1967

Tests of goodness of fit based on discriminatory information

James Robert Gebert
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd
Part of the Statistics and Probability Commons

Recommended Citation
Gebert, James Robert, "Tests of goodness of fit based on discriminatory information " (1967). Retrospective Theses and Dissertations. 3154.
https://lib.dr.iastate.edu/rtd/3154

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
GEBERT, James Robert, 1939-
TESTS OF GOODNESS OF FIT BASED ON DISCRIMINATORY INFORMATION.

Iowa State University of Science and Technology,
Ph.D., 1967
Statistics

University Microfilms, Inc., Ann Arbor, Michigan
TESTS OF GOODNESS OF FIT BASED ON
DISCRIMINATORY INFORMATION

by

James Robert Gebert

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1967
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>INTRODUCTION</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.1. Goodness of Fit</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.2. Goodness of Fit Based on Spacings</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>1.3. Information Statistics</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>1.4. Kimball's and Greenwood's Statistics</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>1.5. Other Tests Based on Spacings</td>
<td>12</td>
</tr>
<tr>
<td>2.</td>
<td>MOMENTS OF THE TEST STATISTICS</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.1. Introduction</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2.2. Some Distributional Properties of Uniform</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Spacings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.3. Moments of Kimball's Statistics and Kale's</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Statistic k</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.4. Darling's Method</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>2.5. Some Integrals Derived from the Beta and</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Gamma Functions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.6. Mean and Variance of k*</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>2.7. Mean and Variance of J</td>
<td>39</td>
</tr>
<tr>
<td>3.</td>
<td>LIMITING DISTRIBUTIONS</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>3.1. Introduction</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>3.2. Asymptotic Normality of k and Kimball's</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>Statistics</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.3. Le Cam's Method</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>3.4. Asymptotic Normality of the Statistic k*</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>3.5. Asymptotic Normality of the Statistic J</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>3.6. Comparison of Moments and Critical Points</td>
<td>56</td>
</tr>
</tbody>
</table>
iii

4. POWER UNDER WEISS' ALTERNATIVES
   4.1. Introduction 61
   4.2. Weiss' Theorem 61
   4.3. Goodness of Fit Problem and Test Statistics 65
   4.4. Power of the Information Statistics 65
   4.5. Power of Kimball's Statistics 93
   4.6. Summary 104
5. A SAMPLING EXPERIMENT 107
6. LITERATURE CITED 110
7. ACKNOWLEDGMENTS 114
1. INTRODUCTION

1.1 Goodness of Fit

A common problem in Statistics is to test whether a random sample $X_1, X_2, \ldots, X_n$ comes from a population having a cumulative distribution function (c.d.f.) $F_0$, where $F_0$ is a completely specified c.d.f.; i.e. we propose to test the hypothesis

$$H_0: F = F_0$$

against the alternatives

$$H_A: F \neq F_0$$

where $F$ is the c.d.f. of the population from which the sample $X_1, X_2, \ldots, X_n$ is drawn. Choosing a test of goodness of fit can be interpreted as deciding in what sense the discrepancy between the population and the sample is to be measured.

Statistics to test $H_0$ are of two types; tests of structure (d) and tests which are not of structure (d). A test of $H_0$ was said to be of structure (d) in Birnbaum (1953) if it is based on $\{U_i = F_0(X_i), i=1, \ldots, n\}$, which under $H_0$ are distributed as Uniform random variables. Such tests are also said to be distribution-free in the sense that the distribution of the test statistic given that $H_0$ is true is the same for all continuous c.d.f.'s $F_0$.

Tests of structure (d) are mainly of two types: those based on empirical distribution functions and those based on
The tests based on empirical distribution functions use some measurement of the distance between the proposed cumulative distribution function \( F_0 \) and the empirical distribution function as its discrepancy measure. Examples are Cramer-von Mises and Kolmogorov-Smirnov tests. These tests yield good minimum power over classes of alternatives \( F^* \) satisfying: distance \( (F^*, F_0) \geq \delta \). As pointed out by Massey (1951), these tests are consistent but biased. A good survey article on Cramer-von Mises and Kolmogorov-Smirnov type tests is Darling (1957).

Goodness of fit tests based on sample spacings will be considered in the next section.

Tests that are not of structure (d) are either \( \chi^2 \)-type tests or tests that are geared to special properties of particular distribution functions. One example of the latter is the test for Normality based on the sample estimate of kurtosis. Since there are many distributions having \( \beta_2 = 3 \), this test would not distinguish between one of these distributions and the Normal distribution. A good survey of \( \chi^2 \)-type tests is Cochran (1952).
1.2. Goodness of Fit Based on Spacings

Let $X_1, X_2, \ldots, X_n$ be a random sample from a distribution with a continuous cumulative distribution function (c.d.f.) $F(x)$. Let $-\infty = X_{(0)} < X_{(1)} < \ldots < X_{(n)} < X_{(n+1)} = \infty$ denote the order statistics obtained by arranging the $X_i$'s in increasing order. Then the sample spacings $\{V_i: i=1, \ldots, n+1\}$ are defined by

$$V_i = F_0(X_{(i)}) - F_0(X_{(i-1)}), \quad i=1, \ldots, n+1.$$  

Then under $H_0$ the $V_i$ are distributed as sample spacings (or sample coverages) from the Uniform distribution.

Although the main interest of this thesis is in spacings as they arise in the context of distribution-free tests of goodness of fit, it should be emphasized that the first studies of Uniform spacings were concerned with the randomness of a series of events, and were motivated by the fact that the intervals between successive events of a Poisson process, conditioned on the number of events in a specified interval, are distributed like Uniform spacings. The earliest studies along this line were by Whitworth (1887), Bortkiewicz (1915) and Morant (1920), a brief discussion of which is contained in Appendix I of Greenwood (1946). Other references along this line are Sukhatme (1936, 1937), who coined the phrase interval analysis to distinguish it from the more frequently used counting analysis. A more modern discussion of series
of events is contained in Cox and Lewis (1966).

The main difference between the tests based on spacings and those based on the empirical distribution, as mentioned in Pyke (1965), lies in the fact that "tests based on the empirical distribution function are only sensitive to significant changes in distribution functions, whereas tests based on spacings are designed to detect differences between density functions (or between rate functions \( \lambda(t) \) in the context of series of events.)" Closeness of the true distribution function to \( F_0 \) in no way implies closeness between the corresponding densities.

The tests that we will consider will be based on ordered spacings for the reason that, just as many tests based on order statistics, or equivalently on the empirical distribution function, are invariant under permutations of the original observations, it is natural to ask that the tests based on spacings should be invariant under permutations of the spacings, the spacings being interchangeable random variables under \( H_0 \).

Most of the theory of spacings crystallized from Greenwood's 1946 invited paper read before the Royal Statistical Society. Greenwood proposed a spacings test for randomness of a series of events in order to decide whether a disease was contagious or not. Some alternative tests were proposed
in the discussion of the paper and it was studies of these tests that constituted the research on spacings until 1953.

In 1953, Darling reviewed all the work obtained between 1946 and 1953. He gave a method for finding any moment of a function of sample spacings. He also presented a unified approach to the distribution theory of Uniform spacings through the mathematically complicated technique of Steepest Descents.

In a series of papers, Weiss (1955-1965) continued the work on the goodness of fit aspect of spacings. He investigated the power of some tests based on spacings. Finally in his 1965 paper he gave a method for finding the power of tests based on spacings against a class of alternatives $F_n$, where $\lim_{n \to \infty} F_n = F_0$, $n$ being the sample size. In an invited address before the Royal Statistical Society, Pyke (1965) delivered a survey article on spacings comparable to Cochran's (1952) $\chi^2$ survey and Darling's (1957) survey article on Kolmogorov-Smirnov and Cramer-von Mises type statistics.

1.3. Information Statistics

Godambe (1961) derived a test for the two sample problem

$H_0: F = G$

$H_A: F \neq G$.

where both $F$ and $G$ are unspecified, everywhere continuous distribution functions, where the alternatives have been
restricted to the class $G = \theta(F)$, where $\theta$ is a function from $(0,1)$ to $(0,1)$ and such that $\theta' = \partial \theta / \partial F$ exists. Then from the fact that the most powerful rank test depends on

$$
\phi(a|v) = \frac{m!}{n+1} \prod_{i=1}^{n+1} \frac{a_i}{v_i} 
$$

where $a = (a_1, \ldots, a_{n+1})$ and $a_i$ is the number of observations from the second sample (the number of $y$'s) lying between $X(i-1)$ and $X(i)$, the test criterion

$$
\phi(a) = \frac{m!}{n+1} \prod_{i=1}^{n+1} \frac{a_i}{m} \prod_{i=1}^{n+1} a_i
$$

with critical region $\phi(a) > \text{constant}$, was proposed.

Then using a technique similar to Moses (1964) we can consider the one sample test derived from the above two sample test by keeping $n$ fixed while we let $m$ tend to infinity.

$$
\log \phi(a) = \log(m!) + \sum_{i=1}^{n+1} a_i \log a_i - m \log m
$$

$$
- \sum_{i=1}^{n+1} \log (a_i!)
$$

Using Sterling's formula

$$
\log m! = \frac{1}{2} \log 2\pi m + m \log m - m + \frac{\theta m}{12m}
$$
where $|\theta_m| < 1$, we get

$$\log \phi(a) = \frac{1}{2} \log 2\pi m - m + \frac{\theta_m}{12m} - \sum_{i=1}^{n+1} \left[ \frac{1}{2} \log 2\pi a_i \right.$$

$$\left. - a_i + \frac{\theta a_i}{12a_i} \right]$$

$$= \frac{1}{2} \log 2\pi + \frac{1}{2} \log m + \frac{\theta m}{12m}$$

$$- \frac{(n+1)}{2} \log 2\pi - \frac{1}{2} \sum_{i=1}^{n+1} \log a_i - \sum_{i=1}^{n+1} \frac{\theta a_i}{12a_i}$$

$$= - \frac{n}{2} \log 2\pi - \frac{n}{2} \log m - \frac{1}{2} \sum_{i=1}^{n+1} \log \frac{a_i}{m} + \frac{\theta m}{12m}$$

$$- \sum_{i=1}^{n+1} \frac{\theta a_i}{12a_i}$$

$$= - \frac{n}{2} \log 2\pi - \frac{n}{2} \log m - \frac{1}{2} \sum_{i=1}^{n+1} \log \frac{a_i}{m} + O \left( \frac{1}{m} \right).$$

\[.\] For $m$ large

$$\log \phi(a) = - \frac{n}{2} \log 2\pi - \frac{n}{2} \log m - \frac{1}{2} \sum_{i=1}^{n+1} \log V_i$$

where $V_i = F(X(i)) - F(X(i-1)) = \lim_{m \to \infty} a_i$.

Godambe (1965) proposed the goodness of fit test

$$k' = \sum_{i=1}^{n+1} \log V_i = \sum_{i=1}^{n+1} \log \{F_o(X(i)) - F_o(X(i-1))\}$$
because for m large with respect to n, the tests \( \phi(a) \) and \( k' \) act equivalently.

Kale (1965) proposed two information statistics

\[
k = \sum_{i=1}^{n+1} \frac{1}{n+1} \log \frac{n+1}{V_i}
\]

and

\[
k* = \sum_{i=1}^{n+1} V_i \log \frac{V_i}{1/n+1}
\]

The justification for these two test criterions lies in the fact that if we consider two multinomial distributions

1: \( P_{in} = \frac{1}{n+1} \quad i=1,2,...,n+1 \)

2: \( P_{in} = V_i = F_0(X(i)) - F_0(X(i-1)) \quad i=1,2,...,n+1, \)

then \( k \) is the mean information per observation from the population hypothesized by \( H_1 \) for discriminating for \( H_1 \) against \( H_2 \) and \( k* \) is the mean information per observation from the population hypothesized by \( H_2 \), for discriminating for \( H_2 \) against \( H_1 \). It should also be mentioned that under

\[ H_0: \quad F = F_0, \quad E(V_i) = 1/(n+1), \quad i=1,2,...,n+1. \]

As stated in Kale (1965) and Kullback (1959), (pp. 113-4); these information statistics \( k* \) and \( k \) are related to Pearson's \( \chi^2 \) and Neyman's \( \chi^{'2} \) respectively. This can be seen from the following result.

For \( a/b > 0 \)
\[
\frac{(a-b)}{a} \leq \log \frac{a}{b} \leq \frac{(a-b)}{b}
\]

where the equalities hold iff \(a=b\). We may therefore use as a first approximation to \(\log (a/b)\), the mean of its upper and lower bounds, that is

\[
\log \frac{a}{b} = \frac{1}{2} \left[ \frac{a^2-b^2}{ab} \right]
\]

the approximation being better the closer \(a/b\) is to 1. Using this approximation with \(k^*\), we get

\[
k^* = \frac{n+1}{n+1} \sum_{i=1}^{n+1} v_i \log \frac{v_i}{n+1} = \frac{n+1}{n+1} \sum_{i=1}^{n+1} \frac{v_i^2 - \left(\frac{1}{n+1}\right)^2}{v_i}
\]

\[
= \frac{(n+1)}{2} \sum_{i=1}^{n+1} \left[v_i^2 - \left(\frac{1}{n+1}\right)^2\right] = \frac{(n+1)}{2} \sum_{i=1}^{n+1} \left(v_i - \frac{1}{n+1}\right)^2
\]

since \(\sum_{i=1}^{n+1} v_i = 1\). Therefore \(k^* = \frac{1}{2} \sum_{i=1}^{n+1} \left(v_i - \frac{1}{n+1}\right)^2\).

It is also worthy of mention that \(\frac{1}{2} \sum_{i=1}^{n+1} \left(v_i - \frac{1}{n+1}\right)^2\) is the leading term in a Taylor series expansion expanding each \(v_i \log v_i\) about \(1/(n+1) = v_0\). We have therefore

\[
k^* = \frac{(n+1)}{2} \sum_{i=1}^{n+1} v_i^2 - \frac{1}{2}
\]  (1.3.1)
On the other hand if we use the approximation

$$\log \frac{a}{b} \approx \frac{1}{2} \left[ \frac{a^2 - b^2}{ab} \right]$$

with the test statistic $k$, we have

$$k = \sum_{i=1}^{n+1} \frac{1}{n+1} \log \frac{\frac{1}{n+1}}{V_i} = \sum_{i=1}^{n+1} \frac{1}{2(n+1)} \left[ \frac{(\frac{1}{n+1})^2 - V_i^2}{V_i} \right]$$

$$= \frac{1}{2} \sum_{i=1}^{n+1} \left( \frac{(\frac{1}{n+1})^2 - V_i^2}{V_i} \right) = \frac{1}{2} \sum_{i=1}^{n+1} \left( \frac{V_i - \frac{1}{n+1}}{V_i} \right)^2 = \chi^2. \quad (1.3.2)$$

The $k$ statistic is a linear function of the statistic

$$k' = \sum_{i=1}^{n+1} \log V_i$$

of Godambe. The statistic $k'$ was first proposed by Darling (1953) who found its moments and asymptotic Normality under $H_0$.

The statistic $J = k + k^*$ has its justification in the information concept of the Divergence between the information indices $k$ and $k^*$. It is also similar in spirit to Kuiper's (1960) statistic

$$d_v(n) = \sqrt{n} \left[ \max_i \left( \frac{i}{n} - U_i \right) + \max_i \left( U_i - \frac{i}{n} \right) \right],$$

which is the sum of two one sided Kolmogorov-Smirnov test statistics.
1.4. Kimball's and Greenwood's Statistics

Kimball (1950) proposed the test statistics

$$G_n = \sum_{i=1}^{n+1} V_i^r$$

for \( r > 0 \). He obtained the limit distribution under \( H_0 : F = F_o \). Weiss (1957) studied the limit distribution under \( H_A : F \neq F_o \). Darling (1953) gave an alternative derivation of the limiting distribution under \( H_0 \).

A special case of Kimball's statistic is Greenwood's (1946) statistic

$$\sum_{i=1}^{n+1} V_i^2$$

The limiting distribution function under \( H_0 \) was obtained by Moran (1947). From (1.3.1), we see that \( k^* \) is approximately a linear function of Greenwood's statistic.

In succeeding chapters the statistics \( k, k^*, J \) and Kimball's statistics will be studied. In chapter 2, the first two moments of these four statistics are given. In chapter 3, the asymptotic Normality of the four statistics is proved. Le Cam's method is used to prove the asymptotic Normality of the statistics \( k^* \) and \( J \). Using the results of chapter 2 along with the asymptotic Normality, critical regions based on these statistics are obtained. In chapter 4, the power of these four statistics is studied against alternatives introduced in Weiss (1965). Finally it is proved that Greenwood's statistic is the most powerful of these four statistics against Weiss' alternatives. In chapter 5, a small
sampling experiment on the power of these four statistics is presented.

1.5. Other Tests Based on Spacings

Most of the tests that have been considered can be written in the form

\[ G_n = \sum_{i=1}^{n+1} g_n(V_i) \]

where \( g_n \) is a Borel-measurable function. Either Darling's or Le Cam's techniques can be used to find the limiting distributions of statistics of this form.

Some of the statistics that have been proposed are

\( g_n(x) = \{x - \frac{1}{n+1}\}^2 \), suggested by Irwin in Greenwood (1946). This statistic is linearly related to Greenwood's statistic and it has the advantage it can be given a \( \chi^2 \) justification.

\( g_n(x) = |x - \frac{1}{n+1}| \), suggested by M.G. Kendall in the discussion of Greenwood (1946), limit distribution under \( H_o \) obtained by Sherman (1950).

\( g_n(x) = \frac{1}{x} \), suggested by Darling (1953), who derived its non-Normal limit distribution. From (1.3.2) we see that

\[ k = \frac{1}{2(n+1)^2} \sum_{i=1}^{n+1} \frac{1}{V_i} - \frac{1}{2}. \]

\( k \) is related to the statistic \( \sum_{i=1}^{n+1} \frac{1}{V_i} \). The difference between the limiting distributions may be attributed to the
nature of the approximation involved and the norming constants involved.
2. MOMENTS OF THE TEST STATISTICS

2.1. Introduction

In this Chapter, the first two moments of the statistics \( k, \ k^*, \ J \) and Kimball's statistic will be given under the hypothesis that \( V_i \) are Uniform Spacings.

