*This research was partially supported by the Air Force Office of Scientific Research under Grant #AFOSR 72-2386.

†A dissertation submitted in partial fulfillment for the requirements for the degree of Doctor of Philosophy under the direction of Professor Gordon D. Simons.

STATISTICAL TESTS BASED ON THE LÉVY AND PROHOROV METRICS†

by

Charles H. Alexander*

Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 965

November, 1974
CHARLES HAYNES ALEXANDER, JR. Statistical Tests Based on the Lévy and Prohorov Metrics. (Under the direction of GORDON D. SIMONS.)

The author develops non-parametric one-sample and two-sample goodness-of-fit tests using as test statistics the Lévy and Prohorov distance between empirical distribution functions. Computational procedures are described for computing the test statistics. Recurrence equations are described for computing the distribution of the two-sample test statistics, using results about the maximal matchings in certain graphs. The asymptotic distribution of the one-sample test statistic is expressed in terms of the distribution of fluctuations in the sample path of the Brownian Bridge stochastic process. Tables of these distributions are given in the appendix. The power of the tests against certain alternatives is discussed, and the results of simulations comparing the power with that of the Kolmogorov-Smirnov test.
ACKNOWLEDGEMENTS

I wish to thank my advisor, Professor G.D. Simons. His suggestions and encouragement combined with a remarkable aversion to mathematical nonsense have assisted the author immensely throughout. More than an advisor, he has proved to be an educator in the broadest sense of the word and it has been a rewarding experience to work with him.

My thanks also go to the members of the faculty of the Department of Statistics who have contributed so much to my graduate education, particularly the members of my examination committee, Professors Hoeffding, Cambanis, Chakravarti, and Wegman.

I am grateful for the financial support provided by the Department of Statistics and also for the assistance of the Department of Mathematics, especially Professors Cannon and Kelly who have been members of my committee.

Let me here add to the praises of June Maxwell who has ably, speedily, and most of cheerfully typed this manuscript under combat conditions.

Finally, I would like to thank my fellow students for their contribution to my education, particularly N.I. Fisher, whose help in the preparation of this manuscript has been most appreciated.
# TABLE OF CONTENTS

Chapter I. Introduction  
1.1 Outline and Summary----------------------------------------------- 1  
1.2 Weak Convergence of Measures-------------------------------------- 4  
1.3 The metrics $d_L$ and $d_P$ ---------------------------------------- 5  
1.4 Asymptotic Distributions Based on the Wiener Process---------- 12  
1.5 Terminology from the Theory of Graphs----------------------------- 14

Chapter 2. Computational Procedures  
2.1 Computation of $d_P(P,Q)$ for General $S$ when $P$ and $Q$ have Finite Support---------------------------------- 16  
2.2 Computation of $d_P(P,Q)$ when $P$ has Finite Support-------- 20  
2.3 Computation of $d_L$ --------------------------------------------- 24

Chapter 3. Null Distribution for the Two-Sample Test Based on $d_P$  
3.1 Introduction------------------------------------------------------ 27  
3.2 Matchings in the Graphs $(X,Y,\Gamma_2)$------------------------------- 29  
3.3 Definition of a Mapping $a: X \rightarrow X$ ------------------------------- 41  
3.4 Classification Scheme for Elements of $X$--------------------- 43  
3.5 Recurrence Equations for $K(u)$, $u \in C$---------------------- 44  
3.6 Reducing the Recurrence Equations------------------------------- 53  
3.7 Computing of $\# D_1(L,s)$-------------------------------------- 55

Chapter 4. Asymptotic Distribution of the Test Statistics  
4.1 Asymptotic Distribution for $d_P$-------------------------------------- 62  
4.2 Asymptotic Distribution for $d_L$----------------------------------- 74  
4.3 Asymptotic Continuity of $f(W^\circ)$------------------------------- 75

Chapter 5. Statistical Properties  
5.1 Related Test Statistics--------------------------------------------- 86  
5.2 Power of the Test Based on $d_P$------------------------------------ 88
Appendices

A.1 Tables of the One Sample Distribution Function-------------------97
A.2 Tables of the Two Sample Distribution Function------------------101
A.3 Bounds on \( \Pr\{f(W^0) \leq c\} \) ---------------------------------110

Bibliography----------------------------------------------------------113
CHAPTER 1.

Introduction

1.1 Outline and Summary

Let \((S, \rho)\) be a metric space. Let \(S\) be the class of Borel subsets of \(S\) and let \(P(S)\) be the space of all probability measures on \((S,S)\). A number of ways of defining a metric on the space \(P(S)\) have been suggested and it seems natural to use such a metric to construct tests of statistical hypotheses. For a one-sample goodness-of-fit test, the metric may be used to measure the distance between the empirical measure of a random sample and the probability measure from which the sample is hypothesized to have been drawn. (By "empirical measure" is meant the measure which puts equal mass at all the sample points and zero mass elsewhere.) The corresponding two-sample test for identical distribution would use as a test statistic the distance between the empirical measures of the two samples. The standard example is the Kolmogorov-Smirnov test, where \(S\) is the real line and the uniform metric on \(P(S)\) is used, i.e., the metric \(d_K(P,Q) = \sup_{-\infty < t < \infty} |P(-\infty,t] - Q(-\infty,t]|\). The Cramér-von Mises test is based on a directed distance between the empirical distribution function \(F_n\) and the null distribution function \(F\), \(d_C(F_n,F) = \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 \psi(F(x)) dF(x)\), where \(\psi(\cdot)\) is a weight function. \(d_C\) is not metric since it is not symmetric in its arguments.) Other tests based on metrics on \(P(S)\) have been suggested in [1] and [26].
The so-called "Levy" metric, \( d_L \), which is defined for \( S = \mathbb{E}^k \) (k-dimensional Euclidean space) and has as its metric topology the topology of weak convergence of probability measures, and the "Prohorov" metric, \( d_P \), which generates the same topology on \( P(S) \) for a more general class of metric spaces \( S \), have been given little discussion in the context of hypothesis testing. Dudley [12] remarks that there exists no practical means for computing the value of the test statistic for the test based on \( d_P \). Zolotarev [32] suggests that a barrier to the use of tests based on \( d_L \) is the lack of knowledge about its characteristic function, which presumably prevents the development of the asymptotic distribution theory along the lines of that for the Kolmogorov-Smirnov test.

An additional difficulty is that while the statistical tests based on \( d_K \) are distribution-free for continuous null distributions, the analogous tests based on \( d_P \) and \( d_L \) are not. In the case \( S = \mathbb{E}^1 \), we can construct a distribution-free one-sample test by using \( d_P \) or \( d_L \) to measure the distance between the uniform measure on \([0,1]\) and the empirical measure of \( \{ F(X_i) \} \), where \( F \) is the null distribution function and \( \{ X_i \} \) denotes the sample. The corresponding two-sample test would use \( d_P \) or \( d_L \) to measure the distance between the empirical measures of \( \{ s_i/N \} \) and \( \{ r_i/N \} \), where \( r_1, \ldots, r_m \) and \( s_1, \ldots, s_n \) are the respective ranks of the two samples among the pooled observations, and \( N = m+n \). Alternatively, the two-sample test statistic could be the distance between the empirical measure of \( \{ s_i/N \} \) and the empirical measure of \( \{ 1/N, \ldots, N/N \} \).
In what follows, we will discuss the computation of these test statistics based on $d_P$ and $d_L$, their distributions under the null hypothesis, and some of their statistical properties. For $d_P$, exact distributions of the two-sample test statistic are computed for the case of equal sample sizes, corresponding to the results of Massey [21] for the Kolmogorov-Smirnov test. We are unable to give similar results for the one-sample test, as are found by Birnbaum and Tingey [6] for the Kolmogorov-Smirnov test. The asymptotic distributions of the test statistics are related to the distribution of particular fluctuations in the sample path of the "Brownian Bridge" process. As with $d_K$, there is a single asymptotic distribution for the one-sample and two-sample test statistics. We have not succeeded in finding a closed-form expression for this distribution, although some crude upper and lower bounds have been computed. The asymptotic distribution of the test statistics based on $d_L$ is closely related to that of the statistics based on $d_K$, and can be given explicitly.

In Chapter 1, we give definitions and a brief description of weak convergence and of the metrics $d_L$ and $d_P$. Results and relationships are collected here which will be of use in later chapters. Chapter 2 contains computational procedures for $d_P$ and $d_L$. One algorithm computes $d_P(P,Q)$ for measures $P$ and $Q$ which each have finite support. For $S = E^1$, a more efficient algorithm computes $d_P(P,Q)$ where $P$ has finite support and $Q$ is arbitrary. In Chapter 3, recurrence equations are derived which are used to compute the exact null distribution of the two-sample test statistic for equal sample sizes. Chapter 4 contains the results on the asymptotic distributions of the test statistics.
Chapter 5 consists of a brief discussion of the statistical properties of the tests, including the results of simulations comparing the power of these tests with the power of the Kolmogorov-Smirnov test against normal location and scale alternatives, and against "Lehmann" nonparametric alternatives. An appendix contains tables of the distribution of the two-sample test statistic obtained using the methods of Chapter 3, and the results of Monte Carlo estimation of the exact distribution of the one-sample test statistic for sample sizes 5, 10, 20, 40, 60, and 80, as well as a description of the computer programming methods used.

1.2 **Weak convergence of measures.**

As before, let \((S,\rho)\) be a metric space and let \(S\) be the \(\sigma\)-field of all Borel subsets of \(S\). Let \(P(S)\) be the space of all probability measures on \((S,S)\). Let \(\{P_n\}\) and \(P\) be elements of \(P(S)\).

**Definition 1.2.1:** (a) \(\{P_n\}\) is said to converge weakly to \(P\) (written \(P_n \Rightarrow P\)) if and only if \(\int_S f \, dP_n \to \int_S f \, dP\) for every bounded continuous real-valued function \(f\) defined on \(S\).

(b) If \(S = E^k\) and \(F(x) = P\{y \leq x\}\) and \(F_n(x) = P_n\{y \leq x\}\), \(n = 1,2,\ldots\), then we say that \(F_n \Rightarrow F\) if and only if \(P_n \Rightarrow P\). (Note: if \(x = (x_1, \ldots, x_k)\) and \(y = (y_1, \ldots, y_k)\), we write \(y \leq x\) if and only if \(y_i \leq x_i\), \(i = 1, \ldots, k\).)

(c) Let \(T\) be a metric space and \(T\) the Borel subsets of \(T\). If \(\{X_n\}\) and \(X\) are random elements (i.e., measurable mappings) from a probability space \((\Omega,A,P)\) to \((T,T)\), we say that \(X_n \Rightarrow X\) if and only if \(P^{X_n^{-1}} \Rightarrow P^{X^{-1}}\) as probability measures in \(P(T)\). In this case we say that \(\{X_n\}\) converges in distribution to \(X\).
Definition 1.2.2: For each $A \in S$, let $\delta A$ denote the boundary of $A$. Let $P \in P(S)$. $A \in S$ is called a $P$-continuity set if and only if $P(\delta A) = 0$. (Note that $\delta A \in S$, since $\delta A$ is closed.)

We will usually use the following equivalent characterization of weak convergence instead of the original definition. The next two Propositions are found in [5], where additional equivalent conditions for weak convergence are given.

Proposition 1.2.3: $P_n \Rightarrow P$ if and only if $P_n(A) \to P(A)$ for all $P$-continuity sets $A$. If $S = E^k$ and $\{F_n\}$ and $F$ are defined as before, then $F_n \Rightarrow F$ if and only if $F_n(x) \to F(x)$ for every point $x \in E^k$ such that $F$ is continuous at $x$.

Proposition 1.2.4: Let $T$ be a separable metric space and $T$ the Borel $\sigma$-field of $T$. Let $\{h_n\}$ and $h$ be measurable functions from $(S,S)$ to $(T,T)$. Let $E$ be the set of all $x \in S$ for which there exists some sequence $\{x_n\}$ of points of $S$ such that $x_n \to x$ but $h_n(x_n) \not\to h(x)$. (Since $T$ is separable, $E \in S$.) If $P_n \Rightarrow P$ and $P(E) = 0$, then $P_n h^{-1} \Rightarrow P h^{-1}$.

Proposition 1.2.4 frequently appears through its corollary that if $h$ is a continuous function from $(S,S)$ to $(T,T)$, then $P_n \Rightarrow P$ implies $P_n h^{-1} \Rightarrow P h^{-1}$.

1.3 The metrics $d_L$ and $d_P$.

Definition 1.3.1: (a) Let $S = E^k$. Define the metric $d_K$ on $P(S)$ by $d_K(P,Q) = \inf\{\varepsilon > 0 : P(-\infty,x] - \varepsilon \leq Q(-\infty,x] \leq P(-\infty,x] + \varepsilon\}$;
for all \( x \in E^k \).

(b) Let \( S = E^k \). Define the metric \( d_L \) on \( \mathcal{P}(S) \) by
\[
d_L(P,Q) = \inf\{\varepsilon > 0 : P(-\infty, x - \varepsilon] \leq Q(-\infty, x] \leq P(-\infty, x + \varepsilon] + \varepsilon; \text{ for all } x \in E^k\}.
\]

(c) Let \((S, \rho)\) be an arbitrary metric space. Define the metric \( d_P \) on \( \mathcal{P}(S) \) by
\[
d_P(P,Q) = \inf\{\varepsilon > 0 : Q(F) \leq P(F^\varepsilon) + \varepsilon; \text{ for all closed sets } F \subseteq S\}, \text{ where } F^\varepsilon = \{y \in S : \text{there exists } x \in F \text{ such that } \rho(x,y) < \varepsilon\}.
\]

(Note: the definition of \( d_P \) is not apparently symmetric in \( P \) and \( Q \). However, suppose that \( Q(F) \leq P(F^\varepsilon) + \varepsilon \) for all closed sets \( F \subseteq S \).

Let \( G \subseteq S \) be closed, then \( G^\varepsilon \) is open. Since \( G \subseteq S - (S - G^\varepsilon) \subseteq G^\varepsilon \) and \( Q(S - G^\varepsilon) \leq P((S - G^\varepsilon)^\varepsilon) + \varepsilon \), we have the inequalities
\[
P(G) \leq P(S - (S - G^\varepsilon)^\varepsilon) \leq Q(G^\varepsilon) + \varepsilon, \text{ hence } d_P(P, Q) = d_P(Q, P).
\]

This proof appears in [30]. Previously the definition of the Prohorov distance between \( P \) and \( Q \) was usually given as what, in our notation would be,
\[
\min(d_P(P, Q), d_P(Q, P)).
\]

This latter definition is more general since when dealing with a larger space of measures so that \( P(S) \neq Q(S) \), \( d_P(P, Q) \) may not equal \( d_P(Q, P) \).

**Proposition 1.3.2:** (a) Let \( S \) be separable and complete. \( P_n \Rightarrow P \)
if and only if \( d_P(P_n, P) \to 0 \).

(b) If \( S = E^k \), \( P_n \Rightarrow P \) if and only if \( d_L(P_n, P) \to 0 \).

The proof of (a) first appears in [24]. A proof of (b) for \( k = 1 \) is given in [14]. Bhattacharya in [3] observes that essentially the same proof works for \( k > 1 \). Prohorov [24] shows that the metric space \((\mathcal{P}(S), d_P)\) is separable and complete when \( S \) is separable and complete.
These assumptions about $S$ can be weakened (see [5]).

1.3.2 shows that $d_P$ and $d_L$ are equivalent metrics, i.e., they generate the same topology on $P(S)$, when $S = E^k$. However, they do not generate the same uniformities, as shown by the following example from [11]. Define $P_n$ and $Q_n$ by $P_n(2j) = Q_n(2j+1) = 1/n$, $j = 1, \ldots, n$. Then $d_L(P_n, Q_n) = 1/n$, while $d_P(P_n, Q_n) = 1/2$. The topologies generated by these metrics are relevant to their statistical properties. Two examples to this effect are suggested by Hampel's remark in [15] that $d_P$ gives a "literal quantitative description" of "rounding of the observations" and "the occurrence of gross errors."

**Example 1.3.3:** Let $x_1, \ldots, x_n$ be real numbers and define $y_1 = x_1 + \varepsilon$, and $y_i = x_i$, for $i = 2, 3, \ldots, n$. Let $P_n$ and $Q_n$ be the empirical measures of $\{x_i\}$ and $\{y_i\}$ respectively. Assume that $0 < \varepsilon < \min |x_i - x_j|$. Then $d_P(P_n, Q_n) = d_L(P_n, Q_n) = \min(\varepsilon, 1/n)$, while $d_K(P_n, Q_n) = 1/n$. In this case, when $\varepsilon < 1/n$, $d_L$ and $d_P$ give a better idea of the size of the "error" in the first observation.

**Example 1.3.4:** Now define $y_i = x_i + \varepsilon$, $i = 1, \ldots, k$, $y_i = x_i$ for $i = k+1, \ldots, n$. Let $\varepsilon > k/n$. Then $d_P(P_n, Q_n) = k/n$, while $d_L(P_n, Q_n) = 1/n$.

