b_{44}}{\sigma_y^4} \right] - \rho \left[ b_{35} \frac{b_{35}}{\sigma_x^3 \sigma_y} + b_{45} \frac{b_{45}}{\sigma_x \sigma_y^3} \right] + \frac{b_{55}}{\sigma_x^2 \sigma_y^2}
$$

where
In this chapter we will also study the distribution function addressed in the ASCLT and connect it with the correlation coefficient to develop a new distribution function. When connecting the ASCLT and the correlation coefficient we get the following distribution function,

\[
\frac{1}{\log N} \sum_{n=1}^{N} \mathbb{1}_{\left\{ \sqrt{n} (r_n - \rho) \leq t \right\}} \xrightarrow{a.s.} N(0, \gamma^2) \tag{1.7}
\]

Our proposal in this thesis is to apply the cumulative distribution function in (1.7) and develop a confidence interval method for the population correlation coefficient. The asymptotic behavior of this distribution function (1.7) will be studied empirically and a new estimation method will be proposed and applied to the correlation coefficient.

In fact, our main goal towards this ASCLT-based theory of confidence interval estimation will be the estimation of the quantiles of (1.7).

In chapter 3 we will use the results of Cramér’s Theorem to prove that the asymptotic distribution of \( \sqrt{n} (g(r_n) - g(\rho)) \) converges to the normal distribution with a mean of 0 and a variance of \( \tau^2 \) where \( g(\rho) = \frac{1}{2} \log \frac{1 + \rho}{1 - \rho} \) and,
\[
\tau^2 = \frac{\rho^2}{4(1 - \rho^2)^2} \left[ \frac{b_{33}}{\sigma_x^4} + 2 \frac{b_{34}}{\sigma_x^2 \sigma_y^2} + \frac{b_{44}}{\sigma_y^4} \right] \\
- \frac{\rho}{(1 - \rho^2)^2} \left[ \frac{b_{35}}{\sigma_x^3 \sigma_y} + \frac{b_{45}}{\sigma_x \sigma_y^3} \right] \left[ \frac{b_{55}}{(1 - \rho^2)^2 \sigma_x^2 \sigma_y^2} \right].
\]

Chapter 3 will also study the distribution function addressed in the ASCLT, Cramér’s Theorem, and connect with the correlation coefficient to develop the following new distribution function,

\[
\lim_{n \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \sqrt{n} \left( \log \frac{1+r_n}{1-r_n} - \log \frac{1+\rho}{1-\rho} \right) \leq t \right\} \xrightarrow{a.s.} N(0, \tau^2). \tag{1.8}
\]

Another proposal in this thesis is to apply the cumulative distribution function in (1.8) and develop an additional confidence interval method for the population correlation coefficient. The asymptotic behavior of this distribution function (1.8) will be studied empirically and a new estimation method will be proposed and applied to the correlation coefficient. In fact, another goal towards this ASCLT-based theory of confidence interval estimation will be the estimation of the quantiles of (1.8).

To develop confidence intervals for the population correlation coefficient, we will need the law of iterated logarithm (see Dudley, 1989) for the correlation statistic. In chapter 3 section 5, we will state and prove the following almost sure result,

\[
\limsup_{n \to \infty} \frac{s_{xy} - n \sigma_{xy}}{\sqrt{2n \log \log n}} \leq \sqrt{b_{55}} + 2|\mu_x|\sigma_y + 2|\mu_y|\sigma_x.
\]

where \(\sigma_{xy}\) is the covariance of \(X\) and \(Y\).
In chapter 4, we will connect the ASCLT result and the correlation coefficient discussed in chapter 3, and develop a confidence interval method for the population correlation coefficient. Before we can state our confidence interval, we first will need to define the inverse cumulative distribution functions. Once the inverse function is known, we can show results of the quantiles of these distributions. Recall that our main goal for the ASCLT-based confidence interval is the estimation of the quantiles of the distribution for the correlation coefficient statistic. The inverse function for the distribution function \( H_N(t) \) stated in (1.7) is denoted as follows (see Thangavelu, 2005):

\[
H_{N}^{-1}(\alpha) = \begin{cases} 
\sup \{ t : H_N(t) = 0 \} & \text{for } \alpha = 0 \\
\sup \{ t : H_N(t) < \alpha \} & \text{for } 0 < \alpha < 1 \\
\inf \{ t : H_N(t) = 1 \} & \text{for } \alpha = 1
\end{cases}
\]

The inverse function for the distribution function \( J_N(t) \) stated in (1.8) is denoted as follows:

\[
J_{N}^{-1}(\alpha) = \begin{cases} 
\sup \{ t : J_N(t) = 0 \} & \text{for } \alpha = 0 \\
\sup \{ t : J_N(t) < \alpha \} & \text{for } 0 < \alpha < 1 \\
\inf \{ t : J_N(t) = 1 \} & \text{for } \alpha = 1
\end{cases}
\]

After the inverse functions were shown for the cumulative distribution functions in (1.7) and (1.8), the results of the quantiles of these functions will be defined as \( t_{\alpha}^{(N)} = H_{N}^{-1}(\alpha) \) or \( t_{\alpha}^{(N)} = J_{N}^{-1}(\alpha) \).

In chapter 4 section 4, a new version of the confidence interval for the population correlation coefficient will be presented. This method uses quantiles from the ASCLT-based distribution function (1.7) to estimate confidence intervals. One key property of
this new method is that the estimation or use of the variance of the observations is not
needed. Therefore this approach uses a variance-free method to estimate the limiting
distribution of the correlation coefficient statistic. The following new result is an
ASCLT-derived confidence interval for $\rho$:

$$I^{(N)}_\alpha = \left[ \hat{\rho} + \frac{t_{1-\alpha}^{(N)}}{\sqrt{N}}, \hat{\rho} + \frac{t_{\alpha}^{(N)}}{\sqrt{N}} \right]$$

(1.9)

where $\hat{\rho}$ is the estimated correlation coefficient.

Recall that Fisher developed a variance stabilizing technique transformation of $r$ that
tends to become quickly normal as $n$ increases. In this variance stabilizing technique,
Fisher used it to construct the confidence interval for the population correlation
coefficient. In chapter 4 section 5, another version of the confidence interval for the
population correlation coefficient will be presented. This method uses quantiles from the
ASCLT-based distribution function (1.8) to estimate confidence intervals. The following
new result is another ASCLT-derived confidence interval for $\rho$ using the variance
stabilizing technique:

$$I^{(N)}_\alpha = \left[ \exp \left\{ 2 z_N + \frac{t_{1-\alpha}^{(N)}}{\sqrt{N}} \right\} - 1 \right] \left[ \exp \left\{ 2 z_N + \frac{t_{\alpha}^{(N)}}{\sqrt{N}} \right\} - 1 \right]$$

$$\left[ \exp \left\{ 2 z_N + \frac{t_{1-\alpha}^{(N)}}{\sqrt{N}} \right\} + 1 \right] \left[ \exp \left\{ 2 z_N + \frac{t_{\alpha}^{(N)}}{\sqrt{N}} \right\} + 1 \right]$$

(1.10)

where $z_N = \log \frac{1 + \hat{\rho}}{1 - \hat{\rho}}$.

To the best of our knowledge, the cumulative distribution functions mentioned in (1.7)
and (1.8) have not been considered before. Even though the property of converging to asymptotic results will be discussed in chapter 3, numerical simulations will be performed in chapter 5 to show how these empirical distribution functions converges to a normal distribution. Chapter 5 will also evaluate the performance of the proposed ASCLT-based confidence intervals in (1.9) and (1.10) by comparing to the nonparametric bootstrap approach (discussed in section 5.8) and to the classic Fisher’s $z$ transformation parametric approach (discussed in section 5.9). All numerical simulations for the empirical distribution functions and confidence interval techniques will be completed for varying values of $n$ and $\rho$. Also, each simulation will be repeated for the bivariate normal, exponential, and poisson distributions.

Before proceeding into this dissertation, we should concisely mention our goals. By using the ASCLT we will be estimating the quantiles from an unknown distribution. In fact, we will present two ASCLT-based confidence interval approaches to drawing inferences for the population correlation coefficient. These new approaches will be developed using an almost sure version of Cramér’s Theorem. This dissertation will conclude by investigating the behavior of the theoretical results presented by numerical simulations.
Chapter 2

The Almost Sure Version of Cramér’s Theorem

From a statistical viewpoint, the almost sure central limit theorem is just a theorem for a special (but important) class of statistics, the mean. In this chapter, we will discuss the ASCLT and how it can be extended to be included with Cramér’s Theorem. This chapter will go into more detail of the ASCLT, introduce Cramér’s Theorem, and finally develop the almost sure version of Cramér’s Theorem.

2.1 Introduction to ASCLT

The almost sure central limit theorem (ASCLT) was developed independently by the researchers Fisher (1987), Brosamler (1988), and Schatte (1988). The theorem presented by these three authors has connected the theory of the central limit theorem to an almost sure version named the almost sure central limit theorem. A version of the ASCLT is presented below.

**Theorem 2.1.** (Almost Sure Central Limit Theorem) Let \( X_1, X_2, X_3, \ldots, X_n \) be i.i.d. random variables with \( S_n = X_1 + X_2 + \cdots + X_n \), being the partial sums. If \( EX_1 = 0, \ EX_1^2 = 1 \), and \( E|X_1|^{2+\delta} \) is finite for some \( \delta > 0 \) then

\[
\frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}} \leq t \right\} \xrightarrow{a.s.} \Phi(t) \text{ for any } t,
\]

(2.1)
where $\Phi(t)$ is the standard normal cumulative distribution function and $I_{i \in A}$ is the indicator function for the set $A$.

Brosamler’s version of the ASCLT presented in Theorem 2.1 is the simplest version since it assumes that $\mu = 0$. However, observe the following version of the ASCLT when the population mean is included:

$$\frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \frac{X_{1} + X_{2} + \cdots + X_{n} - n\mu}{\sqrt{n}} \right\} \overset{a.s.}{\longrightarrow} \Phi_{\sigma}(t) \text{ for any } t, \quad (2.2)$$

where now $\Phi_{\sigma}(t)$ is the normal cumulative distribution function with mean 0 and variance $\sigma^2$. In typical circumstances, the entire problem revolves around hypothesis testing and finding a confidence interval for $\mu$. Including $\mu$ in (2.2) should be carefully considered since this parameter is typically not known in practice. This issue will be addressed later in this thesis.

After the above theorem had been established, during the past decade, there have been several authors investigating the ASCLT and related logarithmic limit theorems for partial sums of independent random variables. In fact, Berkes and Csáki (2001) extended this theory and introduced that not only the central limit theorem, but every weak limit theorem for independent random variables, with minor technical conditions, can have an almost sure version. We will not go into any details surrounding the investigations other authors that have developed advances in the concepts of ASCLT, as these theorems are typically from a mathematical perspective. Our interest moving forward will be dealing with the development of the confidence interval procedures.
Lifshits (2002) extended the ASCLT from random variable to random vectors. He established a sufficient condition for the ASCLT for sums of independent random vectors under minimal moment conditions and assumptions on normalizing sequences. It should be noted that this article was translated from an article from the journal Zapiski Nauchnykh Seminarov (1999).

**Theorem 2.2.** Let \( \{ \xi_j \} \) be a sequence of independent random vectors taking values in \( \mathbb{R}^d \) where \( \delta_x > 0 \) is the probability measure which assigns its total mass to \( x \). Assume the following condition (see Berkes and Dehling; 1993) which is sufficient for the ASCLT, where \( h > 0 \) such that

\[
C = \sup_l E \left[ \log \log \| \xi_l \| \right]^{l+h} < \infty.
\]

Assume also that \( \{ \xi_j \} \in L^2 \). Then the ASCLT holds for random vectors, i.e.

\[
\frac{1}{\log N} \sum_{n=1}^{N} \frac{L}{n} \delta \left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \xi_j - E(\xi_j) \right\} \xrightarrow{a.s.} N(0, \Sigma)
\]

where

\[
\Sigma = \text{cov}(\xi_i, \xi_j)
\]

is a \( dx \)d covariance matrix.

### 2.2 Replace the Log Averaging Term in ASCLT

One issue of the ASCLT as in the form (2.1) or (2.2) is the rate of convergence to a normal distribution. Even for very large sample sizes, the rate of convergence is very
slow. Consider the following quotient:

\[
\sum_{n=1}^{N} \frac{1}{n} \cdot \frac{\log N}{n}
\]

This quotient will converge to 1 for sufficiently large values of \(N\). However, what is a large value of \(N\)? Even with large values on \(N\), this fraction does not equal 1. For example, for \(N = 10^7\), the above ratio is approximately equal to 1.03, and for \(N = 10^{10}\), it is approximately equal to 1.025. In applications of statistics, we typically do not have access to sample sizes of this magnitude. Therefore,

\[
\lim_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}} \leq t \right\}
\]

or

\[
\lim_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sqrt{n}} \leq t \right\}
\]

will not be a distribution function even for very large values of \(N\). Moving forward, we will use a proposal from Thangavelu (2005) that we replace the “\(\log N\)” in the ASCLT with the averaging term \(\sum_{n=1}^{N} \frac{1}{n}\) to create the following,

\[
\lim_{N \to \infty} \frac{1}{\sum_{n=1}^{N} \frac{1}{n}} \sum_{n=1}^{N} \frac{1}{n} \left\{ \frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}} \leq t \right\}
\]

or

\[
\lim_{N \to \infty} \frac{1}{\sum_{n=1}^{N} \frac{1}{n}} \sum_{n=1}^{N} \frac{1}{n} \left\{ \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sqrt{n}} \leq t \right\}.\]
For convenience, subsequent sections we will denote $\sum_{n=1}^{N} \frac{1}{n}$ by $C_N$.

### 2.3 Central Limit Theorem

Consider the following notation and assumptions. For $n \geq 1$, where $n \leq N \in \mathbb{N}$ let the statistic $T_n$ be a sequence of real valued statistics defined on the same measurable space $(\Omega, \beta)$ and $\mathcal{P}$ be a family of probabilities on $\beta$. Let us assume that the statistic $T_n$ satisfies the Central Limit Theorem for each $P \in \mathcal{P}$ where the constants $b_n = n^{-1/2}$ and $a_n(P) = n\mu(P)$ are unknown. Let us also assume that $G_P$ is the unknown normal distribution function (Normal $N(\mu, \sigma^2)$ where $\mu$ and $\sigma^2$ are unknown) that is continuous where $\mu(P) \in \mathbb{R}$ is also unknown:

$$P(\{ \omega \in \Omega : b_n(T_n(\omega) - a_n(P)) \leq t \}) \xrightarrow{a.s.} G_P(t) \quad \text{for } t \in C_G$$

where $C_G$ denotes the set of continuity points of $G_P$. We will be presenting results for a fixed $\omega \in \Omega$, though the results would be applicable to each $\omega \in \Omega$. Moving forward, we will denote $T_n(\omega)$ by $T_n$, $a_n(P)$ by $a_n$, and $\mu(P)$ by $\mu$. These simplifications will hold true for every $P \in \mathcal{P}$.

### 2.4 New Notation and Assumptions of the ASCLT

Again, consider the following notation and assumptions. For $n \geq 1$, where $n \leq N \in \mathbb{N}$, let the statistic $T_n$ be a sequence of real valued statistics defined on the same measurable space $(\Omega, \beta)$ and $\mathcal{P}$ be a family of probabilities on $\beta$. Let us assume that the statistic $T_n$
satisfies the CLT and ASCLT for each $P \in \mathcal{P}$ where the constants $b_n = n^{-1/2}$ and $a_n = n\mu$ are unknown. It follows that

$$
\lim_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} I_{\left\{ b_n (T_n - a_n) \leq t \right\}} = G_P(t), \quad \forall t \in \mathbb{R}^d \text{ a.s.}
$$

(2.3)

where $a_n$ and $b_n$ are non-random sequences. It is known from Brosamler (1988), Schatte (1988), Fisher (1987), and Lacey and Philipp (1990) that the almost sure limit theorem holds for $T_n$ being the mean of $n$ i.i.d. random variables with finite second moment. The case of generalized means, U-statistics, has been considered by Berkes (1993) and later by Holzmann, Koch, and Min (2004). One particular case was established by Thangavelu (2007) for the rank statistics. Other sequences of statistics have not been considered to the best of our knowledge.

So now the following function is defined for each $\omega \in \Omega$ and $t \in \mathbb{R}$:

$$
G_N(t, \omega) = G_N(t) = \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} I_{\left\{ b_n (T_n - a_n) \leq t \right\}}
$$

or

$$
G_N(t, \omega) = G_N(t) = \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} I_{\left(-\infty, t \right]}(b_n (T_n - a_n)).
$$

We will be presenting results for a fixed $\omega \in \Omega$, though the results would be applicable to each $\omega \in \Omega$.

Consider the following two functions that are now defined for each $\omega \in \Omega$ and $t \in \mathbb{R}$ with
replacing the log averaging term with $C_N$:

$$
\widetilde{G}_N(t) = \frac{1}{C_N} \sum_{n=1}^{N} \frac{1}{n} \mathbb{I}\{b_n(T_n-a_n) \leq t\}
$$

$$
= \frac{1}{C_N} \sum_{n=1}^{N} \frac{1}{n} \mathbb{I}_{(\infty, t]}(b_n(T_n-a_n)).
$$

And consider the simpler version of the ASCLT when $\mu = 0$:

$$
\hat{G}_N(t) = \frac{1}{C_N} \sum_{n=1}^{N} \frac{1}{n} \mathbb{I}\{b_nT_n \leq t\}
$$

$$
= \frac{1}{C_N} \sum_{n=1}^{N} \frac{1}{n} \mathbb{I}_{(\infty, t]}(b_nT_n)
$$

**Lemma 2.1.** Let $X_1, X_2, ..., X_n$ be i.i.d. random variables on $(\Omega, \beta, \mathbb{P})$ where $\Omega$ is the sample space, $\beta$ is the Sigma Field, and $\mathbb{P}$ is the probability space (probability of each possible subset in $\Omega$), where $X_i \in (-\infty, +\infty), \mu \in (-\infty, +\infty), \sigma^2 > 0$ and finite.

$\widetilde{G}_N$ and $\hat{G}_N$ are empirical distribution functions. Also, $\tilde{G}_N(t)$ converges to $G_P(t)$ almost surely for every $t \in C_G$.

Observe the following proof (Thangavelu, 2005). Let us first consider $\tilde{G}_N(t)$,

$$
\tilde{G}_N(t) = \frac{1}{C_N} \sum_{n=1}^{N} \frac{1}{n} \mathbb{I}\{\frac{X_1+X_2+...+X_n-n\mu}{\sqrt{n}} \leq t\}
$$

For $t < s \in \mathfrak{R}$, then,

$$
1_{(-\infty, t]}(x) \leq 1_{(-\infty, s]}(x) \quad \text{for all } x \in \mathfrak{R}.
$$

This implies $\tilde{G}_N(t) \leq \tilde{G}_N(s)$ for $n \leq N, N \in \mathbb{N}$ fixed. Therefore, $\tilde{G}_N(t)$ is monotonically
increasing in $t \in \Re$. We also observe that,

$$\lim_{t \to \infty} \frac{1}{C_N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sqrt{n}} \right\} = 0$$

which implies $\lim_{t \to \infty} \tilde{G}_N(t) = 0$. Also,

$$\lim_{t \to +\infty} \frac{1}{C_N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sqrt{n}} \right\} = 1$$

which implies $\lim_{t \to +\infty} \tilde{G}_N(t) = 1$. Further we note that the function $\tilde{G}_N(t)$ is a step function in $t$ and $0 \leq \tilde{G}_N(t) \leq 1$, where $\tilde{G}_N(t)$ is a constant for $t \in (t_{i-1}, t_i]$, for all $i = 2, 3, \ldots, s$.

Also, $t_1 \leq t_2 \leq \ldots \leq t_N$, $\tilde{G}_N(t) \approx 0$ for all $t \leq t_1$, and $\tilde{G}_N(t) \approx 1$ for all $t \geq t_N$. Note that $\tilde{G}_N(t)$ is left continuous, which is a function that is continuous when a point is approaching from the left. Therefore $\tilde{G}_N(t)$ is an empirical distribution function.

Also, since $\hat{G}_N(t)$ is a special case of $\tilde{G}_N(t)$ where $\mu = 0$, all the aforementioned steps in the proof hold true for $\hat{G}_N(t)$; hence it is also an empirical distribution function.

Now that we have established that $\tilde{G}_N(t)$ has the properties of a distribution function, we can now show that the result $\tilde{G}_N(t)$ converges to $G_{p}(t)$ almost surely for all $t \in C_G$. This is a unique case of the Glivenko-Cantelli Theorem, which relates the idea of consistency of statistical distributions. We will now state a version of the Glivenko-Cantelli Theorem (Thangavelu, 2005) below without proof. It should be noted that this theorem establishes
the relationship between the empirical distribution \( \hat{G}_N(t) \) and the theoretical distribution function \( G_P(t) \).

**Theorem 2.3.** (Glivenko-Cantelli Theorem) \( \hat{G}_N(t) \) converges almost surely to \( G_P(t) \) that is,

\[
\lim_{N \to \infty} \sup_{t \in \mathbb{R}} |\hat{G}_N(t) - G_P(t)| = 0.
\]

2.5 Cramér’s Theorem

We will now state Cramér’s Theorem (see (Ferguson, 2002; Lehmann, 1999)) below without proof.