2.2. Some Distributional Properties of Uniform Spacings

Uniform spacings have the singular distribution

\[
f_{V}(v_1, v_2, \ldots, v_{n+1}) = \begin{cases} 
  n! \text{ if } v_i > 0 \text{ and } \sum_{i=1}^{n+1} v_i = 1 \\
  0 \text{ otherwise.}
\end{cases}
\]  

Uniform spacings are interchangeable random variables. This implies the distribution function of any \( V_i \) is equal to that of \( V_1 \) and the joint distribution function of any pair \( (V_i, V_j) \) \( (i \neq j) \), is the same as that of \( (V_1, V_2) \). Therefore for \( x, y > 0 \) and \( x + y < 1 \) we have

\[
F_{V_i}(x) = F_{V_1}(x) = F_X(x) = 1 - (1-x)^n
\]

and

\[
F_{V_i,V_j}(x,y) = 1 - \{(1-x)^n + (1-y)^n - (1-x-y)^n\}.
\]

Equivalently, the corresponding density functions are

\[
f_{V_i}(x) = n(1-x)^{n-1}
\]
\[ f(v_i, v_j)(x, y) = n(n-1)(1-x-y)^{n-2}. \quad (2.2.2) \]

Then we can derive

\[ E(V_i) = \frac{1}{n+1} \quad E(V_i V_j) = \frac{1}{(n+1)(n+2)} \quad (i \neq j) \]

\[ E(V_i^2) = \frac{2}{(n+1)(n+2)} \quad \text{var}(V_i) = \frac{n}{(n+1)^2(n+2)} \]

and for \( i \neq j \)

\[ \text{Cov}(V_i, V_j) = \frac{-1}{(n+1)^2(n+2)} \quad \text{Corr}(V_i, V_j) = -\frac{1}{n}. \]

The following Lemma will be extremely important for Le Cam's method of finding the limiting distribution of a sum

\[ G_n = \sum_{i=1}^{n+1} g_n((n+1) V_i). \]

Lemma 2.2.1

Let \( Y_1, Y_2, \ldots, Y_{n+1} \) be independent Exponential random variables with mean 1 and let

\[ S = Y_1 + Y_2 + \ldots + Y_{n+1}. \]

Then the conditional distribution function of \( (Y_1, Y_2, \ldots, Y_{n+1}) \), given \( S = n+1 \), is the same that of \( n+1 \) normed Uniform spacings.

\[ (n+1)V_1, \ldots, (n+1)V_{n+1}. \]
Proof: 

\[ f(y_1, y_2, \ldots, y_{n+1}) (y_1, y_2, \ldots, y_{n+1}) = e^{-\sum_{i=1}^{n+1} y_i} \]

Therefore

\[ f(y_1, y_2, \ldots, y_n, s) (y_1, y_2, \ldots, y_n, s) = e^{-s} \cdot \]

Let \( \phi_s(t) \) and \( \{\phi_{y_i}(t)\} \) denote the characteristic functions of \( S \) and \( \{y_i\} \) respectively. Then

\[ \phi_s(t) = \prod_{i=1}^{n+1} \phi_{y_i}(t) = (\phi_{y}(t))^{n+1} \]

where

\[ \phi_{y}(t) = \int_0^{\infty} e^{ity} e^{-y} dy = \int_0^{\infty} e^{-y(1-it)} dy = \frac{1}{1-it} \cdot \]

Therefore \( \phi_s(t) = \frac{1}{(1-it)^{n+1}} \) which is recognized as the characteristic function of a Gamma distribution with parameter \( n+1 \). Therefore

\[ f_s(s) = \frac{sn^se^{-s}}{n!} \]

and so

\[ f(y_1, y_2, \ldots, y_n|S) (y_1, y_2, \ldots, y_n|s) = n!s^{-n} \cdot \]

In particular

\[ f(y_1, y_2, \ldots, y_n|S) (y_1, y_2, \ldots, y_{n|n+1}) = n!(n+1)^{-n}. \]

This is just the distribution that can be derived from (2.2.1)
if we let \( z_i = (n+1)v_i \) and make a change of variable.

### 2.3. Moments of Kimball’s Statistics and Kale’s Statistic \( k \)

Kale in his 1965 paper gave the expressions

\[
E(k) = -[\psi(1) - \psi(n+1)] - \log (n+1) \tag{2.3.1}
\]

\[
V(k) = \frac{\psi'(1)}{n+1} - \psi'(n+1)
\]

where

\[
\psi(x) = \frac{d}{dx} \log \Gamma(x),
\]

he also gave the approximations

\[
E(k) = C - \frac{1}{2(n+1)} + o \left( \frac{1}{(n+1)} \right)
\]

\[
\sigma^2 = \frac{1}{n+1} \left( \frac{n^2}{6} - 1 \right) - \frac{1}{2(n+1)^2} + o \left( \frac{1}{(n+1)^2} \right) \tag{2.3.2}
\]

These results could also be derived from Darling’s (1953) results on his statistic \( k' \).

We next consider Kimball’s statistic \( \sum_{i=1}^{n+1} v_i^r \). Darling (1953) gave the following formulae for its two moments.

\[
\mu_1 = \frac{\Gamma(r+1)}{(n+2)^{r-1}} - \frac{\Gamma(r+1) [(r-1)(r-2)]}{2(n+2)^r} + o \left( \frac{1}{n^r} \right)
\]

Expanding in powers of \( \frac{1}{n+1} \) we get
\[ \mu_1 = \frac{\Gamma(r+1)}{(n+1)^{r-1}} \left( 1 + \frac{1}{n+1} \right)^{-r} \left( r \right) \frac{\Gamma(r+1)}{2(n+1)^{r}} \left[ (r-1)(r-2) \right] \]

\[ = \frac{\Gamma(r+1)}{(n+1)^{r-1}} \left( 1 - \frac{r-1}{n+1} + 0 \left( \frac{1}{n+1} \right)^2 \right) \frac{\Gamma(r+1)}{2(n+1)^{r}} \]

\[ = \frac{\Gamma(r+1)}{(n+1)^{r-1}} \left( 1 - \frac{r-1}{n+1} + 0 \left( \frac{1}{n+1} \right)^2 \right) \frac{\Gamma(r+1)}{2(n+1)^{r}} \]

\[ = \frac{\Gamma(r+1)}{(n+1)^{r-1}} \left( 1 - \right) \frac{\Gamma(r+1)}{2(n+1)^{r}} \]

\[ = \frac{\Gamma(r+1)}{(n+1)^{r-1}} \left( 1 - \right) \frac{\Gamma(r+1)}{2(n+1)^{r}} \]

\[ = \frac{\Gamma(r+1)}{(n+1)^{r-1}} \left( 1 - \right) \frac{\Gamma(r+1)}{2(n+1)^{r}} \]

\[ \sigma_n^2 = \frac{1}{(n+1)^{2r-1}} \left( \Gamma(2r+1) - (r^2+1) \right) \frac{\Gamma(r+1)}{2(n+1)^{2r}} \]

\[ + 0 \left( \frac{1}{n+1} \right)^{2r} \]

\[ \sigma_n = \sqrt{\frac{\Gamma(2r+1) - (r^2+1) \Gamma^2(r+1)}{(n+1)^{r-1/2}}} + 0 \left( \frac{1}{n+1} \right)^{r+1/2} \]
In particular for $r=2$, we have

$$E\left(\sum_{i=1}^{n+1} V_i^2\right) = \frac{\Gamma(3)}{(n+1)} - \frac{2\Gamma(3)}{2(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right)$$

$$= \frac{2}{n+1} - \frac{2}{(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right).$$

$$q_n\left(\sum_{i=1}^{n+1} V_i^2\right) = \frac{\sqrt{\Gamma(5) - (5)\Gamma^2(3)}}{(n+1)^{3/2}} + 0\left(\frac{1}{(n+1)^{5/2}}\right)$$

$$= \frac{2}{(n+1)^{3/2}} + o\left(\frac{1}{(n+1)^{5/2}}\right).$$

2.4. Darling's Method

Let $G_n = \sum_{i=1}^{n+1} h(V_i)$. Then Darling (1953) showed that $G_n$ had a characteristic function

$$E(e^{i\xi G_n}) = \frac{n!}{2\pi i} \int_{d-i\infty}^{d+i\infty} e^\omega \left(\int_0^\infty e^{-r\omega + i\xi h(r)} dr\right)^{n+1} d\omega \quad (2.4.1)$$

where the path of integration is the straight line $\text{Re}\,Z = d$ (where $\text{Re}\,Z$ denotes the real part of $Z$). If $\int_0^\infty h^k(r) dr$ is finite it is possible to differentiate (2.4.1) $k$ times under the integral sign with respect to $i\xi$. Differentiating once and putting $\xi = 0$ we obtain

$$\nu_1 = E(\omega_n) = \frac{(n+1)!}{2\pi i} \int_{d-i\infty}^{d+i\infty} e^{\omega - n} \int_0^\infty e^{-r\omega} h(r) dr d\omega$$
\[= (n+1)! \int_0^\infty h(r) \frac{1}{2\pi i} \int_{d-i\infty}^{d+i\infty} e^{\omega(1-r)} \omega^{-n} d\omega dr\]

\[= n(n+1) \int_0^1 (1-r)^{n-1} h(r) dr. \quad (2.4.2)\]

Similarly by differentiating twice and setting \(\xi=0\), we obtain the second moment

\[\mu_2 = n(n+1) \int_0^1 (1-r)^{n-1} h^2(r) dr\]

\[+ n^2(n^2-1) \int_{0 \leq r_1 \leq 1} \int_{r_2 \leq 1} (1-r_1-r_2)^{n-2} h(r_1) h(r_2) dr_1 dr_2.\]

\[r_1 \geq 0, \quad r_2 \geq 0\]

\[\quad (2.4.3)\]

From (2.4.2) and (2.4.3) we can calculate the variance

\[\sigma^2 = \mu_2 - \mu_1^2.\]
2.5. Some Integrals Derived from the Beta and Gamma Functions

\[ \Gamma(x+a) = \int_0^\infty t^{x+a-1} e^{-t} \, dt \text{ where } x+a > 0. \]

Letting

\[ \frac{1}{\Gamma(x+a)} t^{x+a-1} e^{-t} = \frac{1}{\Gamma(x+a)} \exp \{ x \log t + (a-1) \log t - t \} \]

= \phi(x,t), we see that \( \phi(x,t) \) is of the Exponential family and using Theorem 9 (pp. 52-3) of Lehmann (1959) we can take the derivatives of all order with respect to \( x \) under the integral sign. Thus we have

\[ \frac{d}{dx} \Gamma(x+a) = \frac{d}{dx} \int_0^\infty t^{x+a-1} e^{-t} \, dt = \int_0^\infty \frac{d}{dx} (t^{x+a-1} e^{-t}) \, dt \]

\= \int_0^\infty t^{x+a-1} e^{-t} \log t \, dt.

This result agrees with Cramer's (1957) statements (p. 125) that the Gamma function is continuous and has continuous derivatives of all orders

\[ \Gamma^{(r)}(p) = \int_0^\infty x^{p-1} (\log x)^r e^{-x} \, dx \text{ for any } p > 0. \]

Let

\[ \psi(x+a) = \frac{d}{dx} \log \Gamma(x+a) = \frac{1}{\Gamma(x+a)} \frac{d}{dx} \Gamma(x+a) \]
\[ = \frac{1}{\Gamma(x+a)} \int_0^\infty t^{x+a-1} \log t \, e^{-t} \, dt. \]

In particular we have

\[ \psi(x) = \frac{1}{\Gamma(x)} \int_0^\infty t^{x-1} \log t \, e^{-t} \, dt, \quad x > 0. \quad (2.5.1) \]

Since

\[ \Gamma(x) = (x-1) \Gamma(x-1), \quad x > 1 \]

and hence

\[ \log \Gamma(x) = \log (x-1) + \log \Gamma(x-1), \]

we have

\[ \psi(x) = \frac{1}{x-1} + \psi(x-1), \quad x > 1. \quad (2.5.2) \]

It is well known that

\[ \psi(1) = \int_0^\infty \log t \, e^{-t} \, dt = -C \]

where \( C \) is Euler's constant. \( (C \approx 0.5772157) \)

\[ \psi(2) = 1 + \psi(1) = 1 - C \quad (2.5.3) \]

\[ \psi(3) = \frac{1}{2} + \psi(2) = \frac{3}{2} - C. \quad (2.5.4) \]

From (2.5.1)

\[ \int_0^\infty t^{x-1} \log t \, e^{-t} \, dt = \Gamma(x) \, \psi(x) \]
from which it follows that by passing the derivative under the integral sign

$$
\int_0^\infty t^{x-1} (\log t)^2 e^{-t} \, dt = \frac{d}{dx} \left[ \Gamma(x) \psi(x) \right]
$$

$$
= \Gamma(x) \psi'(x) + \psi(x) \frac{d}{dx} \Gamma(x) = \Gamma(x) \left[ \psi'(x) + \psi^2(x) \right],
$$

(2.5.5)

where \( \psi'(x) = \frac{d}{dx} \psi(x) \).

Now it is known, i.e. Pairman (1919), that

$$
\psi'(x) = \sum_{n=1}^{\infty} \frac{1}{(n+x-1)^2}.
$$

For \( x \) an integer, this becomes

$$
\psi'(x) = \sum_{n=1}^{x-1} \frac{1}{2} - \sum_{n=1}^{x-1} \frac{1}{n^2} = \frac{\pi^2}{6} - \sum_{n=1}^{x-1} \frac{1}{n^2}
$$

(2.5.6)

and for large \( x \) (not necessarily an integer)

$$
\psi'(x+1) = \frac{1}{1+x} + \frac{1}{2(1+x)^2} + o \left( \frac{1}{(1+x)^2} \right).
$$

(2.5.7)

Let

$$
I_o(x,y) = \int_0^1 t^{x-1} (1-t)^{y-1} \, dt = \frac{\Gamma(x) \Gamma(y)}{\Gamma(x+y)}
$$

where \( x > 0, y > 0 \). Taking logarithms and differentiating with respect to \( x \), we get
\[ \frac{d}{dx} \log I_0(x, y) = \frac{1}{I_0(x, y)} \frac{d}{dx} I_0(x, y) \]

\[ = \frac{d}{dx} [\log \Gamma(x) + \log \Gamma(y) - \log \Gamma(x+y)] \]

\[ = \psi(x) - \psi(x+y) \]

\[ \therefore \frac{dI_0(x, y)}{dx} = \frac{d}{dx} \int_0^1 t^{x-1} (1-t)^{y-1} \, dt \]

\[ = \int_0^1 (1-t)^{y-1} t^{x-1} \log t \, dt = I_0(x, y) [\psi(x) - \psi(x+y)] \]

\[ = I_1(x, y) \text{ say.} \quad (2.5.8) \]

That the derivative can be taken under the integral sign again follows from the fact that

\[ \frac{\Gamma(x+y)}{\Gamma(x)\Gamma(y)} \frac{x}{x+y} = \frac{\Gamma(x+y)}{\Gamma(x)\Gamma(y)} \exp \{x \log t \}

+ y \log (1-t) - \log t - \log (1-t) \}

which is of the exponential family and as such using Lehmann's theorem quoted earlier, the derivatives of all orders with respect to \( x \) and \( y \) can be computed under the integral sign.

Taking logarithms of (2.5.8) and again differentiating
with respect to $x$ we have

$$\frac{1}{I_1(x,y)} \int_0^1 t^{x-1} (1-t)^{y-1} (\log t)^2 \, dt$$

$$= \frac{d}{dx} \left( \log I_0(x,y) + \log \left[ \psi(x) - \psi(x+y) \right] \right)$$

$$= \psi(x) - \psi(x+y) + \frac{\psi'(x) - \psi'(x+y)}{\psi(x) - \psi(x+y)}.$$ 

Therefore

$$I_2(x,y) = \int_0^1 t^{x-1} (1-t)^{y-1} (\log t)^2 \, dt$$

$$= I_1(x,y) \left[ \frac{\psi'(x) - \psi'(x+y)}{\psi(x) - \psi(x+y)} + (\psi(x) - \psi(x+y)) \right]$$

$$= I_0(x,y) \left[ \psi'(x) - \psi'(x+y) + (\psi(x) - \psi(x+y))^2 \right].$$

(2.5.9)

Taking logarithms of (2.5.8) and differentiating with respect to $y$ we have

$$\frac{1}{I_1(x,y)} \int_0^1 t^{x-1} (1-t)^{y-1} \log t \log (1-t) \, dt$$

$$= \frac{d}{dy} \left[ \log I_0(x,y) + \log(\psi(x) - \psi(x+y)) \right]$$

$$= \psi(y) - \psi(y+x) - \frac{d}{dy} \frac{\psi(x+y)}{\psi(x) - \psi(x+y)}.$$
where analogously \( \psi(y+b) = \frac{d}{dy} \log \Gamma(y+b) \). It is important to remember that the first variable is the one we are taking the derivative with respect to, i.e.,

\[
\frac{d}{dx} \log \Gamma(x+y) = \psi(x+y) \quad \text{and} \quad \frac{d}{dy} \log \Gamma(x+y) = \psi(y+x)
\]

However, the simple consideration

\[
\frac{d}{dx} \Gamma(x+y) = \frac{d\Gamma(x+y)}{d(x+y)} \frac{d(x+y)}{dx} = \frac{d\Gamma(x+y)}{d(x+y)} \frac{d(x+y)}{dy}
\]

shows us that \( \psi(x+y) = \psi(y+x) \). We also have that

\[
\frac{d}{dy} \psi(x+y) = \frac{d(\psi(x+y))}{d(x+y)} \frac{d(x+y)}{dy} = \frac{d(\psi(x+y))}{d(x+y)} \frac{d(x+y)}{dx} = \psi'(x+y)
\]

Therefore

\[
\int_0^1 t^{x-1} (1-t)^{y-1} \log t \log(1-t) \, dt = I_3(x,y)
\]

\[
= I_1(x,y) [- \frac{\psi'(x+y)}{\psi(x)-\psi(x+y)} + (\psi(y) - \psi(y+x))] \\
= I_0(x,y) [- \psi'(x+y) + (\psi(y) - \psi(y+x))(\psi(x) - \psi(x+y))]
\]

(2.5.10)
From (2.5.1), (2.5.3) and (2.5.4) we find

\[ \int_0^\infty t \log t e^{-t} \, dt = \Gamma(2) \psi(2) = 1 - C. \]

and

\[ \int_0^\infty t^2 \log t e^{-t} \, dt = \Gamma(3) \psi(3) = 2 \left( \frac{3}{2} - C \right) = 3 - 2C. \]

(2.5.11)

From (2.5.4), (2.5.5), and (2.5.6), we have

\[ \int_0^\infty t^2 \log^2 t e^{-t} \, dt = \Gamma(3) [\psi'(3) + \psi^2(3)] \]

\[ = 2 \left[ \left( \frac{n^2}{6} - \frac{5}{4} \right) + \left( \frac{3}{2} - C \right)^2 \right] = \frac{n^2}{3} + 2 - 6C + 2C^2. \]

(2.5.12)

2.6. Mean and Variance of k*

Now

\[ k^* = \sum_{i=1}^{n+1} V_i \log V_i + \log(n+1) \]

\[ = \sum_{i=1}^{n+1} h(V_i) + \log(n+1) \]
where

\[ h(r) = r \log r \quad 0 < r < 1 \]

= 0 otherwise.

Since

\[ \int_0^1 h(r) \, dr = \int_0^1 r \log r \, dr = \left( \frac{r^2}{2} \log r - \frac{r^2}{4} \right) \bigg|_0^1 = -\frac{1}{4} \]

and

\[ \int_0^1 h^2(r) \, dr = \int_0^1 r^2 \log^2 r \, dr = \frac{r^3}{3} \log^2 r - \frac{2}{3} \log r \]

\[ + \frac{2}{9} \bigg|_0^1 = \frac{2}{27}, \]

therefore we are justified in using Darling's formulas for \( \mu_1 \) and \( \mu_2 \).

Let \( Z_n = \sum_{i=1}^{n+1} V_i \log V_i \), i.e. \( k^* = Z_n + \log (n+1) \)

\[ E(Z_n) = n(n+1) \int_0^1 (1-x)^{n-1} x \log x \, dx \quad \text{from (2.4.2)} \]

\[ = n(n+1) I_1(2,n) \quad \text{from (2.5.8)} \]

\[ = n(n+1) I_0(2,n) [\psi(2) - \psi(n+2)] \]
\[n(n+1) \left( \frac{(n-1)!}{(n+1)!} \right) [\psi(2) - \psi(n+2)] = \frac{n(n+1)}{(n+1)!} [\psi(2) - \psi(n+2)] = \left[ \psi(2) - \psi(n+2) \right] \] (2.6.1)

which from repeated application of (2.5.2)

\[- \left( \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n+1} \right) = 1 - H_{n+1} \]

where

\[H_{n+1} = \sum_{i=1}^{n+1} \frac{1}{i} \]

Therefore

\[\mu_1 = E(k^*) = \log (n+1) + 1 - H_{n+1}. \]

From Nielsen (1906) we have

\[H_{n+1} = C + \log (n+2) - \frac{1}{2(n+2)} + o \left( \frac{1}{n+2} \right) \]

Therefore

\[\mu_1 = \log (n+1) + 1 - H_{n+1} = \log \left( 1 - \frac{1}{n+2} \right) + 1 - C + \frac{1}{2(n+2)} + o \left( \frac{1}{n+2} \right) \]

\[= \left( 1 - C \right) - \frac{1}{2(n+2)} + o \left( \frac{1}{n+2} \right) \]
\[
E(z_n^2) = n(n+1) \int_0^1 (1-r)^{n-1} r^2 \log^2 r \, dr \\
+ n^2(n^2-1) \int_0^{r_1} \int_{r_2}^{1} (1-r_1-r_2)^{n-2} [(r_1 \log r_1) \\
r_1 \geq 0 \quad r_2 > 0 \\
(r_2 \log r_2)] \, dr_2 \, dr_1 \quad \text{from (2.4.3)}
\]

\[
= n(n+1) \int_0^1 (1-r)^{n-1} r^2 \log^2 r \, dr \\
+ n^2(n^2-1) \int_0^1 \int_0^{1-r_1} (1-r_1-r_2)^{n-2} [(r_1 \log r_1) \\
r_1 \geq 0 \quad r_2 > 0 \\
(r_2 \log r_2)] \, dr_2 \, dr_1
\]

\[
= n(n+1) I_2(3,n) + n^2(n^2-1) \int_0^1 \int_0^{1-r_1} [(1-r_1-r_2)^{n-2} \\
r_1 \quad r_2 \log r_1 \log r_2] \, dr_2 \, dr_1 \quad \text{from (2.5.9)}.
\]

Let

\[
\Delta = \int_0^1 \int_0^{1-r_1} (1-r_1-r_2)^{n-2} r_2 \log r_2 \, dr_2 \, dr_1.
\]

Now

\[
\int_0^{1-r_1} (1-r_1-r_2)^{n-2} r_2 \log r_2 \, dr_2 = P
\]
setting $r_2 = (1-r_1)u$ and $dr_2 = (1-r_1)du$, we have

$$p = \int_0^1 (1-r_1)^{n-2}(1-u)^{n-2} (1-r_1)u \log [(1-r_1)u](1-r_1)du$$

$$= \int_0^1 (1-r_1)^n (1-u)^{n-2}u[\log (1-r_1) + \log u] \, du$$

$$= (1-r_1)^n \left[ \int_0^1 (1-u)^{n-2} \log u \, du \right]$$

$$+ \log (1-r_1) \int_0^1 (1-u)^{n-2} u \, du$$

$$= (1-r_1)^n [I_1(2,n-1) + \log (1-r_1) I_0(2,n-1)]$$

from (2.5.8).