Neither $d_L$ nor $d_P$ generates a distribution-free test. In particular, neither metric is invariant under transformations corresponding to changes of scale in the data. This lack of invariance must hold for any metric which generates the topology of weak convergence on $P(E^k)$. 
The following Glivenko-Cantelli-type result is due to Varadarajan [31].

**Proposition 1.3.5:** Let $P \in P(S)$, where $S$ is separable and complete, and let $P_n$ be the empirical measure of a sample of size $n$ from $P$. Then $P_n \Rightarrow P$ a.s. Hence $d_P(P_n, P) \to 0$ a.s., and for $S = E^k$, $d_L(P_n, P) \to 0$ a.s. \hfill \blacksquare

The definitions which we have given for $d_P$ and $d_L$ are those most often found in the literature. They are actually not the most convenient for our purposes and the following results show that certain modifications can be made. Let $F^\varepsilon = \{y \in S: x \in F \text{ such that } \rho(x, y) \leq \varepsilon\}$.

**Proposition 1.3.6:** (a) Let $P, Q \in P(S)$. Then $d_P(P, Q) = \inf\{\varepsilon > 0: Q(F) \leq P(F^\varepsilon) + \varepsilon; \text{ for all closed } F \subseteq S\}$.

(b) Let $P, Q \in P(E^k)$. Then $d_L(P, Q) = \inf\{\varepsilon > 0: P(-\infty, x-\varepsilon) - \varepsilon \leq Q(-\infty, x] \leq P(-\infty, x+\varepsilon] + \varepsilon; \text{ for all } x \in E^k\}$.

**Proof:** (a) For each closed $F \subseteq S$,

$P(F^\varepsilon) + \varepsilon \leq P(F^\varepsilon) + \varepsilon \leq P(F^\varepsilon') + \varepsilon'$, for each $\varepsilon > 0$ and $\varepsilon' > \varepsilon$.

(b) For each $x \in E^k$,

$P(-\infty, x-\varepsilon') - \varepsilon' \leq P(-\infty, x-\varepsilon) - \varepsilon \leq P(-\infty, x] - \varepsilon$, for each $\varepsilon > 0$ and $\varepsilon' > \varepsilon$. \hfill \blacksquare

**Definition 1.3.7:** For given $P, Q \in P(S)$, define the function $V_P(P, Q; \varepsilon)$ of $\varepsilon$ by $V_P(P, Q; \varepsilon) = \sup(P(F) - Q(F^\varepsilon))$, the supremum being taken over all closed $F \subseteq S$. Define $V_L(P, Q; \varepsilon) = \max\{\sup_{x \in E^k}(P(-\infty, x-\varepsilon) - Q(-\infty, x]), \sup_{x \in E^k}(Q(-\infty, x] - P(-\infty, x+\varepsilon))\}$. \hfill \blacksquare
Proposition 1.3.8: (a) For $c \in [0, 1]$, $d_P(P, Q) \leq c$ if and only if $V_P(P, Q; c) \leq c$.

(b) For $c \in [0, 1]$, $d_L(P, Q) \leq c$ if and only if $V_L(P, Q; c) \leq c$.

Proof: (a) By 1.3.6 (a), if $V_P(P, Q; c) \leq c$, then $d_P(P, Q) \leq c$.

Conversely, $d_P(P, Q) \leq c$ implies for $c' > c$ that $P(F) \leq Q(F^{c'}) + c'$ for all closed $F \subseteq S$. But for each closed set $F$,

$$\lim_{c' \downarrow c} Q(F^{c'}) = Q(F^c),$$

so $P(F) \leq Q(F^c) + c$ for all closed $F \subseteq S$, i.e., $V_P(P, Q; c) \leq c$.

(b) Similarly, $\lim_{c' \downarrow c} P(-\infty, x-c') = P(-\infty, x-c)$ and $\lim_{c' \downarrow c} P(-\infty, x+c') = P(-\infty, x+c)$.

The next two results show that when $S = \mathbb{R}^1$ or $S = [0, 1]$ under the Euclidean metric, then in defining $d_P$ it is not necessary to consider all closed subsets of $S$, but that a smaller class of sets will suffice.

Definition 1.3.9: Let $F$ be the class of all closed subsets of $[0, 1]$ which can be expressed as the union of finitely many disjoint closed intervals. Let $F^{(2\delta)}$ be the class of all members of $F$ such that if we write $F$ as the union of the disjoint intervals $I_j = [b_j, c_j]$, $j = 1, \ldots, m$, where $0 \leq b_1 \leq c_1 < b_2 \leq c_2 < \ldots < b_m \leq c_m \leq 1$, then

$b_i+1 - c_i > 2\delta$ for $i = 1, \ldots, m-1$; and if $b_1 \neq 0$, then $b_1 > \delta$;

and if $c_m \neq 1$, then $c_m < 1 - \delta$.

Proposition 1.3.10: Let $P, Q \in \mathcal{P}([0, 1])$. Then $d_P(P, Q) \leq c$ if and only if

$$\sup_{G \in F^{(2\delta)}} (P(G) - Q(G^c)) \leq c.$$


Proof: Let $\varepsilon > 0$ be arbitrary. It suffices to show that

$$V_{2\varepsilon}(P,Q,c) = \sup_{G \in F(2\varepsilon)} (P(G) - Q(G^c))$$

For each closed set $F \subseteq [0,1]$, let the set $G(F)$ be the union of $F$ with every interval $[c_i, b_{i+1}]$ such that $b_{i+1} - c_i \leq 2\varepsilon$, and with $[0,b_1]$ if $b_1 \leq \delta$, and with $[c_m,1]$ if $c_m \geq 1 - \delta$; where $\{b_i\}$ and $\{c_i\}$ are as in Definition 1.3.9. Then $(G(F))^c = F^c$ and $F \subseteq G(F)$, so $P(F) - Q(F^c) \leq P(G) - Q(G^c)$. But $G(F) \in F(2\varepsilon)$.

The proof of the same result for $P,Q \in P(E^1)$ is similar except that the sets of $F$ are allowed to consist of the unions of finite or semi-infinite intervals. As immediate corollaries to 1.3.10 it follows that $d_{\varepsilon}(P,Q) \leq c$ if and only if $\sup_{F \in F} (P(F) - Q(F^c)) \leq c$, and that $d_{\varepsilon}(P,Q) = \inf\{\varepsilon > 0 : P(F) \leq Q(F^c) + \varepsilon \text{ for all } F \in F\} = \inf\{\varepsilon > 0 : P(F) \leq Q(F^c) + \varepsilon \text{ for all } F \in F(2\varepsilon)\}$.

It follows from Definition 1.3.1 (and the fact that $d_{\varepsilon}(P,Q) = d_{\varepsilon}(Q,P)$) that for $P,Q \in P(E^k)$, $d_{\varepsilon}(P,Q) \leq d_L(P,Q)$ and also that $d_{\varepsilon}(P,Q) \leq d_L(P,Q)$, since the inequalities which must, in the definition of $d_{\varepsilon}$, be satisfied for all closed sets $F$ need only be satisfied for sets of the form $(-\infty,x]$ or $[x,\infty)$ in the definition of $d_L$.

No such general inequality exists between $d_{\varepsilon}$ and $d_L$. However, when dealing with measures in $P([0,1])$ and when $P = U$, the uniform measure, the following relationships hold.

**Proposition 1.3.11:** Let $S = [0,1]$ and $\rho(x,y) = |x-y|$. Let $U$ be Lebesgue measure restricted to $S$. Let $V \in P(S)$ be arbitrary. Then

$$d_{\varepsilon}(U,V)/2 \leq d_{\varepsilon}(U,V) \leq d_L(U,V).$$
Proof: If for some \( x \in [0, 1] \), \(|U(0,x) - V(0,x)| \geq 2\varepsilon > 0\), then either
\[
V([0,x]) \geq U([0,x]^c) + \varepsilon,
\]
or
\[
V([x,1]) \geq U([x,1]^c) + \varepsilon,
\]
and hence \( d_p(U,V) \geq \varepsilon \). Therefore, \( d_p(U,V) \geq d_K(U,V)/2 \).

To show \( d_p(U,V) \leq d_K(U,V) \), suppose \( 0 < \varepsilon < d_p(U,V) \). Then we may choose a closed set \( F \) such that \( V(F) > U(F^c) + \varepsilon \). We may choose \( F \) such that \( F \in F(2\varepsilon) \), as in Proposition 1.3.10. Write \( F \) as the union of the disjoint intervals \( I_j = [b_j,c_j], j=1,\ldots,m \), where \( 0 \leq b_1 < c_1 < b_2 < c_2 < \cdots < b_m < c_m \leq 1 \).

We have one of three cases

(i) \( \varepsilon < b_1 \) and \( c_m < 1 - \varepsilon \);

(ii) \( \varepsilon < b_1 \) and \( c_m = 1 \), or \( b_1 = 0 \) and \( c_m < 1 - \varepsilon \);

(iii) \( 0 = b_1 \) and \( 1 = c_m \).

Then \( U(F^c) = U(F) + \lambda \varepsilon \), where \( \lambda = 2m \) in case (i); \( \lambda = 2m-1 \) in case (ii); \( \lambda = 2m-2 \) in case (ii); and hence
\[
\sum_{j=1}^{m} V(I_j) \geq U(F^c) + \varepsilon = (\lambda + 1)\varepsilon + \sum_{j=1}^{m} U(I_j).
\]

Suppose case (i) holds. Then for some \( j, 1 \leq j \leq m \), \( V(I_j) > U(I_j) + 2\varepsilon \) and hence \( 2\varepsilon < (V-U)(I_j) \leq 2d_K(U,V) \).

Suppose case (ii) holds and for definiteness that \( c_m = 1 \). Then
\[
\text{either } V([b_m,1]) \geq U([b_m,1]) + \varepsilon, \text{ so that } d_K(U,V) \geq \varepsilon, \text{ or }
\]
\[
\sum_{j=1}^{m-1} V(I_j) \geq 2(m-1)\varepsilon + \sum_{j=1}^{m-1} U(I_j),
\]
and hence for some \( j, 1 \leq j \leq m-1 \), \( V(I_j) > U(I_j) + 2\varepsilon \), so that \( 2d_K(U,V) \geq 2\varepsilon \).

Suppose case (iii) holds. Then either \( V([0,c_1]) \geq U([0,c_1]) + \varepsilon \) or \( V([b_m,1]) \geq U([b_m,1]) + \varepsilon \), so that \( d_K(U,V) \geq \varepsilon \), or
\[
\sum_{j=2}^{m-1} V(I_j) \geq 2(m-2)\varepsilon + \sum_{j=2}^{m-1} U(I_j),
\]
and hence for some \( j, 2 \leq j \leq m-1 \), \( V(I_j) \geq U(I_j) + 2\varepsilon \) so that 
\[ 2d_K(U,V) \geq 2\varepsilon. \]
In all three cases \( \varepsilon \leq d_K(U,V) \). Letting \( \varepsilon \downarrow d_P(U,V) \), the proof is complete. \( \square \)

This proof is a modification of that of Dudley in [12], where it is shown that \( d_K(U,V)/2 \leq d_P(U,V) \leq 2d_K(U,V) \). The bounds given are at least very nearly the best possible. Indeed, if \( V \) is the measure in \( P([0,1]) \) which places unit mass at 0, then \( d_K(U,V) = 1 \), while \( d_P(U,V) = 1/2 \). Let \( k \geq 2 \). If \( V^{(k)} \) places mass \( 1/2k \) at the points 0 and 1 and mass \( 1/k \) at \( 1/k, 2/k, \ldots, (k-1)/k \), then 
\[ d_P(V^{(k)},U) = 1/(2k+1), \]
while 
\[ d_K(V^{(k)},U) = 1/2k. \]
Hence as \( k \to \infty \), 
\[ d_P(U,V^{(k)}) / d_K(U,V^{(k)}) \to 1. \] We have no example in which 
\[ d_P(U,V) = d_K(U,V) > 0. \]

1.3.11 shows that if \( U_n \) is the empirical measure of a sample of size \( n \) from \( U \), then \( n^{-1}E(d_P(U_n,U)) \) is bounded and bounded away from zero. Dudley [12] also gives asymptotic upper bounds on \( E(d_P(V_n,V)) \), where \( V \in P(S) \), for a more general class of metric spaces \( S \), where \( V_n \) is the empirical measure of a sample of size \( n \) from \( V \).

1.4 Asymptotic distributions based on the Wiener process.

Definition 1.4.1: (a) The probability measure \( W \) defined on \( C[0,1] \) such that for \( t \in [0,1] \), \( W(t) \) is normally distributed with \( E(W(t)) = 0 \) and \( \text{var}(W(t)) = t \), \( W(0) = 0 \) a.s.) and such that for \( 0 \leq t_0 \leq t_1 \leq \ldots \leq t_k \leq 1 \), the random variables \( W(t_1) - W(t_0), W(t_2) - W(t_1), \ldots, W(t_k) - W(t_{k-1}) \) are independent, will be called Wiener measure.

It follows that \( E(W(t)W(s)) = \min(s,t) \), for \( s, t \in [0,1] \).
(b) The measure $W^o$ on $C[0,1]$ with $W^o(t)$ normal, $E(W^o(t)) = 0$, and $E(W^o(t)W^o(s)) = s(1-t)$ for $s \leq t$ will be called the Brownian Bridge. We may define $W^o$ by $W^o(t) = W(t) - tW(1)$, $0 \leq t \leq 1$.
(Alternatively $W(t)$ may be defined for $t \in [0,\infty)$ and then $W^o(s)$ is defined as $(1-s)W(s/(1-s))$, $0 \leq s \leq 1$.) The "sample path" $W(t)$ is a continuous function on $[0,1]$, and hence $W^o(t)$ is also. $W^o(t)$ may also be thought of as being distributed as $W(t)$ conditioned on the event \{W(1) = 0\}.

$W^o$ has an important role in deriving the asymptotic distribution for a number of test statistics, including the Kolmogorov-Smirnov, making use of Proposition 1.2.4. This was first recognized by Doob [10], and the proof was made rigorous by Donsker [9]. Other examples of the application of this method are [4] and [27]. The following results are found in [5] and [10].

**Proposition 1.4.2:** (a) Let $U$ be the uniform measure on $[0,1]$ and let $\mathcal{U}_n$ be the empirical measure of a random sample of size $n$ from $U$. Then the process $X_n$ defined by $X_n(t) = \frac{1}{n} \sum_{i=1}^{[nt]} U_i$, $0 \leq t \leq 1$, converges weakly to $W^o$ in the Skorohod topology on $D[0,1]$, the space of all functions on $[0,1]$ which are continuous from the right and have left-hand limits. In particular this means by Proposition 1.2.4 that $\frac{1}{n} \sum_{i=1}^{[nt]} U_i \rightarrow W^o(t)$ converges in distribution to $\sup_{0 \leq t \leq 1} |W^o(t)|$.

(b) Let $U_m$ and $V_n$ be the empirical measures of $\{s_i/N\}$ and $\{r_i/N\}$, where $\{s_i\}$ and $\{r_i\}$ are the ranks among the pooled observations of two random samples from the same continuous distribution.
and $N = m+n$, $m$ and $n$ being the respective sample sizes. Then as $m$ and $n$ approach infinity, the process $Y_{m,n}$ defined by

$$Y_{m,n}(t) = (mn/N)^{\frac{1}{2}} (U_m([0,t]) - V_n([0,t]))$$

converges weakly to $W^0$, provided $0 < a \leq m/n \leq b < \infty$.

Proposition 1.4.3: Let $M^0 = \sup_{0 \leq t \leq 1} W^0(t)$ and $m^0 = \inf_{0 \leq t \leq 1} W^0(t)$.

(a) $\Pr(M^0 < b) = 1 - \exp(-2b^2)$, $b > 0$.

(b) $\Pr(\sup_{0 \leq t \leq 1} |W^0(t)| \leq b) = 1 + 2 \sum_{k=1}^{\infty} (-1)^k \exp(-2k^2b^2)$, $b > 0$.

(c) $\Pr(a < m^0 \leq M^0 \leq b) = \sum_{k=-\infty}^{\infty} \exp(-2k^2c^2) - \sum_{k=-\infty}^{\infty} \exp(-2(b+kc)^2)$,

where $c = b-a$, $b > 0$ and $a < 0$.

These results agree with earlier results proved by Kolmogorov and Smirnov. Billingsley proves (c) by arguments about a sequence of random walks which approximate the process $W^0$. Doob uses a more general result about the probabilities of various level crossings by the $W^0$ process, using the Inclusion-Exclusion principle to obtain the expression given for $\Pr(a < m^0 \leq M^0 \leq b)$.

1.5 Terminology from the Theory of Graphs. Let $X$ and $Y$ be finite sets. A finite simple graph, $\Gamma$, is a collection of arcs, or ordered pairs $(x,y)$, with initial vertex $x \in X$ and terminal vertex $y \in Y$. A matching is a subset of $\Gamma$ in which each element of $X$ (resp. $Y$) appears at most once as an initial (resp. terminal) vertex. A maximal matching is a matching of maximum possible cardinality; there is not in
general a unique maximal matching. The number of arcs in a maximal matching in the graph \((X, Y, \Gamma)\) will be denoted \(N(X, Y, \Gamma)\). The 
deficiency of the graph \((X, Y, \Gamma)\) is defined as \(D(X, Y, \Gamma) = \max_{F \subseteq X} (\#(F) - \#(\Gamma(F)))\), where \(\Gamma(F) = \{y \in Y: \text{there exist } x \in F \text{ such that } (x, y) \in \Gamma\}\), and for any set \(S\), \(\#(S)\) denotes the cardinality of \(S\).