**Theorem 2.4.** (Cramér’s Theorem) Let \( g \) be a mapping from \( \mathbb{R}^d \) into \( \mathbb{R}^k \) where \( g \) is differentiable in a neighborhood \( \mu \in \mathbb{R}^d \) and its derivative \( \dot{g}(x) \) is continuous at \( \mu \). Let \( X_n, n \geq 1 \) be a sequence of \( \mathbb{R}^d \)-valued random vectors such that \( \sqrt{n}(X_n - \mu) \xrightarrow{L} X \)
then \( \sqrt{n}(g(X_n) - g(\mu)) \xrightarrow{L} \dot{g}(\mu)X \). In particular, if \( \sqrt{n}(X_n - \mu) \xrightarrow{L} N(0, \Sigma) \) where \( \Sigma \) is a \( d \times d \) covariance matrix, then

\[
\sqrt{n}(g(X_n) - g(\mu)) \xrightarrow{L} N(0, \dot{g}(\mu) \Sigma \dot{g}(\mu)^T)
\]

For a proof of this theorem, please refer to Ferguson (2002).

Cramér’s theorem for smooth functions of sample moments is one of the basic results in statistics to diminish variation and allowing the construction of smaller confidence intervals. In this chapter we extend the result to almost sure convergence towards the limiting distribution (in the sense of the above almost sure weak convergence results.
(2.3)). We use the notation \( x \leq t \) for \( x_i \leq t_i \) for all \( i = 1, \ldots, d \). For a random variable \( X \) we let \( G_X \) denote its cumulative distribution function (c.d.f.) which is a right continuous function.

2.6 Auxiliary Lemmas

Let \( X, X_n \ (n \geq 1) \) be \( d \)-dimensional random vectors. We say that \( X_n \) converges to \( X \) almost surely in the weak topology (for short a.s. weakly) if for any bounded continuous function \( f : \mathbb{R}^d \to \mathbb{R} \),

\[
\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} f(X_k) = Ef(X)
\]

holds. It is necessary and sufficient for this convergence that the relation holds for any function \( f \) of the form

\[
I_{\{x \in \mathbb{R}^d : x_i \leq t_i, i = 1, \ldots, d\}}
\]

where \( t = (t_1, \ldots, t_d) \) is a vector in \( \mathbb{R}^d \) at which the distribution function \( G_X \) of \( X \) is continuous. This follows from the standard fact that these indicator functions form a generating class for all bounded continuous functions. We shall use the notation,

\[
X_n \xrightarrow{a.s.} X
\]

to denote a.s. weak convergence. The set of continuity points will be denoted by \( D(G_X) \).

We first note that Slutsky’s lemma (see (Ferguson, 2002; van der Vaart, 1998; Manoukian, 1963)) can be easily generalized. Slutsky’s lemma is a good approximation tool that can be used in conjunction with the Central Limit Theorem. To simplify
notation, we use $t \leq s$ for vectors $s = (s_1, ..., s_d)$, $t = (t_1, ..., t_d) \in \mathbb{R}^d$ to denote the ordering $t_i \leq s_i$ for $i = 1, ..., d$ of all coordinates. Likewise $t < s$ means that at least one coordinate is strictly smaller.

**Lemma 2.2.**

1. If $a_n \in \mathbb{R}$ $(n \geq 1)$ is a sequence converging to $a \in \mathbb{R}$ and if $X_n \xrightarrow{a.s.} X$, then $a_n X_n \xrightarrow{a.s.} aX$.

2. If $X_n - Y_n \xrightarrow{a.s.} 0$ and if $X_n \xrightarrow{a.s.} X$ then $Y_n \xrightarrow{a.s.} X$.

**Proof:**

(1) Let $a, a_n \in \mathbb{R}^d$ satisfy \( \lim_{n \to \infty} a_n = a \) and let $X_n$ $(n \geq 1)$ be convergent to $X$ almost surely in the weak topology. We use the equivalent definition of this convergence. Let $t \in \mathbb{R}^d$ be a vector at which the distribution function $G_X$ of $X$ is continuous.

If $a = 0$, then $aX$ is zero a.s. $G_{aX}(t)$ vanishes for $t = (t_1, ..., t_d) \in \mathbb{R}^d$ such that $t_i < 0$ for some $1 \leq i \leq d$, and equals 1 for $t \in \mathbb{R}^d$ for which each coordinate is strictly greater than 0. Let $t = (t_1, ..., t_d) \in \mathbb{R}^d$ be a continuity point of $G_0$ such that $G_0(t) = 0$. Note that the set $I = \{ i : 1 \leq i \leq d; t_i < 0 \}$ is non-empty. Let $\eta > 0$. Choose $u, v \in \mathbb{R}^d$ such that $v_i = \infty$ for $i \not\in I$,

$$G_X(v) < \eta \text{ and } I - G_X(u) < \eta.$$
Then

\[
\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\{a_k X_k \leq t\}
\]

\[
= \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\{a_k X_k \leq t\} + \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=n_0+1}^{n} \frac{1}{k} I\{a_k X_k \leq t\}
\]

\[
+ \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=n_0+1}^{n} \frac{1}{k} \prod_{i=1}^{d} I\{X_k(i) \leq t_i / a_k\}
\]

\[
+ \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=n_0+1}^{n} \frac{1}{k} \prod_{i=1}^{d} I\{X_k(i) \geq t_i / a_k\}
\]

\[
\leq \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \prod_{i \in I} I\{X_k(i) \leq v_i\} \prod_{i \notin I} I\{X_k(i) < \infty\}
\]

\[
+ \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \prod_{i \in I} I\{X_k(i) \geq u_i\} \prod_{i \notin I} I\{X_k(i) \geq t_i / a_k\}
\]

\[
= \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\{X_k \leq v\} + \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\{X_k \notin \Pi I(-\infty, u_i)\}
\]

\[
\leq 2\eta
\]

since

\[
\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\{X_k \notin \Pi I(-\infty, u_i)\} = 1 - \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\{X_k \leq u\}.
\]

Let \( t = (t_1, ..., t_d) \in \mathbb{R}^d \) be a continuity point of \( G_0 \) such that \( G_0(t) = 1 \). This means that each \( t_i > 0, i = 1, ..., d \). Let \( \eta > 0 \). Choose \( u, v \in \mathbb{R}^d \) such that \( v_i = \infty \) for \( i \notin I \),

\[
G_{X(v)} \leq \eta \quad \text{and} \quad 1 - G_{X(u)} < \eta.
\]

Since \( \lim_{n \to \infty} a_n = a = 0 \), there is \( n_0 \in \mathbb{N} \) such that for all \( n \geq n_0 \) we have that \( t/a_n \leq u \) in case \( a_n > 0 \) and \( t/a_n \leq v \) in case \( a_n < 0 \). Then
\[
\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} 1\{a_k X_k \leq t\}
\]

\[
= 1 - \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} 1\{a_k X_k \notin \Pi_i (-\infty, t_i]\}
\]

\[
= 1 - \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n_0} \frac{1}{k} 1\{a_k X_k \notin \Pi_i (-\infty, t_i]\}
\]

\[
= \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=n_0 + 1, a_k = 0}^{n} \frac{1}{k} 1\{a_k X_k \notin \Pi_i (-\infty, t_i]\}
\]

\[
= \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=n_0 + 1, a_k > 0}^{n} \frac{1}{k} 1\{X_k \geq t_i / a_k\}
\]

\[
= \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=n_0 + 1, a_k < 0}^{n} \frac{1}{k} 1\{X_k \leq t_i / a_k\}
\]

\[
\geq 1 - \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} 1\{X_k \notin \Pi_i (-\infty, u_i]\}
\]

\[
= \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} 1\{X_k \leq v\}
\]

\[
\geq 1 - 2\eta
\]

since

\[
\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} 1\{X \notin \Pi_i (-\infty, u_i]\}
\]

\[
= 1 - \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} 1\{X \in \Pi_i (-\infty, u_i]\}
\]

\[
= 1 - G_X(u)
\]

\[
< \eta.
\]

It is left to consider the case when \(a \neq 0\). We first show the result for \(a > 0\). We may assume that every \(a_k > 0\) by a similar argument as just used. In this case, let \(t \in \mathbb{R}^d\) be a continuity point of \(G_{aX}\), the distribution function of \(aX\). For every \(k \geq 1\),

\[
a_kX_k \leq t \quad \text{if and only if} \quad X_k \leq t/a_k.
\]
Let $\eta > 0$. Since $t$ is a continuity point and since $G_{aX}(t) = G_X(t/a)$, there exists $\delta > 0$ such that,

$$G_X((t/a) + \delta e) - G_X((t/a) - \delta e) < \eta,$$

where $e = (1, 1, ..., 1) \in \mathbb{R}^d$. Choose $n_0 \in \mathbb{N}$ such that for $n \geq n_0$,

$$\left\| \frac{t}{a_n} - \frac{t}{a} \right\| < \delta$$

where $\| \cdot \|$ denotes the maximum norm in $\mathbb{R}^d$. It follows that

$$\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\{a_k X_k \leq t\} = \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\{X_k \leq t/a_k\} \leq \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\{X_k \leq (t/a) + \delta e\}$$

where $X_k$ is a distribution function of $aX$.

In order to obtain a lower bound, we can calculate

$$\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\{a_k X_k \leq t\} \leq G_X((t/a) + \delta e) \leq G_X(t/a).$$

Now observe that $G_X(t/a)$ is a distribution function of $aX$. In order to obtain a lower bound, we can calculate
Letting $\eta$ tend to zero shows that the lower bound for the limit is as well $G_X(t/a)$, completing the proof in the case when $a > 0$. Finally, the case $a < 0$ is carried out in the same way.

(2) Let $X_n - Y_n$ converge to zero almost surely and $X_n$ converge to $X$ weakly almost surely. Let $t$ be a continuity point of the distribution function $G_X$ of $X$. Let $\eta > 0$. Choose $\delta > 0$ such that 

$$G_X(t + \delta e) - G_X(t - \delta e) < \eta,$$

where $e$ is as in (1). Let $\Omega_0$ be a set of probability one such that for $\omega \in \Omega_0$ we have

$$\lim_{n \to \infty} [X_n(\omega) - Y_n(\omega)] = 0$$

and

$$\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\{X_k(\omega) \leq t\} = G_X(t).$$

Fix $\omega \in \Omega_0$. Choose $n_0$ such that for $n \geq n_0$ we have that $\|X_n(\omega) - Y_n(\omega)\| < \delta$. Then
\[
\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \mathbb{I}\{Y_k \leq t\} = \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{\lfloor n \log 2 \rfloor} \frac{1}{k} \mathbb{I}\{X_k \leq t\} + \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=n_0 + 1}^{n} \frac{1}{k} \mathbb{I}\{X_k \leq t - Y_k + X_k\} \\
\leq \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \mathbb{I}\{X_k \leq t + \delta\}
\]

\[
= G_X(t + \delta) \\
\leq G_X(t) + \eta.
\]

Since \(\eta\) is arbitrary, we obtain

\[
\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \mathbb{I}\{Y_k \leq t\} \leq G_X(t).
\]

For the converse inequality, note that

\[
\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \mathbb{I}\{Y_k \leq t\} = \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{\lfloor n \log 2 \rfloor} \frac{1}{k} \mathbb{I}\{X_k \leq t\} + \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=n_0 + 1}^{n} \frac{1}{k} \mathbb{I}\{X_k \leq t - Y_k + X_k\} \\
\geq \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=n_0 + 1}^{n} \frac{1}{k} \mathbb{I}\{X_k \leq t - \delta\}
\]

\[
= \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \mathbb{I}\{X_k \leq t - \delta\} - \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n_0} \frac{1}{k} \mathbb{I}\{X_k \leq t - \delta\} \\
= G_X(t - \delta) \\
\geq G_X(t) - \eta.
\]

Letting \(\eta \to 0\), claim (2) results.

**Lemma 2.3.** Let \(Z_n\) be \(\mathbb{R}^d\) - valued random variables and \(\Sigma\) be a matrix. If
\[
\lim_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} I\{Z_n \leq t\} = G_Z(t), \quad t \in D(G_Z) \text{ a.s.,}
\]

then

\[
\lim_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} I\{\sum Z_n \leq t\} = G_{\sum Z}(t), \quad t \in D(G_{\sum Z}) \text{ a.s.}
\]

**Proof:** Let \( A = \{ x \in \mathbb{R}^d : \sum x \leq t \} \). Then

\[
\lim_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} I\{\sum Z_n \leq t\}
= \lim_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} I\{Z_n \in A\}
= G_Z(A) = G_{\sum Z}(t) \quad \text{a.s.}
\]

### 2.7 The Almost Sure Version of Cramér’s Theorem

**Theorem 2.5.** Let the function \( g \) be a mapping from \( \mathbb{R}^d \) into \( \mathbb{R}^k \) where \( g \) is differentiable in a neighborhood \( \mu \in \mathbb{R}^d \) and its derivative \( g'(x) \) is continuous at \( \mu \). Let \( X_n, n \geq 1 \) be a sequence of \( \mathbb{R}^d \) – valued random vectors satisfying the almost sure weak convergence property:

\[
\lim_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} I\{n^{-1/2}(X_n - \mu) \leq t\} = G_X(t), \quad t \in D(G_X) \text{ a.s.,} \tag{2.4}
\]

where \( G_X \) is the cumulative distribution function (c.d.f.) of some random variable \( X \) and \( D(G_X) \) is the set of continuity points of \( G_X \). If there exists a sequence \( \{ n_k : k \in \mathbb{N} \} \) of integers such that
\[
\lim_{N \to \infty} \sum_{n=1}^{N} \frac{1}{n} = 0
\]  \hspace{1cm} (2.5)

and

\[
\lim_{k \to \infty} X_{n_k} = \mu \hspace{0.5cm} \text{a.s.}
\]  \hspace{1cm} (2.6)

Then

\[
\lim_{n \to \infty} \log N \sum_{n \in N_o} \frac{1}{n} \iota\{\sqrt{n}(g(X_n) - g(\mu)) \leq t\} = G\dot{g}(\mu)X(t), \hspace{0.5cm} t \in D(G\dot{g}(\mu)X) \hspace{0.5cm} \text{a.s.} \hspace{0.5cm} (2.7)
\]

**Proof**: First note that we may assume that \(X_n \to \mu\) a.s., since (2.5) shows that

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n \in N_o} \frac{1}{n} \iota\{\sqrt{n}(g(X_n) - g(\mu)) \leq t\} = 0.
\]

We shall show the convergence in equation (2.7) on the set of points \(\omega \in \Omega\) for which equations (2.4) and (2.6) hold. By the Taylor’s series expansion,

\[
\sqrt{n}(g(X_n) - g(\mu)) = \int_{0}^{1} \dot{g}(\mu + v(X_n - \mu))dv\sqrt{n}(X_n - \mu),
\]

and by continuity of \(\dot{g}\) at \(\mu\) and (2.6),

\[
\dot{g}(\mu + v(X_n - \mu)) \to \dot{g}(\mu) \hspace{0.5cm} \text{a.s.}
\]

Now let \(\delta > 0\) and \(\varepsilon > 0\). First choose \(t_\delta \in \mathbb{R}^d\) such that

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{n \in N} \frac{1}{n} \iota\{\sqrt{n}(X_n(\omega) - \mu) \leq t_\delta \} \geq 1 - \delta.
\]

Denote by \(N_1 = \{n \in N : \sqrt{n}(X_n(\omega) - \mu) \geq t_\delta \}\), where \(x \geq t\) means that at least one coordinate is larger. It follows that
Next choose \( n_0 \in N \) with \( \| X_n(\omega) - \mu \| < \epsilon \) for all \( n \geq n_0 \). We also may assume \( n_0 \) is so large that

\[
\| \hat{g}(\mu + \nu (X_n(\omega) - \mu)) - \hat{g}(\mu) \| \leq \epsilon \quad \forall 0 \leq \nu \leq 1.
\]

Finally, choose \( N_0 \in N \) with

\[
\frac{\log n_0}{\log N} < \delta \quad \forall N \geq N_0,
\]

and

\[
\sup_{t \in D(G \hat{g}(\mu)X)} \left| \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \{ \sqrt{n}(X_n(\omega) - \mu) \} - G \hat{g}(\mu)X(t) \right| < \delta.
\]

Then for all \( N \geq N_0 \),

\[
\frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \{ \sqrt{n}(g(X_n(\omega)) - g(\mu)) \} \leq t
\]

\[
= \frac{1}{\log N} \sum_{n=1}^{n_0-1} \frac{1}{n} \{ \sqrt{n}(g(X_n(\omega)) - g(\mu)) \} \leq t
\]

\[
+ \frac{1}{\log N} \sum_{n=n_0}^{N} \frac{1}{n} \{ \sqrt{n}(g(X_n(\omega)) - g(\mu)) \} \leq t
\]

\[
+ \frac{1}{\log N} \sum_{n=n_0}^{N} \frac{1}{n} \{ \sqrt{n}(g(X_n(\omega)) - g(\mu)) \} \leq t
\]

\[
\leq \frac{\log n_0}{\log N} + \delta + \frac{1}{\log N} \sum_{n=n_0}^{N} \frac{1}{n} \{ \sqrt{n}(g(X_n(\omega)) - g(\mu)) \} \leq t
\]

30
\[
\leq 2\delta + \frac{1}{\log N} \sum_{n=1\colon n \notin N_1}^N \frac{1}{\log N} \sum_{n=1\colon n \notin N_1}^N \frac{1}{n} \int_0^1 \hat{g}(\mu) \sqrt{n}(X_n(\omega) - g(\mu)) \leq t + \frac{1}{\log N} \sum_{n=1\colon n \notin N_1}^N \frac{1}{n} \int_0^1 \hat{g}(\mu) \sqrt{n}(X_n(\omega) - g(\mu)) \leq t + \epsilon_n \delta \}
\leq G_{\hat{g}(\mu)X}(t + \epsilon_n \delta ) + 3\delta , \quad t \in D(G_{\hat{g}(\mu)X}).
\]

The last inequality uses Lemma 2.3.

In order to continue the proof of the theorem, first let \( \epsilon \to 0 \), then \( \delta \to 0 \) to obtain

\[
\lim_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^N \frac{1}{n} \int \sqrt{n}(g(X_n) - g(\mu)) \leq t \} \leq G_{\hat{g}(\mu)X}(t), \quad t \in \mathbb{R}^k.
\]

The converse inequality is shown in a similar fashion.

\[
\leq \frac{1}{\log N} \sum_{n=1\colon n \notin N_1}^N \frac{1}{n} \int_0^1 \hat{g}(\mu) \sqrt{n}(X_n(\omega) - g(\mu)) \leq t \}
\geq \frac{1}{\log N} \sum_{n=1\colon n \notin N_1}^N \frac{1}{n} \int \sqrt{n}(g(X_n(\omega)) - g(\mu)) \leq t \}
= \frac{1}{\log N} \sum_{n=1\colon n \notin N_1}^N \frac{1}{n} \int \sqrt{n}(X_n(\omega) - g(\mu)) \leq t \}
= \frac{1}{\log N} \sum_{n=1\colon n \notin N_1}^N \frac{1}{n} \int \sqrt{n}(X_n(\omega) - g(\mu)) \leq t + \frac{1}{\log N} \sum_{n=1\colon n \notin N_1}^N \frac{1}{n} \int \sqrt{n}(X_n(\omega) - g(\mu)) \leq t + \epsilon_n \delta \}
\geq \frac{1}{\log N} \sum_{n=1\colon n \notin N_1}^N \frac{1}{n} \int \sqrt{n}(X_n(\omega) - g(\mu)) \leq t - \epsilon_n \delta \}
\]

31
\[
\geq \frac{1}{\log N} \sum_{n=n_0}^{N} \frac{1}{n} \{ \hat{g}(\mu) \sqrt{n} (X_n(\omega) - g(\mu)) \leq t - \varepsilon \delta \} \\
- \frac{1}{\log N} \sum_{n=1}^{n_0-1} \frac{1}{n} \{ \hat{g}(\mu) \sqrt{n} (X_n(\omega) - g(\mu)) \leq t - \varepsilon \delta \} \\
- \frac{1}{\log N} \sum_{n=n_0 \cdot n \in N_1} \frac{1}{n} \{ \sqrt{n} (X_n(\omega) - g(\mu)) \geq t \epsilon \} \\
\leq G_{\hat{g}(\mu)X}(t - \varepsilon \delta) - 3\delta, \quad t \in D(G_{\hat{g}(\mu)X}).
\]

Again, letting \( \varepsilon \to 0 \), then \( \delta \to 0 \) shows that,

\[
\lim_{N \to \infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \{ \sqrt{n} (g(X_n) - g(\mu)) \leq t \} \geq G_{\hat{g}(\mu)X}(t^-), \quad t \in \mathbb{R}^k,
\]

where \( G(t^-) = \lim_{s \uparrow t} G(s) \). Note that \( G \) is taken to be the right continuous version of the c.d.f.
Chapter 3

ASCLT for the Correlation Coefficient

The population correlation coefficient, denoted by $\rho_{xy}$, measures the strength and direction of the linear association between these two quantitative variables from a bivariate distribution. The range of values for the correlation coefficient is from -1 to +1, where the closer $\rho_{xy}$ is to ±1, the stronger the linear relationship. In this chapter, we denote the sample correlation coefficient by $r_{xy}$.