Therefore

$$\Delta = \int_0^1 r_1 \log r_1 \, (1-r_1)^n[I_1(2,n-1)$$

$$+ \log (1-r_1) I_0(2,n-1)]dr_1$$

$$= I_1(2,n-1) \int_0^1 r_1(1-r_1)^n \log r_1 \, dr_1$$

$$+ I_0(2,n-1) \int_0^1 r_1(1-r_1)^n \log r_1 \log (1-r_1) \, dr_1$$
\[ E(Z_n^2) = n(n+1) I_2(3,n) + n^2(n^2-1)[I_1(2,n-1) I_1(2,n+1) + I_0(2,n-1) I_3(2,n+1)] \]

From (2.5.9)
\[ I_2(3,n) = \frac{2}{n(n+1)(n+2)} [\psi'(3) - \psi'(n+3) + (\psi(3) - \psi(n+3))^2] \]

From (2.5.8) we have
\[ I_1(2,n-1) = \frac{1}{n(n-1)} [\psi(2) - \psi(n+1)] \]
and
\[ I_1(2,n+1) = \frac{1}{(n+1)(n+2)} [\psi(2) - \psi(n+3)] \]

From (2.5.10) we have
\[ I_3(2,n+1) = I_0(2,n+1) [-\psi'(n+3) + (\psi(n+1) - \psi(n+3))(\psi(2) - \psi(n+3)) \]
\[ = \frac{1}{n(n+1)(n+2)} [-\psi'(n+3) + (\psi(n+1) - \psi(n+3))(\psi(2) - \psi(n+3))] \]
Therefore

\[ E(Z_n^2) = n(n+1) \left[ \frac{2}{n(n+1)(n+2)} \right] \left\{ \psi'(3) - \psi'(n+3) \right\} + (\psi(3) - \psi(n+3))^2 \]

\[ + \left[ \frac{\psi(2) - \psi(n+1)}{n(n-1)(n+2)(n+1)} \right] \left[ \frac{\psi(2) - \psi(n+3)}{n(n-1)(n+1)(n+2)} \right] \]

\[ + \frac{-\psi'(n+3) + (\psi(n+1) - \psi(n+3)) (\psi(2) - \psi(n+3))}{n(n-1)(n+1)(n+2)} \]

\[ = \frac{2}{n+2} (\psi'(3) - \psi'(n+3)) + \frac{2}{n+2} (\psi(3) - \psi(n+3))^2 \]

\[ + \frac{n}{n+2} \left[ (\psi(2) - \psi(n+1)) (\psi(2) - \psi(n+3)) \right] + \frac{n}{n+2} [-\psi'(n+3) \]

\[ + (\psi(n+1) - \psi(n+3)) (\psi(2) - \psi(n+3)) \].

Therefore

\[ E(Z_n^2) = \frac{2}{n+2} (\psi'(3) - \psi'(n+3)) + \frac{2}{n+2} (\psi(3) - \psi(n+3))^2 \]

\[ + \frac{n}{n+2} \left[ (\psi(2) - \psi(n+1)) (\psi(2) - \psi(n+3)) \right] \]

\[ + \frac{n}{n+2} [-\psi'(n+3)] + \frac{n}{n+2} \left[ (\psi(n+1) - \psi(n+3)) (\psi(2) \right. \]

\[ - \psi(n+3)) \right\] \]

\[ = \frac{2}{n+2} (\psi'(3) - \psi'(n+3)) - \frac{n}{n+2} \psi'(n+3) + \frac{n}{n+2} (\psi(2) \]

\[ - \psi(n+3))^2 + \frac{2}{n+2} (\psi(3) - \psi(n+3))^2. \]
Therefore

\[ V(Z_n) = E(Z_n^2) - E^2(Z_n) = E(Z_n^2) - (\psi(2) - \psi(n+2))^2 \]

\[ = \frac{2}{n+2}(\psi'(3) - \psi'(n+3)) - \frac{n}{n+2} \psi'(n+3) + \frac{2}{n+2} (\psi(3) \]

\[ - \psi(n+3))^2 + \frac{n}{n+2} [\psi(2) - \psi(n+3)]^2 - [\psi(2) - \psi(n+2)]^2 \]

\[ = \frac{2}{n+2} \left( \psi'(3) - \psi'(n+3) \right) - \frac{n}{n+2} \psi'(n+3) \]

\[ + \frac{1}{n+2} \{ 2(\psi(3) - \psi(n+3))^2 + n[\psi(2) - \psi(n+3)]^2 \} \]

\[ - [\psi(2) - \psi(n+2)]^2. \]

Now

\[ \{ 2(\psi(3) - \psi(n+3))^2 + n(\psi(2) - \psi(n+3))^2 \} \]

\[ = \{ 2(\psi(3) - \psi(2) + \psi(2) - \psi(n+3))^2 + n(\psi(2) - \psi(n+3))^2 \} \]

\[ = \{ (n+2)\psi(2) - \psi(n+3))^2 + 2(\psi(3) - \psi(2))^2 \]

\[ + 4[\psi(3) - \psi(2)] \psi(2) - \psi(n+3)] \]

Therefore

\[ V(Z_n) = \frac{2}{n+2} (\psi'(3) - \psi'(n+3)) - \frac{n}{n+2} \psi'(n+3) + [\psi(2) \]


\[- \psi(n+3)^2 + \frac{1}{n+2} \left[ 2(\psi(3) - \psi(2))^2 \right] + \frac{4}{n+2} \left[ \psi(3) \right.\]

\[- \psi(2) \left[ \psi(2) - \psi(n+3) \right] - \left[ \psi(2) - \psi(n+2) \right]^2 \]

\[= \frac{2}{n+2} \left( \psi'(3) - \psi'(n+3) \right) - \frac{n}{n+2} \psi'(n+3) + \left[ \psi(2) \right.\]

\[- \psi(n+3)^2 + \frac{1}{n+2} \left[ 2(\psi(3) - \psi(2))^2 \right] + \frac{4}{n+2} \left[ \psi(3) \right.\]

\[- \psi(2) \left[ \psi(2) - \psi(n+3) \right] - \left[ \psi(2) - \psi(n+3) \right] \]

\[+ \psi(n+3) - \psi(n+2)]^2 \]

\[= \frac{2}{n+2} \left( \psi'(3) - \psi'(n+3) \right) - \frac{n}{n+2} \psi'(n+3) + \left[ \frac{2}{n+2} \left[ \psi(3) \right.\]

\[- \psi'(2)^2 + \frac{4}{n+2} \left[ \psi(3) - \psi(2) \right] \left[ \psi(2) - \psi(n+3) \right] \]

\[- \left[ \psi(2) - \psi(n+3) \right]^2 + 2(\psi(2) - \psi(n+3)) (\psi(n+3) \]

\[- \psi(n+2) + (\psi(n+3) - \psi(n+2)]^2 \]

\[= \frac{2}{n+2} \left( \psi'(3) - \psi'(n+3) \right) - \frac{n}{n+2} \psi'(n+3) \]
\[ + \frac{2}{n+2} [\psi(3) - \psi(2)]^2 \]
\[ + \frac{4}{n+2} [\psi(3) - \psi(2)] [\psi(2) - \psi(n+3)] - 2(\psi(2) - \psi(n+3)) (\psi(n+3) - \psi(n+2)) - (\psi(n+3) - \psi(n+2))^2. \]

From (2.5.2) we have
\[ \psi(3) - \psi(2) = \frac{1}{2} \]
\[ \psi(n+3) - \psi(n+2) = \frac{1}{n+2}. \]

Therefore
\[ V(n,n) = \frac{2}{n+2} (\psi'(3) - \psi'(n+3)) - \frac{n}{n+2} \psi'(n+3) + \frac{2}{n+2} \left( \frac{1}{2} \right)^2 \]
\[ + \frac{4}{n+2} \left( \frac{1}{2} \right) (\psi(2) - \psi(n+3)) - \frac{2}{n+2} (\psi(2) - \psi(n+3)) \]
\[ - \frac{1}{(n+2)^2} \]
\[ = \frac{2}{n+2} \psi'(3) - \psi'(n+3) + \frac{n}{2(n+2)^2}. \]

From (2.5.6) and (2.5.7)
\[ \psi'(3) = \frac{\pi^2}{6} - \frac{5}{4} \]
\[ \psi'(n+3) = \frac{1}{n+3} + \frac{1}{2(n+3)^2} + o \left( \frac{1}{(n+3)^2} \right). \]
Therefore
\[ V(Z_n) = \frac{1}{2(n+2)} - \frac{1}{(n+2)^2} + \frac{2}{n+2} \left( \frac{\pi^2}{6} - \frac{5}{4} \right) - \frac{1}{n+3} \]
\[- \frac{1}{2(n+3)^2} + o \left( \frac{1}{(n+3)^2} \right) \]
\[ = \frac{2}{n+2} \left( \frac{\pi^2}{6} - 1 \right) - \frac{1}{(n+2)^2} - \frac{1}{n+3} - \frac{1}{2(n+3)^2} + o \left( \frac{1}{(n+3)^2} \right). \]

Now since
\[ (1+x)^{-1} = 1 - x + x^2 - x^3 + \ldots \]
\[ (1+x)^{-2} = 1 - 2x + 3x^2 - 4x^3 + \ldots \]
\[ \frac{1}{n+2} = \frac{1}{n+1} (1 + \frac{1}{n+1})^{-1} = \frac{1}{n+1} (1 - \frac{1}{n+1} + \frac{1}{(n+1)^2} - \ldots) \]
\[ = \frac{1}{n+1} - \frac{1}{(n+1)^2} + o \left( \frac{1}{(n+1)^2} \right) \]
\[ \frac{1}{(n+2)^2} = \frac{1}{(n+1)^2} \left[ 1 + \frac{1}{n+1} \right]^{-2} = \frac{1}{(n+1)^2} \left( 1 - \frac{2}{n+1} + \ldots \right) \]
\[ = \frac{1}{(n+1)^2} + o \left( \frac{1}{(n+1)^2} \right) \]
\[ \frac{1}{n+3} = \frac{1}{n+1} \left( 1 + \frac{2}{n+1} \right)^{-1} = \frac{1}{n+1} (1 - \frac{2}{n+1} + \left( \frac{2}{n+1} \right)^2 - \ldots) \]
\[ = \frac{1}{n+1} - \frac{2}{(n+1)^2} + o \left( \frac{1}{(n+1)^2} \right) \]
\[ \frac{1}{(n+3)^2} = \frac{1}{(n+1)^2} \left(1 + \frac{2}{n+1}\right)^{-2} = \frac{1}{(n+1)^2} \left[1 - 2\left(\frac{2}{n+1}\right) + \ldots\right] \]

\[ = \frac{1}{(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right). \]

Therefore

\[ V(Z_n) = 2\left(\frac{\pi^2}{6} - 1\right) \left[\frac{1}{n+1} - \frac{1}{(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right)\right] \]

\[ - \left[\frac{1}{(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right)\right] - \left[\frac{1}{n+1} - \frac{2}{(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right)\right] \]

\[ = \frac{2\left(\frac{\pi^2}{6} - 1\right) - 1}{n+1} + \frac{-\frac{1}{2} - 2\left(\frac{\pi^2}{6} - 1\right) + 2 - 1}{(n+1)^2} \]

\[ + o\left(\frac{1}{(n+1)^2}\right) \]

\[ = \frac{\pi^2}{3} - 3 - \frac{\left(\frac{\pi^2}{3} - \frac{5}{2}\right)}{(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right). \]

Therefore we have

\[ V(k^*) = V(Z_n) = \frac{\pi^2}{3} - 3 - \frac{\left(\frac{\pi^2}{3} - \frac{5}{2}\right)}{(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right). \]
2.7. Mean and Variance of J

We now consider \( J = k + k^* \) and obtain its mean and variance. For this we will first work out the covariance between \( k \) and \( k^* \).

\[
E(V_i \log V_i \log V_j)
\]

\[
= n(n-1) \int_{x>0} \int_{y>0} (\log x) (\log y) (1-x-y)^{n-2} \, dy \, dx
\]

\[
= n(n-1) \int_0^1 \int_0^{1-x} (\log x) (1-x)^n - 2 (1-\frac{y}{1-x})^{n-2} y \log y \, dy \, dx
\]

from (2.2.2)

\[
= n(n-1) \int_0^1 \int_0^1 (\log x) (1-x)^n (1-u)^{n-2} u \log u + \log (1-x) \, du \, dx
\]

\[
= n(n-1) \left[ \int_0^1 (\log x) (1-x)^n \, dx \int_0^1 (1-u)^{n-2} u \log u \, du \right.
\]

\[
+ \int_0^1 \log x \log(1-x) (1-x)^n \, dx \int_0^1 u(1-u)^{n-2} \, du \right]
\]

\[
= n(n-1) \left[ I_1(1,n+1) I_1(2,n-1) + I_0(2,n-1) I_3(1,n+1) \right]
\]

from (2.5.8) and (2.5.10)

\[
= n(n-1) \left[ \frac{1}{n+1} (\psi(1) - \psi(n+2)) \frac{1}{n(n-1)} (\psi(2) - \psi(n+1)) \right]
\]
\[ + \frac{1}{n(n-1)} \left( \frac{1}{n+1} \left[ \psi'(n+2) + (\psi(n+1) - \psi(n+2))(\psi(1) - \psi(n+2)) \right] \right) \]

\[ = \frac{1}{n+1} \left\{ [\psi(1) - \psi(n+2)] [\psi(2) - \psi(n+1)] - \psi'(n+2) \right\} \]

\[ - \frac{1}{n+1} (\psi(1) - \psi(n+2)) \right\} \]

\[ \cdot \]

\[ \therefore \text{Cov} \left( V_i \log V_i, \log V_j \right) = \frac{1}{n+1} \left\{ [\psi(1) - \psi(n+2)] [\psi(2) - \psi(n+1)] - \psi'(n+2) - \frac{1}{n+1} (\psi(1) - \psi(n+2)) \right\} - \frac{1}{n+1} [\psi(2) - \psi(n+2)] [\psi(1) - \psi(n+1)] \text{ from (2.3.1)}, \]

since

\[ k = - \frac{1}{n+1} \sum_{i=1}^{n+1} \log V_i - \log(n+1), \text{ and from (2.6.1)}, \]

\[ E(V_i \log^2 V_i) = n \int_0^1 (1-x)^{n-1} x \log^2 x \, dx \text{ from (2.2.2)} \]

\[ = n I_2(2,n) \text{ from (2.5.9)} \]

\[ = n \left[ \frac{1}{n(n+1)} \left\{ \psi'(2) - \psi'(n+2) + (\psi(2) - \psi(n+2))^2 \right\} \right] \]

\[ = \frac{1}{n+1} \left[ \psi'(2) - \psi'(n+2) + (\psi(2) - \psi(n+2))^2 \right]. \]
Therefore
\[ \text{Cov}(V_i \log V_i, \log V_i) = \frac{1}{n+1} \left[ \psi'(2) - \psi'(n+2) + (\psi(2) - \psi(n+2))^2 \right] - \frac{1}{n+1} [\psi(2) - \psi(n+2)] [\psi(1) - \psi(n+1)]. \]

Therefore
\[ \text{Cov}(k, k^*) = \text{Cov} \left( \sum_{i=1}^{n+1} \frac{V_i}{n+1}, \sum_{i=1}^{n+1} V_i \log V_i \right) = -\frac{1}{n+1} \sum_{i=1}^{n+1} \text{Cov}(\log V_i, V_i \log V_i) \]
\[ - \frac{1}{n+1} \sum_{i \neq j} \text{Cov}(\log V_i, V_j \log V_j) \]
\[ = - \left[ \frac{1}{n+1} \left\{ \psi'(2) - \psi'(n+2) + (\psi(2) - \psi(n+2))^2 \right\} - (\psi(2) - \psi(n+2)) (\psi(1) - \psi(n+1)) \right] \]
\[ - \frac{1}{n+1} \left[ n(n+1) \frac{1}{n+1} \left\{ [\psi(1) - \psi(n+2)] [\psi(2) - \psi(n+1)] \right\} \right. \]
\[ - \left. \psi'(n+2) - \frac{1}{n+1} (\psi(1) - \psi(n+2)) \right] \]
\[ - \left[ [\psi(2) - \psi(n+2)] [\psi(1) - \psi(n+1)] \right]. \]
\[
- \frac{1}{n+1} \left\{ (\psi'(2) - \psi'(n+2) + (\psi(2) - \psi(n+2))^2 \\
- (\psi(2) - \psi(n+2)) (\psi(1) - \psi(n+1)) + n \{ -\psi'(n+2) + (\psi(1) - \psi(n+2)) (\psi(2) - \psi(n+1)) - \frac{1}{n+1} (\psi(1) - \psi(n+2)) \} \\
- \psi(n+2)) - (\psi(2) - \psi(n+2)) (\psi(1) - \psi(n+1)) \right\}
= \psi'(n+2) + (\psi(2) - \psi(n+2)) (\psi(1) - \psi(n+1))
- \frac{1}{n+1} \left\{ (\psi'(2) + (\psi(2) - \psi(n+2))^2 + n(\psi(1) - \psi(n+2)) \right\}
- \psi(n+2)) (\psi(2) - \psi(n+1)) - \frac{n}{n+1} (\psi(1) - \psi(n+2)) \right\}
= \psi'(n+2) + (\psi(2) - \psi(n+2)) (\psi(1) - \psi(n+1))
- \frac{1}{n+1} \left\{ (\psi'(2) + (\psi(2) - \psi(n+2)) (\psi(2) - \psi(1) + \psi(1) - \psi(n+2)) + n(\psi(1) - \psi(n+1)) + \psi(n+1) - \psi(n+2)) (\psi(2) - \psi(n+2) + \psi(n+2) - \psi(n+1)) \right\}
- \frac{n}{n+1} (\psi(1) - \psi(n+2)) \right\}
= \psi'(n+2) + [((\psi(2) - \psi(n+2)) (\psi(1) - \psi(n+1))] - \frac{1}{n+1} \left\{ (\psi'(2) + [((\psi(2) - \psi(n+2)) \right\}