Proposition 1.5.1: \(N(X, Y, \Gamma) = \#(X) - D(X, Y, \Gamma)\). This result is found in [23]. \(\square\)
CHAPTER 2

Computational Procedures

2.1 Computation of $d_P(P,Q)$ for general $S$ when $P$ and $Q$ have finite support.

By Proposition 1.3.8(a), we know that $d_P(P,Q) \leq c$ if and only if $V_P(P,Q;c) \leq c$. For purposes of testing statistical hypotheses, we need only to check whether a single inequality $d_P(P,Q) \leq c_\alpha$ holds, where $c_\alpha$ is the critical value of a test of size $\alpha$. If instead we wish to compute the actual value of $d_P(P,Q)$ to any desired accuracy, we first test whether $d_P(P,Q) \leq 1/2$, then, depending on the result, whether $d_P(P,Q) \leq 1/4$ or $d_P(P,Q) \leq 3/4$, etc., until finally it is determined that $d_P \in (i/2^k, (i+1)/2^k]$ for some value of $i$ between 0 and $2^k-1$, and $k$ chosen large enough to provide the desired accuracy. It is therefore our purpose to compute $V_P(P,Q;c)$ for any given value of $c$.

If $P$ has finite support, i.e., $P(X) = 1$, where $X = \{x_1, \ldots, x_n\}$, then it is clear that $V_P(P,Q;c) = \sup_{F \subseteq X} (P(F) - Q(F^c))$. This supremum is a maximum, so there exists a subset $G \subseteq X$ such that $P(G) - Q(G^c) \geq P(F) - Q(F^c)$ for all $F \subseteq S$. When $S = E^1$, we will be able to determine the set $G$ without placing further restrictions on the measure $Q$. For general $S$, we are able to solve the problem using Proposition 1.5.1, but only when $Q$ also has finite support.
Let $P_n$ put mass $1/n$ at each of the points $x_1, \ldots, x_n$ and let $Q_n$ put mass $1/n$ at each point of $Y = \{y_1, \ldots, y_n\}$. Define the graph $(X, Y, \Gamma_c)$ by letting the arc $(x_i, y_j) \in \Gamma_c$ if and only if $\rho(x_i, y_j) \leq c$, where $\rho$ is the metric on $S$. Then $V_P(P_n, Q_n; c)$ is just $1/n$ times the deficiency of $(X, Y, \Gamma_c)$, and hence by Proposition 1.5.1, $V_P(P_n, Q_n; c) = n^{-1}(n-N(X, Y, \Gamma_c))$. There does not seem to be a method for finding the number of arcs in a maximal matching in a graph except by constructing a maximal matching. There are published algorithms for finding a maximal matching in any graph. The standard method has been the "Hungarian algorithm" (see [2] or [18]). A new and more efficient algorithm has been proposed by Hopcroft and Karp ([16]).

$X$ and $Y$ need not have the same cardinality. Let $P_m$ put mass $1/m$ at each of the points $x_1, \ldots, x_m$ and let $Q_n$ put mass $1/n$ at each of the points $y_1, \ldots, y_n$. Then $V_P(P_m, Q_n; c) = \sup_{F \subseteq X} (m^{-1}\#(F) - n^{-1}\#(\Gamma_c(F)))$, where the graph $(X, Y, \Gamma_c)$ is defined as in the case $m = n$. Suppose $m$ and $n$ have greatest common factor $q$, so that $m = p_1q$ and $n = p_2q$, where $p_1$ and $p_2$ are relatively prime. Then

$$p_1p_2qV_P(P_m, Q_n; c) = \sup_{F \subseteq X} (p_2\#(F) - p_1\#(\Gamma_c(F))).$$

Define the graph $(X', Y', \Gamma'_c)$ as follows. $X'$ and $Y'$ each consist of $p_1p_2q$ vertices,

$$X' = \{x_{11}, x_{12}, \ldots, x_{1p_2}, x_{21}, x_{22}, \ldots, x_{2p_2}, \ldots, x_{m1}, \ldots, x_{mp_2}\}$$

and

$$Y' = \{y_{11}, y_{22}, \ldots, y_{1p_1}, y_{21}, y_{22}, \ldots, y_{2p_1}, \ldots, y_{n1}, \ldots, y_{np_1}\}.$$

Define $\Gamma'_c$ as follows. For each $i$ and $j$, $1 \leq i \leq m$ and $1 \leq j \leq n,$
let \((x_{ir}, y_{js}) \in \Gamma'_c\) for all values of \(r\) and \(s\), \(1 \leq r \leq p_2\) and \(1 \leq s \leq p_1\), if and only if \((x_i, y_j) \in \Gamma_c\). Intuitively, \(X'\) consists of \(p_2\) "copies" of each point of \(X\) and \(Y'\) consists of \(p_1\) copies of each point of \(Y\), and the "copies" are connected by an arc in \(\Gamma'_c\), whenever the " originals" are connected by an arc in \(\Gamma_c\).

**Proposition 2.1.1:** \[
\max_{G \subseteq X'} (\#(G) - \#(\Gamma'_c(G))) = \max_{F \subseteq X} (p_2\#(F) - p_1\#(\Gamma_c(F))).
\]

**Proof:** Suppose \(G_0 \subseteq X\) and \((G_0) - \#(\Gamma'_c(G_0)) = \max_{G \subseteq X'} (\#(G) - \#(\Gamma'_c(G)))\).

If \(G_0\) contains any "copy" of \(x_i\), then \(G_0\) may be assumed to contain all "copies" of \(x_i\), i.e., if \(\{x_{i1}, \ldots, x_{ip_2}\} \cap G_0 \neq \emptyset\), then \(\{x_{i1}, \ldots, x_{ip_2}\} \subseteq G_0\). To see this observe that for each \(r, 1 \leq r \leq p_2\), \(\Gamma'_c(\{x_{i1}, \ldots, x_{ip_2}\}) = \Gamma'_c(\{x_{ir}\})\) and hence, if \(x_{ir} \in G_0\),

\[
\#(G_0 \cup \{x_{i1}, \ldots, x_{ip_2}\}) - \#(\Gamma'_c(G_0 \cup \{x_{i1}, \ldots, x_{ip_2}\})) \geq \#(G_0) - \#(\Gamma'_c(G_0)).
\]

Let \(G\) be the class of all sets \(G \subseteq X\) such that for each \(i\), \(\{x_{i1}, \ldots, x_{ip_2}\} \cap G \neq \emptyset\) if and only if \(\{x_{i1}, \ldots, x_{ip_2}\} \subseteq G\). Then we have just shown that

\[
\max_{G \subseteq X'} (\#(G) - \#(\Gamma'_c(G))) = \max_{G \in G} (\#(G) - \#(\Gamma'_c(G))) = \max_{G \in G} (\#(G) - \#(\Gamma'_c(G)))
\]

For each set \(F^* \subseteq X\), let \(G^* = \bigcup_{\{i: x_i \in F^*\}} \{x_{i1}, \ldots, x_{ip_2}\}\), in this way defining a 1-1 correspondence between \(G\) and the class of all subsets of \(X\). That \(\#(G^*) - \#(\Gamma'_c(G^*)) = p_2\#(F^*) - p_1\#(\Gamma_c(F^*))\) follows from the definition of \(\Gamma'_c\). Hence

\[
\max_{G \in G} (\#(G) - \#(\Gamma'_c(G))) = \max_{F \subseteq X} (p_2\#(F) - p_1\#(\Gamma_c(F))).
\]
We can thus find the value of $p_1p_2qV_P(P_m,P_n,c)$ by applying the Hungarian or Hopcroft-Karp algorithm to $(X',Y',\Gamma')$. However, unless $p_1$ and $p_2$ are relatively small, the sets $X'$ and $Y'$ will contain so many points that the procedure will be excessively time-consuming, even though the special form of $(X',Y',\Gamma'_C)$ allows some steps to be saved in applying the algorithms.

The procedure we have described also extends to the case where $P$ (resp. $Q$) does not give equal probability to the points in $X$ (resp. $Y$), but the probability of each point is a rational number. In this case there exist positive integers $p_1,\ldots,p_m$ and $q_1,\ldots,q_n$ such that $V_P(P_m,P_n,c)$ is equal to a constant times the expression

$$\max_{F \subseteq X} \left| \sum_{i=1}^{m} p_i x_F(x_i) - \sum_{j=1}^{n} q_j x_F(y_j) \right|$$

where $x_F$ denotes the indicator function of the set $F$.

In this case we will let $X'$ consist of $p_i$ "copies" of $x_i$, $i = 1,\ldots,m$, and let $Y'$ consist of $q_j$ "copies" of $y_j$, $j = 1,\ldots,n$, $\Gamma'_C$ being defined as before. With the obvious modifications of the proof of 2.1.1, we can show that for this graph $(X',Y',\Gamma'_C)$,

$$\max_{G \subseteq X'} (\#(G) - \#(\Gamma'_C(G))) = \max_{F \subseteq X} \left( \sum_{i=1}^{m} p_i x_F(x_i) - \sum_{j=1}^{n} q_j x_F(y_j) \right).$$

In this case, depending on the values of $p_1,\ldots,p_m$ and $q_1,\ldots,q_n$, it may be impractical to find a maximal matching in $(X',Y',\Gamma'_C)$ even for small values of $m$ and $n$. In the next section we will treat the special case $S = E^1$ and take advantage of the total ordering on $E^1$ to develop a more efficient and more generally applicable algorithm.
2.2 Computation of $d_P(P,Q)$ when $P$ has finite support.

Let the support of $P$ be the set $X = \{x_1, \ldots, x_n\}$ where $x_1 \leq x_2 \leq \ldots \leq x_n$. Let $Q$ be an arbitrary measure on $E^1$. As described before, we wish to find a set $G \subseteq X$ such that $P(G) - Q(G^c) \geq P(F) - Q(F^c)$ for all $F \subseteq X$. We shall use an equivalent characterization of such a set $G$:

**Definition 2.2.1:** Let $P, Q, c$ be given. Let $G \subseteq X$ and $F \subseteq X$. If $P(G-F) - P(F-G) \geq Q(G^c) - Q(F^c)$, we shall say that $G$ is as bad as $F$, and if the strict inequality holds we shall say $G$ is worse than $F$. If $G$ is as bad as every subset of $X$, we shall call $G$ a worst set.

(Note, if the support of $P$ is infinite, the worst set may not be closed: let $S = E^1$, $c = 1/4$, let $P$ put mass $1/4$ at the points $-1/4$ and $1/4$ and be equal to Lebesgue measure on the set $(-1/4, 1/4)$. Let $Q$ put mass $1/2$ at the points $-1/2$ and $1/2$.) A worst set must maximize $P(G) - Q(G^c)$.

Note that if $F \subseteq G \subseteq X$, then $G$ is as bad as $F$ (resp. worse than $F$) if and only if $Q(G^c) - Q(F^c) \leq n^{-1}(G-F)$ (resp. $n^{-1}(G-F)$). This will be the criterion of relative badness in most of what follows.

Our algorithm for finding a worst set will examine the points $x_1, \ldots, x_n$ in order to see whether they belong in a worst set. The basic idea is the following.

**Proposition 2.2.2:** Let $i$ range from 1 to $n$. Let $i'$ be the least value of $i$ such that either
(a) $Q(\{x_1, \ldots, x_i\}^c) < i/n$, or
(b) $Q(\{x_1, \ldots, x_i\}^c) - \{x_{i+1}\}^c) \geq i/n$.

(If $i = n$, replace $\{x_{i+1}\}^c$ by $\emptyset$.) (Note that either (a) or (b) must hold for $i = n$.) If (a) holds for $i = i'$, then $x_1$ must be included in every worst set; if (b) holds for $i = i'$, then if $G$ is a worst set containing $x_1$, there is a subset of $G$ which does not contain $x_1$ and which is also a worst set.

Proof: Assume (a) holds. Let $H \subseteq X$. Let $I = \{x_1, \ldots, x_i\}$. If $I \cap H = \emptyset$, then (a) implies that $I \cap H$ is worse than $H$. Suppose $I \cap H \neq \emptyset$, but $x_1 \notin H$. Let $x_j$ be the smallest element of $I \cap H$. Since by definition of $i'$, $Q(\{x_1, \ldots, x_j\}^c) - \{x_j\}^c) < (j-1)/n$, $H \cup \{x_1, \ldots, x_{j-1}\}$ is worse than $H$. Thus $x_1 \notin H$ implies that $H$ is not a worst set.

Assume (b) holds. Let $H \subseteq X$. Suppose $x_1 \in H$. If $I \subseteq H$, then clearly $H - I$ is as bad as $H$. If $I \nsubseteq H$, then let $j'$ be the greatest value of $j$, $1 \leq j \leq i'$, such that $x_j \in H$ and

$$\{x_1, \ldots, x_j\}^c \cap [H - \{x_1, \ldots, x_j\}^c] = \emptyset.$$ If no such element exists, let $j' = i'$. Let $J = H - \{x_1, \ldots, x_i\}$.

If $j' = i'$, then $Q(\{x_1, \ldots, x_{i'}\}^c) - J^c \supseteq Q(\{x_1, \ldots, x_{i'}\}^c) - \{x_{i'+1}\}^c)$ $\geq j'/n$, by definition of $i'$, since (b) holds. (This could fail unless $x_1 \ldots x_{i'}$ because we might not have $Q(\{x_1, \ldots, x_i\}^c) \supseteq Q(\{x_1, \ldots, x_j\}^c) - \{x_{i'+1}\}^c).) It follows that $J$ is as bad as $H$, since

$$\# J \geq \# H - j', \text{ while } Q(H^c) - j'/n \geq Q(J^c) - Q(\{x_1, \ldots, x_j\}^c) - J^c) \supseteq Q(J^c), \text{ so } Q(J^c) - Q(H^c) \leq n^{-1}(\# J - \# H).$$
If \( j' < i' \), then \( Q(\{x_1, \ldots, x_j^c\} - J^c) = Q(\{x_1, \ldots, x_j^c\}) \geq j'/n \), since (a) cannot hold for \( j' < i' \), by definition of \( i' \). Again \( J \) is as bad as \( H \). □

We next wish to define a sequence of sets \( G_0 \subseteq G_1 \subseteq \ldots \subseteq G_n \) such that \( G_n \) will be a worst set. Let \( G_0 = \emptyset \). Let \( G_1 = G_0 \) if (b) in 2.2.2 holds for \( i = i' \), and let \( G_1 = \{x_1\} \) if (a) in 2.2.2 holds for \( i = i' \). In general we will have \( G_k \subseteq \{x_1, \ldots, x_k\} \).

Assume inductively that \( G_k \) is contained in any worst set and that if \( G \) is a worst set, then \( G - (\{x_1, \ldots, x_k\} - G_k) \) is also a worst set. I.e., \( G_k \) satisfies the conditions which 2.2.2 shows are satisfied by \( \{x_1\} \). Hence in seeking a worst set, we need only consider points of \( G_k \cup \{x_{k+1}, \ldots, x_n\} \).

Let \( i \) range from 1 to \( n-k+1 \). Let \( i' \) be the least such value of \( i \) such that either

(a') \( Q(\{x_{k+1}, \ldots, x_{k+i}\}^c - G_k^c) < i/n \), or

(b') \( Q(\{x_{k+1}, \ldots, x_{k+i}\}^c - G_k^c \cup \{x_{i+1}\}^c) \geq i/n \).

Then if (a') holds for \( i = i' \), we let \( x_{k+1} \in G_{k+1} \); if (b') holds for \( i = i' \), then \( x_{k+1} \notin G_{k+1} \). ((a) and (b) are really special cases of (a') and (b') since \( G_0 = \emptyset \).)

By the same argument as in 2.2.2 (restricting consideration to sets \( H \) such that \( \{x_1, \ldots, x_k\} \cap H = G_k \)) we can show that \( G_{k+1} \) is contained in any worst set and that there always exists a worst set which contains no element of \( \{x_1, \ldots, x_{k+1}\} - G_{k+1} \). By induction, \( G_n \) will necessarily be a worst set.
Having defined $G_k$, our algorithm will then be to check for each point $x_{k+1}$ whether (a') or (b') holds for $i = i'$, and to construct the set $G_{k+1}$ accordingly. The algorithm can be improved by noting that if (a') holds so that $x_{k+1} \in G_{k+1}$, then we will have $Q([x_{k+2}, \ldots, x_{k+i'-1}]^{c}) - (i'-1)/n \geq (i'-1)/n$ (provided $i' > 1$). Indeed, $i' > 1$ implies that $Q([x_{k+1}]^{c}) - G_k^{c} \geq 1/n$. Hence, $x_{k+1} \in G_n$ implies that $x_{k+2}, \ldots, x_{k+i'-1}$ are all in $G_n$, and they need not be tested individually. This improvement could be significant if the evaluation of the measure $Q$ is time-consuming.