3.1 Introduction to Population Correlation Coefficient

We begin with some more notation. Let $(X_n, Y_n)$ $(n \geq 1)$ be a sequence of independently identically bivariate vector pairs having a joint probability distribution $f(x,y)$ and having the assumption as in Theorem 2.2. We denote $\mu_x = E(X_i)$ and $\mu_y = E(Y_i)$ the expectations of the marginals and by $\sigma^2_x$ and $\sigma^2_y$ their variances. Let the covariance of $X_i$ and $Y_1$ (see Casella and Berger, 2002) be denoted by:

$$\sigma_{xy} = \text{Cov}(X_i, Y_1) = E[(X_i - \mu_x)(Y_1 - \mu_y)] = E(X_i Y_1) - \mu_x \mu_y.$$  

The corresponding sample variance and covariances are:

$$s^2_x = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{x})^2;$$

$$s^2_y = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{y})^2;$$

and
The following is the formula for the population correlation coefficient:

\[ \rho_{xy} = \frac{\text{Cov}(X,Y)}{\sigma_x \sigma_y}, \]

where

\[ \text{Cov}(X,Y) = \begin{cases} \sum \sum (x - \mu_x)(y - \mu_y)f(x,y), & (X,Y) \text{ discrete} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y)f(x,y)dxdy, & (X,Y) \text{ continuous} \end{cases} \]

A positive correlation coefficient indicates a positive relationship between \( X \) and \( Y \), whereas a negative correlation coefficient indicates a negative relationship between \( X \) and \( Y \). If large values of \( X \) tend to be observed with large values of \( Y \), and small values of \( X \) tend to be observed with small values of \( Y \), \( \text{Cov}(X,Y) \) will be positive. This positive relationship can be shown when \( X > \mu_x \), then \( Y > \mu_y \) is likely to be true and the product of \( (X-\mu_x)(Y-\mu_y) \) will be positive. If \( X < \mu_x \), then \( Y < \mu_y \) is also likely to be true and the product of \( (X-\mu_x)(Y-\mu_y) \) will again be positive. However, if large values of \( X \) tend to be observed with small values of \( Y \), or if small values of \( X \) tend to be observed with large values of \( Y \), \( \text{Cov}(X,Y) \) will be negative. The negative relationship can also be shown when \( X > \mu_x \), then \( Y < \mu_y \) (and vice versa) is likely to be true and the product of \( (X-\mu_x)(Y-\mu_y) \) will tend to be negative. Using the fact that \( \text{Cov}(X,Y) = E(XY) - \mu_x \mu_y \), the population correlation coefficient:

\[ \rho_{xy} = \frac{E(XY) - \mu_x \mu_y}{\sigma_x \sigma_y}. \]

The proof of this version of the correlation coeff. can be viewed in Casella and Berger (2002).
3.2 The Sample Correlation Coefficient

Let \((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\) be a random sample of \(n\) pairs from a bivariate distribution. The formula for the sample correlation coefficient is given by

\[
 r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}} = \frac{\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{\sum_{i=1}^{n} x_i^2}{n} \sqrt{\sum_{i=1}^{n} y_i^2 - \frac{\sum_{i=1}^{n} y_i^2}{n}}}}.
\]

Let the sample correlation coefficient \(r_{xy}\) be a point estimate of the population coefficient \(\rho_{xy}\). We can define the above sample correlation coefficient using sample moments. If \(x_1, x_2, x_3, \ldots, x_n\) is a random sample, where the sample moments are defined as

\[
m_X = \frac{1}{n} \sum_{i=1}^{n} x_i, \quad m_{xx} = \frac{1}{n} \sum_{i=1}^{n} x_i^2, \quad m_{xy} = \frac{1}{n} \sum_{i=1}^{n} x_i y_i, \quad \text{etc.}
\]

The following are the steps to define \(r_{xy}\) using sample moments:

\[
r_{xy} = \frac{\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}}{\sqrt{\sum_{i=1}^{n} x_i^2 - \frac{\sum_{i=1}^{n} x_i^2}{n} \sqrt{\sum_{i=1}^{n} y_i^2 - \frac{\sum_{i=1}^{n} y_i^2}{n}}}}.
\]
The definition of the sample correlation coefficient using sample moments will help in the following sections.

3.3 The Asymptotic Distribution of Sample Correlation Coefficient

When developing inferential statistical techniques (i.e. hypothesis testing and confidence intervals) for the population correlation coefficient, the asymptotic distribution of the sample correlation coefficient must be considered (see Ferguson, 2002). First, we let $A = (a_{ij})$ denote a 3x3 matrix with entries

$$
\frac{1}{n} \left[ \sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n} \right]
$$

$$
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i^2 = \frac{\left( \sum_{i=1}^{n} x_i \right)^2}{n} - \frac{1}{\sqrt{n}} \sum_{i=1}^{n} y_i^2 = \frac{\left( \sum_{i=1}^{n} y_i \right)^2}{n}
$$

$$
\frac{\sum_{i=1}^{n} x_i y_i}{n} - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}
$$

$$
\frac{\sum_{i=1}^{n} x_i^2}{n} - \frac{\left( \sum_{i=1}^{n} x_i \right)^2}{n^2} = \frac{\sum_{i=1}^{n} y_i^2}{n} - \frac{\left( \sum_{i=1}^{n} y_i \right)^2}{n^2}
$$

$$
\frac{m_{xy} - m_x \cdot m_y}{\sqrt{m_{xx} - (m_x)^2} \sqrt{m_{yy} - (m_y)^2}}.
$$

The definition of the sample correlation coefficient using sample moments will help in the following sections.
\[
\begin{align*}
a_{11} &= \text{Cov}((X_1 - \mu_x)^2, (X_1 - \mu_x)^2) = E((X_1 - \mu_x)^4) - \sigma_x^4, \\
a_{22} &= \text{Cov}((X_1 - \mu_x)(Y_1 - \mu_y), (X_1 - \mu_x)(Y_1 - \mu_y)) = E((X_1 - \mu_x)^2(Y_1 - \mu_y)^2) - \sigma^2_{xy}, \\
a_{33} &= \text{Cov}((Y_1 - \mu_y)^2, (Y_1 - \mu_y)^2) = E((Y_1 - \mu_y)^4) - \sigma_y^4, \\
a_{12} &= a_{21} = \text{Cov}((X_1 - \mu_x)^2, (X_1 - \mu_x)(Y_1 - \mu_y)) = E((X_1 - \mu_x)^3(Y_1 - \mu_y)) - \sigma_x^2\sigma_{xy}, \\
a_{13} &= a_{31} = \text{Cov}((X_1 - \mu_x)^2, (Y_1 - \mu_y)^2) = E((X_1 - \mu_x)^2(Y_1 - \mu_y)^2) - \sigma_x^2\sigma_y^2, \\
a_{32} &= a_{23} = \text{Cov}((Y_1 - \mu_y)^2, (Y_1 - \mu_y)(X_1 - \mu_x)) = E((X_1 - \mu_x)(Y_1 - \mu_y)^3) - \sigma_{xy}\sigma_y^2. 
\end{align*}
\]

In order to prepare for the following theorem we note the following Lemma.

**Lemma 3.1:** Let \((X_i, Y_i) (i \geq 1)\) be as above and define \(S_n = (S_{n1}, S_{n2}, S_{n3}, S_{n4}, S_{n5})\) by

\[
\begin{align*}
S_{n1} &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - \mu_x); \\
S_{n2} &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (Y_k - \mu_y); \\
S_{n3} &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - \mu_x)^2; \\
S_{n4} &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (Y_k - \mu_y)^2; \\
S_{n5} &= \frac{1}{\sqrt{n}} \sum_{k=1}^{n} (X_k - \mu_x)(Y_k - \mu_y).
\end{align*}
\]

Then we have the almost sure convergence \(S_n \xrightarrow{a.s.} N(b, B)\), where

\[
b = (\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \sigma_{xy})\] and \(B = (b_{ij})\) has the following covariance matrix
\[ b_{11} = \sigma_x^2 \]
\[ b_{22} = \sigma_y^2 \]
\[ b_{33} = \text{Var}( (X_1 - \mu_x)^2 - \sigma_x^2 ) \]
\[ b_{44} = \text{Var}( (Y_1 - \mu_y)^2 - \sigma_y^2 ) \]
\[ b_{55} = \text{Var}( (X_1 - \mu_x)(Y_1 - \mu_y) - \sigma_{xy} ) \]
\[ b_{12} = b_{21} = \sigma_{xy} \]
\[ b_{13} = b_{31} = \text{Cov}(X_1, (X_1 - \mu_x)^2) \]
\[ b_{14} = b_{41} = \text{Cov}(X_1, (Y_1 - \mu_y)^2) \]
\[ b_{15} = b_{51} = \text{Cov}(X_1, (X_1 - \mu_x)(Y_1 - \mu_y)) \]
\[ b_{23} = b_{32} = \text{Cov}(Y_1, (X_1 - \mu_x)^2) \]
\[ b_{24} = b_{42} = \text{Cov}(Y_1, (Y_1 - \mu_y)^2) \]
\[ b_{25} = b_{52} = \text{Cov}(Y_1, (X_1 - \mu_x)(Y_1 - \mu_y)) \]
\[ b_{34} = b_{43} = \text{Cov}( (X_1 - \mu_x)^2, (Y_1 - \mu_y)^2 ) \]
\[ b_{35} = b_{53} = \text{Cov}( (X_1 - \mu_x)^2, (X_1 - \mu_x)(Y_1 - \mu_y)) \]
\[ b_{45} = b_{54} = \text{Cov}( (Y_1 - \mu_y)^2, (X_1 - \mu_x)(Y_1 - \mu_y)) \].

**Proof:** This follows directly from the almost sure central limit theorem for i.i.d. random vectors (Theorem 2.2) applied to

\[ Z_n = (X_n, Y_n, (X_n - \mu_x)^2, (Y_n - \mu_y)^2, (X_n - \mu_x)(Y_n - \mu_y)) \]

Recall that the population correlation coefficient is \( \rho = \rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \), and it is estimated by its empirical version

\[ r = r_{xy} = \frac{s_{xy}}{s_x s_y} \]

The following theorem and outlined proof can be found in Ferguson (2002).

**Theorem 3.1.** Let \( (x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n) \) be a random sample of \( n \) pairs as having been drawn from a bivariate population with finite fourth moments, \( EX^4 \) and \( EY^4 \), then:
(1) The statistic
\[ \sqrt{n} \left[ \begin{pmatrix} s_x^2 \\ s_{xy} \\ s_y^2 \end{pmatrix} - \begin{pmatrix} \sigma_x^2 \\ \sigma_{xy} \\ \sigma_y^2 \end{pmatrix} \right] \]
is asymptotically normal \( N(0, C) \), where
\[ C = \begin{pmatrix} b_{33} & b_{35} & b_{34} \\ b_{35} & b_{55} & b_{45} \\ b_{34} & b_{45} & b_{44} \end{pmatrix} \]

(2) \( \sqrt{n}(r_n - \rho) \xrightarrow{a.s.} N(0, \gamma^2) \), where
\[ \gamma^2 = \frac{1}{4} \rho^2 \left[ \frac{b_{33}}{\sigma_x^4} + 2 \frac{b_{34}}{\sigma_x^2 \sigma_y^2} + \frac{b_{44}}{\sigma_y^4} \right] - \rho \left[ \frac{b_{35}}{\sigma_x^3 \sigma_y} + \frac{b_{45}}{\sigma_x \sigma_y^3} \right] + \left[ \frac{b_{55}}{\sigma_x^2 \sigma_y^2} \right]. \]

**Proof:** Note that the assumption of finite fourth moments ensure the existence and finiteness of all (mixed) moments used below. This follows from Hölder’s inequality (see Dudley, 1989).

(1) Without loss of generality, let us assume \( \mu_x = \mu_y = 0 \) since \( s_{xy}, s_x^2, s_y^2 \) does not depend on location. Now by the aforementioned lemma, \( S_n = (S_{n1}, S_{n2}, S_{n3}, S_{n4}, S_{n5}) \) is a.s. asymptotically normal. Let \( g \) be a mapping where \( g: \mathbb{R}^5 \to \mathbb{R}^3 \) and apply the a.s. Cramér theorem to the function:
\[ g(x) = \begin{pmatrix} g_1(x_1, \ldots, x_5) \\ g_2(x_1, \ldots, x_5) \\ g_3(x_1, \ldots, x_5) \end{pmatrix} = \begin{pmatrix} x_3 - x_1^2 \\ x_5 - x_1 x_2 \\ x_4 - x_2^2 \end{pmatrix}. \]
Note that the mean of $S^*_n = \frac{1}{\sqrt{n}}S_n$ is $\mu = (0,0,\sigma^2_x,\sigma^2_y,\sigma_{xy})$ and

$$\dot{g}(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} g_1(x) & \ldots & \frac{\partial}{\partial x_5} g_1(x) \\ \vdots & \vdots & \vdots \\ \frac{\partial}{\partial x_1} g_3(x) & \ldots & \frac{\partial}{\partial x_5} g_3(x) \end{pmatrix} = \begin{pmatrix} -2x_1 & 0 & 1 & 0 & 0 \\ -x_2 & -x_1 & 0 & 0 & 1 \\ 0 & -2x_2 & 0 & 1 & 0 \end{pmatrix}.$$ 

It should be mentioned that we can prove this theorem by defining the following moment matrix:

$$x_n = \begin{pmatrix} x_1 = m_x \\ x_2 = m_y \\ x_3 = m_{xx} \\ x_4 = m_{xy} \\ x_5 = m_{yy} \end{pmatrix}.$$ 

Please refer to Ferguson (2002) to follow the proof steps that utilizes this $x_n$ matrix.

Observing the derivative of $g(\mu)$ is

$$\dot{g}(\mu) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

and

$$\sqrt{n}(g(S^*_n) - g(\mu)) = \sqrt{n} \begin{pmatrix} s^2_x \\ s_{xy} \\ s^2_y \end{pmatrix} - \begin{pmatrix} \sigma^2_x \\ \sigma_{xy} \\ \sigma^2_y \end{pmatrix} \xrightarrow{a.s.} N(0, \dot{g}(\mu)B\dot{g}(\mu)^T),$$

where the covariance matrix equals

40
Therefore,

$$\sqrt{n}(g(S_n^*) - g(\mu)) = \sqrt{n} \begin{pmatrix} s_x^2 \\ s_{xy} \\ s_y^2 \end{pmatrix} - \begin{pmatrix} \sigma_x^2 \\ \sigma_{xy} \\ \sigma_y^2 \end{pmatrix} \overset{a.s.}{\rightarrow} N \left[ 0, \begin{pmatrix} b_{33} & b_{35} & b_{34} \\ b_{53} & b_{55} & b_{54} \\ b_{43} & b_{45} & b_{44} \end{pmatrix} \right].$$

(2) Let us again assume \( \mu_x = \mu_y = 0 \) since \( r \) does not depend on location. In order to show the second part of the theorem, use part (1) together with Cramér’s theorem applied to the function \( h \). Let \( h \) be a mapping where \( h : \mathbb{R}^3 \to \mathbb{R}^l \) defined by

$$h(y_1, y_2, y_3) = \frac{y_2}{\sqrt{y_1 y_3}};$$

where

$$h(s_x^2, s_{xy}, s_y^2) = \frac{s_{xy}}{s_x s_y};$$

$$h(s_x^2, s_{xy}, s_y^2) = \left( -\frac{1}{2} s_{xy} - \frac{1}{2} s_{xy} \right).$$
and

$$h(\sigma_y^2, \sigma_{xy}, \sigma_x^2) = \left( -\frac{1}{2} \frac{\sigma_{xy}}{\sigma_y^2} - \frac{1}{\sigma_x \sigma_y} - \frac{1}{2} \frac{\sigma_{xy}}{\sigma_y^3} \right).$$

We obtain $\sqrt{n}\left(h(s_x^2, s_{xy}, s_y^2) - h(\sigma_x^2, \sigma_{xy}, \sigma_y^2)\right) \xrightarrow{a.s.} N(0, \gamma^2)$, where

$$\gamma^2 = h(\sigma_x^2, \sigma_{xy}, \sigma_y^2)Ch(\sigma_x^2, \sigma_{xy}, \sigma_y^2)^T$$

$$= \left( -\frac{\sigma_{xy}}{2\sigma_x^3 \sigma_y} - \frac{1}{\sigma_x \sigma_y} - \frac{\sigma_{xy}}{2\sigma_x^2 \sigma_y^2} \right) \begin{pmatrix} b_{33} & b_{35} & b_{34} \\ b_{53} & b_{55} & b_{54} \\ b_{43} & b_{45} & b_{44} \end{pmatrix} \left( -\frac{\sigma_{xy}}{2\sigma_x^3 \sigma_y} - \frac{1}{\sigma_x \sigma_y} - \frac{\sigma_{xy}}{2\sigma_x^2 \sigma_y^2} \right)^T$$

$$= \left( -\frac{\sigma_{xy}}{2\sigma_x^3 \sigma_y} - \frac{1}{\sigma_x \sigma_y} - \frac{\sigma_{xy}}{2\sigma_x^2 \sigma_y^2} \right) \frac{\sigma_{xy}}{2\sigma_x^3 \sigma_y}$$

$$+ \left( -\frac{\sigma_{xy}}{2\sigma_x^3 \sigma_y} - \frac{1}{\sigma_x \sigma_y} - \frac{\sigma_{xy}}{2\sigma_x^2 \sigma_y^2} \right) \frac{1}{\sigma_x \sigma_y}$$

$$+ \left( -\frac{\sigma_{xy}}{2\sigma_x^3 \sigma_y} - \frac{1}{\sigma_x \sigma_y} - \frac{\sigma_{xy}}{2\sigma_x^2 \sigma_y^2} \right) \frac{\sigma_{xy}}{2\sigma_x^3 \sigma_y}$$

$$= \frac{1}{4} \rho^2 \left[ \frac{b_{33}}{\sigma_x^4} + \frac{2b_{34}}{\sigma_x^2 \sigma_y^2} + \frac{b_{44}}{\sigma_y^4} \right] - \rho \left[ \frac{b_{35}}{\sigma_x^2 \sigma_y^3} + \frac{b_{45}}{\sigma_x \sigma_y^2} \right] + \frac{b_{55}}{\sigma_x^2 \sigma_y^2}$$

since

$$h(y_1, y_2, y_3) = \left( -\frac{y_2}{2\sqrt{y_1 y_3}}, \frac{1}{\sqrt{y_1 y_3}}, -\frac{y_2}{2\sqrt{y_1 y_3}} \right).$$
3.4 Connecting the ASCLT and the Correlation Coefficient

In the remaining parts of the chapter, we will discuss the ASCLT and how we can apply it to the estimation of the correlation coefficient. Our proposal in this thesis is to connect these two ideas and converge to a confidence interval method for the population correlation coefficient. In fact, our main goal towards this ASCLT-based theory of confidence interval estimation will be the estimation of the quantiles of the distribution of the correlation coefficient statistic. In this section, the results of a new version of the ASCLT will be stated that will include the correlation coefficient.

Remark 3.1

1) The a.s. weak convergence of $\rho$ means that

$$\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} I\left(\sqrt{n}(r_{k} - \rho) \leq t\right) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi\gamma^2}} \exp\left[-\frac{u^2}{2\gamma^2}\right] du$$

This convergence is not changed when replacing the norming sequence $\frac{1}{\log n}$ by the sequence $a_n$ with $\lim a_n \to 1$. A particular sequence $a_n$ is of the form $\frac{1}{a_n} = \sum_{k=1}^{n} \frac{1}{k}$. This sequence turns the left hand side in the above equation into a distribution function.

2) Another modification of the above result arises when the weights $1/k$ are changed so that $w_k - 1/k \to 0$ as $k \to \infty$. This follows because

$$\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \left| w_k - \frac{1}{k} \right|$$

$$= \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{\frac{m(n)}{1}} \left| w_k + \frac{1}{k} \right| + \lim_{n \to \infty} \frac{1}{\log n} \sum_{k=m(n)+1}^{n} \left| w_k - \frac{1}{k} \right| = 0,$$
by choosing $m(n)$ increasing so slowly that the first limit is zero. In particular, we can choose $w_k = 0$ for $k \leq K$ and equals $1/k$ otherwise. This also shows that one may take different weights for each $n$, i.e. we have

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} w_k^n I\{\sqrt{n}(r_k - \rho) \leq t\} = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi\gamma^2}} \exp\left[-\frac{u^2}{2\gamma^2}\right] du.$$  

If $K(n)$ is increasing sufficiently slow, then we may put $w_k^n = 0$ for $k \leq K(n)$ and

$$w_k^n = \frac{1}{k}$$

otherwise, replacing as well $\log n$ by $\sum_{k=K(n)+1}^{n} \frac{1}{k}$.

3) Since every measurable set of full measure under the distribution of an i.i.d. sequence contains a permutation invariant subset for every permutation $\tau$ of finitely many coordinates, we obtain

$$\lim_{n \to \infty} \frac{1}{m(n) \log n} \sum_{\tau \in M(n)} \sum_{k=1}^{n} \frac{1}{k} I\{\sqrt{n}(r_k(\tau(\omega)) - \rho) \leq t\} = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi\gamma^2}} \exp\left[-\frac{u^2}{2\gamma^2}\right] du$$

where $M(n)$ is a family of maps on the probability space which permutes the values of $X_1, \ldots, X_n$ and which has cardinality $m(n)$. In practice one chooses $M(n)$ at random.