\[
\begin{align*}
& (\psi(1) - \psi(n+1)) + [\psi(2) - \psi(1)] [\psi(2) - \psi(n+2)] \\
& + [\psi(2) - \psi(n+2)] [\psi(n+1) - \psi(n+2)] + n(\psi(1) - \psi(n+1)) \\
& (\psi(2) - \psi(n+2)) + n[(\psi(1) - \psi(n+1)) (\psi(n+2) - \psi(n+1))] \\
& + n(\psi(n+1) - \psi(n+2)) (\psi(2) - \psi(n+2)) \\
& + n(\psi(n+1) - \psi(n+2)) (\psi(n+2) - \psi(n+1)) \\
& = \frac{n}{n+1} (\psi(1) - \psi(n+2)) \\
& = \psi'(n+2) - \frac{1}{n+1} [\psi'(2) + (\psi(2) - \psi(n+2))] \\
& - \frac{1}{n+1} (\psi(2) - \psi(n+2)) + \frac{n}{n+1} (\psi(1) - \psi(n+1)) \\
& - \frac{n}{n+1} (\psi(2) - \psi(n+2)) - \frac{n}{(n+1)^2} - \frac{n}{n+1} (\psi(1) - \psi(n+2)) \\
& = \psi'(n+2) - \frac{1}{n+1} [\psi'(2) + \frac{n}{n+1} (\psi(1) - \psi(n+2))] \\
& + \frac{n}{n+1} (\psi(n+2) - \psi(n+1)) - \frac{n}{(n+1)^2} \\
& - \frac{n}{n+1} (\psi(1) - \psi(n+2)) \\
& = \psi'(n+2) - \psi'(2) \frac{1}{n+1} \\
& = (2.7.1)
\end{align*}
\]
Therefore we find

\[ E(J) = E(k) + E(k*) \]

\[ = - [\psi(1) - \psi(n+1)] + [\psi(2) - \psi(n+2)] \]

from (2.3.1) and (2.6.1)

\[ = -[\psi(1) - \psi(2) + \psi(2) - \psi(n+2) + \psi(n+2) - \psi(n+1)] \]

\[ + [\psi(2) - \psi(n+2)] \]

\[ = [\psi(2) - \psi(1) - (\psi(n+2) - \psi(n+1))] \]

\[ = 1 - \frac{1}{n+1}. \]

\[ V(J) = V(k) + V(k*) + 2 \text{ Cov } (k,k*) \]

\[ = \left[ \frac{\psi'(1)}{n+1} - \psi'(n+1) \right] + \left[ \frac{2}{n+2} \psi'(3) - \psi'(n+3) + \frac{n}{2(n+2)^2} \right] \]

\[ + 2 \left[ \psi'(n+2) - \frac{\psi'(2)}{n+1} \right] \]

from (2.3.1), (2.6.2) and (2.7.1)

\[ = \frac{\pi^2}{6} + \frac{2}{n+2} \left( \frac{\pi^2}{6} - 1 - \frac{1}{4} \right) - \frac{2}{n+1} \left( \frac{\pi^2}{6} - 1 \right) \]
\[ + 2\psi'(n+2) - \psi'(n+1) - \psi'(n+3) + \frac{1}{2(n+2)} - \frac{1}{(n+2)^2} \]

\[ = \frac{-\frac{\pi^2}{6} + 2}{n+1} + \frac{\left(\frac{\pi^2}{3} - 2\right)}{n+1} \left(1 - \frac{1}{n+1} + \ldots\right) - \frac{1}{(n+1)^2} \]

\[ + o\left(\frac{1}{(n+1)^2}\right) \]

\[ = \frac{\pi^2}{6} n+1 - \frac{\left(\frac{\pi^2}{3} - 1\right)}{(n+1)^2} + o\left(\frac{1}{(n+1)^2}\right). \]

Note that of the four statistics, \( J \) is the only one for which the first two moments can be approximated to any degree of accuracy desired.
3. LIMITING DISTRIBUTIONS

3.1. Introduction

The asymptotic Normality of the test statistics under $H_0$ will be proved. Due to the fact that the $V_i$ are dependent random variables, the limiting distributions of the statistics are not readily obtainable. However since for all $i \neq j$,

$$\text{Corr}\ (V_i, V_j) = \frac{1}{n},$$

we might expect that in the limit functions of $V_i$ act as functions of independent and identically distributed random variables and as such might be asymptotically Normal. That this is so under mild restrictions follows from the method of Le Cam. In section 3.3, Le Cam's method of finding the limit distribution of

$$G_n = \sum_{i=1}^{n+1} g_n \{(n+1)V_i\}$$

where $\{g_n, n \geq 1\}$ is a sequence of Borel-measurable functions, will be given. In section 3.4, Le Cam's method will be applied in finding the limiting distribution of $k^*$. In section 3.5, Le Cam's method will be used to find the limiting distribution of the statistic $J$. Finally in section 3.6, the moments by Le Cam's method will be compared with those from Darling's (exact) method.
3.2. Asymptotic Normality of $k$ and Kimball's Statistics

Kale (1965) proved that $k$ was asymptotically Normally distributed with the mean and variance as given in section 2.3. The same results could be obtained from Darling's results concerning the asymptotic Normality of

$$k' = \sum_{i=1}^{n+1} \log V_i$$

which was proved by the complicated method of Steepest Descents on the expression (2.4.1).

Darling also proved that Kimball's statistic is asymptotically Normally distributed with means and variances as given in section 2.3. These results were derived through the method of Steepest Descents. The asymptotic Normality of

$$\sum_{i=1}^{n+1} V_i^r$$

was first derived by Kimball (1950). For the special case $r=2$, i.e., Greenwood's statistic $\sum_{i=1}^{n+1} V_i^2$, the limiting distribution under the assumption that $V_i$ are uniform spacings was obtained by Moran (1947).

3.3. Le Cam's Method

Let $\{g_n; n \geq 1\}$ be a sequence of real-valued Borel-measurable functions defined on $[0, \infty)$ and consider the random variable $G_n = \sum_{i=1}^{n+1} g_n ((n+1) V_i)$. Let $\{Y_i; i \geq 1\}$ be a sequence of independent random variables Exponentially distributed with mean 1. Put
\[ S_n = \frac{1}{n+1} \sum_{i=1}^{n+1} (Y_i - 1) \quad \text{and} \quad B_n = \sum_{i=1}^{n+1} g_n(Y_i). \]

\( B_n \) can be said to be an Exponential copy of \( G_n \). Due to Lemma 2.2.1, the distribution function of \( G_n \) is the same as the conditional distribution function of \( B_n = \sum_{i=1}^{n+1} g_n(Y_i) \), given \( S_n = 0 \).

The method of Le Cam (1958) is to use information about the joint limiting behavior of \( (B_n, S_n) \), to derive the desired conditional limiting distribution of \( B_n \), given \( S_n = 0 \). Since \( B_n \) and \( S_n \) are sums of independent and identically distributed random variables their joint limiting distributions can be obtained through classical limit theorems. The limiting distribution function, if it exists, must be a two dimensional infinitely divisible distribution function. Now \( S_n \) converges in distribution to a \( N(0,1) \) random variable \( S \). If \( B_n \) converges in distribution to a random variable \( B \), then \( B \) must be infinitely divisible and can be written \( B = B^N + B^P \), where \( B \) is split into its independent Normal and non-Normal (or Poisson) parts.

Theorem (Le Cam). If \( (B_n, S_n) \) converges in distribution to a random variable \( (B, S) = (B^N + B^P, S) \), then \( G_n \) converges in distribution to the random variable \( B - \tau S \) where \( \tau = \text{E}(B^N S) \).

In particular if \( B_n \) converges in distribution to \( B \) which only has a normal part, i.e., \( B = B^N \); then \( (B_n, S_n) \) converges
in distribution to the bivariate Normal distribution law of 
\((B,S)\) and \(G_n\) converges in distribution to a univariate normal 
distribution law of \(B - \tau S\) where \(\tau = E(BS)\). Therefore, if

\[
G_n = \sum_{i=1}^{n+1} g((n+1)V_i) \quad \text{(i.e., } g_n((n+1)V_i) = g((n+1)V_i))
\]

then a sufficient condition for \(G_n\) to be asymptotically Normally 
distributed is that \(E(g^2(Y_i)) < \infty\); as under this condition the 
bivariate central limit theorem for \(B_n\) and \(S_n\) holds 
and \((B_n, S_n)\) is asymptotically bivariate Normally distributed 
as \((B,S)\).

3.4. Asymptotic Normality of 
the Statistic \(k^*\)

Let \(T\) be an Exponentially distributed random variable with 
mean 1. Let \(Z = T \log T\). From (2.5.11) and (2.5.12) we have

\[E(Z) = \int_{0}^{\infty} t \log t e^{-t} \, dt = (1-C)\]

and

\[E(Z^2) = \int_{0}^{\infty} t^2 \log^2 t e^{-t} \, dt = \frac{\pi^2}{3} + 2 - 6C + 2C^2.\]

Therefore

\[V(Z) = E(Z^2) - E^2(Z) = \frac{\pi^2}{3} + 1 - 4C + C^2\]

\[= \frac{\pi^2}{3} - 3 + (2-C)^2 = w.\]
Let \( \{T_i \mid i=1, \ldots, n+1\} \) be independent and identically distributed random variables distributed as \( T \).

Let

\[
B_n = \frac{\sum_{i=1}^{n+1} T_i \log T_i - (n+1)(1-C)}{\sqrt{(n+1)w}}
\]

\[
S_n = \frac{\sum_{i=1}^{n+1} (T_i-1)}{\sqrt{n+1}} .
\]

\[
\text{Cov}(B_n, S_n) = \frac{1}{\sqrt{w}} \left[ \int_0^\infty t \log t (t-1) e^{-t} dt \right]
\]

\[
= \frac{1}{\sqrt{w}} \left[ \int_0^\infty t^2 \log t e^{-t} dt - \int_0^\infty t \log t e^{-t} dt \right]
\]

\[
= \frac{1}{\sqrt{w}} [3 - 2C - (1-C)] \text{ from } (2.5.11)
\]

\[
= \frac{(2-C)}{\sqrt{w}} .
\]

Thus \( E(B_n) = 0, V(B_n) = 1, E(S_n) = 0, \) and \( V(S_n) = 1. \) Since \( E(T_i \log T_i - (1-C))^2 = V(Z) = w < \infty, \) by the bivariate central limit theorem \( (B_n, S_n) \) converges to a bivariate Normal distribution with zero means, unit variances, and covariance

\[
\frac{2-C}{\sqrt{w}} = \sigma_{BS} .
\]
Let

\[ G_n = \frac{\sum_{i=1}^{n+1} (n+1)V_i \log [(n+1)V_i] - (n+1)(1-C)}{\sqrt{(n+1)w}} \]

\[ = \frac{\sum_{i=1}^{n+1} (n+1)V_i \left( \log V_i + \log (n+1) \right) - (n+1)(1-C)}{\sqrt{(n+1)w}} \]

\[ = \frac{\sum_{i=1}^{n+1} V_i \log V_i + \log (n+1) - (1-C)}{\sqrt{w/(n+1)}} \]

\[ = \frac{k* - (1-C)}{\sqrt{w/(n+1)}} \]

Now from Le Cam's Theorem, \( G_n \) is asymptotically Normally distributed as \( B - \tau S \).

\[ E(B - \tau S) = 0 \text{ and } V(B - \tau S) = V(B) + \tau^2 V(S) \]

\[ - 2\tau \text{Cov}(B, S) = V(B) - \tau^2 \text{ since } E(B) = E(S) = 0 \]

and therefore \( \text{Cov}(B, S) = E(BS) = \tau \). Therefore

\[ V(B - \tau S) = 1 - \left( \frac{2-C}{\sqrt{w}} \right)^2 = \frac{w - (2-C)^2}{w} \]

\[ = \frac{1}{w} \left[ \frac{n^2}{3} + 1 - 4C + C^2 - (2-C)^2 \right] \]
Therefore $E(k^*) = 1 - C$ and $V(k^*) = \frac{n^2}{3} - 3 \quad \frac{1}{n+1}$ and $k^*$ is asymptotically Normally distributed, with mean $1 - C$ and variance $\frac{n^2}{3} - 3 \quad \frac{1}{n+1}$.

3.5. Asymptotic Normality of the Statistic $J$

Now

$$J = \sum_{i=1}^{n+1} V_i \log \frac{V_i}{1 \quad \frac{1}{n+1}} + \sum_{i=1}^{n+1} \frac{1}{n+1} \log \frac{1}{V_i}$$

$$= \sum_{i=1}^{n+1} V_i \log \frac{V_i}{1 \quad \frac{1}{n+1}} - \sum_{i=1}^{n+1} \frac{1}{n+1} \log \frac{V_i}{1 \quad \frac{1}{n+1}}$$

$$= \sum_{i=1}^{n+1} (V_i - \frac{1}{n+1}) \log \frac{V_i}{1 \quad \frac{1}{n+1}}$$

$$= \sum_{i=1}^{n+1} \frac{(n+1)V_i - 1}{n+1} \log [(n+1)V_i]$$

$$= \sum_{i=1}^{n+1} g_n ((n+1)V_i) = G_n.$$
\[ B_n = \frac{1}{n+1} \sum_{i=1}^{n+1} (T_i - 1) \log T_i , \]

where \( \{T_i \mid i=1, \ldots, n+1\} \) are independent and identically Exponentially distributed with mean 1. Let \( T \) also be Exponentially distributed with mean 1.

Consider

\[ Z = (T-1) \log T, \quad E(Z) = \int_0^\infty (t-1) \log t \, e^{-t} \, dt \]
\[ = \int_0^\infty t \log t \, e^{-t} \, dt - \int_0^\infty \log t \, e^{-t} \, dt = 1 \text{ from (2.5.3)}. \]

\[ E(Z^2) = \int_0^\infty (t-1)^2 \log^2 t \, e^{-t} \, dt = \int_0^\infty \log^2 t \, e^{-t} \, dt \]
\[ - 2 \int_0^\infty t \log^2 t \, e^{-t} \, dt + \int_0^\infty \log^2 t \, e^{-t} \, dt \]
\[ = \Gamma(3) \left[ \psi'(3) + \psi^2(3) \right] - 2\Gamma(2) \left[ \psi'(2) + \psi^2(2) \right] \]
\[ + \Gamma(1) \left[ \psi'(1) + \psi^2(1) \right] \text{ from (2.5.5)}, \]

Since

\[ \psi'(x) = -\frac{1}{(x-1)^2} + \psi'(x-1) \]

and

\[ \psi(x) = \frac{1}{x-1} + \psi(x-1) \]
we have
\[ \psi'(1) = \psi'(2) + 1, \quad \psi'(2) = \psi'(3) + 1/4. \]

Therefore
\[
E(z^2) = \Gamma(3) [\psi'(3) + \psi^2(3)] - 2\Gamma(2) [\psi'(3) + 1/4
\]
\[
+ (\psi(3) - \frac{1}{2})^2] + [\psi'(3) + \frac{5}{4} + (\psi(3) - \frac{3}{2})^2]
\]
\[
= 2[\psi'(3) + \psi^2(3)] - 2[\psi'(3) + \frac{1}{4} + \psi^2(3) - \psi(3)
\]
\[
+ \frac{1}{4}] + [\psi'(3) + \frac{5}{4} + \psi^2(3) - 3\psi(3) + \frac{9}{4}]
\]
\[
= \psi'(3) + \psi^2(3) - \psi(3) - 1 + \frac{14}{4}
\]
\[
= \psi'(3) + (\psi(3) - \frac{1}{2})^2 + \frac{9}{4}
\]
\[
= \frac{\pi^2}{6} - \frac{5}{4} + (1-C)^2 + \frac{9}{4}
\]

from (2.5.4) and (2.5.6)
\[
= \frac{\pi^2}{6} + C^2 - 2C + 2.
\]

\[ V(z) = E(z^2) - E^2(z) = \frac{\pi^2}{6} + C^2 - 2C + 2 - 1
\]
\[
= \frac{\pi^2}{6} + (1-C)^2.
\]
Let \( S^* = (T-1) \).

\[
\text{Cov} (Z, S^*) = \int_0^\infty (t-1)^2 \log t \, e^{-t} \, dt
\]

\[
= \int_0^\infty t^2 \log t \, e^{-t} \, dt
\]

\[
- 2 \int_0^\infty t \log t \, e^{-t} \, dt + \int_0^\infty \log t \, e^{-t} \, dt
\]

\[
= \Gamma(3) \psi(3) - 2\psi(2) + \psi(1)
\]

\[
= 2\psi(3) - 2\psi(2) + \psi(1)
\]

\[
= 2(\psi(1) + \frac{3}{2}) - 2(1 + \psi(1)) + \psi(1)
\]

\[
= 1 + \psi(1) = 1 - \zeta.
\]

Let

\[
S_n = \frac{1}{\sqrt{n+1}} \sum_{i=1}^{n+1} (Y_i - 1),
\]

Consider the bivariate distribution of \((B_n, S_n)\). Since \(B_n\) and \(S_n\) are sums of independent and identically distributed random variables with \(E(Z^2) < \infty\) and \(E(S^2) < \infty\). \((B_n, S_n)\) converges to the bivariate Normal law \((B, S)\) with

\[
E(B) = 1 \quad E(S) = 0
\]
\[ V(B) = \frac{1}{n+1} \left( \frac{\pi^2}{6} + (1-C)^2 \right) \quad V(S) = 1 \]

\[
\text{Cov} (B, S) = \frac{1}{\sqrt{n+1}} (1-C)
\]

and by Le Cam's Theorem, \( G_n \) converges in distribution to the univariate Normal law \( B - \tau S \) where

\[
E(B - \tau S) = E(B) = 1
\]

\[
V(B - \tau S) = V(B) + \tau^2 V(S) - 2\tau \text{Cov} (B, S) = V(B) - \tau^2
\]

\[
= \frac{1}{n+1} \left( \frac{\pi^2}{6} + (1-C)^2 \right) - \frac{1}{n+1} (1-C)^2 = \frac{1}{n+1} \frac{\pi^2}{6}
\]

\[ \therefore G_n \text{ is asymptotically normally distributed} \]

\[ \text{N}(1, \frac{1}{n+1} \frac{\pi^2}{6}) \]

3.6. Comparison of Moments and Critical Points

If we had used Le Cam's method to find the limiting distribution of all test statistics, we would have the results given in Table 3.6.a. Since Le Cam's method is essentially one of finding the limiting distribution of a conditional distribution by taking the conditional distribution of limiting distributions, it is not surprising that the moments given by Le Cam's method are not as accurate as those given
by Darling's exact method.

In Table 3.6.a, the means and variances are given to the corresponding order of approximation as shown.