The algorithm described in this section is sufficiently efficient to permit the use of Monte Carlo methods to estimate the distribution function of test statistics based on $d_P$ when $S = E^1$. For this purpose, the case of a measure $P$ which puts mass $1/n$ at each of $n$ points is sufficiently general. If instead, $P$ places rational-valued probabilities at each of $n$ points, we can compute $d_P(P, Q)$ by in effect replacing $P$ by a measure which places equal mass on an appropriate number of "copies" of each point of the support -- here "copies" are regarded as distinct observations taking the same value. Unlike the case of general $S$, for $S = E^1$ this does not require a significant increase in the amount of computation since all the points in each set of "copies" can be tested at the same time; if any is included in the worst set then all must be. (By contrast, in the general case our procedure was to find a maximal matching -- in which only some of a set of copies might appear as initial vertices -- and not to find a worst set.)
2.3 Computation of $d_L$

Let $S = E^1$, let $P, Q \in \mathcal{P}(S)$, and let the support of $P$ be the set $X = \{x_1, \ldots, x_n\}$ where $x_1 \leq x_2 \leq \ldots \leq x_n$. In this case the computation of $d_L(P, Q)$ is simple. By Proposition 1.3.10, $d_L(P, Q) \leq c$ if and only if $V_L(P, Q; c) \leq c$. Therefore, as with $d_P$, to be able to find out for each $c \in [0, 1]$ whether $V_L(P, Q; c) \leq c$, permits computation of $d_L(P, Q)$ to any desired accuracy.

$V_L(P, Q; c)$ is defined as

$$\max \left\{ \sup_{x \in E^1} (Q(-\infty, x-c) - P(-\infty, x)), \sup_{x \in E^1} (P(-\infty, x) - Q(-\infty, x+c)) \right\}.$$ 

As a function of $x$, $Q(-\infty, x-c) - P(-\infty, x)$ is non-decreasing on the interval $(-\infty, x_1)$ and on each of the intervals $[x_i, x_{i+1})$, $1 \leq i \leq n-1$. $P(-\infty, x) - Q(-\infty, x+c)$ is non-increasing on each of these intervals. Hence

$$\sup_{-\infty < x < \infty} (Q(-\infty, x-c) - P(-\infty, x)) = \max_{1 \leq i \leq n} (Q(-\infty, x_i-c) - P(-\infty, x_i)),$$

and

$$\sup_{-\infty < x < \infty} (P(-\infty, x) - Q(-\infty, x+c)) = \max_{1 \leq i \leq n} (P(-\infty, x_i) - Q(-\infty, x_i+c)).$$

Thus $V_L(P, Q; c)$ can be computed by evaluating the $P$-measure of $2n$ sets and the $Q$-measure of $2n$ sets.

If $P$ is not a discrete measure, it may be possible in some special cases of $P$ and $Q$ to determine points $x$ at which

$$\sup_{-\infty < x < \infty} (P(-\infty, x) - Q(-\infty, x+c)) \quad \text{and} \quad \sup_{-\infty < x < \infty} (Q(-\infty, x-c) - P(-\infty, x))$$

occur, provided that a simple functional form for the distribution functions associated with $P$ and $Q$ can be found. Specific examples can readily be suggested, for example, the case where the two density functions are unimodal. However, a useful general characterization of
cases in which the maximizing values can be readily determined is not apparent.

If \( S = E^k \), we can generalize the method described in this section, although the number of points to be examined increases as \( n^k \), making the method impractical even for moderate values of \( k \). Write the coordinates of \( x_i \) as \( x_i^{(1)}, \ldots, x_i^{(k)} \), for each \( i = 1, \ldots, n \). Let \( y_1^{(j)} \leq y_2^{(j)} \leq \ldots \leq y_n^{(j)} \) denote the ordered values of the set \( \{x_1^{(j)}, \ldots, x_n^{(j)}\} \). Let \( Z \subseteq E^k \) be the set of all points of the form

\[
y = (y_1^{(1)}, y_2^{(2)}, \ldots, y_k^{(k)}) ,
\]

where \( i_1, i_2, \ldots, i_k \) are integers between 1 and \( n \).

\( P(-\infty, x] - Q(-\infty, x+c] \) is non-increasing in each coordinate of \( x \) in any rectangle of the form

\[
[y_1^{(1)}, y_1^{(1)} + 1) \times [y_2^{(2)}, y_2^{(2)} + 1) \times \ldots \times [y_k^{(k)}, y_k^{(k)} + 1) ,
\]

where \( i_1, \ldots, i_k \) are integers between 0 and \( n \), \( y_0 \) being given the meaning of replacing the left-hand endpoint by \( -\infty \). Hence

\[
\sup_{x \in E^k} (P(-\infty, x] - Q(-\infty, x+c]) = \max_{y \in Z} (P(-\infty, y] - Q(-\infty, y+c]) .
\]

Likewise, \( Q(-\infty, x-c] - P(-\infty, x] \) is a non-decreasing function of each coordinate of \( x \) in each of the rectangles described and hence

\[
\sup_{x \in E^k} (Q(-\infty, x-c] - P(-\infty, x]) = \max_{y \in Z} (Q(-\infty, y-c] - P(-\infty, y)) .
\]

As before, the problem is reduced to evaluating the \( P \)-measure and \( Q \)-measure of a finite number of sets. \( #(Z) \) can be as large as \( n^k \).

If \( Q \) has finite support, the problem can be treated by the
methods of Section 2.1, dealing separately with the two quantities

(a) \( \inf\{\varepsilon > 0 : P(-\infty, x] - Q(-\infty, x+c] \leq \varepsilon; \text{ for all } x \in E^k\} \)

and

(b) \( \inf\{\varepsilon > 0 : Q(-\infty, x-c] - P(-\infty, x] \leq \varepsilon; \text{ for all } x \in E^k\} \).

Let \( X = \{x_1, \ldots, x_n\} \) be the support of \( P \) and let \( Y = \{y_1, \ldots, y_n\} \) be the support of \( Q \). We then define a graph \((X,Y,\Gamma_c)\) by \((x_i,y_j) \in \Gamma_c\) if and only if \( y_j \leq x_i + c \), so that \( nD(X,Y,\Gamma_c) = \max_{F \subseteq X} (P(F) - Q(\Gamma_c(F))) \)

\[ = \max_{x \in Z} (P(-\infty, x] - Q(-\infty, x+c]). \]

This deals with (a). For (b), let the arc \((y_j,x_i) \in \Gamma_c\) if and only if \( y_j + c < x_i \), and again \( nD(Y,X,\Gamma_c) = \max_{F \subseteq Y} (Q(F) - P(\Gamma_c(F))) \)

\[ = \max_{x \in Z} (Q(-\infty, x-c] - P(-\infty, x)). \]

In each case, the deficiency may be found by the Hungarian or the Hopcroft-Karp algorithm.

The extension to the cases of unequal numbers of atoms for \( P \) and \( Q \), or for atoms not all the same size but rational-valued, is done in the same way as in Section 2.1.
CHAPTER 3

Null Distribution for the Two-Sample Test Based on \( d_p \)

3.1 Introduction.

As described in Section 1.1, the null distribution of our two-sample test statistic based on \( d_p \) is, for equal sample sizes, the distribution of \( d_p(\mu_n, \nu_n) \) where \( \mu_n \) and \( \nu_n \) are the empirical measures of \( r_1/2n, \ldots, r_n/2n \) and \( s_1/2n, \ldots, s_n/2n \), \( r_1 < \ldots < r_n \) being a random sample without replacement from \( \{1,2,\ldots,2n\} \) and \( s_1 < \ldots < s_n \) being the elements in order of \( \{1,2,\ldots,2n\} - \{r_1,\ldots,r_n\} \). For each \( c \in [0,1] \), we define the graph \((X,Y,\Gamma_c)\) by letting \( X = \{r_1/2n, \ldots, r_n/2n\}, \ Y = \{s_1/2n, \ldots, s_n/2n\} \) and \( (r_i/2n, s_j/2n) \in \Gamma_c \) if and only if \( |r_i/2n - s_j/2n| \leq c \), i.e., if and only if \( |r_i - s_j| \leq i.p. \) \((2nc)\), where "i.p." denotes "integer part of." The following can be stated immediately based on the results of sections 1.3 and 1.5.

Proposition 3.1.1: For positive integers \( p,n,z \), let \( A(p,n,z) \) be the number of distinct samples \( \{r_1, \ldots, r_n\} \) and \( \{s_1, \ldots, s_n\} \) such that \( D(X,Y,\Gamma_z/2n) = p \). Then for \( z/2n \leq c < (z+1)/2n \),

\[
\Pr(\rho(\mu_n, \nu_n) \leq c) = \Pr(D(X,Y,\Gamma_z/2n) \leq z/2) = \binom{2n}{n}^{-1} \sum_{0 \leq p \leq z/2} A(p,n,z).
\]

Proposition 3.1.1 can be generalized to the case of samples which are not real-valued. However, for \( S = \mathbb{E}^1 \), the structure of \( \Gamma_c \) is sufficiently simple to permit the evaluation of \( A(p,n,z) \); the general
case seems unmanageable, as does the case of unequal sample sizes.

In the remainder of this chapter, we will discuss the features of \( \Gamma_c \) which permit this evaluation to be performed using the simple relationship expressed in Proposition 3.1.2 below.

Let \( X \) be a countable set such that each element \( x \) of \( X \) has a unique classification \( \kappa(x) \) in a "classification set" \( C \). We wish to determine for each \( u \in C \) the number \( K(u) \) of elements of \( X \) which are classified as \( u \), i.e., \( K(u) \) is defined as \( \#(\kappa^{-1}(u)) \). To this end we predicate the existence of a mapping \( a: X \to X \) having the property that for each \( u, v \in C \), \( x, y \in \kappa^{-1}(u) \) implies that \( \#(a^{-1}(x) \cap \kappa^{-1}(v)) = \#(a^{-1}(y) \cap \kappa^{-1}(v)) \). Thus for each \( x \in \kappa^{-1}(u) \) the number of elements \( z \in X \) for which \( a(z) = x \) and \( \kappa(z) = v \) depends only on \( u \) and \( v \), not on the choice of \( x \). Call this number \( T(u,v) \). (If \( K(u) = 0 \), let \( T(u,v) = 0 \).

**Proposition 3.1.2:** \( K(v) = \sum_{u \in C} K(u)T(u,v) \), for each \( v \in C \).

**Proof:** Observe that \( K(u)T(u,v) \) is the number of elements \( w \in X \) such that \( \kappa(w) = v \) and \( \kappa(a(w)) = u \). But \( \kappa^{-1}(v) = \sum_{u \in C} \{ w : \kappa(w) = v \text{ and } \kappa(a(w)) = u \} \).

For a fixed \( z \), our problem of computing \( A(p,n,z) \) for various values of \( n \) and \( p \), can now be expressed in terms of Proposition 3.1.2. \( X \) will be the set \( \{ (X,Y) : (X,Y) \text{ is a partition of } \{1,2,\ldots,2n\} \}_{n=0}^{\infty} \) into subsets \( X \) and \( Y \) with \( \#(X) = \#(Y) = n \). The case \( n = 0 \) we take to represent a partition of the empty set, which we will identify with the empty set. Note that the sets \( X \) and \( Y \) are now sets of
integers, not multiples of $1/2n$.

Each element $u$ of the classification set will consist of a 5-tuple of integers $u = (k, n, d, s, t) = (k(u), n(u), d(u), s(u), t(u))$, describing various properties which hold for each partition $(X, Y)$ classified as $u$. In particular $n$ has the meaning of the previous paragraph. $d(u)$ is the deficiency of the graph $(X, Y, \Gamma_z)$. $\Gamma_z$ is similar to $\Gamma_{z/2n}$ and will be defined below, as will the parameters $s$ and $t$.

The mapping $a: X \rightarrow X$ will be defined in such a way that $T(u, v) = 0$ whenever $n(u) \geq n(v)$, except when $u = \kappa(\phi)$, in which case $T(u, u) = 1$, i.e., we define $a(\phi) = \phi$. Only for the classification $u = \kappa(\phi)$ is $n(u) = 0$; we define $K(\kappa(\phi)) = 1$. Hence, if $T(u, v)$ can be computed for all $u$ and $v$ in $C$, then $K(u)$ can be found for all $u \in C$.

We then have

$$A(p, n, z) = \sum_{k} \sum_{s} \sum_{t} K(k, n, p, s, t).$$

In the next section we define a classification set and a mapping $\alpha$ in such a way that $T(u, v)$ can be computed for all $u$ and $v$.

### 3.2 Matchings in the graphs $(X, Y, \Gamma_z)$.

Our ultimate interest is in partitions $(X, Y)$ of $\{1, 2, \ldots, 2n\}$ into sets $X$ and $Y$ of equal size. For use in later proofs we wish to describe properties which apply more generally to partitions of $\{1, 2, \ldots, n+n'\}$ into sets $X$ and $Y$ with $\#(X) = n$ and $\#(Y) = n'$.

The positive integer $z$ will be considered to be fixed and we will hereafter use the convention that for any partition $(X, Y)$ under consideration, $\Gamma_z$ will denote the graph such that $(x, y) \in \Gamma_z$ whenever $|x-y| \leq z$ for $x \in X$ and $y \in Y$. We write $\Gamma_z(X, Y)$ if the partition
is not clear from the context.

We shall deal with $D(X, Y, \Gamma_Z)$ by looking at maximal matchings in $(X, Y, \Gamma_Z)$, using Proposition 1.5.1. Since we are really interested only in $D(X, Y, \Gamma_Z)$, rather than the matchings, it will be convenient to consider a single "canonical" maximal matching, uniquely defined for each partition $(X, Y)$, and having certain special properties, which we will describe in Proposition 3.2.14 following some discussion.

**Definition 3.2.1:** Let $(X, Y)$ partition $\{1, 2, \ldots, n+n'\}$ so that

$\#(X) = n$, $\#(Y) = n'$, and let $M = N(X, Y, \Gamma_Z)$. Let $m \in \Gamma_Z$ be a maximal matching and write $m$ as the set of arcs $\{(x_1(m), y_1(m)), \ldots, (x_M(m), y_M(m))\}$, such that $x_1(m) < x_2(m) < \ldots < x_M(m)$. (The elements $y_1(m), \ldots, y_M(m)$ are not necessarily in order.) Define $X^+(m) = \{x_1(m), \ldots, x_M(m)\}$, $Y^+(m) = \{y_1(m), \ldots, y_M(m)\}$, $X^-(m) = X - X^+(m)$, and $Y^-(m) = Y - Y^+(m)$.

We will be chiefly interested in the sets $X^+$ and $Y^+$ and not in the particular matching which produced them. In fact, we will always be able to assume that the first element of $X^+$ is connected to the first element of $Y^+$, second to second, etc.

**Proposition 3.2.2:** Let $m$ be a maximal matching in $(X, Y, \Gamma_Z)$. Then there exists a maximal matching $m^* \subseteq \Gamma_Z$ such that $X^+(m) = X^+(m^*)$, $Y^+(m) = Y^+(m^*)$, and $y_1(m^*) < y_2(m^*) < \ldots < y_M(m^*)$.

**Proof:** Follows from Lemma 3.2.3.

**Lemma 3.2.3:** Let $y_i, y_j \in Y$ and $x_i, x_j \in X$ be such that $(x_i, y_i) \in \Gamma_Z$ and $(x_j, y_j) \in \Gamma_Z$. If $x_i \leq x_j$ and $y_j \leq y_i$, then
\((x_i, y_j) \in \Gamma_z\) and \((x_j, y_i) \in \Gamma_z\). (The roles of \(X\) and \(Y\) can be interchanged.)

**Proof:** It follows from the inequalities \(|x_i - y_i| \leq z\) and \(|x_j - y_j| \leq z\) and the assumption \(x_i \leq x_j, y_j \leq y_i\), that \(|x_i - y_j| \leq z\) and \(|x_j - y_i| \leq z\).

**Definition 3.2.4:** Let \(M(X, Y, \Gamma_z)\) be the collection of all maximal matchings \(m \leq \Gamma_z\) such that \(y_1(m) < \ldots < y_M(m)\).

A matching in \(M\) is uniquely determined by its sets \(X^+(m)\) and \(Y^+(m)\) and vice versa. However, for a given partition \((X, Y)\), there may be a number of matchings in \(M(X, Y, \Gamma_z)\).

We will introduce a schematic notation to be used in describing the partitions and matchings under consideration and use it to give some simple examples showing considerations involved in choosing a "canonical" matching. \(\text{xyxy}\) will denote a partition of \(\{1,2,3\}\) into the sets \(X = \{2\}\) and \(Y = \{1,3\}\). Here \(n = 1\) and \(n' = 2\). For \(z \geq 1\), we have two matchings, denoted by \(\dddot{xyxy}\) and \(\dddot{yyxx}\), by which we mean respectively \(\{(2,1)\}\) and \(\{(2,3)\}\). In using our notation we always assume that the elements of \(X^+\) and \(Y^+\) are connected in order, i.e., the matching is in \(M\). Of the two matchings given, \(\dddot{xyxy}\) will be chosen as the "canonical" one, because it is "earliest" (the '+'s occur earlier), in a sense to be made precise. From among \(\dddot{yyxx}, \dddot{yyxy}, \text{ and } \dddot{xyxy}\), the matching \(\dddot{yyxy}\) is the "earliest." However, in \(\dddot{yyxx}\) and \(\dddot{yyxy}\) the matched \(x\) and \(y\) are as close together as possible, while in \(\dddot{yyxy}\), there is an intervening element of \(Y^-\). \(\dddot{yyxy}\) will be said not to be "compressed," while the other two matchings will be said to be "compressed," a notion to be made precise below. In this case, the
"canonical" matching will be chosen to be the "earlier" of the two "compressed" matchings, namely $y^xxy$. It will later be seen that "compressed" matchings assign +'s and -'s in a useful pattern. The choice of the "earliest" of the "compressed" matchings is made for definiteness.