Theorem 3.1 is the analogue of the classical asymptotic distribution result for the correlation coefficient. In the non a.s. case it does not provide best confidence intervals because $\gamma^2$ has to be estimated. In the almost sure case this is not necessary; however, as in the classical case we need to investigate the effect of a stabilizing transformation. The corresponding a.s. results (see Ferguson, 2002) are the following:
Corollary 3.1 If the distribution of \((X_t, Y_t)\) is bivariate normal, then \(\gamma^2 = (1-\rho^2)^2\) such that

\[
\sqrt{n}(r_n - \rho) \xrightarrow{a.s.} N(0,(1-\rho^2)^2).
\]

**Proof**: Although the result is well known, we give a proof for completeness. For normal distributions we have

\[
\begin{align*}
 b_{11} &= \sigma_x^2 \\
 b_{22} &= \sigma_y^2 \\
 b_{33} &= \text{Var}((X_1 - \mu_x )^2 - \sigma_x^2 ) = 3\sigma_x^4 - \sigma_x^2 = 2\sigma_x^4 \\
 b_{44} &= \text{Var}((Y_1 - \mu_y )^2 - \sigma_y^2 ) = 2\sigma_y^4 \\
 b_{55} &= \text{Var}((X_1 - \mu_x )(Y_1 - \mu_y ) - \sigma_{xy}) = \sigma_x^2 \sigma_y^2 + 2\sigma_{xy}^2 - \sigma_{xy}^2 = \sigma_x^2 \sigma_y^2 + \sigma_{xy}^2 \\
 b_{12} &= b_{21} = \sigma_{xy} \\
 b_{34} &= b_{43} = \text{Cov}((X_1 - \mu_x )^2,(Y_1 - \mu_y )^2) = \sigma_x^2 \sigma_y^2 + 2\sigma_{xy}^2 - \sigma_x^2 \sigma_y^2 = 2\sigma_{xy}^2 \\
 b_{35} &= b_{53} = \text{Cov}((X_1 - \mu_x )^2,(X_1 - \mu_x )(Y_1 - \mu_y )) = 3\sigma_x^2 \sigma_{xy} - \sigma_x^2 \sigma_{xy} = 2\sigma_x^2 \sigma_{xy} \\
 b_{45} &= b_{45} = \text{Cov}((Y_1 - \mu_y )^2,(X_1 - \mu_x )(Y_1 - \mu_y )) = 3\sigma_{xy}^2 - \sigma_{xy}^2 \sigma_y^2 = 2\sigma_{xy} \sigma_y^2;
\end{align*}
\]

hence

\[
\begin{align*}
\frac{1}{4} \rho^2 \left[ \frac{b_{33}}{\sigma_x^4} + \frac{2b_{34}}{\sigma_x^2 \sigma_y^2} + \frac{b_{44}}{\sigma_y^4} \right] - \rho \left[ \frac{b_{35}}{\sigma_x^2 \sigma_y^2} + \frac{b_{45}}{\sigma_x^3 \sigma_y^2} \right] + \frac{b_{55}}{\sigma_x^2 \sigma_y^2} \\
= \frac{1}{4} \rho^2 \left[ \frac{2\sigma_x^4}{\sigma_x^4} + \frac{4\sigma_{xy}^2}{\sigma_x^2 \sigma_y^2} + \frac{2\sigma_y^4}{\sigma_y^4} \right] - \rho \left[ \frac{2\sigma_x^2 \sigma_{xy}}{\sigma_x^3 \sigma_y^2} + \frac{2\sigma_{xy} \sigma_y^2}{\sigma_x^3 \sigma_y^2} \right] + \frac{\sigma_x^2 \sigma_y^2 + \sigma_{xy}^2}{\sigma_x^2 \sigma_y^2} \\
= \rho^2 + \rho^4 - 4\rho^2 + 1 + \rho^2 = (1 - \rho^2)^2.
\end{align*}
\]

The following theorem can be found in Ferguson (2002) and van der Vaart (1998).
**Theorem 3.2.** Let \((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\) be a random sample of \(n\) pairs as having been drawn from a bivariate population with finite fourth moments, \(EX^4\) and \(EY^4\), and consider the situation in Theorem 3.1, then:

\[
\frac{\sqrt{n}}{2} \left( \log \frac{1 + r_n}{1 - r_n} - \log \frac{1 + \rho}{1 - \rho} \right) \xrightarrow{a.s.} N(0, \tau^2),
\]

where

\[
\tau^2 = \frac{\rho^2}{4(1 - \rho^2)^2} \left\{ \frac{b_{33} + 2b_{34} + b_{44}}{\sigma_x^4} + \frac{2}{\sigma_x^2 \sigma_y^2} + \frac{b_{44}}{\sigma_y^4} \right\} \\
- \frac{\rho}{(1 - \rho^2)^2} \left[ \frac{b_{35}}{\sigma_x^3 \sigma_y} + \frac{b_{45}}{\sigma_x \sigma_y^3} \right] + \left[ \frac{b_{55}}{(1 - \rho^2)^2 \sigma_x^2 \sigma_y^2} \right].
\]

In the case that \((X_i, Y_i)\) is bivariate normal, \(\tau^2 = 1\).

**Proof:** Let \(g(\rho) = \frac{1}{2} \log \frac{1 + \rho}{1 - \rho}\) for \(-1 < \rho < 1\), so \(g\) is differentiable function

\(g : (-1, 1) \to \mathbb{R}\).

We have \(\dot{g}(\rho) = \frac{1}{l - \rho^2}\) and hence by Cramér’s theorem and Theorem 3.1, part (2) we obtain the following:

\[
\sqrt{n}(g(r_n) - g(\rho)) \xrightarrow{a.s.} N(0, \tau^2),
\]

where
\[
\tau^2 = g(\rho)^2 = \frac{\rho^2}{4(1-\rho^2)^2} \left[ \frac{b_{33}}{\sigma_x^4} + 2 \frac{b_{34}}{\sigma_x^2 \sigma_y^2} + \frac{b_{44}}{\sigma_y^4} \right] \\
- \frac{\rho}{(1-\rho^2)^2} \left[ \frac{b_{35}}{\sigma_x^3 \sigma_y} + \frac{b_{45}}{\sigma_x \sigma_y^3} \right] + \frac{b_{55}}{\sigma_x^2 \sigma_y^2} 
\]

**Remark 3.2** The last theorem states that no matter what the asymptotic variance will be, we can estimate the quantiles of the asymptotic statistics from the data. With that in mind, consider the following a.s. result:

\[
\lim_{n \to \infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \left\{ \sqrt{n} \left( \frac{1}{2} \left( \log \frac{1+r_k}{1-r_k} - \log \frac{1+\rho}{1-\rho} \right) \right) \right\} = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi} \tau^2} \exp \left[ -\frac{u^2}{2\tau^2} \right] du
\]

The left hand side for fixed \( n \) serves as an approximation.

### 3.5 Law of Iterated Logarithm for the Correlation Coefficient

The classical law of the iterated logarithm (see (van der Vaart, 1998; Dudley, 1989)) for independent, identically distributed random variables \( Z_n \) with finite variance \( \sigma^2 \) states that almost surely

\[
\limsup_{n \to \infty} \frac{Z_1 + \cdots + Z_n - nEZ_1}{\sqrt{2\sigma^2 n \log \log n}} = 1.
\]

Likewise, the \( \liminf \) is equal to \(-1\). We shall use this result below. In the sequel we will need the law of iterated logarithm for the statistic used for a.s. confidence intervals. In a previous section we dealt with the correlation coefficient statistics

\[
r = r_n = \frac{s_{xy}}{s_x s_y}
\]

and the statistic
\[ \frac{1}{2} \log \frac{1+r_n}{1-r_n}. \]

**Theorem 3.3.** Let \((X_n, Y_n)\) be an i.i.d. sequence of bivariate random vectors with finite fourth moment, \(\mu_x = E(X_1), \mu_y = E(Y_1), \sigma_x^2 = E((X_1 - \mu_x)^2), \sigma_y^2 = E((Y_1 - \mu_y)^2)\), \(\sigma_{xy} = \text{Cov}(X_1, Y_1)\) and \(b_{55}\) as defined in section 3.3. Then almost surely

\[ \limsup_{n \to \infty} \frac{s_{xy} - n\sigma_{xy}}{\sqrt{2n \log \log n}} \leq \sqrt{b_{55} + 2|\mu_x|\sigma_y + 2|\mu_y|\sigma_x}. \]

**Proof:** This follows from the classical law of iterated logarithm for independent identically distributed random variables with finite second moment, together with the definition of the constants \(b_{ij}\) as follows: Let \(m_x = \sum_{k=1}^n X_k\) and \(m_y = \sum_{k=1}^n Y_k\) and recall that:

\[ s_{xy} = \sum_{k=1}^n \left( X_k - \frac{1}{n} m_x \right) \left( Y_k - \frac{1}{n} m_y \right). \]

By the law of iterated logarithm for i.i.d. sequences, applied to \(Z_n = (X_n - \mu_x)(Y_n - \mu_y)\), we obtain

\[ \limsup_{n \to \infty} \frac{\sum_{k=1}^n (X_k - \mu_x)(Y_k - \mu_y) - n\sigma_{xy}}{\sqrt{2nb_{55} \log \log n}} = 1 \]

Moreover, by the same theorem,

\[ \limsup_{n \to \infty} \frac{1}{\sqrt{2n\sigma_x^2 \log \log n}} \sum_{k=1}^n (X_k - \mu_x) = \limsup_{n \to \infty} \frac{m_x - n\mu_x}{\sqrt{2n\sigma_x^2 \log \log n}} = 1 \]
And likewise for the other marginal process $Y_k$. Therefore,

\[
\lim \sum_{n \to \infty} \frac{s_{xy} - n\sigma_{xy}}{\sqrt{2n \log \log n}}
\]

\[
= \lim \sum_{n \to \infty} \frac{\sum_{k=1}^{n} \left( X_k - \frac{1}{n} m_x \right) \left( Y_k - \frac{1}{n} m_y \right) - n\sigma_{xy}}{\sqrt{2n \log \log n}}
\]

\[
= \lim \sum_{n \to \infty} \frac{\sum_{k=1}^{n} \left( X_k - \mu_x \right) \left( Y_k - \mu_y \right) - n\sigma_{xy}}{\sqrt{2n \log \log n}}
\]

\[
+ \frac{\sum_{k=1}^{n} \left( X_k \left( \mu_y - \frac{1}{n} m_y \right) \right) + Y_k \left( \mu_x - \frac{1}{n} m_x \right) + \frac{1}{n^2} m_x m_y - \mu_x \mu_y \right)}{\sqrt{2n \log \log n}}
\]

\[
= \sqrt{b_{55}} + \lim \sum_{n \to \infty} \frac{n}{\sqrt{n \log \log n}}
\]

\[
\leq \sqrt{b_{55}} + 2\mu_x |\sigma_y| + 2\mu_y |\sigma_x|,
\]

where we also used the strong law of large numbers such that almost surely

\[
\lim_{n \to \infty} \frac{1}{n} m_x = \mu_x \quad \text{and} \quad \lim_{n \to \infty} \frac{1}{n} m_y = \mu_y.
\]

The lower bound is shown in similar fashion.

**Corollary 3.2** There exists a $\Gamma > 0$ such that almost surely

\[
\lim sup_{n \to \infty} \frac{\sqrt{n}}{\sqrt{2 \log \log n}} |\rho_n - \rho| \leq \Gamma.
\]

**Proof:** First note that,

\[
s_x^2 = \sum_{i=1}^{n} X_i^2 - n \left( \frac{1}{n} m_x \right)^2 = \sum_{i=1}^{n} (X_i - EX_1)^2 + n(EX_1)^2 - n \left( \frac{1}{n} m_x \right)^2.
\]

By the strong law of large numbers we obtain the following almost sure results,
\[
\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} (X_i - EX_1)^2 = \sigma_x^2
\]
and
\[
\lim_{n \to \infty} \left( (EX_1)^2 - \left( \frac{1}{n} \sum_{k=1}^{n} X_k \right)^2 \right) = 0.
\]
Therefore \( \lim_{n \to \infty} \frac{1}{n} s_x^2 = \sigma_x^2 \) almost surely. Now by replacing \( X \) and \( Y \) variables we also have
\[
\lim_{n \to \infty} \frac{1}{n} s_y^2 = \sigma_y^2;
\]
hence,
\[
\lim_{n \to \infty} \frac{1}{n} s_x s_y = \sigma_x \sigma_y
\]
almost surely. By Theorem 3.3 we have that,
\[
\limsup_{n \to \infty} \frac{1}{\sqrt{2 n \log \log n}} \left| s_{xy} - n \sigma_{xy} \right| \leq \sqrt{\alpha} + 2 \left| \mu_x \right| \left| \sigma_y \right| + 2 \left| \mu_y \right| \frac{\sigma_y}{\sigma_x}
\]
almost surely. Next observe that
\[
\limsup_{n \to \infty} \frac{1}{\sqrt{2 \log \log n}} |r_n - \rho|
\]
\[
= \limsup_{n \to \infty} \frac{1}{\sqrt{2 n \log \log n}} \left| \frac{s_{xy}}{\sigma_x \sigma_y} - \frac{n \sigma_{xy}}{\sigma_x \sigma_y} \right|
\]
\[
\leq \limsup_{n \to \infty} \frac{1}{\sqrt{2 n \log \log n}} \left( \frac{s_{xy}}{\sigma_x \sigma_y} - \frac{n \sigma_{xy}}{\sigma_x \sigma_y} \right) + \frac{\sigma_{xy}}{\sqrt{2 \log \log n}} \frac{s_x s_y - n \sigma_x \sigma_y}{l n \frac{s_x s_y \sigma_x \sigma_y}{l n}}.
\]
Almost surely, for \( \eta > 0 \) there exists \( n_0 \in N \) such that for \( n \geq n_0 \)
\[
n(2 \sigma_x \sigma_y - \eta) \leq s_x s_y + n \sigma_x \sigma_y,
\]
thus
\[ n(2\sigma_x\sigma_y - \eta)(s_x s_y - n \sigma_x \sigma_y) \leq (s_x s_y + n \sigma_x \sigma_y)(s_x s_y - n \sigma_x \sigma_y) = s_x^2 s_y^2 - n^2 \sigma_x^2 \sigma_y^2. \]

Since \( s_x^2 s_y^2 = \sum_{k,l=1}^{n} X_k^2 Y_l^2 \) is a non-degenerate von Mises statistic with \( \frac{1}{n} s_x^2 s_y^2 = \sigma_x^2 \sigma_y^2 \),
it follows from the bounded law of iterated logarithm for von Mises statistics (see Dehling, Denker, Philipp (1984)), that there exists a constant \( C \) such that,
\[
\limsup_{n \to \infty} \frac{1}{n^{3/2} \sqrt{\log \log n}} \left| s_x^2 s_y^2 - n^2 \sigma_x^2 \sigma_y^2 \right| \leq C \leq \infty.
\]

It follows for sufficiently large \( n \),
\[
\left(2\sigma_x \sigma_y - \eta\right) \frac{s_x s_y - n \sigma_x \sigma_y}{\sqrt{n \log \log n}} \leq \frac{|s_x s_y - n \sigma_x \sigma_y|}{n^{3/2} \sqrt{\log \log n}} \leq C + \eta.
\]

Setting
\[
\Gamma = \frac{\sqrt{b_{55}}}{\sigma_x \sigma_y} + \frac{2|\mu_x| \, 2|\mu_y|}{\sigma_x^3} + \frac{2\sigma_{xy}}{2\sigma_x \sigma_y \sigma_y \sigma_y} C
\]
completes the proof.

**Theorem 3.4** There exists a constant \( \Lambda \) such that
\[
\limsup_{n \to \infty} \frac{\sqrt{n}}{\sqrt{2 \log \log n}} \left[ \log \frac{1+r_n}{1-r_n} - \log \frac{1+\rho}{1-\rho} \right] \leq \Lambda < \infty.
\]

**Proof:** First use the Taylor series expansion of \( \log \frac{1+t}{1-t} \) around \( t = \rho \) up to the first derivative to obtain the following:
\[
\log \frac{1+t}{1-t} = \log \frac{1+\rho}{1-\rho} + (t-\rho) \frac{d}{dt} \log \frac{1+t}{1-t} \bigg|_{t=\xi}
\]
\[
= \log \frac{1+\rho}{1-\rho} + (t-\rho) \frac{2}{1-\xi^2}.
\]
where $\xi = \xi(t)$ is a point in the interval from $t$ to $\rho$. Note that we may assume that $\rho \neq \pm 1$ because otherwise the random variables $X_1$ and $Y_1$ are collinear. Taking $t = r_n$ it follows that $1 - \xi(r_n)^2 \to 1 - \rho^2 \neq 0$ and hence

$$\limsup_{n \to \infty} \frac{\sqrt{n}}{\sqrt{2 \log \log n}} \left| \log \frac{1 + r_n}{1 - r_n} - \log \frac{1 + \rho}{1 - \rho} \right|$$

$$= \limsup_{n \to \infty} \frac{\sqrt{n}}{\sqrt{\log \log n}} \frac{1}{1 - \xi(r_n)^2} |r_n - \rho|$$

$$\leq A := \sqrt{2} \Gamma \left(1 - \rho^2 \right)$$

using Corollary 3.2.
Chapter 4

ASCLT-based Confidence Interval for the Correlation Coefficient

In this chapter, we will now define the empirical and true distribution function using the notation introduced in chapter 2. This notation will be applied to the development of the ASCLT-based theory of confidence interval estimation for the population correlation coefficient.

4.1 Defined Cumulative Distribution Functions

In this section, we will observe some of the previously mentioned distribution functions and develop additional ones. After this has been completed we can continue developing our confidence intervals.

Recall the ASCLT where we defined

\[ G_N(t) = \frac{1}{\log N} \sum_{n=1}^{N} \left\{ \frac{1}{n} \right\} \{ b_n(T_n - a_n) \leq t \}. \]

Due to the slow nature of convergence of the distribution function, the following was introduced where we denoted \( \sum_{n=1}^{N} \frac{1}{n} \) by \( C_N \):

\[ \tilde{G}_N(t) = \frac{1}{C_N} \sum_{n=1}^{N} \left\{ \frac{1}{n} \right\} \{ b_n(T_n - a_n) \leq t \}. \]

We introduced the following function in Remark 3.1. I will define this function \( H_N(t) \):
\[ H_N(t) = \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \{ \sqrt{n}(r_n - \rho) \leq t \}, \]

where this function had an a.s. weak convergence to the \( N(0, \gamma^2) \). However, according to Corollary 3.1, if the distribution of \((X_i, Y_i)\) is bivariate normal, then this function had an a.s. weak convergence to the \( N(0, (1-\rho^2)^2) \).

We can re-write the above function as:

\[ \tilde{H}_N(t) = \frac{1}{C_N} \sum_{n=1}^{N} \frac{1}{n} \{ \sqrt{n}(r_n - \rho) \leq t \}, \]

We introduced the following function in Remark 3.2. I will define this function \( J_N(t) \):

\[ J_N(t) = \frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \sqrt{n} \left( \frac{1}{2} \left( \frac{\log \left( \frac{1+r_n}{1-r_n} \right) - \log \left( \frac{1+r_n}{1-\rho} \right) }{1-\rho} \right) \right) \leq t \}, \]

where this function had an a.s. weak convergence to the \( N(0, \tau^2) \). However, according to Theorem 3.2, if the distribution of \((X_i, Y_i)\) is bivariate normal, then this function had an a.s. weak convergence to the \( N(0, 1) \).

We can re-write the above function as:

\[ \tilde{J}_N(t) = \frac{1}{C_N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \sqrt{n} \left( \frac{1}{2} \left( \frac{\log \left( \frac{1+r_n}{1-r_n} \right) - \log \left( \frac{1+r_n}{1-\rho} \right) }{1-\rho} \right) \right) \leq t \}, \]

4.2 Defined Inverse Cumulative Distribution Functions

In the previous section of this chapter, we have defined empirical distribution functions
In this section, we will define their corresponding inverse functions which are based on the definition from Thangavelu (2005). Once the inverse function is known, we can show results of the quantiles of these distributions. However, before we can define the quantiles, we must first define the inverse function of our distribution function.