The critical points of the four statistics are given in Table 3.6.b. Since the information statistics are all greater than or equal to zero, with equality if and only if all \( V_i = \frac{1}{n+1} \), it is the large values of \( k, k^* \) or \( J \) that show departure from \( H_0 \). Similarly \( \sum_{i=1}^{n+1} V_i^r \) is a minimum when \( V_i = \frac{1}{n+1} V_i \) and so it is large values of \( \sum_{i=1}^{n+1} V_i^r \) that show departure from \( H_0 \).
Table 3.6.a. Moments of test statistics derived by Darling's and Le Cam's methods

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Le Cam's Mean</th>
<th>Le Cam's Variance</th>
<th>Darling's Mean</th>
<th>Darling's Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kimball's</td>
<td>( \frac{\Gamma(r+1)}{(n+1)r-1} )</td>
<td>( \frac{\Gamma(2r+1)}{(n+1)^2r-1} )</td>
<td>( \frac{\Gamma(r+1)}{(n+1)r-1} )</td>
<td>( \frac{\Gamma(2r+1)-(r^2+1)r^2(r+1)}{(n+1)^2r-1} )</td>
</tr>
<tr>
<td>Greenwood</td>
<td>( \frac{2}{n+1} )</td>
<td>( \frac{4}{(n+1)^3} )</td>
<td>( \frac{2}{n+1} - \frac{2}{(n+1)^2} )</td>
<td>( \frac{4}{(n+1)^3} + o \left( \frac{1}{(n+1)^3} \right) )</td>
</tr>
<tr>
<td>k</td>
<td>c</td>
<td>( \frac{\pi^2 - 1}{6n+1} )</td>
<td>( c - \frac{1}{2(n+1)} )</td>
<td>( \frac{\pi^2 - 1}{6n+1} - \frac{1}{2(n+1)^2} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( + o \left( \frac{1}{(n+1)^2} \right) )</td>
<td>( + o \left( \frac{1}{(n+1)^2} \right) )</td>
<td>( + o \left( \frac{1}{(n+1)^2} \right) )</td>
</tr>
<tr>
<td>Statistic</td>
<td>Le Cam's Mean</td>
<td>Le Cam's Variance</td>
<td>Darling's Mean</td>
<td>Darling's Variance</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------</td>
<td>------------------</td>
<td>---------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>(k^*)</td>
<td>1 - C</td>
<td>(\frac{\pi^2}{3} - 3) (\frac{1}{n+1})</td>
<td>1 - C - (\frac{1}{2(n+1)})</td>
<td>(\frac{\pi^2}{3} - 3) - (\frac{\pi^2}{3} - \frac{5}{2}) (\frac{1}{(n+1)^2})</td>
</tr>
<tr>
<td>(J)</td>
<td>1</td>
<td>(\frac{\pi^2}{6}) (\frac{1}{n+1})</td>
<td>1 - (\frac{1}{n+1})</td>
<td>(\frac{\pi^2}{6}) - (\frac{\pi^2}{3} - 1) (\frac{1}{(n+1)^2}) + (\frac{1}{(n+1)^2})</td>
</tr>
</tbody>
</table>
Table 3.6.b. Critical points of the test statistics*  

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Critical Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kimball's</td>
<td>[ \frac{r(r+1)}{(n+1)r-1} - \frac{r(r-1)\Gamma(r+1)}{2(n+1)} + \xi_{\alpha} \sqrt{\frac{\Gamma(2r+1)-(r^2+1)\Gamma^2(r+1)}{(n+1)^2r-1}}} ]</td>
</tr>
<tr>
<td>Greenwood's</td>
<td>[ \frac{2}{n+1} - \frac{2}{(n+1)^2} + \xi_{\alpha} \frac{2}{(n+1)^{3/2}} ]</td>
</tr>
<tr>
<td>k</td>
<td>[ C - \frac{1}{2(n+1)} + \xi_{\alpha} \sqrt{\frac{n^2}{6} - 1} ]</td>
</tr>
<tr>
<td>k*</td>
<td>[ 1 - C - \frac{1}{2(n+1)} + \xi_{\alpha} \sqrt{\frac{n^2}{3} - 3} ]</td>
</tr>
<tr>
<td>J</td>
<td>[ 1 - \frac{1}{n+1} + \xi_{\alpha} \sqrt{\frac{n^2}{6}} ]</td>
</tr>
</tbody>
</table>

*The test procedure is to reject $H_0$ if the observed value of the statistics exceeds the corresponding critical point. $\phi(\xi_{\alpha}) = 1-\alpha$, $\phi(x)$ being the standard normal integral.
4. POWER UNDER WEISS' ALTERNATIVES

4.1. Introduction

In this chapter the power of the test statistics will be studied against a wide class of alternatives. These are the alternatives \( \{ F_n(x) \} \) where

\[
F_n(x) = F_0(x) + \frac{1}{n^\delta} \int_0^{F_0(x)} r(y) dy
\]

where

\[
\int_0^1 r(y) dy = 0, \quad |r''(y)| < D \text{ for } 0 < y < 1 \text{ and } \frac{1}{4} < \delta < \frac{1}{2}.
\]

It will be shown that Greenwood's test is the most powerful among tests considered here against these alternatives.

4.2. Weiss' Theorem

Suppose \( X_1, X_2, \ldots, X_n \) are independent and identically distributed random variables with probability density function \( f_n(x) \) defined over the interval \((0, 1)\)

\[
f_n(x) = 1 + \frac{r(x)}{n^\delta} \quad \text{for } 0 < x < 1
\]

where

\[
\int_0^1 r(x) dx = 0, \quad |r''(x)| < D < \infty \text{ for } 0 < x < 1, \delta > 0
\]

and

\[
F_n(x) = \int_0^x f_n(u) du. \quad \text{Let } Y_1 < Y_2 < \ldots < Y_n \text{ denote the}
\]
ordered values of \(X_1, X_2, \ldots, X_n\) and let \(Y_0 = 0, Y_{n+1} = 1\).
Define \(V_i\) as \(Y_i - Y_{i-1}\) for \(i = 1, \ldots, n+1\). Let
\[g_n(v_1, \ldots, v_n)\]
denote the joint probability density function for \(V_1, V_2, \ldots, V_n\).

Let \(U_1, U_2, \ldots, U_{n+1}\) be independent and identically distributed random variables each with probability density function \(e^{-u}\) for \(u > 0\), i.e. each \(U_i\) is exponentially distributed.

Define \(W_i\) as \(a_i U_i\) where
\[a_i = \frac{1}{f_n[F_n^{-1}(\frac{i}{n+1})]} \quad i = 1, 2, \ldots, n+1. \tag{4.2.1}\]

Let \(T_n^i = \sum_{i=1}^{n+1} W_i\) and \(Z_i = W_i/T_n^i\) for \(i=1, 2, \ldots, n+1\).

Let \(h_n(z_1, z_2, \ldots, z_n)\) denote the joint probability density for \(Z_1, Z_2, \ldots, Z_n\).

Then Weiss' Theorem states:

For each \(n\), let \(R_n\) be any measurable set in \(n\) dimensional space. Then
\[
\lim_{n \to \infty} \int_{R_n} \cdots \int_{R_n} g_n(v_1, v_2, \ldots, v_n) \, dv_1 \, dv_2 \ldots \, dv_n = 0
\]

This theorem enables us to investigate the asymptotic power of a test based on the \(V_i\)'s for a class of alternatives.
determined by \( f_n(x) \) through the independent random variables \( Z_i \)'s.

Weiss used this theorem to derive the asymptotic power of the test statistic \( \sum_{i=1}^{n+1} V_i^2 \) for the test \( H_0: f(x) = 1, \ 0 < x < 1 \) against the sequence of alternatives \( \{ f_n(x) \} \)

where \( f_n(x) = 1 + \frac{r(x)}{n^\delta} \) for \( 0 < x < 1 \), \( \int_0^1 r(x) \, dx = 0 \) and \( |r''(x)| < D < \infty \) for \( 0 < x < 1 \). The technique was to replace

\[
\sum_{i=1}^{n+1} V_i^2 \]

by

\[
\sum_{i=1}^{n+1} Z_i^2 = \sum_{i=1}^{n+1} W_i^2 = \frac{A'_n}{T'_n^2},
\]

where

\[
A'_n = \sum_{i=1}^{n+1} W_i^2.
\]

The critical region is then \( \frac{A'_n}{T'_n^2} > \frac{2}{n} + \frac{2\xi_\alpha}{n^{3/2}} \) where \( \xi_\alpha \) is such that \( \phi(\xi_\alpha) = 1 - \alpha \). \( \phi \) being the standard normal integral and \( \alpha \) is the level of significance. This was shown to be

equivalent to \( \frac{\tau A_n}{\sqrt{n+1}} + \frac{m T_n}{\sqrt{n+1}} \) + polynomial in \( \frac{\tau A_n}{\sqrt{n+1}} \) and \( \frac{m T_n}{\sqrt{n+1}} \ > \frac{\xi_\alpha}{\sqrt{n+1}} \)

+ term of order \( \frac{1}{n^{2\delta}} + \) terms in powers of \( \frac{1}{n} \), where \( A_n \) and \( T_n \) are standardized \( A'_n \) and \( T'_n \) respectively. Thus to some order of approximation we can approximate the ratio \( \frac{A'_n}{T'_n^2} \) by a linear
combination of variables $A_n$, $T_n$ which are bivariately Normally distributed. We therefore require that no higher powers of

$$\frac{A_n}{\sqrt{n+1}} \quad \text{or} \quad \frac{T_n}{\sqrt{n+1}}$$

or crossproducts enter into the expansion for otherwise the approximate evaluation of the power function in terms of $\phi(x)$ would be impossible. We must therefore neglect terms of order $\frac{1}{n}$. Therefore if we chose $\delta < \frac{1}{2}$,

all higher powers of $\frac{A_n}{\sqrt{n+1}}$ and $\frac{T_n}{\sqrt{n+1}}$ are $o\left(\frac{1}{n^{2\delta}}\right)$. Hence $\delta < \frac{1}{2}$.

We chose $\delta > \frac{1}{4}$ so that terms $\frac{A_n}{\sqrt{n+1}}$ and $\frac{T_n}{\sqrt{n+1}}$ come into the expansion. Therefore $\delta$ is restricted to $(1/4, 1/2)$. Weiss did not give sufficient consideration to the possible values of $\delta$. Weiss gave the asymptotic power of the $\sum_{i=1}^{n+1} V_i^2$ test as

$$1 - \Phi(\xi_\alpha - \frac{1}{\sqrt{\frac{n}{2\delta - 1/2}}} \chi(x) dx + \Delta_n),$$

where $\Delta_n$ converges stochastically to zero with $n$. This result is correct if the sequence of alternatives $\{f_n(x)\}$ are those of the form $f_n(x) = 1 + \frac{x}{n^{\delta}}$ where our restrictions are as before but now $\frac{1}{4} < \delta < \frac{1}{2}$ and $\Delta_n$ is of the form $o\left(\frac{1}{n^{2\delta - 1/2}}\right)$.

Weiss' technique will be used in sections 4.4 and 4.5 where we will again elaborate on the restriction that must be placed on $\delta$. 
4.3. Goodness of Fit Problem and Test Statistics

A common problem is that of testing the hypothesis that the common unknown distribution of the independent random variables

$$F_0(X_1), F_0(X_2), \ldots, F_0(X_n)$$

is the uniform distribution over (0,1), i.e.,

$$H_0 : F(X_i) = F_0(X_i).$$

For this test it would be of interest to know its asymptotic power against the sequence of alternatives \( \{F(x)\} \) where

$$F_n(x) = F_0(x) + \frac{1}{n^\delta} \int_0^x r(x') dx'.$$

where \( \int_0^1 r(x') dx' = 0, |r''(x')| < D \) for \( 0 < x' < 1 \) and \( \frac{1}{4} < \delta < \frac{1}{2} \). This is equivalent to the testing problem

$$H_0 : f(x) = 1 \quad 0 < x < 1$$

against Weiss' sequence of alternatives with the restriction

$$\frac{1}{4} < \delta < \frac{1}{2}.$$  

4.4. Power of the Information Statistics

The test \( k' \) rejects when

$$- \sum_{i=1}^{n+1} \frac{\log V_i}{n+1} > C - \frac{1}{2(n+1)} + \log(n+1) + \xi \sqrt{\frac{1}{(n+1)[\frac{n}{6} - 1]}}$$
from Table 3.6.b, where \( C \) is Euler's constant. Forming the exponential copy of \( k' \) we get

\[
\frac{-k'}{n+1} = - \sum_{i=1}^{n+1} \frac{\log Z_i}{n+1} = - \frac{1}{n+1} \sum_{i=1}^{n+1} \log \frac{W_i}{T_n}
\]

\[
= - \frac{1}{(n+1)} \sum_{i=1}^{n+1} \log W_i + \log T_n = - \frac{A_{1n}'}{n+1} + \log T_n
\]

where

\[
\sum_{i=1}^{n+1} \log W_i = A_{1n}'.
\]

\[
E(A_{1n}') = E(\sum_{i=1}^{n+1} (\log a_i + \log U_i))
\]

\[
= \sum_{i=1}^{n+1} \log a_i + (n+1) E(\log U)
\]

\[
= \sum_{i=1}^{n+1} \log a_i - (n+1)C
\]

since

\[
E(\log U) = \int_0^\infty \log u e^{-u} du = - C.
\]

\[
V(A_{1n}') = (n+1) V(\log U) = (n+1) [E(\log^2 U) - C^2],
\]

Now

\[
E(\log^2 U) = \frac{\pi^2}{6} + C^2 \text{ from (2.5.5) and (2.5.6).}
\]
Therefore
\[ V(A_n) = (n+1) \frac{\pi^2}{6}. \]

Let
\[ A_n = \frac{\sum_{i=1}^{n+1} \log a_i}{\sqrt{(n+1) \frac{\pi^2}{6}}}. \]

Since
\[ E(\sum_{i=1}^{n+1} w_i) = E(\sum_{i=1}^{n+1} a_i u_i) = \sum_{i=1}^{n+1} a_i E(u_i) = \sum_{i=1}^{n+1} a_i. \]

\[ V(\sum_{i=1}^{n+1} w_i) = V(\sum_{i=1}^{n+1} a_i u_i) = \sum_{i=1}^{n+1} a_i^2 V(u_i) = \sum_{i=1}^{n+1} a_i^2. \]

Setting
\[ C_n(b) = \sum_{i=1}^{n+1} a_i b = \sum_{i=1}^{n+1} \{f_n[F_n^{-1}(\frac{i}{n+1})]\} - b \]

we then have
\[ E(T_n') = C_n(1) \quad \text{and} \quad V(T_n') = C_n(2). \]

Let
\[ T_n = \frac{T_n' - C_n(1)}{\sqrt{C_n(2)}} \]

\[ \text{Cov}(T_n', A_n) = \frac{Cov(T_n', A_n)}{\sqrt{C_n(2)} \sqrt{(n+1) \frac{\pi^2}{6}}}. \]
\[
= \frac{1}{\sqrt{C_n(2)}} \sqrt{\frac{n}{6}} \text{Cov} \left( \sum a_i U_i, \sum \log a_i U_i \right)
\]

\[
= \frac{1}{\sqrt{(n+1) \frac{\pi^2}{6}}} \sqrt{C_n(2)} \left[ \sum a_i \text{Cov}(U_i, \log U_i) \right]
\]

\[
= \frac{1}{\sqrt{C_n(2)}} \sqrt{\frac{n}{6}} \sqrt{(n+1) \frac{\pi^2}{6}} \left[ \sum a_i \left( E(U_i \log U_i) \right) \right]
\]

\[
- E(U_i) E(\log U_i)
\]

\[
= \frac{C_n(1)}{\sqrt{C_n(2)}} \frac{n}{\sqrt{(n+1) \frac{\pi^2}{6}}} \left[ E(U \log U) - E(U) E(\log U) \right]
\]

\[
= \frac{C_n(1)}{\sqrt{C_n(2)}} \sqrt{\frac{n}{6}} \sqrt{(n+1) \frac{\pi^2}{6}}
\]

since

\[
E(U \log U) = 1 - C \text{ from (2.5.11)} \quad \text{and} \quad E(\log U) = - C.
\]
Now

\[
\frac{C_n(b)}{n+1} = \frac{1}{n+1} \sum_{i=1}^{n+1} \left( f_n \left[ F_n^{-1} \left( \frac{i}{n+1} \right) \right] \right)^{-b} \]

\[
= \frac{1}{n+1} \sum_{i=1}^{n+1} \left( 1 + \frac{r(F_n^{-1}(\frac{i}{n+1}))}{n^\delta} \right)^{-b} \]

\[
= \frac{1}{n+1} \sum_{i=1}^{n+1} \left( 1 - \frac{br(F_n^{-1}(\frac{i}{n+1}))}{n^\delta} \right) \]

\[
+ \frac{b(b+1)}{2n^{2\delta}} r^2 (F_n^{-1}(\frac{i}{n+1})) + 0 \left( \frac{1}{n^{3\delta}} \right) \]

\[
= 1 - \frac{b}{n^\delta} \left\{ \sum_{i=1}^{n+1} \frac{r(F_n^{-1}(\frac{i}{n+1}))}{n+1} \right\} \]

\[
+ \frac{b(b+1)}{2n^{2\delta}} \left\{ \sum_{i=1}^{n+1} \frac{r^2(F_n^{-1}(\frac{i}{n+1}))}{n+1} \right\} + 0 \left( \frac{1}{n^{3\delta}} \right) \]

\[
= 1 - \frac{b}{n^\delta} \left\{ \int_0^1 r(x) \, dx + 0 \left( \frac{1}{n} \right) \right\} \]

\[
+ \frac{b(b+1)}{2n^{2\delta}} \left\{ \int_0^1 r^2(x) \, dx + 0 \left( \frac{1}{n^2} \right) + 0 \left( \frac{1}{n^{1+\delta}} \right) \right\} \]

\[
= 1 + \frac{b(b+1)}{2n^{2\delta}} \int_0^1 r^2(x) \, dx + 0 \left( \frac{1}{n^{3\delta}} \right) + 0 \left( \frac{1}{n^{1+\delta}} \right) \]
\[ + 0 \left( \frac{1}{n^{1+2\delta}} \right) \]

\[ = 1 + \frac{b(b+1)}{2n^{2\delta}} d + 0 \left( \frac{1}{n^{3\delta}} \right) + 0 \left( \frac{1}{n^{1+\delta}} \right) \]

\[ + 0 \left( \frac{1}{n^{1+2\delta}} \right) \text{ where } d = \int_0^1 r^2(x) dx . \]

In Weiss (1965), we are given

\[ \frac{C_n(b)}{n+1} = 1 + \frac{b(b+1)}{2n^{2\delta}} d + \frac{K_n(b)}{n^{1+\delta}} \text{ where } |K_n(b)| < K(b) < \infty . \]

Note there is an obvious misprint in Weiss' formula of \((b-1)\) instead of \((b+1)\). Moreover, there is a more subtle mistake. For suppose that we have \(\frac{1}{4} < \delta < \frac{1}{2}\), then \(3\delta < 1 + \delta\) which implies \(n^{3\delta} < n^{1+\delta}\) which implies \(\frac{1}{n^{1+\delta}} < \frac{1}{n^{3\delta}}\). Therefore rather than Weiss' result we would have to have the expansion

\[ \frac{C_n(b)}{n+1} = 1 + \frac{b(b+1)d}{2n^{2\delta}} + \frac{K_n(b)}{n^{3\delta}} = 1 + \frac{b(b+1)d}{2n^{2\delta}} + 0 \left( \frac{1}{n^{3\delta}} \right) \]

where \(|K_n(b)| < K(b) < \infty\).

Alternatively, we can simplify our considerations of order of approximation if we consider this expansion as

\[ \frac{C_n(b)}{n+1} = 1 + \frac{b(b+1)d}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) . \quad (4.4.2) \]
Now
\[
\text{Cov} \left( T_n, A_{1n} \right) = \frac{C_n(1)}{\sqrt{n+1}} \left( n+1 \right)^{\frac{\Pi^2}{6}} = \frac{\frac{C_n(1)}{n+1}}{\sqrt{n+1}} \left( n+1 \right)^{\frac{\Pi^2}{6}}
\]

\[
= \left( 1 + \frac{d}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) \right) \left( 1 + \frac{3d}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) \right)^{-1/2} \left( \sqrt{\frac{\Pi^2}{6}} \right)^{-1}
\]

\[
= \frac{1}{\sqrt{\frac{\Pi^2}{6}}} (1 + \frac{d}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right)) (1 + \frac{3d}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right))
\]

\[
= \frac{1}{\sqrt{\frac{\Pi^2}{6}}} + 0 \left( \frac{1}{n^{2\delta}} \right) \cdot \quad (4.4.3)
\]

The bivariate central limit theorem shows that asymptotically $A_{1n}$ and $T_n$ have a joint normal distribution with zero means, unit variances and covariance $1/\sqrt{\frac{\Pi^2}{6}}$ to the order of approximation $0(\frac{1}{n^{2\delta}})$

\[
- \frac{A_{1n}}{n+1} = - \frac{A_{1n}}{n+1} \log \frac{T_n}{n+1} + \log (n+1)
\]

\[
= - \frac{A_{1n}}{\sqrt{n+1}} \sqrt{\frac{\Pi^2}{6}} + C - \sum_{i=1}^{n+1} \frac{\log a_i}{n+1} + \log (n+1)
\]
\[ + \log \left( \frac{T_n}{\sqrt{n+1}} \right) \leq \frac{C_n(2)}{n+1} + \frac{C_n(1)}{n+1} \]

\[ = - \frac{A_{1n}}{\sqrt{n+1}} \sqrt{\frac{2}{6}} + C \]

\[ - \sum_{i=1}^{n+1} \frac{\log a_i}{n+1} + \log (n+1) + \log \left[ 1 + \frac{d}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) \right] \]

\[ + \frac{T_n}{\sqrt{n+1}} \left( 1 + \frac{3d}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) \right) \]

\[ = - \frac{A_{1n}}{\sqrt{n+1}} \sqrt{\frac{2}{6}} + C \]

\[ - \sum_{i=1}^{n+1} \frac{\log a_i}{n+1} + \log (n+1) + \frac{d}{n^{2\delta}} + \frac{T_n}{\sqrt{n+1}} + o \left( \frac{1}{n^{2\delta}} \right). \]

Note: The higher powers of \( \frac{T_n}{\sqrt{n+1}} \) do not come into the expression since \( \frac{T_n^2}{n+1} \) is \( o \left( \frac{1}{n^{2\delta}} \right) \) for, if \( \frac{1}{4} < \delta < \frac{1}{2} \), \( \frac{1}{n+1} < \frac{1}{n^{2\delta}} \).