Corresponding to the notion of earliest matching, we now introduce a total ordering on $M(X, Y, \Gamma_z)$ which we shall call the lexicographic ordering. In this ordering we say that $m < m'$ if the vector $(x_1(m), \ldots, x_M(m), y_1(m), \ldots, y_M(m))$ is lexicographically smaller than $(x_1(m'), \ldots, x_M(m'), y_1(m'), \ldots, y_M(m'))$, i.e., if $x_1(m) < x_1(m')$, or $x_1(m) = x_1(m')$ but $x_2(m) < x_2(m')$, or $x_1(m) = x_1(m')$ and $x_2(m) = x_2(m')$ but $x_3(m) < x_3(m')$, etc.

**Proposition 3.2.5**: Let $m$ be the least element of $M(X, Y, \Gamma_z)$ with respect to the ordering just defined and let $m'$ be any element of $M$. Then $x_1(m) \leq x_1(m'), \ldots, x_M(m) \leq x_M(m')$, $y_1(m) \leq y_1(m'), \ldots, y_M(m) \leq y_M(m')$. (Note that the definition of lexicographic minimality only implies that $x_1(m) \leq x_1(m')$. If $m \not\in M$ the result may not hold, for example if $m$ is $y^xxy$, $m'$ could be $\{(3,5), (4,1)\}$ provided $z \geq 3$.)

**Proof**: Necessarily $x_1(m) \leq x_1(m')$. Further, $y_1(m) \leq y_1(m')$. Indeed, if $y_1(m) > y_1(m')$, then by Lemma 3.2.3, we can in $Y^+(m)$ replace $y_1(m)$ by $y_1(m')$ (since $y_1(m') \in Y^-(m)$) and get a matching in $M$ which is lexicographically smaller than $m$.

Let $j$ be the least index, $2 \leq j \leq M$, such that $x_j(m) > x_j(m')$, or $y_j(m) > y_j(m')$. We wish to arrive at a contradiction by assuming the existence of such a $j$. Suppose first that $x_j(m) > x_j(m')$. Since $j$
is the first index of its kind, \( x_{j-1}(m) \leq x_{j-1}(m') < x_j(m') < x_j(m) \).

Hence \( x_j(m') \in X^-(m) \). Similarly, \( y_{j-1}(m') \leq y_{j-1}(m) < y_j(m) \).

We have two cases:

a) \( y_j(m) > y_j(m') \);

b) \( y_j(m) \leq y_j(m') \).

In case (b), we can apply Lemma 3.2.3 to show that \( (x_j(m'), y_j(m)) \in \Gamma_z \), because \( x_j(m) > x_j(m') \) and \( y_j(m) \leq y_j(m') \). Hence we can in \( X^+(m) \) replace \( x_j(m) \) by \( x_j(m') \), which is in \( X^-(m) \), to get a lexicographically smaller matching in \( M \), contradicting the definition of \( m \). In case (a), we contradict the maximality of \( m \), since in this case we must have \( y_j(m') \in Y^-(m) \), as was shown above for \( x_j(m') \). Since either (a) or (b) leads to a contradiction, we must have \( x_j(m) \leq x_j(m') \).

So assume that \( x_j(m) \leq x_j(m') \), but \( y_j(m) > y_j(m') \). Then as before \( y_j(m') \in Y^-(m) \). By Lemma 3.2.3 we may replace the element \( y_j(m) \) in \( Y^+(m) \) by \( y_j(m') \), obtaining an earlier matching than \( m \), contradicting the definition of \( m \).

\[ \square \]

We now define the class of compressed matchings.

**Definition 3.2.6:** Let \( M'(X, Y, \Gamma_z) \) be the class of all \( m \in (X, Y, \Gamma_z) \) such that \( X^+(m) + Y^+(m) = \bigcup_{i=1}^{\infty} [\min(x_i(m), y_i(m)), \max(x_i(m), y_i(m))] \), where \([i, j]\) denotes the integers between \( i \) and \( j \) inclusive. (I.e., none of the intervals \( [\min(x_i(m), y_i(m)), \max(x_i(m), y_i(m))] \) contains any point of \( X^-(m) + Y^-(m) \).) \[ \square \]

The lexicographic ordering on \( M \) induces an ordering on \( M' \) as a subset of \( M \). We wish to show that the least element \( m \) of \( M' \) in this
ordering satisfies the stronger conditions of Proposition 3.2.5 when compared to any other element of $M'$. To do this we first give an algorithm which associates with any element $m \in M$ a unique element $m^c$ of $M'$.

**Definition 3.2.7:** Let $m \in M(X, Y, \Gamma_Z)$. We define the element $m^c \in M'(X, Y, \Gamma_Z)$ by the following algorithm.

Let $m^{(0)} = m$. Suppose $m^{(0)}, \ldots, m^{(i)}$ have been defined. If $m^{(i)} \in M'$, let $m^c = m^{(i)}$; we have completed the algorithm. If $m^{(i)} \notin M'$, we define $m^{(i+1)}$ as follows. Let $j$ be the least index, $1 \leq j \leq M$, such that $[\min(x_j(m^{(i)}), y_j(m^{(i)})), \max(x_j(m^{(i)}), y_j(m^{(i)}))] \notin X^+(m^{(i)}) + Y^+(m^{(i)})$. We have one of four mutually exclusive cases:

a) there exists an $x \in X^-(m^{(i)})$ such that $x_j(m^{(i)}) < x < y_j(m^{(i)})$

b) there exists an $x \in X^-(m^{(i)})$ such that $y_j(m^{(i)}) < x < x_j(m^{(i)})$

c) there exists an $y \in Y^-(m^{(i)})$ such that $x_j(m^{(i)}) < y < y_j(m^{(i)})$

d) there exists an $y \in Y^-(m^{(i)})$ such that $y_j(m^{(i)}) < y < x_j(m^{(i)})$.

The cases are mutually exclusive because if, for example, (a) and (c) both hold, then $(x, y) \in \Gamma_Z$, contradicting the maximality of $m$.

$x$ (or $y$) need not be unique. Let $x^*$ (resp. $y^*$) be the largest such $x$ (resp. $y$) in case (b) (resp. (c)) and the smallest such $x$ (resp. $y$) in case (a) (resp. (d)). Define $m^{(i+1)}$ in cases (a) and (b) (resp. (c) and (d)) by replacing, in $X^-(m^{(i)})$ (resp. $Y^-(m^{(i)})$), $x_j(m^{(i)})$ by $x^*$ (resp. $y_j(m^{(i)})$ by $y^*$).

We must show that the procedure described in 3.2.7 eventually produces an element of $M'$. To do this, define the function $g$ on $M$...
by
\[ g(m) = \sum_{i=1}^{M} |x_i(m) - y_i(m)|. \]

We claim that (unless \( m^{(i)} \in M' \))
\[ g(m^{(i)}) > g(m^{(i+1)}); \]

it will follow that since the function \( g \) is bounded below by 0, the algorithm will continue until \( g \) can no longer be decreased.

Our claim is an immediate consequence of the following result.

**Lemma 3.2.8:** (i) In case (a) above, let \( \ell \geq 0 \) be the integer such that
\[ x_1^{(m^{(i)})} < \ldots < x_j^{(m^{(i)})} < \ldots < x_{j+\ell}^{(m^{(i)})} < x < x_{j+\ell+1}^{(m^{(i)})} < \ldots < x_M^{(m^{(i)})}. \]

Then \( m^{(i+1)} \) consists of the pairs \((x_1^{(m^{(i)})}, y_1^{(m^{(i)})})\), \ldots,
\((x_{j-1}^{(m^{(i)})}, y_{j-1}^{(m^{(i)})})\) and if \( \ell > 0 \), \((x_{j+1}^{(m^{(i)})}, y_{j}^{(m^{(i)})})\), \ldots,
\((x_{j+\ell}^{(m^{(i)})}, y_{j+\ell-1}^{(m^{(i)})})\), and (for \( \ell \geq 0 \)), \((x^o, y_j^{(m^{(i)})})\) and
\((x_{j+\ell}^{(m^{(i)})}, y_{j+\ell+1}^{(m^{(i)})})\), \ldots, \((x_M^{(m^{(i)})}, y_M^{(m^{(i)})})\).

(ii) In case (b) above, \( m^{(i+1)} \) is identical to \( m^{(i)} \) except that the arc \((x_j^{(m^{(i)})}, y_j^{(m^{(i)})})\) is replaced by \((x^o, y_j^{(m^{(i)})})\).

The cases (c) and (d) interchange the role of \( X \) and \( Y \) but are otherwise identical.

**Proof:** (i) is obvious when one follows through Definition 3.2.7.

(ii) Because \( j \) is the least integer such that there is an element \( x \in X^-(m^{(i)}) \) such that \( y_j^{(m^{(i)})} < x < x_j^{(m^{(i)})} \), we cannot have \( x < x_{j-1}^{(m^{(i)})} < x_j^{(m^{(i)})} \), because then \( y_{j-1}^{(m^{(i)})} < x < y_j^{(m^{(i)})} \).

The result is then obvious. \( \square \)
Proposition 3.2.9: Let $m^o \in M'(X,Y,\Gamma_Z)$ be the lexicographically minimal element of $M'$. Let $m' \in M'$ be any other element. Then

$$x_1(m^o) \leq x_1(m'), \ldots, x_M(m^o) \leq x_M(m'), y_1(m^o) \leq y_1(m'), \ldots, y_M(m^o) \leq y_M(m').$$

Proof: The result will be an immediate corollary of the following.

Lemma 3.2.10: Let $m' \in M'$ and $m \in M$ such that $x_i(m) \leq x_i(m')$ and $y_i(m) \leq y_i(m')$, $i=1,\ldots,M$. Then $x_i(m^c) \leq x_i(m')$ and $y_i(m^c) \leq y_i(m')$, $i=1,\ldots,M$.

Proof: It suffices to prove the conclusion for $m^{(1)}$ instead of $m^c$, since $m^{(2)} = (m^{(1)})^{(1)}$, etc. We prove this result for $m^{(1)}$ in the cases (a) and (b) above, the cases (c) and (d) following by interchanging the roles of $X$ and $Y$.

Suppose (a) obtains. We distinguish two further cases:

A) $x_j(m') < y_j(m')$

B) $x_j(m') > y_j(m')$.

Suppose (A) holds. Let $\mathcal{L}$ be as in 3.2.8(a). Since $x^o$ is the smallest element of $X^-(m)$ between $x_i(m)$ and $y_j(m)$, there are $\mathcal{L}+1$ elements of $X$ in the half-open interval $(x_j(m), x^o]$, namely $x_{j+1}(m), \ldots, x_{j+\mathcal{L}}(m)$, and $x^o$. Since $x^o < y_j(m) \leq y_j(m')$, and because $m' \in M'(X,Y,\Gamma_Z)$, if $x_j(m') < x^o$, it follows that all elements of $X$ in the interval $(x_j(m'), x^o]$ must be elements of $X^+(m')$. But because $x_{p+\mathcal{L}+1}(m') \geq x_{p+\mathcal{L}+1}(m) > x^o$, we have at most $\mathcal{L}$ elements of $X^+(m')$, and hence of $X$ in the interval $(x_j(m'), x^o]$. Hence, if $x_j(m') < x^o$, then $x_j(m) < x_j(m')$. Of course, if $x_j(m') \geq x^o$, then $x_j(m) < x_j(m')$.
In either case, then, \( x_{j+1}(m) \leq x_j(m') \), \( \ldots, x_{j+\ell}(m) \leq x_{j+\ell-1}(m') \), and \( x^o \leq x_{j+\ell}(m') \). Therefore, in making the replacements of arcs described in 3.2.8(i), we have

\[
x_j(m(1)) \leq x_j(m'), \ldots, x_{j+\ell}(m(1)) \leq x_{j+\ell}(m'),
\]

and

\[
Y^+(m) = Y^+(m(1)).
\]

Suppose (B) holds. Since \( x^o < y_j(m) \leq y_j(m') < x_j(m') \), we again have

\[
x_j(m(1)) \leq x_j(m'), \ldots, x_{j+\ell}(m(1)) \leq x_{j+\ell}(m').
\]

Suppose that case (b) obtains. Only one arc, \((x_j, y_j)\), is changed in going from \( m \) to \( m(1) \). But \( x_j(m(1)) = x_j(m) \) and \( y_j(m(1)) = y_j(m) \), so \( x_j(m(1)) \leq x_j(m') \) and \( y_j(m(1)) = y_j(m') \).

Proposition 3.2.11: \( m^o = m^c \), where \( m \) is the minimal matching in \( M(X, Y, \Gamma_z) \).

Proof: Follows immediately from 3.2.10.

The matching \( m^o \) will be our canonical matching for the partition \((X, Y)\). For each \((X, Y)\), the sets \( X^+(m^o), Y^+(m^o), X^-(m^o), Y^-(m^o) \) are uniquely determined. The choice of \( m^o \) as canonical matching establishes a certain pattern in the arrangement of the elements of these sets. For example, let \( z = 2 \) and consider \( \ldots xxxx+y+y+y+y+y+y+xxy+y \). The elements of \( X^+ + Y^+ \) are arranged in two segments of consecutive elements, namely \( \{5,6,7,8\} \) and \( \{10,11,12,13\} \). In general, such segments must each contain as many elements of \( X^+ \) as of \( Y^+ \), since otherwise an element of \( X^- + Y^- \) would lie in one of the intervals \([\min(x_i(m^o), y_i(m^o)), \max(x_i(m^o), y_i(m^o))]\). The segments are separated by one or more elements of \( X^- \) or \( Y^- \), but clearly an element of \( X^- \) cannot be adjacent to an
element of $Y^-$. Other remarks can be made, but we wish to proceed more formally and collect the results in Proposition 3.2.14.

**Definition 3.2.12:** For each partition $(X,Y)$ of \{1,2,\ldots,n+n'\} define the function \( f = f_{(X,Y)} \) on \{0,1,\ldots,n+n'\} by \( f(0) = 0 \) and

\[ f(j) = #(X\cap\{1,2,\ldots,j\}) - #(Y\cap\{1,2,\ldots,j\}), \quad \text{for } j=1,2,\ldots,n+n'. \]

Let \( I \) be the interval of integers \([j_0+1, j_0+\ell+\ell']\) consisting of \( \ell+\ell' \) consecutive integers from the set \{1,2,\ldots,n+n'\}, \( \ell \) of \( X \) and \( \ell' \) of \( Y \), where \( 0 \leq j_0 \leq n+n' - 1, \ell \geq 0 \) and \( \ell' \geq 0 \). \( I \) will be said to be **non-switching** if either

\[ f(j) - f(j_0) \geq 0, \text{ for all } j \in [j_0+1, j_0+\ell+\ell'] \] or if

\[ f(j) - f(j_0) \leq 0, \text{ for all } j \in [j_0+1, j_0+\ell+\ell'] \]

The interval \( I \) will be said to be **strictly non-switching** if the strict inequality holds for all \( j \in [j_0+1, j_0+\ell+\ell'] \).

\[ \square \]

**Examples:**

<table>
<thead>
<tr>
<th>switching</th>
<th>xxyyx</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-switching (not strict)</td>
<td>xxyxy</td>
</tr>
<tr>
<td>non-switching (strict)</td>
<td>xxyyy</td>
</tr>
</tbody>
</table>

Note that if \([j_0+1, j_0+2\ell]\) is a non-switching interval containing \( \ell \) elements of \( X \) and \( \ell \) elements of \( Y \), then \( j_0+1 \in X \) if and only if \( j_0+2\ell \in Y \).

**Definition 3.2.13:** Let \((X,Y)\) partition \{1,2,\ldots,n+n'\} as before. The elements 1,2,\ldots,n+n' will be divided into the sets \( A_1 \leq B_1 \leq A_2 \leq \ldots \leq A_{r+1} \leq A_r \leq B_r \leq A_{r+1} \), for some value of \( r \), where each set \( A_i \leq X^- + Y^- \) and each set \( B_i \leq X^+ + Y^+ \). Let \( q_1 \) be the least element of \( X^+ + Y^- + (m^0) \). Let \( B_1 \) be the largest non-switching interval of
\( X^+(m^o) + Y^+(m^o) \) which has \( q_1 \) as its left-hand endpoint. Let \( q_2 \) be the first element (if one exists) of \( X^+(m^o) + Y^+(m^o) \) which is greater than the right-hand endpoint of \( B_1 \) and let \( B_2 \) be the largest non-switching interval of \( X^+(m^o) + Y^+(m^o) \) which has \( q_2 \) as its left-hand endpoint. Let \( q_i \) and \( B_i \) be defined in this manner until the union \( B_1 + B_2 + \ldots + B_r \) contains all the elements of \( X^+(m^o) + Y^+(m^o) \).