**Definition 4.1.** Let \( H_N(t), \tilde{H}_N(t), J_N(t), \text{ and } \tilde{J}_N(t) \) be empirical distribution functions that converge to the true distribution \( G_P(t) \). For a fixed \( N \in \mathbb{N} \), let the inverse of our distribution functions be denoted by the functions \( H_N^{-1}(\alpha), \tilde{H}_N^{-1}(\alpha), J_N^{-1}(\alpha), \) and \( \tilde{J}_N^{-1}(\alpha) \), where \( \alpha \) is between 0 and 1, inclusive. The inverse functions are denoted as follows:

\[
G_P^{-1}(\alpha) = \begin{cases} \sup \{ t \mid G_P(t) = 0 \} & \text{for } \alpha = 0 \\ \sup \{ t \mid G_P(t) < \alpha \} & \text{for } 0 < \alpha < 1 \\ \inf \{ t \mid G_P(t) = 1 \} & \text{for } \alpha = 1 \end{cases}
\]

\[
H_N^{-1}(\alpha) = \begin{cases} \sup \{ t \mid H_N(t) = 0 \} & \text{for } \alpha = 0 \\ \sup \{ t \mid H_N(t) < \alpha \} & \text{for } 0 < \alpha < 1 \\ \inf \{ t \mid H_N(t) = 1 \} & \text{for } \alpha = 1 \end{cases}
\]

\[
\tilde{H}_N^{-1}(\alpha) = \begin{cases} \sup \{ t \mid \tilde{H}_N(t) = 0 \} & \text{for } \alpha = 0 \\ \sup \{ t \mid \tilde{H}_N(t) < \alpha \} & \text{for } 0 < \alpha < 1 \\ \inf \{ t \mid \tilde{H}_N(t) = 1 \} & \text{for } \alpha = 1 \end{cases}
\]

\[
J_N^{-1}(\alpha) = \begin{cases} \sup \{ t \mid J_N(t) = 0 \} & \text{for } \alpha = 0 \\ \sup \{ t \mid J_N(t) < \alpha \} & \text{for } 0 < \alpha < 1 \\ \inf \{ t \mid J_N(t) = 1 \} & \text{for } \alpha = 1 \end{cases}
\]
and

\[
\tilde{J}_N^{-1}(\alpha) = \begin{cases} 
\sup\{ t \mid \tilde{J}_N(t) = 0 \} & \text{for } \alpha = 0 \\
\sup\{ t \mid \tilde{J}_N(t) < \alpha \} & \text{for } 0 < \alpha < 1 \\
\inf\{ t \mid \tilde{J}_N(t) = 1 \} & \text{for } \alpha = 1 
\end{cases}
\]

### 4.3 Empirical $\alpha$-Quantiles

In the previous sections, we have defined our empirical distribution functions and their inverse functions. Now that the inverse functions are known, we can show results of the empirical quantiles of these distributions. In this section, we base our definitions from Thangavelu (2005).

**Definition 4.2.** Let $H_N(t)$, $\tilde{H}_N(t)$, $J_N(t)$, and $\tilde{J}_N(t)$ be empirical distribution functions that converge to the true distribution $P$. For a fixed $N \in \mathbb{N}$, let the inverse of our distribution functions be denoted by the functions $H_N^{-1}(\alpha)$, $\tilde{H}_N^{-1}(\alpha)$, $J_N^{-1}(\alpha)$, and $\tilde{J}_N^{-1}(\alpha)$. For $0 \leq \alpha \leq 1$, the empirical $\alpha$-quantiles for our statistics, where $n \leq N \in \mathbb{N}$, are defined as follows:

\[
t_\alpha^{(N)} = H_N^{-1}(\alpha) \\
\bar{t}_\alpha^{(N)} = \tilde{H}_N^{-1}(\alpha)
\]

or

\[
t_\alpha^{(N)} = J_N^{-1}(\alpha) \\
\bar{t}_\alpha^{(N)} = \tilde{J}_N^{-1}(\alpha)
\]
4.4 Confidence Intervals for $\rho$

Recall in the beginning of this thesis we discussed how the ASCLT could be used to develop an estimation technique for the correlation coefficient. However, our main goal was to work towards an ASCLT-based theory of confidence interval estimation for our parameter of interest. Up to this point, we have defined empirical distribution functions $H_N(t)$, $\tilde{H}_N(t)$, $J_N(t)$, and $\tilde{J}_N(t)$, the inverse functions $H^{-1}_N(\alpha)$, $\tilde{H}^{-1}_N(\alpha)$, $J^{-1}_N(\alpha)$, and $\tilde{J}^{-1}_N(\alpha)$, and the estimated $\alpha$-quantiles for the empirical distribution functions. In this section, a new version of the confidence interval for the population correlation coefficient will be presented where the confidence level $= 1-2\alpha$.

Definition 4.3. Let $(x_1, y_1), (x_2, y_2), \ldots$ be a sample from a bivariate distribution with finite fourth moments, $EX^4$ and $EY^4$. For $n \geq 1$ let the statistic $r_n$ be a sequence of real valued correlation coefficient statistics defined on the same measurable space $(\Omega, \beta)$ and $\mathcal{P}$ be a family of probabilities on $\beta$. Recall that the distribution function $\tilde{H}_N(t)$ converges almost surely for any $t$ to a $N(0, \gamma^2)$ distribution. The following is the ASCLT-derived confidence interval for $\rho$ at a significance level of $2\alpha$:

$$I^{(N)}_\alpha = \left[ \hat{\rho} + \frac{\tilde{\gamma}^{(N)}_{1-\alpha}}{\sqrt{N}}, \hat{\rho} + \frac{\tilde{\gamma}^{(N)}_{\alpha}}{\sqrt{N}} \right]. \quad (4.5)$$

Here we will estimate the population correlation coefficient with $\hat{\rho}$. Notice that the above asymptotic confidence interval does not include the variance. The material presented in
this chapter show procedures developed from the ASCLT to derive confidence intervals that are completely independent from the variance estimation.

4.5 Confidence Intervals for \( \rho \) (Variance Stabilizing Technique)

Fisher developed a variance stabilizing technique transformation of \( r \) that tends to become quickly normal as \( n \) increases. In this case \( n \geq 1 \) where \( n \leq N \in \mathbb{N} \). In this variance stabilizing technique, Fisher used it to construct the confidence interval for the population correlation. Recall for normal populations,

\[
\sqrt{n}(r - \rho) \xrightarrow{L} N(0,(1 - \rho^2)^2).
\]

The variance stabilizing technique seeks a transformation, \( g(r) \) and \( g(\rho) \), such that

\[
\sqrt{n}(g(r) - g(\rho)) \xrightarrow{L} N(0,1).
\]

This transformation relies on Cramer’s Theorem where

\[
\sqrt{n}(g(r) - g(\rho)) \xrightarrow{L} N(0,\hat{g}(\rho)^2(1 - \rho^2)^2).
\]

To find a function that satisfies this equation, we must solve the following differential equation:

\[
\hat{g}(\rho)^2(1 - \rho^2)^2 = 1;
\]

\[
\hat{g}(\rho) = \frac{1}{(1 - \rho^2)}.
\]

The following solution is called a Fisher’s transformation:

\[
g(\rho) = \int \frac{1}{(1 - \rho^2)^2} \, d\rho = \int \left[ \frac{1/2}{(1 - \rho)} + \frac{1/2}{(1 + \rho)} \right] \, d\rho = \frac{1}{2} \left[ -\ln(1 - \rho) + \ln(1 + \rho) \right] + \frac{1}{2} \left[ \ln(1 + \rho) \right] = \frac{1}{2} \ln \left[ \frac{1 + \rho}{1 - \rho} \right]
\]

and

\[
g(r) = \frac{1}{2} \ln \left[ \frac{1 + r}{1 - r} \right].
\]

**Definition 4.4.** Let \( J_N(t) \) or \( \tilde{J}_N(t) \) be empirical distribution functions that converge to
the true distribution $G_P(t)$. For a fixed $N \in \mathbb{N}$, let the inverse of our distribution functions be denoted by the functions $J_N^{-1}(t)$ or $\bar{J}_N^{-1}(t)$. For $0 \leq \alpha \leq 1$, the empirical $\alpha$-quantiles of the statistic $z_n = \log \frac{1 + r_n}{1 - r_n}$, where $n \leq N \in \mathbb{N}$, were defined in (4.3) and (4.4).

Also, define $z_N = \log \frac{1 + \hat{\rho}}{1 - \hat{\rho}}$ where the estimate of the population correlation coefficient will be denoted by $\hat{\rho}$. Although the procedure to develop the confidence interval is well known using the variance stabilizing technique, for completeness we will develop the lower and upper bound for $\rho$.

**Lower Bound CI for $\rho$:**

\[
\begin{align*}
\frac{z_N + \bar{t}_{1-\alpha}(N)}{\sqrt{N}} &\leq \frac{1}{2} \ln \left( \frac{1 + \rho}{1 - \rho} \right) \\
\Rightarrow \exp \left[ 2 \left( z_N + \bar{t}_{1-\alpha}(N) \right) \right] &\leq \frac{1 + \rho}{1 - \rho} \\
\Rightarrow \exp \left[ 2 \left( z_N + \bar{t}_{1-\alpha}(N) \right) \right] - \exp \left[ 2 \left( z_N + \bar{t}_{1-\alpha}(N) \right) \right] \rho &\leq 1 + \rho \\
\Rightarrow -\exp \left[ 2 \left( z_N + \bar{t}_{1-\alpha}(N) \right) \right] \rho - \rho &\leq 1 - \exp \left[ 2 \left( z_N + \bar{t}_{1-\alpha}(N) \right) \right] \\
\Rightarrow \rho \left[ \exp \left[ 2 \left( z_N + \bar{t}_{1-\alpha}(N) \right) \right] + 1 \right] &\geq \exp \left[ 2 \left( z_N + \bar{t}_{1-\alpha}(N) \right) \right] - 1 \\
\Rightarrow \rho &\geq \frac{\exp \left[ 2 \left( z_N + \bar{t}_{1-\alpha}(N) \right) \right] - 1}{\exp \left[ 2 \left( z_N + \bar{t}_{1-\alpha}(N) \right) \right] + 1}.
\end{align*}
\]
Upper Bound CI for $\rho$:

$$z_N + \frac{\tilde{t}^{(N)}_{\alpha}}{\sqrt{N}} \geq \frac{1}{2} \ln \left( \frac{1 + \rho}{1 - \rho} \right)$$

$$\Rightarrow \exp \left[ 2 \left( z_N + \frac{\tilde{t}^{(N)}_{\alpha}}{\sqrt{N}} \right) \right] \geq \frac{1 + \rho}{1 - \rho}$$

$$\Rightarrow \exp \left[ 2 \left( z_N + \frac{\tilde{t}^{(N)}_{\alpha}}{\sqrt{N}} \right) \right] \geq 1 + \rho$$

$$\Rightarrow -\exp \left[ 2 \left( z_N + \frac{\tilde{t}^{(N)}_{\alpha}}{\sqrt{N}} \right) \right] \rho \geq 1 - \exp \left[ 2 \left( z_N + \frac{\tilde{t}^{(N)}_{\alpha}}{\sqrt{N}} \right) \right] \rho$$

$$\Rightarrow -\rho \left[ \exp \left[ 2 \left( z_N + \frac{\tilde{t}^{(N)}_{\alpha}}{\sqrt{N}} \right) \right] - 1 \right] \leq \exp \left[ 2 \left( z_N + \frac{\tilde{t}^{(N)}_{\alpha}}{\sqrt{N}} \right) \right] - 1$$

$$\Rightarrow \rho \leq \frac{\exp \left[ 2 \left( z_N + \frac{\tilde{t}^{(N)}_{\alpha}}{\sqrt{N}} \right) \right] - 1}{\exp \left[ 2 \left( z_N + \frac{\tilde{t}^{(N)}_{\alpha}}{\sqrt{N}} \right) \right] + 1}.$$ 

Therefore the following is the ASCLT-derived confidence interval for $\rho$ using the variance stabilizing technique at a significance level of $2\alpha$:

$$I^{(N)}_{\alpha} = \left[ \frac{\exp \left[ 2 \left( z_N + \frac{\tilde{t}^{(N)}_{1-\alpha}}{\sqrt{N}} \right) \right] - 1}{\exp \left[ 2 \left( z_N + \frac{\tilde{t}^{(N)}_{\alpha}}{\sqrt{N}} \right) \right] + 1}, \frac{\exp \left[ 2 \left( z_N + \frac{\tilde{t}^{(N)}_{\alpha}}{\sqrt{N}} \right) \right] - 1}{\exp \left[ 2 \left( z_N + \frac{\tilde{t}^{(N)}_{1-\alpha}}{\sqrt{N}} \right) \right] + 1} \right]. \quad (4.6)$$
Chapter 5

Numerical Applications

5.1 Introduction

A considerable amount of time was used in the previous chapters to develop the theoretical results. In this chapter, the main objective will be to validate the theoretical results presented earlier by numerical simulations. We are also interested in how the methods presented earlier work for finite samples. It has been emphasized earlier that all the theoretical results that utilize the almost sure central limit theorem are asymptotic in nature. Also, these asymptotic results converge at a very slow rate. These issues will be addressed in this chapter. The following paragraphs will outline what will be presented in this chapter.

Consider the following ASCLT distribution function that was presented in Chapter 2:

\[
\frac{1}{\log N} \sum_{n=1}^{N} \mathbb{I} \left\{ \frac{X_1 + X_2 + \cdots + X_n - n \mu}{\sqrt{n}} \leq t \right\} \overset{a.s.}{\to} \Phi_{\sigma}(t) \quad \text{for any } t, \tag{5.1}
\]

where \( \Phi_{\sigma}(t) \) is the normal distribution function with mean 0 and variance \( \sigma^2 \). This distribution has been proven to converge therefore we will not present any simulated pictures.

Consider the following distribution functions that were presented in Section 3.4 where the correlation coefficient was included in the ASCLT:
\[
\frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \mathbb{I}_{n} \left( \sqrt{n} (r_n - \rho) \leq t \right) \xrightarrow{a.s.} \Phi_\gamma(t) \text{ for any } t \ (5.2)
\]

and

\[
\frac{1}{\log N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \frac{\sqrt{n}}{2} \left( \log \frac{1 + r_n}{1 - r_n} - \log \frac{1 + \rho}{1 - \rho} \right) \leq t \right\} \xrightarrow{a.s.} \Phi_\tau(t) \text{ for any } t \ (5.3)
\]

In subsequent sections, we will refer to the ASCLT-based distribution function method 1 in (5.2) as “ASCLT #1” and method 2 in (5.3) as “ASCLT #2.”

For (5.2), \( \Phi_\gamma(t) \) is the normal distribution function with mean 0 and variance was

\[
\gamma^2 = \frac{1}{4} \rho^2 \left[ \frac{b_{33}}{\sigma_x^4} + 2 \frac{b_{34}}{\sigma_x^2 \sigma_y^2} + \frac{b_{44}}{\sigma_y^4} \right] - \rho \left[ \frac{b_{35}}{\sigma_x^3 \sigma_y} + \frac{b_{45}}{\sigma_x \sigma_y^3} \right] + \frac{b_{55}}{\sigma_x^2 \sigma_y^2}.
\]

For (5.3), \( \Phi_\tau(t) \) is the normal distribution function with mean 0 and variance was

\[
\tau^2 = \hat{g}(\rho)^2 \gamma^2 = \frac{\rho^2}{4(1 - \rho^2)^2} \left[ \frac{b_{33}}{\sigma_x^4} + 2 \frac{b_{34}}{\sigma_x^2 \sigma_y^2} + \frac{b_{44}}{\sigma_y^4} \right] - \frac{\rho}{(1 - \rho^2)^2} \left[ \frac{b_{35}}{\sigma_x^3 \sigma_y} + \frac{b_{45}}{\sigma_x \sigma_y^3} \right] + \frac{b_{55}}{(1 - \rho^2)^2 \sigma_x^2 \sigma_y^2},
\]

where

\[
\hat{g}(\rho) = \frac{l}{l - \rho^2}.
\]

To the best of our knowledge, these empirical distributions also have not been considered before. Even though the property of converging to asymptotic results were derived theoretically, numerical studies based on computer simulations will be presented to show how these empirical distribution functions converges to a normal distribution. It should
be noted that the above empirical distribution functions include the unknown parameter $\rho$. We will have to replace this unknown quantity with a suitable approximation.

As mentioned previously, it will be shown that the empirical distribution functions that combine the ASCLT and the correlation coefficient converge to a normal distribution. However, our true goal is to estimate the quantiles from these distribution functions. For interval estimation purposes, we will be interested in the quantiles on the tails of these empirical distributions. Due to the asymptotic nature of the ASCLT-based confidence intervals for the population coefficient, the quantiles $\tilde{t}_\alpha$ and $\tilde{t}_{1-\alpha}$ defined earlier would require large sample sizes to estimate accurately the true quantiles, $t_\alpha$ and $t_{1-\alpha}$. In the following sections of this chapter, we will present techniques that will lead to the estimation of these quantiles, but with moderate sample sizes.

In practical data analysis problems, it is not uncommon to deal with small to moderate sample situations. Therefore, we must develop a procedure to help connect the asymptotic results for the proposed distribution functions, quantiles, and confidence interval estimation procedures mentioned earlier for small and moderate sample sizes. These procedures will be clearly shown and validated based on simulation studies. In our simulation based studies, we will be estimating the quantiles for different sized samples. To overcome the problem of asymptotic results that converge at a very slow rate, we propose the following:
• Do random permutations of the full sample in the process of estimating the
distribution function and its corresponding quantiles. We will be able to observe
how the results of the permuted samples affect the rate of convergence.

• As previously mentioned, we proposed the replacing \( \log N \) with the quantity
\[
C_N = \sum_{n=1}^{N} \frac{1}{n}, \text{ where } N \text{ is the sample size.}
\]

Before these simulations are observed, we should keep in mind that any proposed
ASCLT-based confidence intervals that adjust for finite-sample cases should satisfy
certain interval estimation properties. For example, any interval estimator should have a
probability of \((1-2\alpha)\) that includes the true value of the population parameter. Recall the
two confidence interval techniques that were presented in chapter 4. For the population
correlation coefficient,

\[
P\left[ \rho \in I^{(N)}_{\alpha} \right] = 1 - 2\alpha. \tag{5.4}
\]

To validate that this statement in (5.4) is true, this chapter will rely on the long-run
frequency interpretation of probability. To determine if the true population parameter is
in this interval estimate with \((1-2\alpha)\%\) confidence, numerical studies based on computer
simulations will be used to check this property. Each simulation to estimate the
confidence interval will be repeated and a long-run determination of the percentage of
these intervals that contain the true parameter will be observed.

In this chapter, the following confidence intervals for the population correlation
coefficient will be considered:
Confidence intervals (4.5) and (4.6) that were introduced earlier. These interval estimates uses the derived quantiles generated from the distribution functions $\tilde{H}_N(t)$ and $\tilde{J}_N(t)$, respectively.

Confidence interval (5.9) derived from the classic bootstrap method which will be discussed in the section 5.8.

Confidence interval (5.10) derived from the classic Fisher's z transformation which will also be presented in section 5.9.

All the above confidence intervals will be compared assuming the bivariate normal distributions and repeated with bivariate distributions that are non-normal.

It should be mentioned that, being a new confidence interval approach for the correlation coefficient, the ASCLT-based intervals will need many simulation-based testing and evaluations to determine its performance and comparison over existing methods. Many simulations, with different samples sizes and permutations were performed and results evaluated. An adequate number of results will be presented to represent our general findings.

5.2 Bivariate Normal Random Deviate Generation

Many of the simulation studies observed for this research will assume bivariate normal distributions. The preferred method of generating standard normal random deviates is the Box and Muller method (see Lange (1999)). This method generates two independent standard normal deviates $X$ and $Y$ by starting with independent uniform deviates $U$ and $V$ on $[0,1]$. The Box and Muller method transforms from the random Cartesian coordinates...
(X, Y) in the plane to random polar coordinates (Θ, R). If Θ is uniformly distributed on [0, 2π] and R² is exponentially distributed with mean 2, where Θ = 2πU and R² = -2lnV, our defined independent standard normal random deviates will be defined as X = R cos Θ and Y = R sin Θ. It should be noted that if we are interested in generating independent $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ random deviates, simply do the transformations $Z_1 = \mu_1 + X\sigma_1$ and $Z_2 = \mu_2 + Y\sigma_2$.

Now suppose we want to simulate dependent bivariate normal random deviates with desired normal population parameters and population correlation coefficient. The simplest way to generate these bivariate random deviates is to complete the following algorithm (see Gentle (2003)):

Step 1: Use the Box Muller technique to generate the matrix $Z$ (n x 2) that consist of independent standard normal random deviates $X$ and $Y$.

Step 2: Create the following covariance matrix $Σ$ where the covariance of two standard normal variables is equal to the correlation coefficient:

$$Σ = \begin{bmatrix} 1 & \sigma_{xy} \\ \sigma_{xy} & 1 \end{bmatrix} \begin{bmatrix} 1 & \rho_{xy} \\ \rho_{xy} & 1 \end{bmatrix}.$$ 

Step 3: Generate the Choleski decomposition from numerical linear algebra for the covariance matrix $Σ$. This produces a lower triangular matrix, $L$, so that $LL^t = Σ$. This procedure works for any positive definite $Σ$ matrix.

Step 4: Post multiply the generated $Z$ matrix by the Choleski decomposition of $Σ$. 

66
After these steps have been completed, we have generated random deviates from a bivariate standard normal distribution with a specified population correlation. We can again easily transform these random variables by multiplying the standard deviations and adding the means.

### 5.3 Bivariate Exponential Deviate Generation

Many of the simulation studies observed for this research will assume bivariate exponential distributions. These bivariate random variables will be generated assuming there is dependence between the variables $X$ and $Y$. For a more complete explanation of copulas, please refer to Nelson (1999). Copulas will be used as a starting point for simulating dependent exponentially distributed random variables. The study of copulas and their applications in statistics is a rather modern phenomenon. What are copulas? From one point of view, copulas are functions that join or couple multivariate distribution functions to their one-dimensional marginal distribution functions.