This is why we require the restriction that \( \frac{1}{4} < \delta < \frac{1}{2} \), so as to ensure no terms of higher powers of \( T_n \), \( A_{1n} \) or cross-products, \( A_{1n} T_n \), would come into the approximation. Thus with an order of approximation \( o \left( \frac{1}{n^{2\delta}} \right) \), we ensure this.
Now
\[ - \sum_{i=1}^{n+1} \frac{\log a_i}{n+1} = \sum_{i=1}^{n+1} \log \frac{1}{a_i} \]

where from (4.2.1)

\[ a_i^{-1} = f_n[F_n^{-1}(\frac{i}{n+1})] = 1 + \frac{r(F_n^{-1}(\frac{i}{n+1}))}{n^\delta} \]

\[ \log \frac{1}{a_i} = \frac{r(F_n^{-1}(\frac{i}{n+1}))}{n^\delta} - \frac{1}{2} \frac{r^2(F_n^{-1}(\frac{i}{n+1}))}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) \]

\[ \therefore \sum_{i=1}^{n+1} \frac{\log a_i}{n+1} = \frac{1}{n^\delta} \sum_{i=1}^{n+1} \frac{r(F_n^{-1}(\frac{i}{n+1}))}{n+1} \]

\[ - \frac{1}{2n^{2\delta}} \sum_{i=1}^{n+1} \frac{r^2(F_n^{-1}(\frac{i}{n+1}))}{n+1} + o \left( \frac{1}{n^{2\delta}} \right) \]

\[ = \frac{1}{n^{\delta}} \left( \int_0^1 r(x) \, dx + 0 \left( \frac{1}{n} \right) \right) - \frac{1}{2n^{2\delta}} \left( \int_0^1 r^2(x) \, dx \right) + 0 \left( \frac{1}{n^{\delta}} \right) + o \left( \frac{1}{n^{2\delta}} \right) \]

\[ = - \frac{d}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) \]
Therefore
\[- \frac{k'}{n+1} > C - \frac{1}{2(n+1)} + \log (n+1) + \xi \sqrt{\frac{1}{n+1} \left[ \frac{n^2}{6} - 1 \right]}\]
this implies
\[- \frac{A_{1n}}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{6}} + C - \frac{d}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) + \log (n+1) + \frac{d}{n^{2\delta}}\]
\[+ \frac{T_n}{\sqrt{n+1}} + o \left( \frac{1}{n^{2\delta}} \right) > C + \log (n+1) + \xi \sqrt{\frac{1}{n+1} \left[ \frac{n^2}{6} - 1 \right]} + o \left( \frac{1}{n^{2\delta}} \right)\]
Since
\[V(- A_{1n} \sqrt{\frac{\pi^2}{6}} + T_n) = \frac{\pi^2}{6} + 1 - 2 \sqrt{\frac{\pi^2}{6}} \text{ Cov} (A_{1n}, T_n)\]
\[= \frac{\pi^2}{6} + 1 - 2 \sqrt{\frac{\pi^2}{6}} \left( \frac{1}{\sqrt{\frac{\pi^2}{6}}} \right) + 0 \left( \frac{1}{n^{2\delta}} \right) \text{ from (4.4.3)}\]
\[= \frac{\pi^2}{6} - 1 + 0 \left( \frac{1}{n^{2\delta}} \right)\]
therefore
\[- A_{1n} \sqrt{\frac{\pi^2}{6}} + T_n > \xi \sqrt{\frac{\pi^2}{6} - 1} + \frac{1}{n^{2\delta - 1/2}} \frac{d}{2 \sqrt{\frac{\pi^2}{6} - 1}} + o \left( \frac{1}{n^{2\delta - 1/2}} \right)\]
We find that the asymptotic power of the test $k'$ against $f_n(x)$ is by Weiss' Theorem

$$1 - \phi[\xi_0 - \frac{1}{2} \frac{d}{\sqrt{n2\delta - 1/2}} + o \left( \frac{1}{n^{2\delta - 1/2}} \right)]$$

Since

$$k = \sum_{i=1}^{n+1} \frac{\log \frac{n+1}{V_i}}{n+1} = - \sum_{i=1}^{n+1} \frac{\log V_i}{n+1} - \log (n+1),$$

the asymptotic power of the test $k$ is the same as the asymptotic power of the test $k'$.

The test $k^*$ rejects when

$$k^* = \sum_{i=1}^{n+1} V_i \log V_i + \log (n+1) > (1-C) - \frac{1}{2(n+1)}$$

$$+ \xi_0 \sqrt{\frac{\frac{1}{2}}{n+1}}$$

from Table 3.6.b.

Let us form the exponential copy of $k^*$,

$$\hat{k}^* = \sum_{i=1}^{n+1} Z_i \log Z_i + \log (n+1) = \sum_{i=1}^{n+1} \frac{W_i}{T_n^i} \log \frac{T_n^i}{T_n}$$

$$+ \log (n+1) = \frac{T_n^i}{T_n} \sum_{i=1}^{n+1} W_i \log W_i - \log \frac{T_n^i}{n+1}$$

$$= \frac{A_{2n}}{T_n} - \log \frac{T_n}{n+1}$$

where $A_{2n} = \sum_{i=1}^{n+1} W_i \log W_i$. 
\[ E(A_{2n}') = E(\sum_{i=1}^{n+1} a_i U_i \log (a_i U_i)) = E(\sum_{i=1}^{n+1} a_i U_i [\log a_i + \log U_i]) \]
\[ = \sum_{i=1}^{n+1} a_i \log a_i + E(U \log U) C_n(1) \]
\[ = \sum_{i=1}^{n+1} a_i \log a_i + C_n(1) (1-C) \text{ from (2.5.11)} \]
\[ = b \text{ say.} \quad (4.4.4) \]

\[ V(A_{2n}') = V(\sum_{i=1}^{n+1} (a_i \log a_i) U_i + \sum_{i=1}^{n+1} a_i U_i \log U_i) \]
\[ = \sum_{i=1}^{n+1} a_i^2 \log^2 a_i V(U_i) + \sum_{i=1}^{n+1} a_i^2 V(U_i \log U_i) \]
\[ + 2 \sum_{i=1}^{n+1} a_i^2 \log a_i \text{ Cov}(U_i, U_i \log U_i) \]
\[ = \sum_{i=1}^{n+1} a_i^2 \log^2 a_i + C_n(2) V(U \log U) \]
\[ + 2 \text{ Cov}(U, U \log U) \sum_{i=1}^{n+1} a_i^2 \log a_i , \]
Now
\[ V(U \log U) = E(U^2 \log^2 U) - E(U) E(U \log U) \]

\[ = \Gamma(3)(\psi'(3) + \psi^2(3)) - (1-C)^2 \text{ from (2.5.5) and (2.5.11)} \]

\[ = 2\left(\frac{\pi^2}{6} + 1 - 3C + C^2\right) - (1-C)^2 \]

\[ = \frac{\pi^2}{3} + 1 - 4C + C^2. \]

\[ \text{ Cov } (U, U \log U) = E(U^2 \log U) - E(U) E(U \log U) \]

\[ = (3-2C) - (1-C) = 2 - C \text{ from (2.5.1) and (2.5.11)}. \]

Therefore
\[ V(A_{2n}) = \sum_{i=1}^{n+1} a_i^2 \log^2 a_i + \left(\frac{\pi^2}{3} + 1 - 4C + C^2\right) C_n(2) \]

\[ + 2(2-C) \sum_{i=1}^{n+1} a_i^2 \log a_i = \tau^2 \text{ say}. \]

Let \( A_{2n} \) denote \( \frac{A_{2n} - b}{\tau} \) and \( T_n \) denote \( \frac{T_n - C_n(1)}{\sqrt{C_n(2)}} \).

Since
\[ \text{ Cov } (A_{2n}, T_n) = \text{ Cov } \left( \frac{1}{\tau} A_{2n}, \frac{1}{\sqrt{C_n(2)}} T_n \right) \]
Thus the bivariate central limit theorem shows that asymptotically $A_{2n}$ and $T_n$ have a joint normal distribution with zero means, unit variances and covariance $h$ which to the order of approximation $O\left(\frac{1}{n^{2\delta}}\right)$ will be shown to be

$\sqrt{\frac{11}{3}} + 1 - 4C + c^2$

Now

$\hat{f}^* = \frac{A_{2n}'}{n+1} / \frac{T_n'}{n+1} - \log \frac{T_n'}{n+1}$
\[
\begin{align*}
\frac{\tau A_{2n} + b}{n+1} & - \left( \frac{T_n}{\sqrt{n+1}} \right) \sqrt{\frac{C_n(2)}{n+1}} \\
& + \left( \frac{C_n(1)}{n+1} \right) \log \left( \frac{T_n}{\sqrt{n+1}} \right) \left[ \frac{C_n(2)}{n+1} + \frac{C_n(1)}{n+1} \right]^{-1} \\
& = \frac{\tau A_{2n} + b}{n+1} - (1 + \frac{d}{n^{2\delta}} + o\left( \frac{1}{n^{2\delta}} \right)) + \frac{T_n}{\sqrt{n+1}} \\
& + o\left( \frac{1}{n^{2\delta}} \right) \log \left( 1 + \frac{d}{n^{2\delta}} + o\left( \frac{1}{n^{2\delta}} \right) + \frac{T_n}{\sqrt{n+1}} \right) \\
& + o\left( \frac{1}{n^{2\delta}} \right) \left[ 1 + \frac{d}{n^{2\delta}} + \frac{T_n}{\sqrt{n+1}} + o\left( \frac{1}{n^{2\delta}} \right) \right]^{-1} \\
& = \frac{\tau A_{2n} + b}{n+1} - (1 + \frac{d}{n^{2\delta}} + \frac{T_n}{\sqrt{n+1}} + o\left( \frac{1}{n^{2\delta}} \right)) \left( \frac{d}{n^{2\delta}} + \frac{T_n}{\sqrt{n+1}} \right) \\
& + o\left( \frac{1}{n^{2\delta}} \right) \left\{ 1 - \frac{d}{n^{2\delta}} - \frac{T_n}{\sqrt{n+1}} + o\left( \frac{1}{n^{2\delta}} \right) \right\}
\end{align*}
\]

(Again we see that \( \delta \) should satisfy \( \frac{1}{4} < \delta < \frac{1}{2} \) in order that no higher powers of \( \frac{T_n}{\sqrt{n+1}} \) come into the expansion from expanding the logarithm in powers of \( n \).)
\[ \frac{r^{2}}{n+1} = \frac{1}{n+1} \sum_{i=1}^{n+1} a_{i}^{2} \log^{2} a_{i} + \left( \frac{n^{2}}{3} + 1 - 4C + C^{2} \right) \frac{C_{n+1}^{2}}{n+1} \]

+ \frac{2(2-C)}{n+1} \sum_{i=1}^{n+1} a_{i}^{2} \log a_{i}.

Now

\[ -a_{i} \log a_{i} = \frac{\log \frac{1}{a_{i}}}{\frac{1}{a_{i}}} = \left( \frac{r(F_{n}^{-1}(\frac{i}{n+1}))}{n^\delta} \right) \]

\[ = \frac{1}{2} \frac{r^{2}(F_{n}^{-1}(\frac{i}{n+1}))}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) \left( 1 - \frac{r(F_{n}^{-1}(\frac{i}{n+1}))}{n^\delta} + o \left( \frac{1}{n^\delta} \right) \right) \]

\[ = \frac{r(F_{n}^{-1}(\frac{i}{n+1}))}{n^\delta} - \frac{3}{2} \frac{r^{2}(F_{n}^{-1}(\frac{i}{n+1}))}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right). \]

Therefore

\[ \frac{n+1}{n+1} \sum_{i=1}^{n+1} a_{i} \log a_{i} = \frac{3d}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right). \]

\[ -a_{i}^{2} \log a_{i} = \frac{\log \frac{1}{a_{i}}}{\frac{1}{a_{i}}} = \left( \frac{r(F_{n}^{-1}(\frac{i}{n+1}))}{n^\delta} \right) \]
\[
- \frac{1}{2} \frac{r^2(F_n^{-1}(\frac{i}{n+1}))}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) (1 - \frac{2r(F_n^{-1}(\frac{i}{n+1}))}{n^\delta} + o \left( \frac{1}{n^\delta} \right))
\]

\[
= \frac{r(F_n^{-1}(\frac{i}{n+1}))}{n^{\delta}} - \frac{5}{2} \frac{r^2(F_n^{-1}(\frac{i}{n+1}))}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right).
\]

Therefore

\[
- \sum_{i=1}^{n+1} \frac{a_i^2 \log a_i}{n+1} = - \frac{5d}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right)
\]

Therefore

\[
- \sum_{i=1}^{n+1} \frac{a_i^2 \log^2 a_i}{n+1} = - \frac{d}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right).
\]

Therefore

\[
\frac{\tau^2}{n+1} = \frac{d}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) + \left( \frac{\pi^2}{3} + 1 - 4C + C^2 \right) \left( 1 + \frac{3d}{n^{2\delta}} \right)
\]

\[
+ o \left( \frac{1}{n^{2\delta}} \right) + 2(2-C) \left( \frac{5d}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) \right)
\]

\[
= \frac{\pi^2}{3} + 1 - 4C + C^2 + o \left( \frac{1}{n^{2\delta}} \right).
\]
\[
\sqrt{\frac{n^2}{3}} + 1 - 4C + C^2 + 0 \left( \frac{1}{n^{2\delta}} \right)
\]

\[
\frac{b}{n+1} = \frac{n+1}{n+1} \sum_{i=1}^{n+1} \frac{a_i \log a_i}{n+1} + (1-C) \frac{C_n(1)}{n+1}
\]

\[
= \frac{3d}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) + (1-C) \left( 1 + \frac{d}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) \right)
\]

\[
= (1-C) + \left( \frac{5}{2} - C \right) \frac{d}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right)
\]

\[
\therefore \ k^* > (1-C) - \frac{1}{2(n+1)} + \xi_a \sqrt{n+1} \frac{n^{2\delta} - 3}{n+1}
\]

implies

\[
\left\{ \sqrt{\frac{n^2}{3}} + 1 - 4C + C^2 \frac{A_{2n}}{\sqrt{n+1}} + o \left( \frac{1}{n^{2\delta}} \right) + (1-C) + \left( \frac{5}{2} - C \right) \frac{d}{n^{2\delta}} 
\right. 
\]

\[
+ o \left( \frac{1}{n^{2\delta}} \right) - \frac{d}{n^{2\delta}} - \frac{T_{n}}{\sqrt{n+1}} + o \left( \frac{1}{n^{2\delta}} \right) \right\} \left( 1 - \frac{d}{n^{2\delta}} \right)
\]

\[
- \frac{T_{n}}{\sqrt{n+1}} + o \left( \frac{1}{n^{2\delta}} \right) \right\} > (1-C) - \frac{1}{2(n+1)} + \xi_a \sqrt{n+1} \frac{n^{2\delta} - 3}{n+1}
\]

which implies

\[
\left\{ \sqrt{\frac{n^2}{3}} + 1 - 4C + C^2 \frac{A_{2n}}{\sqrt{n+1}} - (2-C) \frac{T_{n}}{\sqrt{n+1}} \right\} > \xi_a \sqrt{\frac{n^{2\delta} - 3}{n+1}}
\]

\[
- \frac{d}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right).
\]
Now

\[ V(\sqrt{\frac{n^2}{3} + 1 - 4C + C^2} \cdot \frac{A_{2n}}{\sqrt{n+1}} - (2-C) \cdot \frac{T_n}{\sqrt{n+1}}) \]

\[ = \frac{1}{(n+1)} \left[ \frac{n^2}{3} + 1 - 4C + C^2 + (2-C)^2 - 2(2-C) \right] \]

\[ \left( \frac{n^2}{3} + 1 - 4C + C^2 \right) \left( \text{Cov} \left( T_n, A_{2n} \right) \right) \]

\[ = \frac{1}{n+1} \left[ \frac{n^2}{3} - 3 + 2(2-C)^2 - 2(2-C)^2 + 0 \left( \frac{1}{n^{2\delta}} \right) \right] \]

\[ = \frac{n^2}{3n+1} - 3 + 0 \left( \frac{1}{n^{1+2\delta}} \right) \]

since from (4.4.5)

\[ \text{Cov} \left( A_{2n}, T_n \right) = \frac{1}{\sqrt{n+1}} \cdot \sqrt{C_n(2)} \cdot \left[ \frac{1}{n+1} \sum_{i=1}^{n+1} a_i^2 \log a_i \right] \]

\[ + (2-C) \cdot \frac{C_n(2)}{n+1} \]

\[ = \left[ \frac{5d}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) + (2-C) \left( 1 + \frac{3d}{n^{2\delta}} \right) \right] \]
\[
+ o \left( \frac{1}{n^{2\delta}} \right) \left[ \sqrt{\frac{n^2}{3}} + 1 - 4C + C^2 + 0 \left( \frac{1}{n^{2\delta}} \right) \right]^{-1}
\]

\[
(1 - \frac{3d}{2n^2\delta} + o \left( \frac{1}{n^{2\delta}} \right))
\]

\[
= \frac{2 - C}{\sqrt{\frac{n^2}{3}} + 1 - 4C + C^2} + 0 \left( \frac{1}{n^{2\delta}} \right)
\]

(4.4.6)

So

\[
\hat{k}^* \geq (1-C) - \frac{1}{2(n+1)} + \xi \sqrt{\frac{n^2}{3}} - 3
\]

implies

\[
\sqrt{\frac{n^2}{3}} + 1 - 4C + C^2 \frac{A_{2n}}{2 - C} \frac{T_n}{\sqrt{\frac{n^2}{3}} - 3} > \xi \sqrt{\frac{n^2}{3}} - 3
\]

\[
+ o \left( \frac{1}{n^{2\delta-1/2}} \right). \text{ The power is thus}
\]

\[
1 - \phi(\xi - \frac{1}{\sqrt{n^{2\delta-1/2}} \frac{d}{2^{\frac{n^2}{3}} - 3}} + o \left( \frac{1}{n^{2\delta-1/2}} \right))
\]

The test statistic J has critical region
\[ J > 1 - \frac{1}{n+1} + \xi_a \sqrt{\frac{n^2}{6(n+1)}} \]

from Table 3.6.b.