Let \( A_1 \) be defined as the interval \([1,q_1]\) and let \( A_i \), for \( i = 2, 3, \ldots, r \), be the interval of elements of \( X^- + Y^- \) which lie between the right-hand endpoint of \( B_{i-1} \) and the left-hand endpoint of \( B_i \), and let \( A_{r+1} \) be the interval of elements to the right of \( B_r \). Any or all of the sets \( A_1, A_2, \ldots, A_{r+1} \) may be empty.

Proposition 3.2.14:

(a) For \( i = 1, \ldots, r \), \( B_i \) consists of the same number of elements of \( X \) as of \( Y \).

(b) For \( i = 2, \ldots, r \), if \( A_i = \emptyset \), then \( q_i \in X \) if and only if \( q_{i-1} \in Y \).

(c) Either \( A_i \subseteq X^-(m^o) \) or \( A_i \subseteq Y^-(m^o) \) for \( i = 1, 2, \ldots, r+1 \).

(d) If \( A_i \neq \emptyset \), then \( q_i \in X \) if and only if \( A_i \subseteq X \), for \( i = 1, 2, \ldots, r \).

Proof: (a) and (c) are as noted above. (b) holds because otherwise \( B_{i-1} \) would not be the largest non-switching interval starting at \( q_{i-1} \). For (d) we need a preliminary Lemma.

Definition 3.2.15: Let the interval \( I = [j_0^+1, j_0^+2L] \) consist of \( L \) elements of \( X \) and \( L \) of \( Y \). \( I \) will be said to be completely matchable under \( \Gamma_z \) if the graph \( (I \cap X, I \cap Y, \Gamma_z) \) has zero deficiency.
Lemma 3.2.16: Let \( I \) be a strictly non-switching interval which is completely matchable under \( \Gamma_z \). Then \([j_0+2, j_0+2z-1]\) is completely matchable under \( \Gamma_z \).

Proof: Assume for definiteness that \( j_0+1 \in X \). Let \( f \) be as in 3.2.9. Let \( a_1 < a_2 < \ldots < a_{\ell} \) be the elements of \( L \cap X \) and \( b_1 < b_2 < \ldots < b_{\ell} \) be the elements of \( L \cap Y \). Then for \( i=2, \ldots, \ell \), \( a_i < b_{i-1} \); otherwise \( f(b_{i-1}) - f(j_0) \leq 0 \) and \( I \) would not be strictly non-switching. But for \( i=2, \ldots, \ell \), \( |a_i - b_{i-1}| \leq z \), since \( I \) is completely matchable under \( \Gamma_z \). Hence, since \( a_{i+1} < b_i < b_{i-1} < a_i \), \( (a_{i+1}, b_{i-1}) \in \Gamma_z \). The matching \( \{(a_{i+1}, b_{i-1}), 1 \leq i \leq \ell-1\} \) is a complete matching on \([j_0+2, j_0+2\ell-1]\). \( \square \)

Continuation of 3.2.14: Assume that (d) does not hold for some \( i \) and for definiteness that \( q_i \in X \) and \( A_i \subseteq Y \). Let \( J \) be the largest strictly non-switching interval starting at \( q_i \). Then \( J \subseteq B_i \).

We can form a matching \( m \) which is the same as \( m^o \) except that \( q_i \) is matched to the last element of \( A_i \), the last element of \( J \) is unmatched, and the remaining elements of \( J \) are matched to each other as in Lemma 3.2.16. Then \( m \in M'(X, Y, \Gamma_z) \), but \( m \) is lexicographically smaller than \( m^o \), contradicting the definition of \( m^o \). \( \square \)

(As an illustration of this last point, take \( r=2, i=1 \) and \( z=3 \), and consider \( yyyyyyyyy \). The matching shown is not \( m^o \) because \( yyyyyyyyy \) is a lexicographically smaller compressed matching.)
3.3 Definition of a mapping \( a: X \rightarrow X \).

Let \( X_n = \{(X,Y) : (X,Y) \) is a partition of \( \{1,2,\ldots,2n\} \) such that \( \#(X) = \#(Y) = n \). Then \( X = \bigcup_{n=0}^{\infty} X_n \), where \( X_0 = \{\phi\} \). Let \( x = (X,Y) \in X_n \) and let \( j = \#(XnB_r(X,Y,\Gamma_z)) = \#(YnB_r(X,Y,\Gamma_z)) \), where \( B_r(X,Y,\Gamma_z) \) is the set \( B_r \) defined in Definition 3.2.13. Then \( a(x) \) will be that element of \( X_{n-j} \) which is the same as \( (X,Y) \) except that \( B_r \) has been "removed" and the elements of \( A_{r+1} \) have been "pushed forward" and "relabeled" in the obvious manner. (In our schematic notation if \( z = 2 \) and \( x \) is \( \underline{\underline{\underline{\underline{\underline{\underline{\underline{x}}}}}}xxxxyyyyxyyxyy \) with \( B_r \) underlined, then \( a(x) \) will be the element \( xxxxyyyyyyxyy \).

In describing the relationship of \( (X,Y) \) to \( a(X,Y) \), it will be convenient to adopt some suggestive terminology. \( a(X,Y) \) will be said to have been obtained by removing \( B_r \) from \( (X,Y) \). Conversely, \( (X,Y) \) will be said to have been obtained by inserting \( B_r \) (or really a copy of \( B_r \)) into \( a(X,Y) \) after the integer \( q_r(X,Y,\Gamma_z) - 1 \). For purposes of proof we shall wish to consider in the same sense "removals" and "insertions" of arbitrary numbers of \( X \)-type and \( Y \)-type elements, not just completely matchable intervals.

We now return to Proposition 3.1.2. \( a^{-1}(X,Y) \) will consist precisely of every element \( (X',Y') \) of \( X \) which we can get by inserting an interval into \( (X,Y) \) such that the inserted interval coincides exactly with \( B_r(X',Y',\Gamma_z) \). Our interest will be in sets of elements of \( X \) of the form \( a^{-1}(X,Y) \circ \kappa^{-1}(v) \), in particular in how the deficiency of each element of \( a^{-1}(X,Y) \), which is given by one of the parameters of its classification \( v \), relates to the deficiency of elements of \( \kappa^{-1}(\kappa(X,Y)) \), all of which have the same deficiency.
There is one further complication. Consider the inverse procedure of going from \(a(X',Y')\) to \((X',Y')\) by insertion of an interval. Depending on the location of the insert, it may be the case that not just any interval can be inserted so as to coincide with \(B_r(X',Y')\). For example, we may insert \(xxyy\) into \((X,Y) = xxxxyyy\) after 6 to get (for \(z = 2\)) \((X',Y') = \overline{xxxxyyyyyy}\), and have \(a(X',Y') = (X,Y)\). But if (for \(z = 2\)) we insert \(xy\) after 6 to get \(xxxxyxxxxyy\), we then have \(a(X',Y') = xxxyyy\), which is not the original element \((X,Y)\), i.e., \(B_r(X',Y',\Gamma_z)\) does not coincide with the inserted interval.

In this example, \(xxyy\) is intuitively "long enough" to separate element 6 (in \(X\)) from element 7 (in \(Y\)) so that after the insertion these elements cannot be part of the same matched segment, while \(xy\) is not long enough. However, length of the inserted interval is not a sufficient condition for the interval to be inserted in such a location.

**Definition 3.3.1:** A partition \((X,Y)\) of \(\{1,2,\ldots,2\ell\}\) with \(#(X) = #(Y)\) will be said to be **separating** if

a) \((X,Y)\) is completely matchable under \(\Gamma_z\).

b) The partition \((X',Y')\) obtained by inserting \((X,Y)\) after 1 into \(xy\) (assuming for definiteness that 1 \(\in X\); if 1 \(\in Y\), then insert into \(yx\)) is not completely matchable under \(\Gamma_z\). \(\Box\)

Thus for \(z = 2\), the interval \(xxyy\) is separating because \(\overline{xxxxyyyyyy}\) is not completely matchable. \(xy\) is not separating because \(xxyy\) is completely matchable. (Note that for \(z = 3\), \(xxxxyy\) (and also \(xyxyy\))
are separating but xyxyxy is not, although it is of the same length.}

3.4 Classification scheme for elements of $X$.

Suppose that

a) $(X, Y) \in X_n$ ,

b) $(X, Y, \Gamma_z)$ has deficiency $d$, 

c) $B_r(X, Y, \Gamma_z) \sigma$ ends with $s$ consecutive elements of either

$X$ or $Y$, $s \geq 1$,

d) $\#A_{r+1}(X, Y, \Gamma_z) = t$.

Then $\kappa(X, Y) = u = (k, n, d, s, t)$, where $X$ is divided into three sub-
sets corresponding to $k = 0, 1$, or $2$ according to which of the following 
conditions is satisfied.

A) For some $i^*$, $1 \leq i^* \leq r(X, Y, \Gamma_z)$, $A_{i^*} \neq \phi$ and $A_{i^*+1} = \phi, \ldots$,

$A_r = \phi$ and $B_{i^*}(X, Y, \Gamma_z)$ is a separating interval.

B) $A_{r+1} = \phi$, or $A_{r+1} \neq \phi$ and $A_{r+1} \subseteq X$ if and only if the final 
element of $B_r$ is classified as $X$.

C) $A_{r+1} \neq \phi$ and $A_{r+1} \subseteq X$ if and only if the final element of $B_r$ 
is classified as $Y$.

D) $A_i = \phi$ for all $i$, $1 \leq i \leq r$.

$k = 0$ when (D) holds or (A) and (B) hold.

$k = 1$ when (A) and (C) hold.

$k = 2$ when (C) holds, but (A) does not hold.

Examples: ($B_r$ is underlined) $z = 3$.

$k = 0$: yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy yyyy (A) and (B) hold. $i^* = 2$.

$k = 1$: yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy yyyy (A) and (C) hold. $i^* = 2$.

$k = 2$: yyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyyy yyyy (C) holds, (A) does not.
The remaining case, namely (A) and (D) fail to hold but (B) holds, is not possible because of the nature of \( m^0 \). To see this, let \( A_1 \neq \emptyset \) and \( A_{i+1} = \emptyset, \ldots, A_r = \emptyset \). If \( i^* = r \), then if \( B_{i^*} \) is not separating, the interval consisting of the last element of \( A_{i^*} \), all the elements of \( B_{i^*} \) and the first element of \( A_{r+1} \), is a completely matchable interval. Thus we can construct a matching of larger cardinality than \( m^0 \). If \( i^* < r \), then the interval consisting of the last element of \( A_{i^*} \), all the elements of \( B_{i^*} \) and the first element of \( B_{i^*+1} \), is a completely matchable interval. Then, using Lemma 3.2.16, we may add the last element of \( A_{i^*} \) to \( X^* + Y^* \) and take out the last element of \( B_{i^*+1} \), matching instead the elements of \( B_{i^*+1} \) except for the first and last, to each other only. The result is a matching in \( M' \) which is lexicographically smaller than \( m^0 \). This is a contradiction and so the remaining case cannot occur.

The value of \( k \) is crucial in determining what can happen to \( m^0(X,Y,\Gamma_2) \) if \( B_r(X,Y,\Gamma_2) \) is removed, in particular, whether the deficiency will change, as will be seen in the next section.

3.5 Recurrence equations for \( K(u), u \in C \).

Definition 3.5.1: Let \( D_0(\mathcal{I},s) \) be the set of all non-switching intervals consisting of \( \mathcal{I} \) elements of \( X \) and \( \mathcal{I} \) elements of \( Y \), which start with an element of \( X \) and end with \( s \) consecutive elements of \( Y \). Let \( D_1(\mathcal{I},x) \) be the set of members of \( D_0(\mathcal{I},s) \) which are separating and let \( D_2(\mathcal{I},s) = D_0(\mathcal{I},s) - D_1(\mathcal{I},s) \). Let \( D_0'(\mathcal{I},s), D_1'(\mathcal{I},s) \), and \( D_2'(\mathcal{I},s) \) be the corresponding sets of intervals which start with an element of \( Y \).

As an example, let \( z = 2 \) and \( (X,Y) = xxxxxxxyyyyyy \). Then \( \kappa(X,Y) = (0,6,4,2,4) \). In this case any element of \( D_1'(3,1) \) may be
inserted immediately after the 10th position to give an element \((X', Y')\) of \(\kappa^{-1}(1, 9, 4, 1, 2)\) such that \(a(X', Y') = (X, Y)\). Distinct elements of 
\(D'_{1}(3, 1)\) yield distinct elements \((X', Y')\). Further, for any element 
\((X', Y')\) in \(\kappa^{-1}(1, 9, 4, 1, 2) \cap a^{-1}(\kappa^{-1}(0, 6, 4, 2, 4))\), it is clear that 
\(B_{r}(X', Y', \Gamma_{z}) \in D'_{1}(3, 1)\), because \(B_{r}(X', Y', \Gamma_{z})\) must be separating when 
\(\kappa(X', Y') = 1\). It follows that there will be a 1-1 correspondence between 
\(\kappa^{-1}(1, 9, 4, 1, 2) \cap a^{-1}(\kappa^{-1}(0, 6, 4, 2, 4))\) and \(D'_{1}(3, 1)\), i.e., 
\(T(u, v) = \#D'_{1}(3, 1)\), where \(u = (0, 6, 4, 2, 4)\) and \(v = (1, 9, 4, 1, 2)\). The general situation is

the following.

**Proposition 3.5.2:** Let \((X, Y) \in \kappa^{-1}(u)\), with the last element in \(Y\) (for definiteness). Let \(v \in C\) with \(n(v) > n(u)\). Let \(Z = n(v) - n(u)\).

a) Suppose \(d(u) = d(v) > 0\), and \(t(u) > t(v)\). Then if \(k(v) = 1\) (resp. \(k(v) = 2\)), 
\(T(u, v) = \#D'_{1}(Z, s(v))\) (resp. \(\#D'_{2}(Z, s(v))\)). This corresponds to the previous example.

b) Let \(d(u) = d(v) > 0\), and \(t(u) = t(v) > 0\). Then if \(k(u) = 0\) (resp. \(k(u) = 1\)) and \(k(v) = 1\) (resp. \(k(v) = 0\)), 
\(T(u, v) = \#D_{0}(Z, s(v))\) (resp. \(\#D'_{0}(Z, s(v))\)). If \(t(u) = t(v) = 0\) and \(d(u) = d(v) > 0\), then 
\(T(u, v) = \#D_{0}(Z, s(v))\). For example, \((X, Y) = \text{xxxxxxyyyyyy}\) , insert an 
element of \(D'_{0}(Z, s(v))\) after the 8th position.

c) Let \(d(u) = d(v) = 0\). Then if \(k(u) = k(v) = 0\), 
\(T(u, v) = \#D'_{0}(Z, s(v))\). In this case \((X, Y)\) is completely matched and we insert 
an element of \(D'_{0}(Z, s(v))\) (or \(D_{0}(Z, s(v))\) as appropriate) after the 
final position. The resulting \((X', Y')\) will have \(k(v) = 0\), since 
\(d(v) = 0\).

d) Let \(d(v) = d(u) + s(u)\) and \(t(v) = t(u) + s(u)\) and \(k(u) = k(v) = 0\). In this case 
\(T(u, v) = \#D_{1}(Z, s(v))\). For example if
\((X,Y) = \ldots \ldots \ldots \), insert a separating interval after the 6th position, i.e., immediately after the last element of \(B_r(X,Y) \cap X\).

e) If none of the above relationships holds between the parameters of \(u\) and \(v\), then \(a^{-1}(X,Y) \cap \kappa^{-1}(v) = \emptyset\), i.e., there is no element \((X',Y')\) of \(X\) which is classified as \(v\) from which we can remove \(B_r(X',Y',\Gamma_{Z'})\) to get an element \((X,Y)\) in \(\kappa^{-1}(u)\). We postpone the proof of 3.5.2 until after the following corollary of 3.5.2 and 3.1.2.

**Proposition 3.5.3:**

\[
K(0,n,d,s,t) = \sum_{l=1}^{n-1} \# D_0(l,s) \sum_{i=1}^{z} K(1,n-l,d,i,t) \\
\quad + \sum_{l=1}^{n-1} \# D_1(l,s) \sum_{i=1}^{\min(z,t)} K(0,n-l,d-i,i,t-i) \\
\quad + (\text{if } d = 0) \\
\sum_{l=1}^{n} \# D_0(l,s) \sum_{i=0}^{z} K(0,n-l,0,i,0) .
\]

(The first line includes classifications such that (b) of 3.5.2 holds; the second includes classifications such that (d) holds; and the third and fourth such that (c) holds.)

We will write below \(K(\ast,n,d,s,t) = K(0,n,d,s,t) + K(1,n,d,s,t) + K(2,n,d,s,t)\).

\[
K(1,n,d,s,t) = \sum_{l=1}^{n-1} \# D_0(l,s) \sum_{i=0}^{z} K(0,n-l,d,i,t) \\
\quad + \sum_{l=1}^{n-1} \# D_1(l,s) \sum_{t' > t}^{z} \sum_{i=1}^{l} K(\ast,n-l,d,i,t') .
\]

(The first line comes from classifications such that (b) holds. The second comes from those for which (a) holds.)
\[ K(2,n,d,s,t) = \sum_{l=1}^{n-1} \# D_2(l,s) \sum_{t'>t} \sum_{i=1}^{z} K(\cdot, n-l, d, i, t'). \]

(Corresponds to the case when (a) holds.)