The word copula is a Latin noun which means “a link, tie, bond,” and is used in grammar and logic to describe “that part of a proposition which connects the subject and predicate.” The word copula was first used in a statistical sense by Abe Sklar in 1959 in a theorem that describes functions which “join together” one-dimensional distribution functions to form multivariate distribution functions. Consider the following theorem that bears Sklar’s name.

**Sklar’s Theorem.** Let $H$ be a joint distribution function with marginals $F$ and $G$. Then
there exists a copula $C$ such that for all $x, y$ in $\mathbb{R}$,
\[ H(x,y) = C(F(x), G(y)). \]

If $C$ is a copula and $F$ and $G$ are distribution functions, then the function $H$ is a joint
distribution function with marginals $F$ and $G$. For a thorough overview of copulas, we refer to Nelsen (1999).

The univariate exponential distribution plays an important role in statistics since it is a
distribution of time between events in a typical Poisson process. The following example,
first described by Marshall and Olkin, describes the role in a two-dimensional Poisson
process with bivariate exponential interarrival times. Consider a two component system –
such as a two engine aircraft, or a computer with dual CPU co-processors. The
components are subject to “shocks” which are fatal to one or both of the components.
For example, one of the two aircraft engines may fail, or a massive explosion could
destroy both engines simultaneously; or the CPU or a co-processor could fail, or a power
surge could eliminate both simultaneously. Let $X$ and $Y$ denote the lifetimes of
components 1 and 2, respectively. The “shocks” to the two components are assumed to
form three independent Poisson processes with positive parameters $\lambda_1, \lambda_2,$ and $\lambda_{12}$. These
parameters are dependent on whether the shock kills only component 1, component 2, or
both components. The times $Z_1, Z_2, Z_{12}$ of the occurrence of these three shocks are
exponential random variables with parameters $\lambda_1, \lambda_2,$ and $\lambda_{12}$.

Now suppose we want to simulate dependent bivariate exponential random deviates with
desired exponential population parameters and population correlation coefficient. The
The following algorithm developed by Devroye (1986) generates random variates \((X,Y)\) from the Marshall-Olkin bivariate exponential distribution with parameters \(\lambda_1, \lambda_2, \text{ and } \lambda_{12}\):

**Step 1:** Generate three independent uniform deviates \(r, s, \text{ and } t\) on \([0,1]\).

**Step 2:** Set \(X = \min\left[ \frac{-\ln(r)}{\lambda_1}, \frac{-\ln(t)}{\lambda_{12}} \right]\), and \(Y = \min\left[ \frac{-\ln(s)}{\lambda_2}, \frac{-\ln(t)}{\lambda_{12}} \right]\).

**Step 3:** The desired pair is \((X,Y)\).

The ordinary correlation coefficient of \(X\) and \(Y\) is given by:

\[
\rho = \frac{\lambda_{12}}{\lambda_1 + \lambda_2 + \lambda_{12}}.
\]

### 5.4 ASCLT for the Correlation Coefficient Simulations

In this section we will show how the ASCLT-based distribution function method 1 in (5.2) that includes the correlation coefficient converges to a normal distribution. For our estimated distribution function in (5.2), the simulations in this section will be completed for the following scenarios:

- Bivariate normal distributions for different values of \(\rho\).
- Bivariate exponential distributions for different values of \(\lambda_1, \lambda_2, \text{ and } \lambda_{12}\).
- Bivariate poisson distribution for different values of \(\lambda_i\).

In chapter 3, we spent time developing the theoretical results, now we would like to validate these results with computer based simulations. As mentioned in the introduction, due to the slow rate of convergence of the theoretical proposal of the ASCLT-based distribution, we applied modifications to the theoretical results in order to make them more applicable to small to moderate sample size situations. This adjustment will be accomplished through random permutations of the original sample and by replacing \(\log N\)
with the quantity $C_N$.

Since our function includes the unknown parameter $\rho$, we propose replacing this unknown quantity with the suitable approximation $\hat{\rho}$. With this in mind, by replacing the log averaging term and by Corollary 3.1, all the simulations in this section will be developed using the following distribution function:

$$
\tilde{H}_N(t) = \frac{1}{C_N} \sum_{n=1}^{N} I\{\sqrt{n}(r_n - \hat{\rho}) \leq t\} 
$$

where this function has an a.s. weak convergence to the normal distribution with mean 0 and variance $(1-\rho^2)^2$. Also, $\hat{\rho}$ will be the estimated population correlation coefficient for the entire sample $N$.

In order to overcome the problem of small to moderately sized samples, we propose using random permutations of the entire sample in the process of estimating the empirical distribution function. The following are the steps to complete this permutation process:

Step 1: Use the techniques mentioned previously to generate the random bivariate normal sample $(X_N, Y_N)$ where $N$ is our desired sample size.

Step 2: Let the $i^{th}$ permuted sample vector be denoted and given by

$$(X^{*i}, Y^{*i}) = \text{permute}(X_N, Y_N), \quad i = 1, \ldots, n_{\text{per}},$$

where $n_{\text{per}}$ represents the total number of permutations of the sample we are interested in.

Step 3: Now for each permuted sample, we compute the distribution function:
\[
\tilde{H}_N^*(t) = \frac{1}{C_N} \sum_{n=1}^{N} \int_{\left\{ \sqrt{n}(r^*_n - \hat{\rho}) \leq t \right\}} \frac{1}{n} \, dF_n, \quad n = I, \ldots, N,
\]

where \( r^*_n \) denotes the correlation coefficient of the partial, \( i^{th} \) permuted sample \( ((x_1, y_1), \ldots, (x_n, y_n)), n = I, \ldots, N \). For every permuted sample, each value of \( \sqrt{n}(r^*_n - \hat{\rho}) \) will be included in the function \( \tilde{H}_N(t) \).

Step 4: For each value of \( t \), the estimated distribution function that will be observed is

\[
\overline{\tilde{H}}_N(t) = \frac{\sum_{i=1}^{n_{per}} \tilde{H}_N^*(t)}{n_{per}}.
\]

Monte Carlo simulations were performed to observe the ASCLT-based distribution function for method 1. For the bivariate normal and poisson distributions, the following are the common set-up for each of these simulations:

1. The number of simulated random deviates were \( N_{sim} = 10,000 \).
2. The number of permutations completed were \( n_{per} = 100 \).

However, for the bivariate exponential distribution, the simulation set-up changed:

1. The number of simulated random deviates were \( N_{sim} = 50,000 \).
2. The number of permutations completed were \( n_{per} = 10,000 \).

The skewed nature of the parent populations substantially affects our rate of convergence, which will be presented later.

For the bivariate normal distribution, the following will give the reader some insight into
the progression of simulations to our final estimated distribution function. Each plot will superimpose the estimated distribution function with the true normal distribution. The idea here is to evaluate the convergence of the empirical distribution to the true distribution function. Within this framework, we will observe the simulation progression to arrive at the final estimated distribution function. In the following sequence, independent standard normal bivariate random variables were generated.

The initial simulation results are presented in Figure 5.1. Observe how for finite random samples, our estimated function is not a distribution function. It does not satisfy the condition \( \lim_{t \to \infty} H_N(t) = 1 \).

Figure 5.1: Estimated distribution function \( H_N(t) \) for simulated samples from bivariate standard normal distributions.
In the next iteration presented in Figure 5.2, we replace $\log N$ with the quantity

$$C_N = \sum_{n=1}^{N} \frac{1}{n},$$

where $N$ is the sample size. As discussed in chapter 2, this should help with the violated condition mentioned earlier and increase the rate of convergence.

Figure 5.2: Estimated distribution function $\tilde{H}_N(t)$ with $C_N$ for simulated samples from bivariate standard normal distributions.

Observe in Figure 5.2 how for finite random samples, our estimated function satisfies the condition $\lim_{t \to \infty} H_N(t) = 1$. Replacing the log averaging term with $C_N$ helped this condition; however, notice the jump when the t-value is approximately 0. This function does not satisfy the condition $H_N(t)$ is right continuous, that is, for every number $t_0$, $\lim_{t \downarrow t_0} H_N(t) = H_N(t_0)$. To address this issue, we refer to Remark 3.1 and the Law of Iterated Logarithm discussed in Section 3.5. When estimating the distribution using...
(5.5), notice that for small values of $n$, the weights of the function is determined by the quotient $1/n$. Notice how more weight is applied to the estimated function for smaller values of $n$. To address this issue, when performing the computer simulations, we will start this function at $n=m(n)$ and end at $N=max(n)$.

Observe in Figure 5.3 that our empirical distribution function through simulation is approaching the normal distribution with mean 0 and standard deviation $(1-\rho^2)^2$. The sequence of events to arrive at Figure 5.3 validates our theoretical assumptions.

Figure 5.3: Estimated distribution function $\tilde{H}_N(t)$ with $C_N$ from $n=m(n)$ to $N=max(n)$. 

![Graph of distribution function](image)
It may be interesting to observe how the ASCLT-based distribution function performs when generating from dependent normal distributions for various strengths in the correlation parameter. When observing this distribution function, we may assume the means are 0 and the variances are 1, because the correlation coefficient is independent of a change in location and scale in \( X \) or \( Y \). The following charts in Figures 5.4-5.6 were created from a bivariate normal distribution where \( \rho=0.3, \rho=0.5, \) and \( \rho=0.7 \).

Figure 5.4: Estimated distribution function \( \widetilde{H}_N(t) \) for simulated samples from dependent bivariate normal distributions when \( \rho=0.3 \).
Figure 5.5: Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from dependent bivariate normal distributions when $\rho = 0.5$.

Figure 5.6: Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from dependent bivariate normal distributions when $\rho = 0.7$. 
Observe in Figure 5.4 that our empirical distribution function through simulation is also approaching the normal distribution with mean 0 and standard deviation \((1-\rho^2)^2\), even if there exists a correlation between the two normal random variables. However, as the strength of the relationship between \(X\) and \(Y\) increase, the rate of convergence decreases. Notice the progression in the ASCLT-based distribution function as the correlation increased.

It may be interesting to observe how the ASCLT-based distribution function performs when generating from non-normal distributions. Observe the following charts in Figures 5.7-5.10 created from bivariate exponential distributions for various strengths in the correlation parameter.

**Figure 5.7**: Estimated distribution function \(\hat{H}_N(t)\) for simulated samples from bivariate exponential distributions with parameters \(\lambda_1 = \frac{1}{6}, \lambda_2 = \frac{1}{4}, \lambda_{12} = 1 (\rho = 0.7)\).
Figure 5.8: Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from bivariate exponential distributions with parameters $\lambda_1 = \frac{1}{6}, \lambda_2 = \frac{1}{3}, \lambda_{12} = \frac{1}{2} (\rho = 0.5)$.

ASCLT #1 for the Correlation Coefficient - Dependent Exponential RV's ($\rho=0.5$)

$$\tilde{H}_N(t) = \frac{1}{C_N} \sum_{n=m(n)}^{\text{max}(n)} \frac{1}{n} |\sqrt{n}(n-\rho) \leq t|$$

Figure 5.9: Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from bivariate exponential distributions with parameters $\lambda_1 = \frac{1}{2}, \lambda_2 = \frac{1}{2}, \lambda_{12} = \frac{1}{3} (\rho = 0.25)$.

ASCLT #1 for the Correlation Coefficient - Dependent Bivariate Exponential ($\rho=0.25$)

$$\tilde{H}_N(t) = \frac{1}{C_N} \sum_{n=m(n)}^{\text{max}(n)} \frac{1}{n} |\sqrt{n}(n-\rho) \leq t|$$
Figure 5.10: Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from independent bivariate exponential distributions with parameters $\lambda_i=3, i=1, 2$.

Observe the progression of the ASCLT-based distribution function method 1 in the above figures for exponential distributions as the correlation changes. The skewed nature of the parent populations and the dependent structure of the variables cause the rate of convergence to substantially decrease. In fact, when the strength of the correlation decreases for the exponentially distributed variables, the rate of convergence to a normal distribution increases. It should be noted that any simulations based on non-normal simulations converge to a normal distribution with a mean of 0 and a variance that was defined in Theorem 3.1.

It may be interesting to observe how the ASCLT-based distribution function method 1 performs when generating from non-normal discrete distributions. Observe the following charts in Figures 5.11-5.13 created from independent bivariate poisson distributions for different parameters.
Figure 5.11: Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from bivariate Poisson distributions with parameters $\lambda_i=1$, $i=1, 2$.

Figure 5.12: Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from bivariate Poisson distributions with parameters $\lambda_i=3$, $i=1, 2$. 
Figure 5.13: Estimated distribution function $\tilde{H}_N(t)$ for simulated samples from bivariate poisson distributions with parameters $\lambda_i = 10, i = 1, 2$.

![Graph of ASCLT #1 for the Correlation Coefficient - Independent Poisson RV's](image)

It does not appear that the discrete nature of the parent population affects the convergence. The above figures indicate that the independent poisson distributions converge to a normal distribution, regardless of the poisson parameters.

5.5 ASCLT for the Correlation Coefficient Simulations – Variance Stabilizing Technique

In this section we will show how the ASCLT-based distribution function method 2 in (5.3) that includes the correlation coefficient, Cramér’s Theorem, and Theorem 3.2, converges to a normal distribution. For our estimated distribution function in (5.3), the simulations in this section will also be completed for the same scenarios as in the previous section:
• Bivariate normal distributions for different values of \( \rho \).
• Bivariate exponential distributions for different values of \( \lambda_1, \lambda_2, \) and \( \lambda_{12} \).
• Bivariate poisson distribution for different values of \( \lambda_i \).

We will again propose the same modifications to the original theoretical results in order to make them more applicable in small sample situations. In this section, all the simulations will be developed using the following distribution function:

\[
\tilde{J}_N(t) = \frac{1}{C_N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \sqrt{n} \left( \log \frac{1+r_n}{1-r_n} - \log \frac{1+\hat{\rho}}{1-\hat{\rho}} \right) \leq t \right\}
\]  

(5.6)

As mentioned in Remark 3.2, this function has an a.s. weak convergence to the normal distribution with mean 0 and variance \( \tau^2 \). Also, \( \hat{\rho} \) will be the estimated population correlation coefficient for the entire sample \( N \).

In order to overcome the problem of small to moderately sized samples, we will again use permutations of the entire sample in the process of estimating the empirical distribution function. We will use the same steps as in the previous section; however, in Step 3 the used estimated distribution for each permutation will be:

\[
\tilde{J}_{N}^{*i}(t) = \frac{1}{C_N} \sum_{n=1}^{N} \frac{1}{n} \left\{ \sqrt{n} \left( \log \frac{1+r_{n}^{*i}}{1-r_{n}^{*i}} - \log \frac{1+\hat{\rho}}{1-\hat{\rho}} \right) \leq t \right\}, \quad n = 1, \ldots, N,
\]

and for each value of \( t \), the estimated distribution function that will be observed:

\[
\overline{J}_N(t) = \frac{\sum_{i=1}^{nper} \tilde{J}_N^{*i}(t)}{nper}.
\]
The following will give the reader some insight how the empirical distribution function (5.6) converges to the true distribution function. Each plot will superimpose the estimated distribution function with the true normal distribution. The idea here is to evaluate the convergence of the empirical distribution to the true distribution function, regardless of the distribution on $X$ and $Y$.

In Figures 5.14-5.17, observe how the ASCLT-based distribution method 2 converges when generated from normal populations with different strengths in the correlation parameter.

Figure 5.14: Estimated distribution function $\tilde{F}_N(t)$ for simulated samples from dependent bivariate normal distributions with $\rho=0.7$. 

\[
\tilde{F}_N(t) = \frac{1}{C_N} \sum_{n=m+1}^{max(n)} \frac{1}{n} \left\{ \frac{\sqrt{n} \left( \log \frac{1+\rho}{1-\rho} \right)}{2} \right\}
\]
Figure 5.15: Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from dependent bivariate normal distributions with $\rho=0.5$.

\[
\tilde{J}_N(t) = \frac{1}{C_N} \sum_{n=m(n)}^{\max(n)} \frac{1}{n} \left\{ \sqrt{n} \left( \log \left( \frac{1+r_n}{1-r_n} \right) - \log \left( \frac{1+\hat{\rho}}{1-\hat{\rho}} \right) \right) \leq t \right\}
\]

Figure 5.16: Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from dependent bivariate normal distributions with $\rho=0.3$.

\[
\tilde{J}_N(t) = \frac{1}{C_N} \sum_{n=m(n)}^{\max(n)} \frac{1}{n} \left\{ \sqrt{n} \left( \log \left( \frac{1+r_n}{1-r_n} \right) - \log \left( \frac{1+\hat{\rho}}{1-\hat{\rho}} \right) \right) \leq t \right\}
\]
The simulation presented in Figures 5.14-5.17 indicates that our empirical distribution function through simulation is approaching the normal distribution with mean 0 and standard deviation 1, which validates our theoretical assumptions. Even if there exists a correlation between the two normal random variables, regardless of the strength, convergence still occurs for the distribution function (5.3).

Figures 5.18-5.21 illustrates how the ASCLT-based distribution function performs when generating from non-normal distributions. In fact, the following graphs were created from a bivariate exponential distribution for various strengths in the correlation parameter.
Figure 5.18: Estimated distribution function $\bar{J}_N(t)$ for simulated samples from bivariate exponential distributions with parameters $\lambda_1 = \frac{1}{6}, \lambda_2 = \frac{1}{4}, \lambda_{12} = 1$ ($\rho = 0.7$).

Figure 5.19: Estimated distribution function $\bar{J}_N(t)$ for simulated samples from bivariate exponential distributions with parameters $\lambda_1 = \frac{1}{6}, \lambda_2 = \frac{1}{3}, \lambda_{12} = \frac{1}{2}$ ($\rho = 0.5$).
Figure 5.20: Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from bivariate exponential distributions with parameters $\lambda_1 = \frac{1}{2}, \lambda_2 = \frac{1}{2}, \lambda_{12} = \frac{1}{3}$ ($\rho = 0.25$).

Figure 5.21: Estimated distribution function $\tilde{J}_N(t)$ for simulated samples from bivariate exponential distributions with parameter $\lambda_i = 3, i=1,2$.
It appears in the Figure 5.18 that the skewed nature of the parent populations and the strong dependence of the variables caused the rate of convergence to substantially decrease. However, as the strength of the correlation decreases from Figure 5.19 to 5.21, the rate of convergence increases. Again, our empirical distribution function $\tilde{F}_N(t)$ is again approaching the normal distribution with mean 0 and variance that was defined in Theorem 3.1.

It may be interesting to observe how the ASCLT-based distribution function method 2 performs when generating from non-normal discrete distributions. Figures 5.22-5.24 were created from independent bivariate poisson distributions for different parameters.

Figure 5.22: Estimated distribution function $\tilde{F}_N(t)$ for simulated samples from bivariate poisson distributions with parameters $\lambda_i=1, i=1, 2$. 

\[
\tilde{J}_N(t) = \frac{1}{C_N} \sum_{n \in m(n)} \frac{1}{n} \left\{ \sqrt{n} \log \left( \frac{1 + r_n}{1 - r_n} \right) \left( \frac{1 + \rho}{1 - \rho} \right) \leq t \right\}
\]
Figure 5.23: Estimated distribution function \( \tilde{J}_N(t) \) for simulated samples from bivariate Poisson distributions with parameters \( \lambda_i = 3, i = 1, 2 \).

Figure 5.24: Estimated distribution function \( \tilde{J}_N(t) \) for simulated samples from bivariate Poisson distributions with parameters \( \lambda_i = 10, i = 1, 2 \).
It does not appear that the discrete nature of the parent population affects the convergence for the distribution function (5.3). The above figures indicate that the independent poisson distributions converge to a normal distribution, regardless of the poisson parameters.

5.6 ASCLT-based Confidence Interval for Permuted Samples (Technique #1)

In this section, for the distribution function (5.5) we will define the quantile estimates for each permuted sample. For $\alpha \in (0,1)$, the quantiles $\tilde{t}_{\alpha i}^{*(N)}$ and $\tilde{t}_{1-\alpha i}^{*(N)}$ are now estimated via the relationship defined in (4.2) in chapter 4:

$$\tilde{t}_{\alpha i}^{*(N)} = \max \left\{ t \left| C_N^{-1} \sum_{n=1}^{N} I_{\left\{ \sqrt{n} \left( t_{ni}^* - \hat{\rho} \right) \leq t \right\}} \right\} \leq \alpha \right\},$$

and

$$\tilde{t}_{1-\alpha i}^{*(N)} = \max \left\{ t \left| C_N^{-1} \sum_{n=1}^{N} I_{\left\{ \sqrt{n} \left( t_{ni}^* - \hat{\rho} \right) \leq t \right\}} \right\} \leq 1 - \alpha \right\},$$

where $C_N = \sum_{n=1}^{N} 1/n$.

The following will be the ASCLT-based confidence interval estimates for permuted samples:

$$I_{\alpha}^{(N)} = \left[ \hat{\rho} + \frac{\tilde{t}_{1-\alpha}}{\sqrt{N}}, \hat{\rho} + \frac{\tilde{t}_{\alpha}}{\sqrt{N}} \right],$$

(5.7)

where $\tilde{t}_{1-\alpha} = \frac{\sum_{i=1}^{n_{\text{per}}} t_{1-\alpha i}^{*(N)}}{n_{\text{per}}}$ and $\tilde{t}_{\alpha} = \frac{\sum_{i=1}^{n_{\text{per}}} t_{\alpha i}^{*(N)}}{n_{\text{per}}}$. 