Writing the exponential copy of \( J \), we get

\[ \hat{J} = \sum_{i=1}^{n+1} \left( z_i - \frac{1}{n+1} \right) \log z_i = \sum_{i=1}^{n+1} \left( \frac{w_i}{T_n} \right) \log \frac{w_i}{T_n} - \frac{1}{n+1} \log \frac{w_i}{T_n} \]

\[ = \sum_{i=1}^{n+1} \left( \frac{w_i}{T_n} \right) \log w_i = \frac{1}{T_n} \sum_{i=1}^{n+1} w_i \log w_i \]

\[ - \frac{1}{n+1} \sum_{i=1}^{n+1} \log w_i = \frac{A_{2n}}{T_n} - \frac{1}{n+1} A_{1n} \]

\[ = \frac{A_{2n}/(n+1)}{T_n/(n+1)} - \frac{1}{n+1} A_{1n} \]

From the power considerations of the tests \( k \) and \( k^* \) we know from (4.4.1) and (4.4.4) that

\[ A_{2n} = \frac{\sum_{i=1}^{n+1} w_i \log w_i - (1-C) C_n(1) - \sum_{i=1}^{n+1} a_i \log a_i}{\tau} \]

\[ T_n = \frac{T_n - C_n(1)}{\sqrt{C_n(2)}} \]

\[ = \sum_{i=1}^{n+1} \log w_i + (n+1) C - \sum_{i=1}^{n+1} \log a_i \]

\[ A_{1n} = \frac{\sqrt{(n+1)n^2/6}}{\sqrt{(n+1)n^2/6}} \]
are asymptotically normally distributed with zero means, unit variances and covariances

\[
\text{Cov} (A_{2n}, T_n) = \frac{(2-C)}{\sqrt{n^2 - 3 + (2-C)^2}} + 0 \left( \frac{1}{n^{2\delta}} \right)
\]

from (4.4.6)

\[
\text{Cov} (A_{1n}, T_n) = \frac{C_n(1)}{\sqrt{C_n(2)(n+1)\frac{\Pi^2}{6}}} = \frac{1}{\sqrt{\frac{\Pi^2}{6}}} + 0 \left( \frac{1}{n^{2\delta}} \right)
\]

from (4.4.3)

\[
\text{Cov} (A_{1n}, A_{2n}) = \frac{1}{\sqrt{(n+1)\frac{\Pi^2}{6}}} \text{Cov} \left( \sum_{i=1}^{n+1} \log W_i, \sum_{i=1}^{n+1} W_i \log W_i \right)
\]

\[
= \frac{1}{\sqrt{(n+1)\frac{\Pi^2}{6}}} \text{Cov} \left( \sum_{i=1}^{n+1} \log a_i + \sum_{i=1}^{n+1} \log U_i, \right)
\]

\[
= \frac{1}{\sqrt{(n+1)\frac{\Pi^2}{6}}} \left[ \text{Cov} \left( \sum_{i=1}^{n+1} \log U_i, \sum_{i=1}^{n+1} a_i \log a_i U_i \right) \right]
\]
\[+
\sum_{i=1}^{n+1} a_i U_i \log U_i)\]

\[= \frac{1}{\sqrt{(n+1)^2 \tau}} \left[ \text{Cov}(U, \log U) \sum_{i=1}^{n+1} a_i \log a_i \right.\]

\[+ \text{Cov}(\log U, U \log U) C_n(1)\]

\[= \frac{1}{\sqrt{(n+1)^2 \tau}} \left[ \sum_{i=1}^{n+1} a_i \log a_i \right.\]

\[+ C_n(1) [\psi'(2) + \psi^2(2) + C(1-C)]\]

from (2.5.5)

\[= \frac{1}{\sqrt{(n+1)^2 \tau}} \left[ \sum_{i=1}^{n+1} a_i \log a_i + C_n(1) (\frac{\pi^2}{6} - C) \right]\]

from (2.5.3, 2.5.5, 2.5.6)

\[= \frac{1}{\sqrt{(n+1)^2 \tau}} \left[ \frac{1}{n+1} \sum_{i=1}^{n+1} a_i \log a_i + (\frac{\pi^2}{6} - C) \frac{C_n(1)}{n+1} \right]\]
\[ J = \left\{ \frac{A_{2n}}{\sqrt{n+1}} \cdot \left( \frac{\pi}{\sqrt{n+1}} + \frac{b}{n+1} \right) \cdot \left( \frac{T_n}{\sqrt{n+1}} \cdot \sqrt{\frac{C_n(2)}{n+1}} + \frac{C_n(1)}{n+1} \right) \right\}^{-1} \]

\[ = \frac{A_{2n}}{\sqrt{n+1}} \cdot \sqrt{\frac{\pi^2}{6} - c} + \frac{1}{n+1} \sum_{i=1}^{n+1} \log a_i \]

\[ = \frac{A_{2n}}{\sqrt{n+1}} \cdot \sqrt{\frac{\pi^2}{3} - 3 + (2-c)^2 + o\left(\frac{1}{n^{2\delta}}\right) + (1-c) + \left(\frac{5}{2} - c\right) \frac{\delta}{n^{2\delta}}} \]

\[ + o\left(\frac{1}{n^{2\delta}}\right) \cdot \left(1 + \frac{c}{n^{2\delta}} + \frac{T_n}{\sqrt{n+1}} + o\left(\frac{1}{n^{2\delta}}\right) \right)^{-1} \]
\[-\left\{\frac{A_{1n}}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{6}} - C + \frac{d}{2n^{2\delta}} + o\left(\frac{1}{n^{2\delta}}\right)\right\}\]

\[= \left\{\frac{A_{2n}}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{3}} - 3 + (2-C)^2 + (1-C) + \left(\frac{5}{2} - C\right) \frac{d}{n^{2\delta}}\right\}

+ o\left(\frac{1}{n^{2\delta}}\right) \{1 - \frac{d}{n^{2\delta}} - \frac{T_n}{\sqrt{n+1}} + o\left(\frac{1}{n^{2\delta}}\right)\}

- \left\{\frac{A_{1n}}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{6}} - C + \frac{d}{2n^{2\delta}} + o\left(\frac{1}{n^{2\delta}}\right)\right\}\]

(Again we see that \(\frac{1}{4} < \delta < \frac{1}{2}\) is necessary so that no higher powers of \(T_n/\sqrt{n+1}\) enter the expansion.)

\[= \left\{\frac{A_{2n}}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{3}} - 3 + (2-C)^2 - \frac{T_n}{\sqrt{n+1}} (1-C) - (1-C) \frac{d}{n^{2\delta}}\right\}

+ (1-C) + \left(\frac{5}{2} - C\right) \frac{d}{n^{2\delta}} + o\left(\frac{1}{n^{2\delta}}\right) - \left\{\frac{A_{1n}}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{6}}\right\}

- C + \frac{d}{2n^{2\delta}} + o\left(\frac{1}{n^{2\delta}}\right)\]
Again we see that we need $\frac{1}{4} < \delta < \frac{1}{2}$ in order that no cross-products $A_{2n} T_n / (n+1)$ come into the expansion.

\[
\hat{J} = \frac{A_{2n}}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{3} - 3 + (2-C)^2} - \frac{T_n}{\sqrt{n+1}} (1-C) - \frac{A_{1n}}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{6}} \\
+ 1 + \frac{d}{n^{2\delta}} [(C-1) + \left( \frac{5}{2} - C \right) - \frac{1}{2}] + o \left( \frac{1}{n^{2\delta}} \right)
\]

\[
= \frac{A_{2n}}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{3} - 3 + (2-C)^2} - \frac{T_n}{\sqrt{n+1}} (1-C) - \frac{A_{1n}}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{6}} \\
+ 1 + \frac{d}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right).
\]

Now

\[
\hat{J} > 1 - \frac{1}{n+1} + \xi_n \sqrt{\frac{\pi^2}{6(n+1)}}
\]

implies

\[
\frac{A_{2n}}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{3} - 3 + (2-C)^2} - (1-C) \frac{T_n}{\sqrt{n+1}} - \sqrt{\frac{\pi^2}{6}} \frac{A_{1n}}{\sqrt{n+1}} \\
> \xi_n \sqrt{\frac{\pi^2}{6(n+1)}} - \frac{d}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right).
\]
Now

\[ V \left( \sqrt{\frac{\pi^2}{3} - 3 + (2-C)^2} A_{2n} - (1-C) T_n - \frac{\pi^2}{6} A_{1n} \right) \]

\[ = \frac{\pi^2}{3} - 3 + (2-C)^2 + (1-C)^2 + \frac{\pi^2}{6} \]

\[- 2(1-C) \sqrt{\frac{\pi^2}{3} - 3 + (2-C)^2} \text{ Cov} (A_{2n}, T_n) \]

\[ + 2(1-C) \sqrt{\frac{\pi^2}{6} \text{ Cov}(T_n, A_{1n})} - 2 \frac{\pi^2}{6} \sqrt{\frac{\pi^2}{3} - 3 + (2-C)^2} \]

\[ \text{[Cov}(A_{1n}, A_{2n})] \]

\[ = \frac{\pi^2}{3} - 3 + (2-C)^2 + (1-C)^2 + \frac{\pi^2}{6} \]

\[- 2(1-C) \sqrt{\frac{\pi^2}{3} - 3 + (2-C)^2} \left( \frac{\pi^2}{3} - 3 + (2-C)^2 \right) \]

\[ + 0 \left( \frac{1}{n^{2\delta}} \right) + 2(1-C) \sqrt{\frac{\pi^2}{6}} \left( \frac{1}{\sqrt{\frac{\pi^2}{6}}} + 0 \left( \frac{1}{n^{2\delta}} \right) \right) \]
\[-2\sqrt{\frac{\pi^2}{6}} - \sqrt{\frac{\pi^2}{3}} - 3 + (2-C)^2 \left( \frac{\pi^2}{6} - C \right) \sqrt{\frac{\pi^2}{3}} - 3 + (2-C)^2 \sqrt{\frac{\pi^2}{6}} + 0 \left( \frac{1}{n^{2\delta}} \right) \]

\[= \frac{\pi^2}{3} - 3 + (2-C)^2 + (1-C)^2 + \frac{\pi^2}{6} \]

\[-2(1-C)(2-C) + 2(1-C) - 2\left( \frac{\pi^2}{6} - C \right) + 0 \left( \frac{1}{n^{2\delta}} \right) \]

\[= \frac{\pi^2}{3} - 3 + 4 - 4C + C^2 + 1 - 2C + C^2 + \frac{\pi^2}{6} - 4 \]

\[+ 6C - 2C^2 + 2 - 2C - \frac{\pi^2}{3} + 2C + 0 \left( \frac{1}{n^{2\delta}} \right) \]

\[= \frac{\pi^2}{6} + 0 \left( \frac{1}{n^{2\delta}} \right) \]

\[\therefore J > \xi_\alpha \sqrt{\frac{\pi^2}{6(n+1)}} + 1 - \frac{1}{n+1} \rightarrow \]

\[\frac{A_{2n}}{\sqrt{n+1}} - \sqrt{\frac{\pi^2}{3}} - 3 + (2-C)^2 - (1-C) \frac{T_n}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{6}} - \frac{\xi_\alpha}{\sqrt{n+1}} \sqrt{\frac{\pi^2}{6}} \]

\[\xi_\alpha - \frac{1}{\sqrt{\frac{\pi^2}{6}}} \frac{d}{n^{2\delta-1/2}} + o \left( \frac{1}{n^{2\delta-1/2}} \right). \text{ Therefore the power is} \]
1 - \phi(\xi_\alpha - \frac{1}{\sqrt{\frac{2}{\pi}}} \frac{d}{n^{2\delta-1/2}})
+ o \left(\frac{1}{n^{2\delta-1/2}}\right).

4.5. Power of Kimball's Statistics

The critical region for Kimball's Statistic is from Table 3.6.b.

\[ \sum_{i=1}^{n+1} V_i^r > \frac{\Gamma(r+1)}{(n+1)r-1} - \frac{r(r-1)\Gamma(r+1)}{2(n+1)^r} \]

\[ + \xi_\alpha \sqrt{\frac{\Gamma(2r+1)-(r^2+1)\Gamma^2(r+1)}}{(n+1)^{2r-1}} \]

where \( \xi_\alpha \) is such that \( \phi(\xi_\alpha) = 1 - \alpha \). \( \phi \) being the standard normal integral and \( \alpha \) is the level of significance.

Let us consider the exponential copy of Kimball's statistic, i.e., \( \sum_{i=1}^{n+1} (Z_i)^r \) instead of \( \sum_{i=1}^{n+1} V_i^r \). Now

\[ \sum_{i=1}^{n+1} z_i^r = \sum_{i=1}^{n+1} \left(\frac{W_i}{T_n}\right)^r = \frac{1}{T_n^r} \sum_{i=1}^{n+1} W_i^r = \frac{A'_n}{T'_n} \]

where

\[ A'_n = \sum_{i=1}^{n+1} W_i^r \quad T'_n = \sum_{i=1}^{n+1} W_i. \]
Now

\[
E(A'_n) = \sum_{i=1}^{n+1} a_i^r E(U_i^r) = \sum_{i=1}^{n+1} a_i^r (r+1) = (r+1) C_n(r).
\]

\[
V(A'_n) = \sum_{i=1}^{n+1} a_i^{2r} V(U_i^r) = \sum_{i=1}^{n+1} a_i^{2r} [E(U_i^{2r}) - E^2(U_i^r)]
\]

\[
= [\Gamma(2r+1) - r^2(r+1)] C_n(2r).
\]

Let

\[
A_n = \frac{A'_n - (r+1) C_n(r)}{\sqrt{C_n(2r) [\Gamma(2r+1) - r^2(r+1)]}}
\]

and

\[
T_n = \frac{T'_n - C_n(1)}{\sqrt{C_n(2)}}.
\]

Then by the bivariate central limit theorem \(A_n\) and \(T_n\) have asymptotically a joint normal distribution with zero means, unit variance and will be shown to have covariance

\[
\frac{r\Gamma(r+1)}{\sqrt{[\Gamma(2r+1) - r^2(r+1)]}} + o \left( \frac{1}{n^{2\delta}} \right).
\]

\[
\text{Cov}(A'_n, T'_n) = \sum_{i=1}^{n+1} \text{Cov}(W_i^r, W_i) = \sum_{i=1}^{n+1} [E(W_i^{r+1})]
\]

\[
- E(W_i^r) E(W_i)] = \sum_{i=1}^{n+1} [a_i^{r+1} \Gamma(r+2) - a_i^r \Gamma(r+1) \cdot a_i].
\]
\[ \begin{align*}
= C_n (r+1) \left[ r \Gamma (r+1) \right] \\
\therefore \text{Cov} \left( A_n, T_n \right) &= \frac{C_n (r+1) \left[ r \Gamma (r+1) \right]}{\sqrt{C_n(2)} \sqrt{C_n(2r)} \left[ \Gamma (2r+1) - r^2 (r+1) \right]} \\
= \frac{r \Gamma (r+1)}{\sqrt{[\Gamma (2r+1) - r^2 (r+1)]}} \\
&\frac{C_n (r+1) \left[ \frac{C_n(2)(n+1)}{n+1} \right]^{-1/2}}{\left[ \frac{C_n(2)(n+1)}{n+1} \right]^{-1/2}} \\
&\left[ 1 + \frac{(r+1)(r+2)}{2} \frac{d}{n^2} + o \left( \frac{1}{n^2} \right) \right] \left[ 1 - \frac{3d}{2n^2} \right] \\
&\left[ 1 + o \left( \frac{1}{n^2} \right) \right] \left[ 1 - \frac{r(2r+1)}{2n^2} + o \left( \frac{1}{n^2} \right) \right] \\
&= \frac{r \Gamma (r+1)}{\sqrt{[\Gamma (2r+1) - r^2 (r+1)]}} + o \left( \frac{1}{n^2} \right) \\
\text{Now} \\
\frac{A_n'}{T_n'} &= \frac{1}{(n+1)^{r-1}} \left[ \frac{A_n'}{(n+1)} \right]^{-r} = \frac{1}{(n+1)^{r-1}} \left\{ \frac{A_n}{\sqrt{n+1}} \right\} \\
&\sqrt{[\Gamma (2r+1) - r^2 (r+1)]} \left[ \frac{C_n(2)(n+1)}{n+1} + r(r+1) \frac{C_n(r)(n+1)}{n+1} \right] \\
\end{align*} \]
\[
\left( \frac{T_n}{\sqrt{n+1}} \right)^r \left( \frac{C_n(2)}{n+1} \right) + \left( \frac{C_n(1)}{n+1} \right)^r = \frac{1}{(n+1)^{r-1}} \left( \frac{A_n}{\sqrt{n+1}} \right)^r \sqrt{\left[ \Gamma(2r+1) - \Gamma^2(r+1) \right] + \Gamma(r+1)}
\]

\[
+ o \left( \frac{1}{n^{2\delta}} \right) + \Gamma(r+1) \left( 1 + \frac{r(r+1)d}{2n^2\delta} \right)
\]

\[
+ o \left( \frac{1}{n^{2\delta}} \right) \left( 1 + \frac{d}{n^{2\delta}} + \frac{T_n}{\sqrt{n+1}} + o \left( \frac{1}{n^{2\delta}} \right) \right)^r
\]

\[
= \frac{1}{(n+1)^{r-1}} \left( \frac{A_n}{\sqrt{n+1}} \right)^r \sqrt{\left[ \Gamma(2r+1) - \Gamma^2(r+1) \right] + \Gamma(r+1)}
\]

\[
(1 + \frac{r(r+1)d}{2n^2\delta}) + o \left( \frac{1}{n^{2\delta}} \right) \left( 1 - \frac{rd}{n^{2\delta}} - \frac{T_n}{\sqrt{n+1}} + o \left( \frac{1}{n^{2\delta}} \right) \right)^r
\]

(Again we see that we need \(\frac{1}{4} < \delta < \frac{1}{2}\) to ensure no higher terms of \(T_n/\sqrt{n+1}\) come into the expansion)

\[
= \frac{1}{(n+1)^{r-1}} \left( \frac{A_n}{\sqrt{n+1}} \right)^r \sqrt{\left[ \Gamma(2r+1) - \Gamma^2(r+1) \right] - r\Gamma(r+1) \frac{T_n}{\sqrt{n+1}}}
\]

\[
+ \Gamma(r+1) \left( 1 + \frac{r(r+1)d}{2n^2\delta} \right) + o \left( \frac{1}{n^{2\delta}} \right) - \frac{r\Gamma(r+1)d}{n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right)
\]

\[
= \frac{1}{(n+1)^{r-1}} \left( \frac{A_n}{\sqrt{n+1}} \right)^r \sqrt{\left[ \Gamma(2r+1) - \Gamma^2(r+1) \right]}
\]
\[- r \Gamma(r+1) \frac{T_n}{\sqrt{n+1}} + \Gamma(r+1) + \frac{r(r-1) \Gamma(r+1)}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right) \]

\[ \frac{A_n}{T_{n}^r} > \frac{\Gamma(r+1)}{(n+1)^{r-1}} - \frac{r(r-1) \Gamma(r+1)}{2(n+1)^r} + \xi_a \sqrt{\frac{1}{(n+1)^{2r-1}}} \]

\[ \sqrt{\{ \Gamma(2r+1) - (r^2+1) \Gamma^2(r+1) \}} \rightarrow \frac{A_n}{\sqrt{n+1}} \sqrt{\{ \Gamma(2r+1) - \Gamma^2(r+1) \}} \]

\[- r \Gamma(r+1) \frac{T_n}{\sqrt{n+1}} > \xi_a \sqrt{\frac{\Gamma(2r+1) - (r^2+1) \Gamma^2(r+1)}{n+1}} \]

\[- \frac{r(r-1) \Gamma(r+1)}{2n^{2\delta}} + o \left( \frac{1}{n^{2\delta}} \right). \]

Since

\[ V \left( \frac{A_n}{\sqrt{n+1}} \sqrt{\{ \Gamma(2r+1) - \Gamma^2(r+1) - r \Gamma(r+1) \frac{T_n}{\sqrt{n+1}} \}} \right) \]

\[ = \frac{1}{n+1} \left[ \Gamma(2r+1) - r^2(r+1) + r^2 \Gamma^2(r+1) \right. \]

\[- 2r \Gamma(r+1) \sqrt{\{ \Gamma(2r+1) - \Gamma^2(r+1) \}} \right] \left( \text{Cov} \left( A_n, T_n \right) \right) \]

\[ = \frac{1}{n+1} \left[ \Gamma(2r+1) - r^2(r+1) + r^2 \Gamma^2(r+1) - 2r^2 \Gamma^2(r+1) \right] \]

\[ + o \left( \frac{1}{n^{1+2\delta}} \right) \]
\[ \frac{A_n'}{T_n'} > \frac{\Gamma(r+1)}{(n+1)^{r-1}} - \frac{r(r-1)\Gamma(r+1)}{2(n+1)^r} + \xi_n \sqrt{\frac{\Gamma(2r+1) - (r^2+1)\Gamma^2(r+1)}{(n+1)^{2r-1}}} \]

\[ \frac{A_n}{T_n} > \frac{\Gamma(r+1)}{(n+1)^{r-1}} - \frac{r(r-1)\Gamma(r+1)}{2(n+1)^r} + \xi_n \sqrt{\frac{\Gamma(2r+1) - (r^2+1)\Gamma^2(r+1)}{(n+1)^{2r-1}}} \]

The power of the test \( T_n \) is then

\[ 1 - \phi(\xi_n) - \frac{r(r-1)\Gamma(r+1)d}{2 \sqrt{\Gamma(2r+1) - (r^2+1)\Gamma^2(r+1)} n^{2\delta-1/2}} + o \left( \frac{1}{n^{2\delta-1/2}} \right) \]

In the particular case \( r = 2 \) this reduces to the properties

\[ E(\sum_{i=1}^{n+1} W_i^2) = 2C_n(2), V(\sum_{i=1}^{n+1} W_i^2) = 20C_n(4). \]

The power of the test \( \sum V_i^2 \) is then

\[ 1 - \phi(\xi_n) - \frac{d}{n^{2\delta-1/2}} + o \left( \frac{1}{n^{2\delta-1/2}} \right) \]

Now it is interesting to investigate which value of \( r \) in Kimball's statistic is optimum against Weiss' alternatives.
Let
\[
\phi(r) = \frac{r(r-1)\Gamma(r+1)}{2\sqrt{\Gamma(2r+1) - (r^2+1)\Gamma^2(r+1)}}.
\]

Then from the power of Kimball's statistic against Weiss' alternatives we see that we should choose that value of \( r \) which maximizes \( \phi(r) \).