Proof of 3.5.2.

(a) For \( u \) and \( v \) related by \( k(v) = 1, d(u) = d(v), \) and \( t(u) > t(v) \), we wish to show that for each \( (X,Y) \in \kappa^{-1}(u) \), there is only one way to insert an interval into \( (X,Y) \) to get an element \( (X',Y') \) of \( \kappa^{-1}(v) \) such that the inserted interval coincides with \( B_T(X',Y',\Gamma_z) \), i.e., such that \( a(X',Y') = (X,Y) \). This is to insert an element of \( D_1^t(l,s(v)) \) into \( (X,Y) \) immediately in front of the last \( t(v) \) elements of \( A_{r+1}(X,Y,\Gamma_z) \).

It will be apparent from the argument that any element of \( D_1^t(l,s(v)) \) can be so inserted into \( (X,Y) \). Necessarily for distinct elements of \( D_1^t(l,s(v)) \) inserted at the same location into \( (X,Y) \), the resulting elements \( (X',Y') \) are distinct. These statements hold independently of the choice of \( (X,Y) \) for any \( (X,Y) \in \kappa^{-1}(u) \), given the choice of \( v \). This will demonstrate a 1-1 correspondence between \( a^{-1}(X,Y) \cap \kappa^{-1}(v) \) and \( D_1^t(l,s(v)) \).

Given the values \( t(u) \) and \( t(v) \), provided that the inserted element does actually correspond to \( B_T(X,Y,\Gamma_z) \), it must have been inserted after the first \( t(u) - t(v) \) elements of \( A_{r+1}(X,Y,\Gamma_z) \) and it must be in \( D_1^t(l,s(v)) \); if it is in \( D_0(l,s(v)) \), we would violate 3.2.14(b). If it is in \( D_2^t(l,s(v)) \) we would violate (A) above, since by the next Lemma the inserted elements are in this case matched only to each other by \( m^o(X,Y,\Gamma_z) \) (and hence \( A_T(X',Y',\Gamma_z) \) is nonempty.)

Lemma 3.5.4: (i) Let \( (X,Y) \) partition \( \{1,2,\ldots,n+n'\} \). Let \( I = [j_0+1, j_0+2l] \) be a non-switching interval containing \( l \) elements of
$X$ and $Z$ of $Y$, and starting with an element of $Y$ (we could equally well choose $X$). If no elements of $X \cap [j_0+2Z+1,n+n']$ are matched by $m^o$ to elements of $I \cap Y$, then no elements of $I \cap X$ are matched by $m^o$ to elements of $Y \cap [1,j_0]$. (For our purposes, we will apply the Lemma to $I = B_\gamma(X',Y',\Gamma_2)$ -- the inserted interval.)

(ii) If no elements of $Y \cap [1,j_0]$ are matched by $m^o$ to elements of $I \cap X$, then no elements of $I \cap Y$ are matched by $m^o$ to elements of $Y \cap [j_0+2Z+1,n+n']$.

Proof: Let $x_j$ be the first element of $I \cap X$. Then $x_j > j_0 + 1$. If $x_j$ is matched to $j_0$ or some earlier element of $Y$, then we must have $j_0 + 1 \in Y^+(m^o)$, since $m^o$ is compressed. But then $j_0 + 1$ must be matched to an element $x_j$ of $I \cap X$ such that $x_j > x_j$. Repeating this argument using $x_j$, in place of $x_j$, it will follow that since $I$ is non-switching, each element of $I \cap Y^+$ must be in $Y^+(m^o)$. But because $m^o$ matches $X^+$ to $Y^+$ in order, the elements of $I \cap Y$ can only be matched to elements of $I \cap X$ greater than $x_j$, since by hypothesis no elements of $X \cap [j_0+2Z+1,n+n']$ are mapped to $I \cap Y$. This implies that $I \cap Y$ is matched to a proper subset of $I \cap X$, which is impossible. This contradiction establishes (i).

(ii) If the interval $I$ is non-switching as defined above, as we run through the elements starting at the left-hand endpoint, then it is also "non-switching" as we run through the elements in reverse order, starting with the right-hand endpoint. (ii) is thus the same as (i), except that in the argument take the elements in reverse order, i.e., going in the direction $n+n',...,2,1$. □
It follows that in case (a) the inserted elements can be matched only to each other and therefore will comprise $B_r(X', Y', \Gamma_z)$. Indeed, there are no elements of $X$ after the inserted interval, since it is inserted in the middle of $A_{r+1}(X, Y, \Gamma_z)$. This is true for intervals in $D_2^r(\mathcal{L}, s(v))$ as well as $D_1^r(\mathcal{L}, s(v))$, hence the result when $k(v) = 2$.

(b) We need a preliminary Lemma.

Lemma 3.5.5: Let $(X, Y)$ partition \{1, 2, ..., n+n'\} into n elements of $X$ and $n'$ of $Y$. Suppose $n+n' \in Y^-(m^o(X, Y, \Gamma_z))$. Let $(X', Y')$ be formed by inserting $j$ $Y$-type elements at the end of $(X, Y)$. Then $m^o(X, Y, \Gamma_z) = m^o(X', Y', \Gamma_z)$.

Proof: It suffices to prove the result for $j = 1$; for $j > 1$ insert the elements one at a time, using the result for $j = 1$.

If $N(X', Y', \Gamma_z) = N(X, Y, \Gamma_z)$, then $m^o(X, Y, \Gamma_z) = m^o(X', Y', \Gamma_z)$ by Proposition 3.2.9. So assume $N(X', Y', \Gamma_z) = N(X, Y, \Gamma_z) + 1$. (It is easy to see that $N(X', Y', \Gamma_z) \leq N(X, Y, \Gamma_z) + 1$, since only one $Y$-type element has been added to $(X, Y)$.)

If $m^o(X, Y, \Gamma_z) \neq m^o(X', Y', \Gamma_z)$, then $n+n'+1 \in Y^+(m^o(X', Y', \Gamma_z))$ and hence since $m^o \in M', (x^*, n+n'+1) \in m^o(X', Y', \Gamma_z)$, where $x^*$ is the $n$-th and final element of $X$, and hence since $X = X'$, of $X'$.

The matching $m'$ defined as $m^o - \{(x^*, n+n'+1)\}$ is then a maximal matching in $M(X, Y, \Gamma_z)$ (but of course not in $M(X', Y', \Gamma_z)$), and $m'$ does not include $x^*$ as an initial vertex. By Proposition 3.2.5, the minimal matching $m$ in $M(X, Y, \Gamma_z)$ also does not include $x^*$ as an initial vertex. Then $[x^*+1, n+n'] \subseteq Y^+(m)$ since $|x^* - (n+n')| \leq z$. 
and if one of those elements were in $Y^-(m)$, it could be matched to
$x^* \in X^-(m)$ to increase the cardinality of $m$. In particular
$n+n' \in Y^+(m)$.

By Proposition 3.2.11, $m^o(X, Y, \Gamma_z) = m^c$ and therefore since $m$
is lexicographically minimal in $M$ and so there are no elements of
$Y^-(m)$ in the interval $[m(n+n'), n+n']$ (where $m(n+n')$ denotes the
initial vertex of the arc in $m$ with terminal vertex $n+n'$), then in
constructing $m^c$ (see Definition 3.2.7), $n+n'$ will never be replaced
by another element of $Y$ and hence $n+n' \in Y^+(m^o)$. This contradicts our
hypothesis.

As in (a), if the inserted elements do indeed form $B^*_{(X', Y', \Gamma_z)}$
we can conclude that the insertion must be made at the location claimed,
i.e., right after $B_{(X, Y, \Gamma_z)}$. It will thus suffice to show that the
elements of an interval inserted at this location are mapped only to
each other by $m^o(X', Y', \Gamma_z)$.

Consider now the elements of $(X, Y)$ up through and including
$B_{i^*}(X, Y, \Gamma_z)$. Suppose for definiteness that $A_{i^*} \subseteq X$, so that $B_{i^*}$ begins
with an element of $X$ and ends with an element of $Y$. Then in $(X', Y', \Gamma_z)$,
the element after $B_{i^*}(X, Y, \Gamma_z)$ will be an element of $Y'$ (either the
first element of $B_{i^*+1}(X, Y, \Gamma_z)$ or the first element of the inserted
interval, depending on whether $i^* = r$). By 3.5.4 (ii), if this element
of $Y'$ were matched to an earlier element of $X'$ by $m^o$, some element
of $A_{i^*}$ would have to be also matched to an element of $B_{i^*}$, violating
the assumption that $B_{i^*}$ is separating. Hence by 3.5.4 (i) no $Y'$-
type element of $B_{i^*}$ or subsequent to $B_{i^*}$ can be matched by $m^o$ to
any element to the left of $B_{i^*}$. But by 3.5.5, since $A_{i^*} \neq \emptyset$,
no element of $X \cap B_{i*}$ can be matched by $m^\circ$ to any elements to the left of $B_{i*}$. Hence the elements of $B_{i*}, B_{i*+1}, \ldots, B_{T}(X,Y, \Gamma_z)$ and the inserted interval are matched only to each other, this being the minimal compressed matching for the whole right-hand portion of the partition $(X,Y)$ starting with $B_{i*}$.

The existence of a 1-1 correspondence between $a^{-1}(X,Y) \cap \kappa^{-1}(v)$ and either $D_0(L,s(v))$ (if $k(v) = 1$) or $D_0'(L,s(v))$ (if $k(v) = 0$) follows.

(c) In this case, both $(X,Y)$ and $(X', Y')$ are completely matched, so any completely matchable nonswitching interval starting with the right type of element can be inserted at the end of $(X,Y)$.

(d) In the previous cases, the insertion of an interval has left the rest of the matching unchanged. In case (d) the situation is more complicated.

We first note that there are four possibilities for $(X', Y')$:

(i) $A_T(X', Y', \Gamma_z)$ and $A_{T+1}(X', Y', \Gamma_z)$ are of the same type, i.e., both $\subseteq X'$ or both $\subseteq Y'$, or $A_T \neq \emptyset$ and $A_{T+1} = \emptyset$.

(ii) $A_i = \emptyset$ for $i = 1, \ldots, r+1$.

(iii) $A_T = \emptyset$, but some earlier set $A_i \neq \emptyset$. $A_{T+1}$ is arbitrary in this case.

(iv) $A_T$ and $A_{T+1}$ are of opposite type, and both are nonempty.

Case (ii) corresponds to (c) above. Case (i) corresponds to (a) and case (iii) corresponds to (b). Only in case (iv) have we not seen how the parameters of $\kappa(X', Y')$ and $\kappa(X, Y)$ are related. (Note: (e) will have been proved when we show that (iv) corresponds to (d)).
In case (iv), when \( B_r(X', Y', \Gamma_z) \) is removed, some of the elements of \( A_{r+1}(X', Y', \Gamma_z) \) will be matched to elements of \( A_r(X', Y', \Gamma_z) \), and will be the last \( s(u) \) elements of \( B_r(X, Y, \Gamma_z) \). Hence the inserted interval which is assumed to become \( B_r(X', Y', \Gamma_z) \) must be inserted in the location described in (d), i.e., just before the last \( s(u) \) elements of \( B_r(X, Y, \Gamma_z) \). The inserted interval must be separating if it is to coincide with \( B_r(X', Y', \Gamma_z) \), because \( A_{r+1}(X', Y', \Gamma_z) \) and \( A_r(X', Y', \Gamma_z) \) are of opposite type. Further, we must have \( k(u) = 0 \) and \( t(v) = t(u) + s(u) \). ((iv) implies that \( k(v) = 0 \), by 3.2.14 (d).)

It remains to be shown that \( d(v) = d(u) + s(u) \) and that when a separating interval starting with the appropriate type element is inserted in the location described, the inserted elements are matched to each other. The following example shows what might happen when such an insert is made. Let \( z = 3 \).

\[
\begin{align*}
(X, Y) & = +++++------xxyyyyyyyx
\end{align*}
\]

\[
\begin{align*}
(X', Y') & = +++++++xxyyyyyyyxyyyyyy
\end{align*}
\]

Suppose for definiteness that \( B_r(X, Y, \Gamma_z) \) starts with an element of \( X \), so that \( A_{r+1}(X, Y, \Gamma_z) \) \( \subseteq Y \). Let \( I \) denote the interval of elements of \( B_r(X, Y, \Gamma_z) \) up to and including the last element of \( B_r \cap X \). We argue

(1) Since the inserted interval, which we will call \( J \), is separating, no element of \( Y \cap J \) can be matched by \( m^s(X', Y', \Gamma_z) \) to an element earlier than \( J \). This follows from 3.5.4 (ii), as shown in (b) above, since \( X A_{r+1}(X', Y', \Gamma_z) = \emptyset \).

(2) Consequently, as in 3.5.4(i), we can see that no element of \( I \cap Y \) can be matched to an element earlier than \( I \).

(3) If \( A_r(X, Y, \Gamma_z) \neq \emptyset \), then it follows by 3.5.5 that elements of
$B_r(X,Y,\Gamma_z)$ and later are matched only to each other. Thus the elements of $J$, the $\#B_r(X,Y,\Gamma_z)/2 - s(u)$ elements of $Y \cap I$ and the $\#B_r(X,Y,\Gamma_z)/2$ elements of $X \cap I$ are mapped only to each other by $m^o(X',Y',\Gamma_z)$, and hence the deficiency of $m^o(X',Y',\Gamma_z)$ is increased by $s(u)$ over that of $m^o(X,Y,\Gamma_z)$ (the part of the matching earlier than $B_r$ remains unchanged.) Since $m^o$ is the minimal compressed matching, it is clear that the last element of $X \cap I \in X^-(m^o(X',Y',\Gamma_z))$, and that the elements of the inserted interval are matched only to each other.

(4) If $A_r(X,Y,\Gamma_z) = \emptyset$, then the sets $B_{i*}, B_{i+1*}, \ldots, B_r$ are adjacent to each other. But in this case as in (b) above, since $B_{i*}$ is separating, no elements of $B_{i*}$ or later elements can be matched by $m^o$ to elements of $A_{i*}$ or earlier, and hence the elements of $B_{i*}, B_{i+1*}, \ldots, B_{r-1}$ are matched only to each other, giving the same conclusion as in (3).

\[ \square \]

### 3.6 Reducing the recurrence equations.

The following result places limits on the range of the indices of summation in the equations of Proposition 3.5.2.

**Proposition 3.6.1:** Let $(X,Y)$ partition $\{1,2,\ldots,2n\}$, where $n \geq 2$. Then $D(X,Y,\Gamma_z) \leq n - z$, i.e., $N(X,Y,\Gamma_z) \geq z$, $1 \leq z \leq n$.

**Proof:** Suppose $N(X,Y,\Gamma_z) < z$. Then $X^+(m^o) + Y^+(m^o)$ must contain at least one separating interval in order to keep the elements of $X^-$ and $Y^-$ from being matched to each other. Using 3.2.16, we can, as before, show that there must be a non-switching separating interval, $I$. Since $N(X,Y,\Gamma_z) < z$, $\#(I) < 2z$. 

However, a non-switching interval I of \( z \) elements of X and \( z \) elements of Y cannot be separating if \( z < z \). To see this, let the elements of \( \text{In}X \) be \( a_1 < \ldots < a_z \) and those of \( \text{In}Y \) be \( b_1 < \ldots < b_z \), for \( z < z \). Suppose for definiteness that \( a_1 < b_1 \). For each \( i, 2 \leq i \leq z \), there are at least \( i - 2 \) elements of I to the left of \( a_i \) and at least \( z - i \) to the right of \( b_i \). So \( |a_{i-1} - b_i| \leq 2z - (z-i) - (i-2) = z + 1 \leq z \). Hence \( (a_{i-1}, b_i) \in \Gamma_z \). If we insert an X-type element in front of I and a Y-type element at the end of I, then we can match (in \( \Gamma_z \)) \( b_1 \) to the new X-type element and \( a_z \) to the new Y-type element. Hence I is not separating.

We can reduce the number of variables appearing in 3.5.2 by collecting terms in the following manner.