90
5.7 ASCLT-based Confidence Interval for Permuted Samples (Technique #2)

In this section, for the distribution function (5.6) we will define the quantile estimates for each permuted sample. For \( \alpha \in (0,1) \), the quantiles \( \tilde{t}_{\alpha},(N) \) and \( \tilde{t}_{1-\alpha}\) are now estimated via the relationship defined in (4.4) in chapter 4:

\[
\tilde{t}_{\alpha},(N) = \max_t \left\{ C_N^{-1} \sum_{i=1}^{N} \frac{1}{n} \left[ \sqrt{n} \left( \frac{\log \frac{1+r_n^*}{1-r_n^*}}{\log \frac{1+\rho^{*i}}{1-\rho}} \right) \leq t \right] \right\} \leq \alpha
\]

and

\[
\tilde{t}_{1-\alpha} = \max_t \left\{ C_N^{-1} \sum_{i=1}^{N} \frac{1}{n} \left[ \sqrt{n} \left( \frac{\log \frac{1+r_n^*}{1-r_n^*}}{\log \frac{1+\rho^{*i}}{1-\rho}} \right) \leq t \right] \right\} \leq 1-\alpha
\]

where \( C_N = \sum_{n=1}^{N} 1/n \).

The following will be the ASCLT-based confidence intervals estimates for permuted samples:

\[
I_{\alpha}^{(N)} = \begin{bmatrix}
\exp \left\{ 2 \left( z_N + \frac{\tilde{t}_{1-\alpha}^{(N)}}{\sqrt{N}} \right) \right\} - 1 & \exp \left\{ 2 \left( z_N + \frac{\tilde{t}_{\alpha}^{(N)}}{\sqrt{N}} \right) \right\} - 1 \\
\exp \left\{ 2 \left( z_N + \frac{\tilde{t}_{1-\alpha}^{(N)}}{\sqrt{N}} \right) \right\} + 1 & \exp \left\{ 2 \left( z_N + \frac{\tilde{t}_{\alpha}^{(N)}}{\sqrt{N}} \right) \right\} + 1
\end{bmatrix},
\]

where \( \tilde{t}_{1-\alpha} = \frac{\sum_{i=1}^{nper} \tilde{t}_{1-\alpha}^{*i,(N)}}{nper} \) and \( \tilde{t}_{\alpha} = \frac{\sum_{i=1}^{nper} \tilde{t}_{\alpha}^{*i,(N)}}{nper} \).
5.8 Bootstrap Confidence Interval for Population Coefficient

In this section, we will introduce the confidence interval technique for \( \rho \) that utilizes a bootstrapping technique. In later sections, we will compare this confidence interval approach to the ASCLT-based intervals presented in chapter 4. Bootstrap resampling methods are a recently developed technique for making statistical inferences. In fact, with the recent development of more powerful computers, bootstrap techniques have been extensively explored. This is because it requires the modern computer power to develop the bootstrapped samples. Bootstrapping is a statistical method for estimating the sampling distribution of a statistic by sampling with replacement from the original sample. These samples are then used in the computation of the relevant bootstrap estimates. This technique will be used to develop estimates of the standard errors and confidence intervals of a population correlation coefficient. This technique can also be used for the parameter estimation for the mean, median, proportion, odds ratio, and regression coefficients. It can also be used in the development of hypothesis tests. For a thorough overview of bootstrap methods, we refer to Davison and Hinkley (1997) or to Efron and Tibshirani (1993).

For interval estimation, the following resampling scheme is proposed to compute the bootstrap confidence interval for \( \rho \). The bootstrap algorithm for the confidence interval for the correlation coefficient is:

Step 1: Generate i.i.d. bivariate samples \((X, Y)\) from an empirical distribution. Calculate the correlation coefficient \( \hat{\rho} \) from this original sample.
Step 2: Generate $B_1$ independent bootstrap samples $(X^{(b_1)}, Y^{(b_1)})$ with replacement from the original sample $(X, Y)$, $b_1 = 1, ..., B_1$.

Calculate the correlation coefficient $\hat{\rho}^{(b_1)}$ for each bootstrap sample.

Step 3: Calculate the standard error for the $B_1$ bootstrap samples:

$$se^{(b_1)} = \sqrt{\frac{1}{B_1 - 1} \sum_{b_1=1}^{B_1} \left( \hat{\rho}^{(b_1)} - \bar{\rho}^{(b_1)} \right)^2}$$

where $\bar{\rho}^{(b_1)} = \frac{1}{B_1} \sum_{b_1=1}^{B_1} \hat{\rho}^{(b_1)}$.

Step 4: Generate $B_2$ independent bootstrap sub-samples $(X^{(b_2)}, Y^{(b_2)})$ with replacement from the original sample $(X^{(b_1)}, Y^{(b_1)})$, $b_2 = 1, ..., B_2$.

Calculate the correlation coefficient $\hat{\rho}^{(b_2)}$ for each bootstrap sub-sample. The bootstrap estimate of the standard error for each bootstrap sub-sample:

$$se^{(b_2)} = \sqrt{\frac{1}{(B_2 - 1)} \sum_{b_2=1}^{B_2} \left( \hat{\rho}^{(b_2)} - \bar{\rho}^{(b_2)} \right)^2}$$

where $\bar{\rho}^{(b_2)} = \frac{1}{B_2} \sum_{b_2=1}^{B_2} \hat{\rho}^{(b_2)}$.

Step 5: Compute the quantiles from the empirical distribution for:

$$t^*_b = \frac{\hat{\rho}^{(b_1)} - \hat{\rho}}{se^{(b_2)}} \quad b = 1, ..., B_1$$
Step 6: The $\alpha$th quantile of $t_b^*$ is estimated by the value $t^{(\alpha)}$ such that:

$$\alpha = \frac{\# \{ t_b^* \leq t^{(\alpha)} \}}{B_1}$$

Step 7: Compute the confidence interval for the correlation coefficient:

$$I_\alpha = [\overline{\rho}^{(\ast b_1)} - t^{(\alpha)} \cdot se^{(\ast b_1)}, \overline{\rho}^{(\ast b_1)} + t^{(1-\alpha)} \cdot se^{(\ast b_1)}].$$  \hspace{1cm} (5.9)

This algorithm was developed directly using S-plus.

### 5.9 Variance Stabilizing Transformation for the Confidence Interval for Population Coefficient (Classical Technique)

In this section, we will introduce the classical confidence interval technique for $\rho$ that utilizes a variance stabilizing transformation. In later sections, we will compare this classical confidence interval approach to the ASCLT-based interval presented in (5.8) and (5.9). For normal populations, Theorem 3.2 stated that $\sqrt{n}(r - \rho) \overset{L}{\rightarrow} N(0, (1 - \rho^2)^2)$. From Theorem 3.3, Fisher created a transformation, $g(r)$, such that

$$\sqrt{n}(g(r) - g(\rho)) \overset{L}{\rightarrow} N(0, 1),$$

where the function $g$ is:

$$g(x) = \frac{1}{2} \log \frac{1 + \rho}{1 - \rho}$$

In this case, $g(r)$ tends to become normal quickly as the sample size increases, and the variance of this transformation, $\frac{1}{n-3}$, is independent of the population parameter. Furthermore, even if the distribution of $g(r)$ is not strictly normal; it tends to be normal rapidly as the sample size increases for any values of $\rho$.

**Remark 5.1** The confidence interval for the population correlation coefficient by using the variance stabilizing transformation proposed by Fisher is:
\[ I^{(N)}_\alpha = \left[ \frac{\exp(2 \cdot \zeta_l) - 1}{\exp(2 \cdot \zeta_u) + 1}, \frac{\exp(2 \cdot \zeta_u) - 1}{\exp(2 \cdot \zeta_l) + 1} \right], \quad (5.10) \]

where

\[ \zeta_l = \frac{1}{2} \log \frac{1 + r_n}{1 - r_n} - z_\alpha \sqrt{\frac{1}{n - 3}}, \]

and

\[ \zeta_u = \frac{1}{2} \log \frac{1 + r_n}{1 - r_n} + z_\alpha \sqrt{\frac{1}{n - 3}}. \]

For a \((1-\alpha)\)% confidence level, \(z_\alpha\) is the upper percentage point of the standard normal distribution. For an overview of this confidence interval method, we refer to Devore (2004).

### 5.10 Simulation Results for the Correlation Coefficient Confidence Interval

Monte-carlo simulations were performed to observe the confidence interval performance of the ASCLT-based methods 1 and 2, bootstrap, and classic procedures. In this section, the idea is to evaluate the confidence interval procedures when the samples come from different population distributions. It will also be interesting to observe confidence interval performance by changing the sample sizes, confidence levels, and correlation strengths. As mentioned earlier, one of the major advantages of the ASCLT-based confidence intervals is that the procedures do not involve the variance estimation. A focused concern when the variance estimation is not involved is how well the derived confidence intervals perform when estimating the correlation coefficient. The confidence interval methods detailed in this chapter will be observed via monte-carlo simulations in order to compare the results of
these procedures. The confidence interval results will then be compared and evaluated by completing the following:

- Confidence interval widths will be compared for each method. An important concept that is embodied in confidence interval estimation is precision. The shorter the confidence interval, the more precise the estimation of the population correlation coefficient.

- Confidence interval accuracy will be compared for each method. For each procedure, multiple simulations to develop the confidence interval will be accomplished, and a determination of the proportion of these intervals that contain the true parameter will be observed.

As mentioned in the beginning of this chapter, only a few simulation results will be presented here since the general trend remains similar over several settings. Numerical results will be presented using tables and figures, and also summarized with discussion throughout. The following are the simulation scenarios:

- Random deviates will be generated for the normal, exp., and poisson distributions.

- For the normal populations, \( \rho \) will take on the values 0.7, 0.5, 0.3, and 0.0.

- For the exponential populations, \( \rho \) will take on the values 0.7, 0.5, 0.25, and 0.0.

- For the poisson populations, we will assume independence between the variables, but the poisson parameter will take on the values \( \lambda = 1, 3, \) and 10.

- The total number of random deviates generated will be \( N_{sim} = 10,000 \) and \( nper = 100, N_{sim} = 1,000 \) and \( nper = 500, \) and \( N_{sim} = 100 \) and \( nper = 1000. \)

- The confidence levels that will be observed are \( 2\alpha = 90\%, 2\alpha = 95\%, \) and \( 2\alpha = 99\%. \)

- For the bootstrap confidence intervals, \( B_1 = 1000 \) and \( B_2 = 25. \)
5.11 Simulation Results based on Bivariate Normal Distributions

The idea here is to evaluate the performance of the ASCLT-based confidence intervals versus other conventional methods when samples come from bivariate normal distributions. Within this framework, we consider the scenarios when the samples originate from bivariate populations with varying dependencies. In the Appendix, the outcomes of 144 simulations will be presented in Table 5.1. Also, results of selected simulations will be displayed and conclusions presented throughout this section.

Figure 5.25: Confidence interval results for $N=10,000$, $CL=90\%$, and $\rho=0.7$.

Simulations evaluating the property of precision and accuracy are presented in Figure 5.25 for $N=10,000$, $CL=90\%$, and $\rho=0.7$. It is clear in this scenario that the ASCLT-based confidence intervals for method 2 yielded the most precise interval. Even though the accuracy did not reach the desired confidence level, this difference is not drastic enough to be concerned.
Figure 5.26: Confidence interval results for $N=1,000$, $CL=90\%$, and $\rho=0.7$.

![Confidence Interval Comparison When $\rho=0.7$](image)

Figure 5.27: Confidence interval results for $N=1,000$, $CL=95\%$, and $\rho=0.5$.

![Confidence Interval Comparison When $\rho=0.5$](image)
Simulations evaluating the property of precision and accuracy are presented in Figure 5.26 for \( N=1,000, \ CL=90\% \), and \( \rho=0.7 \), and in Figure 5.27 for \( N=1,000, \ CL=95\% \), and \( \rho=0.5 \). It is clear in these scenarios how the confidence interval precision changes as the correlation strength changes. The stronger the correlation, the ASCLT-based confidence intervals for method 2 yielded the most precise interval. As the correlation decreased to 0.5, the ASCLT-based confidence intervals for method 1 yielded the most precise interval.

Figure 5.28: Confidence interval results for \( N=100, \ CL=90\% \), and \( \rho=0.7 \).

Simulations evaluating the property of precision and accuracy are presented in Figure 5.28 for \( N=100, \ CL=90\% \), and \( \rho=0.7 \). In this scenario, the ASCLT-based confidence intervals for method 2 yielded the most precise, however the accuracy level of 87% is still not drastic enough to be concerned.
Figure 5.29: Confidence interval results for $N=100$, $CL=95\%$, and $\rho=0.3$.

Figure 5.30: Confidence interval results for $N=100$, $CL=99\%$, and $\rho=0.0$. 
Simulations evaluating the property of precision and accuracy are presented in Figure 5.29 for $N=100$, $CL=95\%$, and $\rho=0.3$, and in Figure 5.30 for $N=100$, $CL=99\%$, and $\rho=0.0$. It is clear in these weak correlation scenarios that the ASCLT-based confidence intervals for method 1 yielded the most precise interval while upholding the accuracy criteria.

The main findings from the results presented in tables and figures can be summarized as follows.

- All the confidence intervals presented in this section were repeated and a long-run determination of the percentage of these intervals that contained the true population parameter was observed. At each confidence level observed, all methods appropriately contained the population correlation coefficient. That is, whenever a simulation was repeated 100 times, nearly $100(1 - 2\alpha)\%$ of these intervals from each procedure contained the true parameter.

- For normal random variables with strong correlations, the ASCLT-based confidence intervals for method 2 consistently yielded the most precise interval.

- For normal random variables with moderate to weak correlations, the ASCLT-based confidence intervals for method 1 consistently yielded the most precise interval. However, this method 1 produced slightly more conservative intervals with strong correlations compared to the classic Bootstrap and Fisher’s variance stabilizing transformation procedures. Recall that one of goals in this research was to develop confidence intervals for small to moderate sample sizes. Future research may show that this interval technique may be more precise as the sample sizes increase.
5.12 Simulation Results based on Bivariate Exponential Distributions

The idea here is to evaluate the performance of the ASCLT-based confidence intervals versus other conventional methods when samples come from distributions that are non-normal. Therefore we want to see how the confidence interval results presented in the chapter change if the assumption of normality is violated. Within this framework, we consider the scenarios when the samples originate from bivariate exponential populations with varying dependencies. We will use the Marshall-Olkin technique presented earlier when simulating from dependent exponential distributions. In the Appendix, the outcomes of 144 simulations will be presented in Table 5.2. Also, results of selected simulations will be displayed and conclusions presented throughout this section.

Figure 5.31: Confidence interval results for $N=10,000$, $CL=95\%$, and $\rho=0.7$. 

![Confidence Interval Comparison When $\rho=0.7$](image)
Simulations evaluating the property of precision and accuracy are presented in Figure 5.31 for \( N=10,000, \ CL=95\% \), and \( \rho=0.7 \). It is clear in this scenario the classic confidence interval yielded a precise interval, however the accuracy was poor. However, the ASCLT-based confidence intervals for method 2 yielded a precise interval with a high accuracy rate.

Figure 5.32: Confidence interval results for \( N=1,000, \ CL=95\% \), and \( \rho=0.7 \).

Simulations evaluating the property of precision and accuracy are presented in Figure 5.32 for \( N=1,000, \ CL=95\% \), and \( \rho=0.7 \). Again, the classic technique had a low accurate, high precision interval. The ASCLT-based confidence intervals for method 2 yielded a precise interval with a low accuracy rate. Also, the ASCLT-based confidence intervals for method 1 yielded a more accurate interval.
Figure 5.33: Confidence interval results for $N=100$, $CL=95\%$, and $\rho=0.7$.

Figure 5.34: Confidence interval results for $N=100$, $CL=99\%$, and $\rho=0.0$. 
Simulations evaluating the property of precision and accuracy are presented in Figure 5.33 for $N=100$, $CL=95\%$, and $\rho=0.7$, and in Figure 5.34 for $N=100$, $CL=99\%$, and $\rho=0.0$. It is clear in these scenarios that for a small sample size and a strong correlation, there are issues with both precision and accuracy for all confidence interval techniques. However, as the strength of the correlation decreases the accuracy of each interval technique increases. For the independent case, the ASCLT-based confidence intervals for method 1 yielded the most precise interval. The main findings from the results presented in tables and figures can be summarized as follows.

- The classic confidence interval technique had the poorest results when the population distributions were exponentially distributed. Even though this method had the most precise intervals, it had very inaccurate results. When the random variables had a strong correlation, the long-run relative frequency did not obtain the desired confidence level. However, as the correlation strength decreased to a moderate to weak level, the accuracy increased.

- The bootstrap confidence interval technique consistently had the highest percentage of intervals containing the true population correlation coefficient. However, this method also consistently had the least precise interval estimates.

- For exponential random variables with strong correlations, the ASCLT-based confidence intervals for method 2 consistently had more precise intervals compared to method 1 and the bootstrap method, however there were issues with accuracy for this method. As the strength of the correlation decreased to moderate to weak levels, the accuracy and precision of method 1 was consistently better compared to method 2 and the bootstrap method.
5.13 Simulation Results based on Bivariate Poisson Distributions

The idea here is to evaluate the performance of the ASCLT-based confidence intervals versus other conventional methods when samples come from distributions that are discrete and non-normal. Therefore we want to see how the confidence interval results presented in the chapter change if the assumption of normality is violated. Within this framework, we consider the scenarios when the samples originate from independent bivariate poisson populations. In the Appendix, the outcomes of 36 simulations will be presented in Table 5.3. Also, results of selected simulations will be displayed and conclusions presented throughout this section. Consider the following three simulation scenarios.

Figure 5.35: Confidence interval results for $N=10,000$, $CL=95\%$, and $\rho=0.0$. 

![Confidence Interval Comparison When $\rho=0.0$](image)

<table>
<thead>
<tr>
<th>Method</th>
<th>Ave CI Width</th>
<th>% Containing ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASCLT #1</td>
<td>0.03792344</td>
<td>92%</td>
</tr>
<tr>
<td>ASCLT #2</td>
<td>0.04155826</td>
<td>91%</td>
</tr>
<tr>
<td>Bootstrap</td>
<td>0.04008428</td>
<td>96%</td>
</tr>
<tr>
<td>Classic</td>
<td>0.03920124</td>
<td>95%</td>
</tr>
</tbody>
</table>
Figure 5.36: Confidence interval results for $N=1,000$, $CL=95\%$, and $\rho=0.0$.

Figure 5.37: Confidence interval results for $N=100$, $CL=95\%$, and $\rho=0.0$. 

![Graph showing confidence interval comparison when $\rho=0.0$.]
It is clear in the three simulations poisson examples presented above that the ASCLT-based confidence intervals for method 1 yielded the most precise interval. Even though the accuracy did not reach the desired confidence level when \( N=10,000 \), this difference is not drastic enough to be concerned.

The main findings from the results presented in tables and figures can be summarized as follows.

- All the confidence intervals presented in this section were repeated and a long-run determination of the percentage of these intervals that contained the true population parameter was observed. At each confidence level observed, all methods appropriately contained the population correlation coefficient. That is, whenever a simulation was repeated 100 times, nearly \( 100(1 – 2\alpha)\% \) of these intervals from each procedure contained the true parameter.

- For independent poisson random variables, overall the ASCLT-based confidence intervals for method 1 yielded the most precise interval, regardless of the sample size.

- The ASCLT-based confidence intervals for method 2 consistently produced less precise interval estimates when compared to the bootstrap and classic methods.
Chapter 6

Conclusion

Throughout this thesis, we considered the balance between the asymptotic theory and real-life, finite sample approximations. With this in mind, we tried to present all the ideas with a mixture of mathematical theory with supporting simulation based evaluation. This thesis presented some existing theoretical results on the Almost Sure Central Limit Theorem. After we introduced the ASCLT, two distribution functions that included the ASCLT and the correlation coefficient were presented. We immediately discussed ways of modifying these distribution functions to address the rate of convergence issue. These modifications were presented and evaluated with simulations. This was then followed with proposals for estimating the quantiles of the distribution of our correlation coefficient statistic using these two unique methods.

Throughout this thesis, wherever appropriate, there have been conclusions stated. In this final chapter, we will summarize our conclusions, and mention possible future research ideas.

6.1 Summary

The following bullet points summarize the conclusions about the ASCLT-based distribution functions that were presented throughout this thesis.
• For the bivariate normal distribution, both ASCLT-based distribution function methods converged to a normal distribution. However for method 1, as the strength of the relationship between $X$ and $Y$ increased, the rate of convergence decreased. For method 2, even if there existed a correlation between the two normal random variables, regardless of the strength, the rate of convergence was not affected.

• For the bivariate exponential distribution, both ASCLT-based distribution function methods converged to a normal distribution very slowly when a correlation between the two normal random variables existed. The skewed nature of the parent populations and the dependent structure of the variables cause the rate of convergence to substantially decrease.