Now
\[
\log \phi(r) = \log r + \log (r-1) + \log \Gamma(r+1) - \log 2
- \frac{1}{2} \log (\Gamma(2r+1) - (r^2+1)\Gamma^2(r+1)).
\]

\[
\frac{d}{dr} \log \phi(r) = \frac{1}{r} + \frac{1}{r-1} + \frac{d}{dr} \log \Gamma(r+1)
- \frac{1}{2}\frac{d}{dr} \log (\Gamma(2r+1) - (r^2+1)\Gamma^2(r+1))
= \frac{1}{r} + \frac{1}{r-1} + \frac{1}{2}\frac{\Gamma'(2r+1) - 2r\Gamma^2(r+1) - 2(r^2+1)\Gamma(r+1)\frac{d}{dr} \Gamma(r+1)}{\Gamma(2r+1) - (r^2+1)\Gamma^2(r+1)}
= \frac{1}{r} + \frac{1}{r-1} + \frac{1}{2}\frac{\Gamma'(2r+1) - 2r\Gamma^2(r+1) - 2(r^2+1)}{\Gamma(2r+1) - (r^2+1)\Gamma^2(r+1)}
\]

\[
= \frac{1}{r} + \frac{1}{r-1} + \psi(r+1)
\]
\[
\frac{2\Gamma(2r+1)\psi(2r+1)-2r\Gamma(r+1)-(r^2+1)\Gamma^2(r+1)\psi(r+1)}{2[\Gamma(2r+1)-(r^2+1)\Gamma^2(r+1)]}
\]

where

\[
\psi(r) = \frac{d}{dr} \log \Gamma(r), \text{ i.e. see (2.5.1)}.
\]

\[
\frac{d \log \phi(r)}{dr} = \frac{1}{r} + \frac{1}{r-1} + \frac{2\Gamma(2r+1)(\psi(r+1)-\psi(2r+1)) + 2r^2(r+1)}{2[\Gamma(2r+1)-(r^2+1)\Gamma^2(r+1)]}
\]

\[
= \frac{1}{r} + \frac{1}{r-1} - \frac{\Gamma(2r+1)(\psi(2r+1)-\psi(r+1)) - r\Gamma^2(r+1)}{[\Gamma(2r+1)-(r^2+1)\Gamma^2(r+1)]}
\]

\[
\frac{d^2 \log \phi(r)}{dr^2} = -\frac{1}{r^2} - \frac{1}{(r-1)^2} - \frac{1}{[\Gamma(2r+1)-(r^2+1)\Gamma^2(r+1)]^2}
\]

\[
\left[\Gamma(2r+1)-(r^2+1)\Gamma^2(r+1)\right]\left[\Gamma(2r+1)(2\psi'(2r+1)
\right.
\]

\[
-\psi'(r+1)) + 2\Gamma(2r+1)(\psi(2r+1)-\psi(r+1))\psi(2r+1)
\]

\[
-\Gamma^2(r+1)-2r\Gamma^2(r+1)\psi(r+1)) - \{\Gamma(2r+1)(\psi(2r+1)
\]

\[
-\psi(r+1)) - r\Gamma^2(r+1))\{2\Gamma(2r+1)(\psi(2r+1)-2r\Gamma^2(r+1)
\]

\[
-2(r^2+1)\Gamma^2(r+1)\psi(r+1)\}.
\]

In particular we have
\[ \frac{d}{dr} \log \phi(r) \bigg|_{r=2} = \frac{1}{2} + 1 - \frac{1}{4} \left[ 24(\psi(5) - \psi(3)) - 8 \right] \]

\[ = \frac{3}{2} - \frac{1}{4} \left[ 24 \left( \frac{1}{4} + \frac{1}{3} \right) - 8 \right] = 0 \]

since

\[ \psi(z) = \frac{1}{z-1} + \psi(z-1) \text{ from (2.5.2).} \]

\[ \frac{d^2}{dr^2} \log \phi(r) \bigg|_{r=2} = -\frac{1}{4} - 1 - \frac{1}{16} \left[ 96(2\psi'(5) - \psi'(3)) + 112\psi(5) - 16 - 64\psi(3) - 288\psi(5) + 96 \\
+ 240\psi(3) \right] = -\frac{1}{4} - 1 - \frac{1}{16} \left[ 96\psi'(5) - \psi'(3) \right] \\
+ 176\psi(3) - 176\psi(5) + 80 \right] = -\frac{1}{4} - 1 - \frac{1}{16} \left[ 96\psi'(5) \\
+ 96(\psi'(5) - \psi'(3)) - 176 \left( \frac{7}{12} \right) + 80 \right], \]

Because

\[ \psi'(z) = -\frac{1}{(z-1)^2} + \psi'(z-1) \]
and
\[ \psi'(1) = \frac{\pi^2}{6} \]

\[
\frac{d^2 \log \phi(r)}{dr^2} \bigg|_{r=2} = -\frac{1}{4} - 1 - \frac{1}{16} \left[ 96 \left( \frac{\pi^2}{6} - 1 - \frac{1}{2^2} \right) 
- \frac{1}{3^2} - \frac{1}{4^2} \right] + 96 \left( \frac{25}{144} \right) - 176 \left( \frac{7}{12} \right) + 80 \]

\[ = - \frac{5}{4} - \frac{1}{16} \left[ 16\pi^2 - 96 \left( \frac{230}{144} \right) - 176 \left( \frac{7}{12} \right) + 80 \right] \]

\[ = - \frac{5}{4} - \frac{1}{16} \left[ 16\pi^2 - \frac{36864}{144} + 80 \right] = - \frac{5}{4} - \frac{1}{16} \left[ 16\pi^2 - 256 + 80 \right] = 9.75 - \pi^2 \]

\[ = - .1196. \]

Therefore \( \phi(r) \) has a local maximum at \( r = 2 \).

Let us investigate which positive integer \( n \geq 2 \) maximizes \( \phi(r) \). Since \( \phi(r) \) has a local maximum at \( r = 2 \), with \( \phi(2) = 1 \), it would seem sensible to investigate whether \( \phi(n) \leq 1 \) for all positive integers \( n \geq 2 \). Let us set up the induction hypothesis \( \phi(n) \leq 1 \) for a fixed \( n \geq 2 \). Now \( \phi(2) = 1 \) and

\[ \phi(n) \leq 1 + \frac{n(n-1)\Gamma(n+1)}{2^7\Gamma(2n+1) - (n^2+1) \Gamma^2(n+1)} \leq 1 \]

\[ + n^2(n-1)^2 \Gamma^2(n+1) \leq 4 \left[ \Gamma(2n+1) - (n^2+1) \Gamma^2(n+1) \right] \]
Consider

\[
\frac{4\Gamma(2(n+1)+1)}{r^2((n+1)+1)} = \frac{4\Gamma(2n+3)}{r^2(n+2)} = \frac{4(2n+2)(2n+3)\Gamma(2n+1)}{(n+1) r^2(n+1)}
\]

\[
= \frac{2(2n+1)}{(n+1)} \frac{4\Gamma(2n+1)}{r^2(n+1)} \geq \frac{2(2n+1)}{n+1} [n^2(n-1)^2 + 4(n^2+1)]
\]

\[
= \frac{4n^5 - 6n^4 + 16n^3 + 10n^2 + 16n + 8}{n+1}.
\]

Now

\[
(n+1)^2 ((n+1)-1)^2 + 4((n+1)^2+1) = n^4 + 2n^3 + 5n^2
\]

\[+ 8n + 8 = \frac{n^5 + 3n^4 + 7n^3 + 13n^2 + 16n + 8}{n+1}.
\]

Now

\[
(4n^5 - 6n^4 + 16n^3 + 10n^2 + 16n + 8) - (n^5 + 3n^4 + 7n^3 + 13n^2 + 16n + 8) = 3n^5 - 9n^4 + 9n^3 - 3n^2
\]

\[= 3n^2 (n^3 - 3n^2 + 3n - 1) = 3n^2 (n-1)^3 \geq 0
\]

for \(n > 1\).

Therefore \(\phi(n) \leq 1\) for \(n > 1\) implies

\[
\frac{4\Gamma(2(n+1)+1)}{r^2((n+1)+1)} \geq \frac{4n^5 - 6n^4 + 16n^3 + 10n^2 + 16n + 8}{n+1}
\]
\[(n+l)^2 (n+l-1)^2 + 4 ((n+l)^2 + 1)] + \phi(n+1) \leq 1.\]

Therefore \(\phi(n) \leq 1\) for all integers \(n \geq 2\) and in particular we have that \(n = 2\) maximizes \(\phi(r)\) among integral values of \(r \geq 2\).

The induction proof above generalizes to the fact that if \(\phi(r) \leq 1\) for \(r > 1\) then \(\phi(r + m) \leq 1\) for all positive integers \(m\). Therefore if we could show that \(\phi(r) \leq 1\) for \(1 < r < 2\) then \(\phi(r) \leq 1\) for \(r > 1\). The values of \(\phi(r)\) for \(1 < r < 2\) given in Table 4.5 seem to indicate this to be true. Table 4.5 also shows that \(\phi(r)\) is extremely flat around \(r = 2\).

4.6. Summary

From Table 4.6, we see that Greenwood's statistic is the most powerful of the four statistics considered against Weiss' alternatives. However the statistic \(k^*\) is almost as powerful. \(k^*\) is the most powerful of the information statistics against Weiss' alternatives.
Table 4.5. Values of $\phi(r)$ for $1 < r < 2$

<table>
<thead>
<tr>
<th>$r$</th>
<th>$\phi(r)$</th>
<th>$r$</th>
<th>$\phi(r)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>0.9434</td>
<td>1.91</td>
<td>0.9995</td>
</tr>
<tr>
<td>1.2</td>
<td>0.9561</td>
<td>1.92</td>
<td>0.9996</td>
</tr>
<tr>
<td>1.3</td>
<td>0.9670</td>
<td>1.93</td>
<td>0.9997</td>
</tr>
<tr>
<td>1.4</td>
<td>0.9762</td>
<td>1.94</td>
<td>0.9998</td>
</tr>
<tr>
<td>1.5</td>
<td>0.9837</td>
<td>1.95</td>
<td>0.99985</td>
</tr>
<tr>
<td>1.6</td>
<td>0.9898</td>
<td>1.96</td>
<td>0.9999</td>
</tr>
<tr>
<td>1.7</td>
<td>0.9944</td>
<td>1.97</td>
<td>0.99995</td>
</tr>
<tr>
<td>1.8</td>
<td>0.9976</td>
<td>1.98</td>
<td>0.99997</td>
</tr>
<tr>
<td>1.9</td>
<td>0.9994</td>
<td>1.99</td>
<td>0.99999</td>
</tr>
</tbody>
</table>
Table 4.6. Power of the test statistics against Weiss' alternatives

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Power against Weiss' alternatives ( F_n(X) ) where</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kimball's: ( \sum_{i=1}^{n+1} V_i^r )</td>
<td>( 1 - \phi[\xi_a - \frac{r(r-1)r(r+1)d}{2n^{2\delta-1/2} \sqrt{[\Gamma(2r+1) - (r^2+1) \Gamma^2(r+1)]}} + o \left( \frac{1}{n^{2\delta-1/2}} \right) ]</td>
</tr>
<tr>
<td>Greenwood's: ( \sum_{i=1}^{n+1} V_i^2 )</td>
<td>( 1 - \phi[\xi_a - \frac{d}{n^{2\delta-1/2}} + o \left( \frac{1}{n^{2\delta-1/2}} \right) ]</td>
</tr>
<tr>
<td>( k: - \log (n+1) - \frac{1}{n+1} \sum_{i=1}^{n+1} \log V_i )</td>
<td>( 1 - \phi[\xi_a - \frac{6226d}{n^{2\delta-1/2}} + o \left( \frac{1}{n^{2\delta-1/2}} \right) ]</td>
</tr>
<tr>
<td>( k*: \log (n+1) + \sum_{i=1}^{n+1} V_i \log V_i )</td>
<td>( 1 - \phi[\xi_a - \frac{9287d}{n^{2\delta-1/2}} + o \left( \frac{1}{n^{2\delta-1/2}} \right) ]</td>
</tr>
<tr>
<td>( J: \sum_{i=1}^{n+1} V_i \log V_i - \frac{1}{n+1} \sum_{i=1}^{n+1} \log V_i )</td>
<td>( 1 - \phi[\xi_a - \frac{7797d}{n^{2\delta-1/2}} + o \left( \frac{1}{n^{2\delta-1/2}} \right) ]</td>
</tr>
</tbody>
</table>
5. A SAMPLING EXPERIMENT

In order to compare the performance of these four test procedures we apply these test procedures to Durbin's (1961) sampling experiment. We have five random samples of size 50 each from each of the following populations

1. Exponential: \( \exp \left\{ -(x+1) \right\} x \geq -1 \)
2. Normal: \( \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2} \right\} -\infty < x < \infty \)
3. Laplace \( \frac{1}{2\sqrt{2}} \left\{ \exp -\frac{|x|}{2} \right\} -\infty < x < \infty \)

Each of the above distributions has mean zero and variance one. Random deviates from these distributions have been tabulated by Quenouille (1959) and the five sets of three samples in this experiment are the values on the first five pages of Quenouille's tables. Quenouille's tables are given to two decimal places. When ties occurred the values were adjusted according to the following scheme:

\[
\begin{align*}
\text{x.yz} \quad &\rightarrow \quad \text{x.y(z-1) 9} \\
\text{x.yz} \quad &\rightarrow \quad \text{x.y} \quad z \quad 1 \\
\text{x.yz} \quad &\rightarrow \quad \text{x.y(z-1) 9} \\
\text{x.yz} \quad &\rightarrow \quad \text{x.yz} \\
\text{x.yz} \quad &\rightarrow \quad \text{x.yz} \quad 1
\end{align*}
\]

and where the obvious adjustment is made if \( z = 0 \).

Table 5.1 shows the results of this experiment. \( k_0 \) denotes the Kolmogorov-Smirnov test, \( M_{25} \) and \( k_n \) are modifi-
cations of the Kolmogorov-Smirnov test proposed by Durbin (1961). k and k* are the information indices proposed by Kale (1965), J is the divergence between the information indices k and k*, while G is Greenwood's test statistic. The critical regions of the last four tests are those given in Table 3.6.b. If the critical regions given by application of Le Cam's theorem are used the only change is that now the 1st normal sample is just barely not significant at the five per cent level when the test k* is applied.

We see that once again k* and Greenwood's statistic have acted similarly.

Table 5.1. Power of test statistics

<table>
<thead>
<tr>
<th>Population</th>
<th>Test</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential</td>
<td>$\chi^2$</td>
<td>17.20*</td>
<td>24.00**</td>
<td>22.80</td>
<td>38.00</td>
<td>24.80</td>
</tr>
<tr>
<td></td>
<td>$k_0$</td>
<td>0.16</td>
<td>0.18</td>
<td>0.18</td>
<td>0.26</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>$M_{25}$</td>
<td>2.20**</td>
<td>1.71*</td>
<td>2.48</td>
<td>2.44</td>
<td>1.61</td>
</tr>
<tr>
<td></td>
<td>$k_n$</td>
<td>0.23**</td>
<td>0.22**</td>
<td>0.25</td>
<td>0.22</td>
<td>0.18</td>
</tr>
<tr>
<td></td>
<td>$k$</td>
<td>1.02**</td>
<td>0.82*</td>
<td>0.97**</td>
<td>0.92**</td>
<td>0.91**</td>
</tr>
<tr>
<td></td>
<td>$k^*$</td>
<td>0.69**</td>
<td>0.71**</td>
<td>0.74**</td>
<td>0.65**</td>
<td>0.66**</td>
</tr>
<tr>
<td></td>
<td>$J$</td>
<td>1.71**</td>
<td>1.53**</td>
<td>1.70**</td>
<td>1.57**</td>
<td>1.58**</td>
</tr>
<tr>
<td></td>
<td>$G$</td>
<td>0.060**</td>
<td>0.065**</td>
<td>0.062**</td>
<td>0.056**</td>
<td>0.059**</td>
</tr>
</tbody>
</table>

*Significance at the 5% level.

**Significance at the 1% level.
Table 5.1 (Continued)

<table>
<thead>
<tr>
<th>Population</th>
<th>Test</th>
<th>( \chi^2 )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laplace</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_0</td>
<td>0.19*</td>
<td>0.15</td>
<td>0.10</td>
<td>0.14</td>
<td>0.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M_{25}</td>
<td>2.38**</td>
<td>1.23</td>
<td>1.19</td>
<td>0.91</td>
<td>1.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_n</td>
<td>0.25**</td>
<td>0.17*</td>
<td>0.08</td>
<td>0.17*</td>
<td>0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>0.88**</td>
<td>0.69</td>
<td>0.56</td>
<td>0.70</td>
<td>0.75*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k*</td>
<td>0.70**</td>
<td>0.56*</td>
<td>0.43</td>
<td>0.46</td>
<td>0.56*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>1.58**</td>
<td>1.25</td>
<td>0.99</td>
<td>1.16</td>
<td>1.32*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>0.063**</td>
<td>0.050*</td>
<td>0.039</td>
<td>0.039</td>
<td>0.049*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_0</td>
<td>0.17</td>
<td>0.10</td>
<td>0.15</td>
<td>0.14</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M_{25}</td>
<td>1.62*</td>
<td>1.24</td>
<td>0.93</td>
<td>0.70</td>
<td>1.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k_n</td>
<td>0.16</td>
<td>0.12</td>
<td>0.05</td>
<td>0.12</td>
<td>0.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k</td>
<td>0.66</td>
<td>0.65</td>
<td>0.48</td>
<td>0.56</td>
<td>0.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k*</td>
<td>0.54*</td>
<td>0.48</td>
<td>0.38</td>
<td>0.36</td>
<td>0.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J</td>
<td>1.21</td>
<td>1.13</td>
<td>0.86</td>
<td>0.92</td>
<td>1.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>0.048*</td>
<td>0.042</td>
<td>0.036</td>
<td>0.034</td>
<td>0.040</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The critical points are, for

<table>
<thead>
<tr>
<th>( \chi^2 )</th>
<th>k_0</th>
<th>M_{25}</th>
<th>k_n</th>
<th>k</th>
<th>k*</th>
<th>J</th>
<th>G</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>16.92</td>
<td>0.188</td>
<td>1.609</td>
<td>0.170</td>
<td>.7504</td>
<td>.537</td>
<td>1.276</td>
</tr>
<tr>
<td>1%</td>
<td>21.67</td>
<td>0.226</td>
<td>1.967</td>
<td>0.211</td>
<td>.8268</td>
<td>.588</td>
<td>1.398</td>
</tr>
</tbody>
</table>
6. LITERATURE CITED


7. ACKNOWLEDGMENTS

The author wishes to acknowledge his appreciation to Dr. B. K. Kale for his suggestion of the problem and his valuable suggestions and criticisms while directing this study. The author also wishes to thank Mrs. Mary Ann Carney and Mrs. Gretchen Snowden of the Numerical Analysis Programming Group, Iowa State University, for programming assistance.

The National Institutes of Health (Training Program in Biometry, GM-34-09) are also acknowledged for the financial support of this dissertation.