• For the independent bivariate poisson distribution, both ASCLT-based distribution function methods converged to a normal distribution. The discrete nature of the parent population (nor the poisson parameter) did not affect the convergence.

For the bivariate normal distribution, the following bullet points summarize the conclusions about the confidence intervals constructed for the population correlation from large sample simulations.

• For strong, moderate, weak, and no variable correlations, simulations show that the accuracy for all methods were reasonable. However, the classical procedure was the only method to consistently meet the accuracy criteria set by each confidence level. For strong to moderate variable correlations, the ASCLT-based
confidence interval method 2 had the most precise intervals. However, for weak to no variable correlations, the ASCLT-based confidence interval method 1 had the most precise intervals.

For the bivariate normal distribution, the following bullet points summarize the conclusions about the confidence intervals constructed for the population correlation from moderate sample simulations.

• For strong, moderate, weak, and no variable correlations, simulations show that the accuracy for all methods were reasonable. However, the bootstrap procedure was the only method to consistently meet the accuracy criteria set by each confidence level. For strong to moderate variable correlations, the ASCLT-based confidence interval method 2 typically had the most precise intervals. However, for weak to no variable correlations, the ASCLT-based confidence interval method 1 had the most precise intervals.

For the bivariate normal distribution, the following bullet points summarize the conclusions about the confidence intervals constructed for the population correlation from small sample simulations.

• For moderate, weak, and no variable correlations, simulations show that the accuracy for all methods were reasonable. However, for strong correlations, simulations show that all methods had accuracy issues, except for the classical procedure. For strong variable correlations, the ASCLT-based confidence interval method 2 had the most precise intervals. However, for moderate, weak,
and no variable correlations, the ASCLT-based confidence interval method 1 had the most precise intervals.

After observing the simulation results for the bivariate normal distribution, consider the following concluding statements. For strong variable correlations, after considering both accuracy and precision collectively, the ASCLT-based confidence interval for method 2 had the “best” intervals by comparison. However as the correlation weakened, and after considering both accuracy and precision collectively, the ASCLT-based confidence interval for method 1 had the “best” intervals by comparison.

For the bivariate exponential distribution, the following bullet points summarize the conclusions about the confidence intervals constructed for the population correlation from large and moderate sample simulations.

• For strong, moderate, weak, and no variable correlations, the ASCLT-based confidence interval method 1 consistently had the most precise intervals. However, this method occasionally did not meet the accuracy requirements set by the confidence level. Though the bootstrap method did not have the most precise intervals, it typically had the most accurate intervals.

• For strong, moderate, and weak variable correlations, simulations show that the classical confidence interval procedure was severely inaccurate.

For the bivariate exponential distribution, the following bullet points summarize the conclusions about the confidence intervals constructed for the population correlation
from small sample simulations.

- For strong variable correlations, simulations show that all confidence interval procedure were severely inaccurate with imprecise intervals.
- For moderate, weak, and no correlations, simulations show that the ASCLT-based confidence interval for method 1 had the most reasonable intervals. However, all methods show issues with accuracy.

After observing the simulation results for the bivariate exponential distribution, consider the following concluding statements. For strong, moderate, and weak variable correlations, all methods had issues with accuracy regardless of the sample size. The only scenario where accuracy occurred was when the simulated variables were independent. Also for small sample sizes, the estimated intervals for each method were both inaccurate and imprecise. For strong to moderate variable correlations, after considering both accuracy and precision collectively, the ASCLT-based confidence interval for method 1 (more precise) or the bootstrap method (more accurate) had the “best” intervals by comparison. However as the correlation weakened, and after considering both accuracy and precision collectively, the ASCLT-based confidence interval for method 1 had the “best” intervals by comparison.

For the bivariate poisson distribution, the following bullet points summarize the conclusions about the confidence intervals constructed for the population correlation.

- For any confidence level and sample size, simulations show that the accuracy for all methods were reasonable. For the independent bivariate poisson distribution,
the ASCLT-based confidence interval for method 1 consistently had the most precise intervals.

After observing the accuracy and precision for the bivariate Poisson distribution, the ASCLT-based confidence interval for method 1 had the “best” intervals by comparison.

6.2 Future Research Ideas

Significant numerical and theoretical investigations have been completed, however there are topics that were not considered in this dissertation. Possible examples to consider: additional ways to speed up the convergence for the ASCLT-based distribution function methods 1 and 2 when the parent population is severely skewed; additional ways to make the interval estimation for methods 1 and 2 more precise for small sample sizes; develop methods to perform a test of hypothesis about the correlation coefficient (i.e. $\rho \neq 0$) using the ASCLT-based distribution function.
## Appendix

### Table 5.1: Simulation Results for the Bivariate Normal Distribution

<table>
<thead>
<tr>
<th>N</th>
<th>p</th>
<th>Method</th>
<th>Ave CI Width</th>
<th>% Containing</th>
<th>90% CI Ave CI</th>
<th>% Containing</th>
<th>95% CI Ave CI</th>
<th>% Containing</th>
<th>99% CI Ave CI</th>
<th>% Containing</th>
</tr>
</thead>
<tbody>
<tr>
<td>10,000</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.016760</td>
<td>87%</td>
<td>0.020813</td>
<td>94%</td>
<td>0.026037</td>
<td>95%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.015521</td>
<td>85%</td>
<td>0.015665</td>
<td>90%</td>
<td>0.015619</td>
<td>89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.016836</td>
<td>85%</td>
<td>0.020551</td>
<td>95%</td>
<td>0.027465</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.016782</td>
<td>91%</td>
<td>0.019998</td>
<td>96%</td>
<td>0.026285</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.50</td>
<td>ASCLT #1</td>
<td>0.044480</td>
<td>88%</td>
<td>0.039731</td>
<td>98%</td>
<td>0.03818</td>
<td>98%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.025593</td>
<td>92%</td>
<td>0.031285</td>
<td>96%</td>
<td>0.033630</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.025223</td>
<td>92%</td>
<td>0.029911</td>
<td>97%</td>
<td>0.040547</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.024673</td>
<td>93%</td>
<td>0.029401</td>
<td>96%</td>
<td>0.038641</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.30</td>
<td>ASCLT #1</td>
<td>0.029419</td>
<td>87%</td>
<td>0.035011</td>
<td>96%</td>
<td>0.045979</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.08149</td>
<td>100%</td>
<td>0.037799</td>
<td>95%</td>
<td>0.04944</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.030416</td>
<td>90%</td>
<td>0.036181</td>
<td>94%</td>
<td>0.04892</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.029935</td>
<td>92%</td>
<td>0.035670</td>
<td>96%</td>
<td>0.04687</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.00</td>
<td>ASCLT #1</td>
<td>0.032436</td>
<td>89%</td>
<td>0.037919</td>
<td>95%</td>
<td>0.048353</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.053520</td>
<td>89%</td>
<td>0.041733</td>
<td>97%</td>
<td>0.057913</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.032520</td>
<td>90%</td>
<td>0.040211</td>
<td>95%</td>
<td>0.05381</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.032900</td>
<td>93%</td>
<td>0.039202</td>
<td>95%</td>
<td>0.051519</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>p</th>
<th>Method</th>
<th>Ave CI Width</th>
<th>% Containing</th>
<th>90% CI Ave CI</th>
<th>% Containing</th>
<th>95% CI Ave CI</th>
<th>% Containing</th>
<th>99% CI Ave CI</th>
<th>% Containing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.053097</td>
<td>89%</td>
<td>0.062554</td>
<td>96%</td>
<td>0.078919</td>
<td>97%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.049309</td>
<td>93%</td>
<td>0.050033</td>
<td>88%</td>
<td>0.049905</td>
<td>90%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.035373</td>
<td>91%</td>
<td>0.065382</td>
<td>97%</td>
<td>0.08709</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.053167</td>
<td>93%</td>
<td>0.063068</td>
<td>96%</td>
<td>0.08283</td>
<td>97%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.50</td>
<td>ASCLT #1</td>
<td>0.076189</td>
<td>88%</td>
<td>0.093893</td>
<td>93%</td>
<td>0.11920</td>
<td>97%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.081026</td>
<td>90%</td>
<td>0.099779</td>
<td>97%</td>
<td>0.106508</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.081574</td>
<td>93%</td>
<td>0.095971</td>
<td>96%</td>
<td>0.128602</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.078449</td>
<td>89%</td>
<td>0.093740</td>
<td>96%</td>
<td>0.12183</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.30</td>
<td>ASCLT #1</td>
<td>0.092195</td>
<td>88%</td>
<td>0.109691</td>
<td>94%</td>
<td>0.14092</td>
<td>98%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.098833</td>
<td>90%</td>
<td>0.121687</td>
<td>98%</td>
<td>0.15486</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.097220</td>
<td>92%</td>
<td>0.116314</td>
<td>95%</td>
<td>0.15469</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.094838</td>
<td>94%</td>
<td>0.112935</td>
<td>95%</td>
<td>0.14801</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.00</td>
<td>ASCLT #1</td>
<td>0.109044</td>
<td>94%</td>
<td>0.118205</td>
<td>94%</td>
<td>0.14759</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.109168</td>
<td>92%</td>
<td>0.134456</td>
<td>96%</td>
<td>0.18742</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.106691</td>
<td>91%</td>
<td>0.127053</td>
<td>96%</td>
<td>0.17033</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.104088</td>
<td>92%</td>
<td>0.124020</td>
<td>94%</td>
<td>0.16293</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>p</th>
<th>Method</th>
<th>Ave CI Width</th>
<th>% Containing</th>
<th>90% CI Ave CI</th>
<th>% Containing</th>
<th>95% CI Ave CI</th>
<th>% Containing</th>
<th>99% CI Ave CI</th>
<th>% Containing</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.168134</td>
<td>84%</td>
<td>0.208358</td>
<td>91%</td>
<td>0.23466</td>
<td>92%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.155859</td>
<td>87%</td>
<td>0.154536</td>
<td>89%</td>
<td>0.16490</td>
<td>89%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.179894</td>
<td>88%</td>
<td>0.218637</td>
<td>94%</td>
<td>0.31215</td>
<td>98%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.172117</td>
<td>93%</td>
<td>0.208611</td>
<td>95%</td>
<td>0.27464</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>0.50</td>
<td>ASCLT #1</td>
<td>0.232041</td>
<td>94%</td>
<td>0.290611</td>
<td>93%</td>
<td>0.34472</td>
<td>94%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.253401</td>
<td>95%</td>
<td>0.301355</td>
<td>90%</td>
<td>0.30513</td>
<td>97%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.262222</td>
<td>88%</td>
<td>0.316835</td>
<td>96%</td>
<td>0.43985</td>
<td>98%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.252295</td>
<td>91%</td>
<td>0.301170</td>
<td>97%</td>
<td>0.40037</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>0.30</td>
<td>ASCLT #1</td>
<td>0.281420</td>
<td>93%</td>
<td>0.332225</td>
<td>94%</td>
<td>0.41782</td>
<td>98%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.302794</td>
<td>91%</td>
<td>0.380784</td>
<td>94%</td>
<td>0.47672</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.315035</td>
<td>90%</td>
<td>0.375174</td>
<td>92%</td>
<td>0.53353</td>
<td>96%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.303604</td>
<td>88%</td>
<td>0.360629</td>
<td>96%</td>
<td>0.46958</td>
<td>97%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.00</td>
<td>ASCLT #1</td>
<td>0.306676</td>
<td>92%</td>
<td>0.356251</td>
<td>94%</td>
<td>0.43818</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.336245</td>
<td>92%</td>
<td>0.421554</td>
<td>95%</td>
<td>0.57336</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.350691</td>
<td>89%</td>
<td>0.417566</td>
<td>93%</td>
<td>0.58080</td>
<td>98%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.330683</td>
<td>89%</td>
<td>0.39293</td>
<td>92%</td>
<td>0.51566</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5.2: Simulation Results for the Bivariate Exponential Distribution

<table>
<thead>
<tr>
<th>N</th>
<th>ρ</th>
<th>Method</th>
<th>Ave CI Width</th>
<th>% Containing ρ</th>
<th>Ave CI Width</th>
<th>% Containing ρ</th>
<th>Ave CI Width</th>
<th>% Containing ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>10,000</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.032065</td>
<td>86%</td>
<td>0.038767</td>
<td>89%</td>
<td>0.051489</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.033221</td>
<td>88%</td>
<td>0.035670</td>
<td>98%</td>
<td>0.060015</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.038090</td>
<td>88%</td>
<td>0.045206</td>
<td>93%</td>
<td>0.025911</td>
<td>75%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.016440</td>
<td>57%</td>
<td>0.019685</td>
<td>58%</td>
<td>0.025911</td>
<td>75%</td>
</tr>
<tr>
<td>0.50</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.036373</td>
<td>82%</td>
<td>0.043257</td>
<td>92%</td>
<td>0.055872</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.041423</td>
<td>92%</td>
<td>0.050924</td>
<td>90%</td>
<td>0.060220</td>
<td>94%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.041533</td>
<td>96%</td>
<td>0.050045</td>
<td>96%</td>
<td>0.066624</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.024627</td>
<td>68%</td>
<td>0.029448</td>
<td>78%</td>
<td>0.038657</td>
<td>91%</td>
</tr>
<tr>
<td>0.25</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.036366</td>
<td>91%</td>
<td>0.042552</td>
<td>88%</td>
<td>0.055353</td>
<td>98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.040480</td>
<td>83%</td>
<td>0.048773</td>
<td>96%</td>
<td>0.067998</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.040269</td>
<td>86%</td>
<td>0.048058</td>
<td>94%</td>
<td>0.065025</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.030865</td>
<td>77%</td>
<td>0.036735</td>
<td>86%</td>
<td>0.048334</td>
<td>96%</td>
</tr>
<tr>
<td>0.00</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.032495</td>
<td>87%</td>
<td>0.038286</td>
<td>97%</td>
<td>0.048393</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.034292</td>
<td>91%</td>
<td>0.041746</td>
<td>97%</td>
<td>0.056116</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.033691</td>
<td>88%</td>
<td>0.040075</td>
<td>97%</td>
<td>0.053782</td>
<td>98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.032899</td>
<td>94%</td>
<td>0.039202</td>
<td>98%</td>
<td>0.051517</td>
<td>99%</td>
</tr>
<tr>
<td>1,000</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.092277</td>
<td>78%</td>
<td>0.113406</td>
<td>92%</td>
<td>0.150398</td>
<td>98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.092671</td>
<td>84%</td>
<td>0.103837</td>
<td>87%</td>
<td>0.108079</td>
<td>91%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.122329</td>
<td>90%</td>
<td>0.141078</td>
<td>92%</td>
<td>0.198151</td>
<td>96%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.052848</td>
<td>50%</td>
<td>0.062540</td>
<td>57%</td>
<td>0.082296</td>
<td>69%</td>
</tr>
<tr>
<td>0.50</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.107056</td>
<td>82%</td>
<td>0.126401</td>
<td>87%</td>
<td>0.162764</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.127498</td>
<td>82%</td>
<td>0.155255</td>
<td>96%</td>
<td>0.177400</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.134615</td>
<td>93%</td>
<td>0.163741</td>
<td>95%</td>
<td>0.217820</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.077547</td>
<td>67%</td>
<td>0.093360</td>
<td>79%</td>
<td>0.122973</td>
<td>93%</td>
</tr>
<tr>
<td>0.25</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.110484</td>
<td>87%</td>
<td>0.127638</td>
<td>94%</td>
<td>0.164030</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.126770</td>
<td>92%</td>
<td>0.153679</td>
<td>95%</td>
<td>0.205492</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.132823</td>
<td>93%</td>
<td>0.154037</td>
<td>96%</td>
<td>0.205607</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.097068</td>
<td>78%</td>
<td>0.116112</td>
<td>91%</td>
<td>0.152433</td>
<td>95%</td>
</tr>
<tr>
<td>0.00</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.101428</td>
<td>93%</td>
<td>0.118141</td>
<td>94%</td>
<td>0.148526</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.107980</td>
<td>92%</td>
<td>0.134109</td>
<td>94%</td>
<td>0.169666</td>
<td>97%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.108364</td>
<td>90%</td>
<td>0.129499</td>
<td>96%</td>
<td>0.173068</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.104082</td>
<td>91%</td>
<td>0.124023</td>
<td>96%</td>
<td>0.162963</td>
<td>100%</td>
</tr>
<tr>
<td>100</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.222324</td>
<td>67%</td>
<td>0.288143</td>
<td>77%</td>
<td>0.335237</td>
<td>86%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.214935</td>
<td>70%</td>
<td>0.225367</td>
<td>73%</td>
<td>0.262033</td>
<td>75%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.448585</td>
<td>84%</td>
<td>0.555880</td>
<td>86%</td>
<td>0.836366</td>
<td>96%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.167347</td>
<td>51%</td>
<td>0.204889</td>
<td>50%</td>
<td>0.261185</td>
<td>72%</td>
</tr>
<tr>
<td>0.50</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.297121</td>
<td>87%</td>
<td>0.352425</td>
<td>92%</td>
<td>0.417689</td>
<td>92%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.330797</td>
<td>89%</td>
<td>0.387804</td>
<td>91%</td>
<td>0.445493</td>
<td>92%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.419721</td>
<td>90%</td>
<td>0.527652</td>
<td>88%</td>
<td>0.802283</td>
<td>96%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.256149</td>
<td>66%</td>
<td>0.303418</td>
<td>85%</td>
<td>0.386621</td>
<td>87%</td>
</tr>
<tr>
<td>0.25</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.314982</td>
<td>92%</td>
<td>0.368330</td>
<td>96%</td>
<td>0.442280</td>
<td>91%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.371677</td>
<td>88%</td>
<td>0.433830</td>
<td>87%</td>
<td>0.547076</td>
<td>94%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.429236</td>
<td>91%</td>
<td>0.517461</td>
<td>93%</td>
<td>0.713506</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.309720</td>
<td>83%</td>
<td>0.369116</td>
<td>95%</td>
<td>0.484182</td>
<td>95%</td>
</tr>
<tr>
<td>0.00</td>
<td>0.70</td>
<td>ASCLT #1</td>
<td>0.330087</td>
<td>90%</td>
<td>0.350719</td>
<td>92%</td>
<td>0.411019</td>
<td>94%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.330448</td>
<td>93%</td>
<td>0.387270</td>
<td>93%</td>
<td>0.478400</td>
<td>93%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.335042</td>
<td>89%</td>
<td>0.448995</td>
<td>94%</td>
<td>0.643136</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.330028</td>
<td>86%</td>
<td>0.393099</td>
<td>96%</td>
<td>0.515515</td>
<td>98%</td>
</tr>
</tbody>
</table>
### Table 5.3: Simulation Results for the Bivariate Poisson Distribution

<table>
<thead>
<tr>
<th>N</th>
<th>ρ</th>
<th>Method</th>
<th>90% CI Ave Cl Width</th>
<th>% Containing ρ</th>
<th>95% CI Ave Cl Width</th>
<th>% Containing ρ</th>
<th>99% CI Ave Cl Width</th>
<th>% Containing ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>10,000</td>
<td>0.00</td>
<td>ASCLT #1</td>
<td>0.032594</td>
<td>87%</td>
<td>0.037923</td>
<td>92%</td>
<td>0.048187</td>
<td>98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.034468</td>
<td>93%</td>
<td>0.041558</td>
<td>91%</td>
<td>0.057794</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.033253</td>
<td>91%</td>
<td>0.040084</td>
<td>96%</td>
<td>0.053505</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.032899</td>
<td>87%</td>
<td>0.039201</td>
<td>95%</td>
<td>0.051517</td>
<td>99%</td>
</tr>
<tr>
<td>1,000</td>
<td>0.00</td>
<td>ASCLT #1</td>
<td>0.101176</td>
<td>87%</td>
<td>0.118072</td>
<td>95%</td>
<td>0.147901</td>
<td>98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.108233</td>
<td>87%</td>
<td>0.133967</td>
<td>91%</td>
<td>0.185813</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.106671</td>
<td>89%</td>
<td>0.129034</td>
<td>96%</td>
<td>0.170738</td>
<td>99%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.104085</td>
<td>92%</td>
<td>0.123997</td>
<td>93%</td>
<td>0.162918</td>
<td>100%</td>
</tr>
<tr>
<td>100</td>
<td>0.00</td>
<td>ASCLT #1</td>
<td>0.305960</td>
<td>91%</td>
<td>0.358225</td>
<td>95%</td>
<td>0.433465</td>
<td>98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ASCLT #2</td>
<td>0.333065</td>
<td>86%</td>
<td>0.423958</td>
<td>95%</td>
<td>0.535977</td>
<td>98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bootstrap</td>
<td>0.342812</td>
<td>90%</td>
<td>0.424730</td>
<td>93%</td>
<td>0.592401</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classic</td>
<td>0.330644</td>
<td>91%</td>
<td>0.394338</td>
<td>97%</td>
<td>0.516767</td>
<td>100%</td>
</tr>
</tbody>
</table>
Bibliography


Galton, F., Regression towards mediocrity in hereditary stature, Journal of the Anthropological Institute, 15, 246-263, 1885.


Fisher, R.A., Frequency distribution of the values of the correlation coefficient in samples of an indefinitely large population, Biometrika, 10, 507–521, 1915.

Fisher, R.A., On the `probable error' of a coefficient of correlation deduced from a small sample, Metron, 1, 3–32, 1921.