Define

\[
V_1(n, d, t) = \sum_{t' \geq t} \sum_{i=1}^{z} K(\cdot, n, d, i, t')
\]

\[
V_2(n, d, t) = \sum_{i=1}^{z} K(0, n, d, i, t)
\]

\[
V_3(n, d, t) = \sum_{i=1}^{z} K(1, n, d, i, t) 
\]

\[
V_4(n, d, t) = \sum_{i=1}^{\min(z, t)} K(0, n, d - i, i, t - i)
\]

The equations of 3.5.2 then become:

\[
K(0, n, d, s, t) = \sum_{z=1}^{n-1} \#D_0(z, s) V_3(n-z, d, t)
\]

\[
+ \sum_{z=1}^{n-1} \#D_1(z, s) V_4(n-z, d, t)
\]

\[
+ (\text{if } d=0) \sum_{z=0}^{n} \#D_0(z, s) V_2(n-z, 0, 0)
\]
\[ K(1,n,d,s,t) = \sum_{\ell=1}^{n-1} \#D_0(\ell,s)V_2(n-\ell,d,t) + \sum_{\ell=1}^{n-1} \#D_1(\ell,s)V_1(n-\ell,d,t) \]

\[ K(2,n,d,s,t) = \sum_{\ell=1}^{n-1} \#D_2(\ell,s)V_1(n-\ell,d,t) \]

Our initial conditions are \( V_2(0,0,0) = 1 \), corresponding to the case \( n = 0 \) and \( (X,Y) = \emptyset \), and \( V_1(0,0,0) = 0 \), \( V_3(0,0,0) = 0 \), \( V_4(0,0,0) = 0 \). If \( V_i(p,d,t) \) are known for \( i=1,2,3,4 \) and all \( d \) and \( t \) and all \( p \leq n-1 \), for some \( n \), then once \( K(k,n,d,s,t) \) have been computed for all \( k,d,s, \) and \( t \), we can compute \( V_i(n,d,t) \) for \( i=1,2,3,4 \) and all values of \( d \) and \( t \). Thus in evaluating the recurrence equations, it will never be necessary to store the values of \( K(n,k,d,s,t) \). This reduction from four indices to three makes evaluation by electronic computer feasible for values of \( n \) up to several hundred, once the values of \( \#D_i(\ell,s) \) are known.

3.7 Computation of \( \#D_i(\ell,s) \).

Using an argument based on 3.1.2 we will be able to find \( \#D_i(\ell,s) \), \( i=0,1,2, \) for values of \( \ell \) up to several hundred and for values of \( z \) up to 12. The limit of \( z = 12 \) is established by computer storage facilities. The method will suffice to construct tables of our distribution function up to about \( n = 70 \), after which we cannot compute the upper tail.

After Sections 3.1 - 3.6, we have remaining only the problem of enumerating for a given value of \( z \), the completely matchable non-switching partitions \( (X,Y) \) of \( \{1,2,\ldots,2\ell\} \) such that \( \#(X) = \#(Y) \) and
such that the interval is separating, and also those such that the interval is non-separating. The "separating" question will be difficult to handle directly. However, we can instead enumerate \( \#D_0(\mathcal{I},s) \) and \( \#D_3(\mathcal{I},s) \), where \( D_3(\mathcal{I},s) \) is defined as the subset of \( D_0(\mathcal{I},s) \) consisting of strictly non-switching partitions.

**Proposition 3.7.1:** \( \#D_2(\mathcal{I},s) = \#D_3(\mathcal{I}+1,s+1) \).

**Proof:** For \( (X,Y) \in D_2(\mathcal{I},s) \), let \( t(X,Y) \) be the partition of \( \{1,2,\ldots,2(\mathcal{I}+1)\} \) formed by inserting an \( X \)-type element at the beginning of \( (X,Y) \) and a \( Y \)-type element at the end. Then, since \( (X,Y) \) is non-separating, \( t(X,Y) \in D_3(\mathcal{I}+1,s+1) \). \( t \) is clearly an injective mapping onto a subset of \( D_3(\mathcal{I}+1,s+1) \). In fact, \( t \) is onto all of \( D_3(\mathcal{I}+1,s+1) \), with \( t^{-1} \) being defined by removing the first and last elements of \( (X',Y') \in D_3(\mathcal{I}+1,s+1) \). So defined, \( t^{-1}(X',Y') \) is completely matchable, by 3.2.16, and is clearly non-switching (though not necessarily strictly non-switching). \( t^{-1}(X',Y') \) is not separating since \( (X',Y') \) is completely matchable. Hence \( t^{-1}(X',Y') \in D_2(\mathcal{I},s) \). \( \square \)

If we can find \( \#D_0(\mathcal{I},s) \) and \( \#D_3(\mathcal{I},s) \) for all \( \mathcal{I} \) and \( s \), then we find \( \#D_2(\mathcal{I},s) = \#D_3(\mathcal{I}+1,s+1) \) and \( \#D_1(\mathcal{I},s) = \#D_0(\mathcal{I},s) - \#D_2(\mathcal{I},s) \).

Let us first address the problem of finding for a given value of \( z \), \( \#D_0(\mathcal{I},s) \) for all \( \mathcal{I} \) and \( s \). We use the ideas of 3.1.2. Let \( Y_{\mathcal{I}} \) be the class of all non-switching completely matchable partitions of \( \{1,2,\ldots,2\mathcal{I}\} \) such that \( \#(X) = \#(Y) \) and \( 1 \in X \). Let \( Y = \bigcup_{\mathcal{I}=0}^{\infty} Y_{\mathcal{I}} \).

Define the mapping \( a: Y_{\mathcal{I}} \to Y_{\mathcal{I}-1} \) (for \( \mathcal{I} \geq 1 \)) by letting \( a(X',Y') \) be
the element of \( V_{\ell-1} \) obtained by removing from \((X',Y')\) the last element of \(X'\) and the last element of \(Y'\). We let \( V_0 = \{ \emptyset \} \) and \( a(\emptyset) = \emptyset \).

Define the classification set \( C \) for elements of \( Y \) as follows. Let \((X,Y) \in V_{\ell}\). Suppose \((X,Y)\) ends with \(s\) successive elements of \(Y\), namely \(y_{\ell-s+1}, y_{\ell-s+2}, \ldots, y_{\ell}\). Define \(r_1, \ldots, r_s\) by \(r_1 = |x_{\ell-s+1} - y_{\ell-s+1}|, r_2 = |x_{\ell-s+2} - y_{\ell-s+2}|, \ldots, r_s = |x - y_\ell|\). For example, if \(z = 4\) and \((X,Y) = xxxyxyxxyy\), then \(\ell = 5, s = 3, r_1 = 4, r_2 = 4, r_3 = 3\). Then let \((X,Y) = (\ell, s, r_1, \ldots, r_s)\).

**Proposition 3.7.2:** Let \(u = (\ell, s, r_1, \ldots, r_s)\) be a vector of positive integers. The following conditions are necessary and sufficient for \(u \in C\).

1. \(\ell \geq s, s \leq z\).
2. \(r_i \leq z\), for \(1 \leq i \leq s-1\), and \(r_s = s\).
3. \(\ell \geq r_1 \geq r_2 \geq \ldots \geq r_s\).

**Proof:** The necessity of (1) and (2) are obvious as is the necessity of \(r_1 \geq r_2 \geq \ldots \geq r_s\). \(\ell \geq r_1\), since between \(x_{\ell-s+1}\) and \(y_{\ell-s+1}\) there are \(s-1\) elements of \(X\) and \(r_1 - r_s\) elements of \(Y\). So we have located \(r_1 - r_s\) elements of \(Y\) not including \(y_{\ell-s+1}, \ldots, y_{\ell}\). Hence \(r_1 \leq \#(Y) = \ell\).

To show sufficiency, we describe an algorithm by which an element \((X,Y) \in V_{\ell}\) can be constructed with \(u\) as its classification vector. Start with an \(X\)-type element, followed by \(s\) \(Y\)-type elements. Insert in front of these elements an \(X\)-type element followed by \(r_{s-1} - r_s\) \(Y\)-type elements. In front of these elements, insert an \(X\)-type element
followed by \( r_{s-2} - r_{s-1} \) Y-type elements,..., in front of these elements insert an X-type element followed by \( r_1 - r_2 \) Y-type elements. We now have assembled \( r_1 \) Y-type elements and \( s \) X-type elements. Insert \( r_1 - s \) X-type elements in front of these. If \( l > r_1 \), insert an element of \( Y_{l-r_1} \) at the beginning. The resulting partition of \( \{1,2,...,2l\} \) is clearly non-switching and has classification vector \( u \).

We need only show that it is completely matchable. This follows from (2): clearly the initial element of \( Y_{l-r_1} \) is completely matchable (to itself); the first \( s \) X-type elements inserted, one at a time, by the algorithm can be matched to the right-most \( s \) Y-type elements; since \( x_{l-s+1} \) and \( y_{l-s+1} \) are matched, then the \( r_1 \)-\( s \) elements of \( X \) inserted at the same time (which will be immediately in front of \( x_{l-s+1} \)) can be matched respectively to \( y_{l-r_1+1},...,y_{l-s} \).

For \((X',Y') \in Y, u,v \in C\), we have \( \# a(X',Y') = 1 \). For \((X,Y) \in \kappa^{-1}(v)\), we may obtain every element of \( a^{-1}(X,Y) \) by inserting an X-type element after one of the last \( s(u) \) elements of \( Y \) and inserting a Y-type element at the end. It will be clear that for each \( v \in C \), \( T(u,v) = \# a^{-1}(X,Y) \cap \kappa^{-1}(v) = 0 \) or 1, depending on \( u \).

**Proposition 3.7.3:** Let \( u,v \in C \). The following are necessary and sufficient conditions for \( T(u,v) = 1 \).

1. \( \ell(v) = \ell(u) + 1 \)
2. \( s(v) \leq s(u) + 1 \)
3. For \( i=1,2,...,s(v)-1 \), we have \( r_{s(v)-i+1}(u) + 1 = r_{s(v)-i}(v) \).
Proof: (1) and (2) are clearly necessary (consider the definition of 
)

a). Note that in performing the insertion described to get an element 
\((X',Y') = \alpha^{-1}(X,Y) \cap s^{-1}(v)\), for each of the elements 
\(y_\ell, y_{\ell-1}, \ldots, y_{\ell-s(v)+1}\) of \((X,Y)\), we have inserted exactly one \(X\)-type element between that element and the corresponding element \(x, x_{\ell-1}, \ldots, x_{\ell-s(v)+1}\) of \(X\). Then given \((X,Y) \in \kappa^{-1}(u)\), we can make the appropriate insertion to obtain an element \((X',Y') = \alpha^{-1}(X,Y) \cap s^{-1}(v)\), if and only if (1), (2) and (3) are satisfied.

\[\text{Proposition 3.7.4: Let } \ell \geq z. \text{ Let } U = \{u \in C: s(u) = s_0\} \text{ for some } s_0, 1 \leq s_0 \leq \min(\ell, z). \text{ Then } \#(U) = \min(\ell, z) - s_0 \left( \frac{\sum_{j=1}^{\ell} j}{j} \right).\]

Proof: \(\#(U) = \sum_{j=s_0}^{\ell} \#\{u \in C: s(u) = s_0 \text{ and } r_1(u) = j\} \).

Thus for given \(j, s_0 \leq j \leq \ell\), we wish to evaluate \(\#(U_j)\), where 
\(U_j = \{u \in C: s(u) = s_0, r_1(u) = j\}\). Let \(u = (\ell, s_0, j, r_2, \ldots, r_{s_0})\). Let \(j_1 = j - r_2, j_2 = r_2 - r_3, \ldots, j_{s_0-1} = r_{s_0-1} - s_0\). Then 
\[\sum_{i=1}^{s_0-1} j_i = j - s_0; \text{ and } j_1 \geq 0, \ldots, j_{s_0-1} \geq 0.\]

It is clear that, conversely, for any integers \(j_1, \ldots, j_{s_0-1}\) which satisfy (*), the vector 
\[(\ell, s_0, j, s_0 + \sum_{i=2}^{s_0} j_i, s_0, j_1, \ldots, s_0 + j_{s_0-1}, s_0) \in U_j,\]

by Proposition 3.7.2. It follows that \(\#(U_j)\) is the number of sequences 
\(j_1, \ldots, j_{s_0-1}\) satisfying (*). This number is \(\left\{\frac{j+s_0+2}{j}\right\}\), (see, for example, [15]).
If \( \ell < z \), then we need only consider \( j \) such that \( s_0 \leq j \leq z \).

Using the notation of 3.1.2,

\[
\#D_0(\ell_0, s_0) = \sum_{\{u \in C : s(u) = s_0, \ell(u) = \ell_0\}} K(u).
\]

Once the values of \( T(u,v) \) are known for all \( u \) and \( v \), we can find \( K(u) \) for all \( u \in C \) by starting with \( K(\kappa^{-1}(\emptyset)) = 1 \). We will not actually compute the values \( T(u,v) \), but will use them implicitly as follows.

Let \( C_s \) be the set of all possible values of \((r_1(u), \ldots, r_s(u))\) where \( u \in C \) such that \( s(u) = s \). For two elements \( w = (r_1, \ldots, r_s) \) and \( w' = (r'_1, \ldots, r'_s) \) of \( C_s \), we say that \( w < w' \) when the vector \((r_1, \ldots, r_s)\) is lexicographically smaller than \((r'_1, \ldots, r'_s)\). For each \( u \in C \), let \( j(u) \) denote the rank of \((r_1, \ldots, r_s(u))\) in \( C_s(u) \) in this lexicographic ordering. A classification \( u \in C \) can now be described by three parameters, \((\ell(u), s(u), j(u))\). Let \( v \in C \) and let \( j^*(j(v), s_0, s(v)) \) be the rank in \( C_{s_0} \) of that vector \((r_1, \ldots, r_{s_0})\) such that \( u = (\ell(v)-1, s_0, r_1, \ldots, r_{s_0}) \). By 3.7.3 it can be seen that there is at most one such vector; if \( T(u,v) = 0 \), there will be no such vector.

Then

\[
K(\ell(v), s(v), j(v)) = \sum_{s = s(v)-1}^z K(\ell(v)-1, s, j^*(j(v), s, s(v))),
\]

with the convention that \( K(0,0,0) = 1 \) and \( K(\ell, s, 0) = 0 \) otherwise.

The array \( j^*(j', s, s') \) for \( 1 \leq s' \leq z \) and \( s'-1 \leq s \leq z \) and \( 1 \leq j' \leq \#(U_{s'}) \) \((U_s', \text{ as in 3.7.4})\) was determined using 3.7.3 essentially by brute force, though systematically. The pattern established by the lexicographic ordering permits this. The practical limitation
on the value of $z$ which can be handled is due to the size of this array.

To evaluate $\#D_3(l, s)$, we use the same procedure except that in 3.7.2 we have the additional condition that $s > 1$ if $l > 1$. To see this, consider the sequence $a(X, Y), a_2(X, Y), \ldots, a_{l-1}(X, Y), \emptyset$. If $(X, Y)$ is strictly non-switching, then no element of this sequence except $a_{l-1}(X, Y)$ may be a partition which ends with $\ldots xy$, i.e., such that $s(u) = 1$. It can then be seen that $s > 1$ if $l > 1$ is also a sufficient condition for $(X, Y) \in D_3$.

The problem of enumerating $D_0(l, s)$ can be rephrased as a problem of enumerating the set of matrices of 0's and 1's which satisfy certain conditions. However, this equivalent problem does not produce a more satisfactory solution, and we will simply give the statement of this problem.

To enumerate the set of $l \times l$ matrices $M$ such that:

1. $M$ consists only of 0's and 1's;
2. The non-zero elements in each row are consecutive, with the last 1 being on the main diagonal;
3. For $i=2, \ldots, l$, the first non-zero element in row $i$ never appears in an earlier column than the first element of row $i-1$;
4. The sum of the elements in each row is less than or equal to $z$.

Equivalently we may replace (4) by:

1. $M$ is symmetric;
2. For $i=1, \ldots, l$, the total number of non-zero elements in the $i$-th row and $i$-th column (no element being counted twice) is less than or equal to $z$. 

CHAPTER 4

Asymptotic Distribution of the Test Statistics

4.1 Asymptotic distribution for $d_p$.

We use results from Sections 1.3 and 1.4 to get some preliminary lemmas. Let $F$ be defined as in 1.3.9. If $F \in F$ and

$F = [b_1, c_1] \cup \ldots \cup [b_k, c_k], 0 \leq b_1 \leq c_1 < b_2 \leq c_2 \cdots < b_k \leq c_k \leq 1,$

then the intervals $[b_i, c_i]$ will be referred to as the "component" intervals of $F$. For each $F \in F$, we are interested in the total number of left-hand endpoints, excluding 1, of component intervals, plus the number of right-hand endpoints, excluding 0. Specifically we define

$$q(F) = \lim_{\delta \to 0} \frac{\mu(F_\delta - F)}{\delta},$$

where $\mu$ denotes uniform measure on $[0,1]$. Let $F^{(2\delta)}$ be defined as in 1.3.9. Note that $F^{(2\delta)} = \{F \in F : q(F) = q(F_\delta)\}$.

Define the function $f$ on $D[0,1]$ by

$$f(y) = \sup_{F \in F} \frac{1}{q(F)+1} \int_F dy(t), \text{ for all } y \in D[0,1].$$

(Note: for $F = \bigcup_{i=1}^{k} [b_i, c_i]$, we define $\int_F dy(t) = \sum_{i=1}^{k} (y(c_i) - y(b_i)).$)

For each $\delta > 0$, define the function $f^{(\delta)}$ on $D[0,1]$ by
\[ f(\delta)(y) = \sup_{F \in F(2\delta)} \frac{1}{q(F)+1} \int_{F} dy(t). \]

**Lemma 4.1.1:** Let \( X \) and \( X \) be probability measures on \( I[0,1] \) with \( X(C[0,1]) = 1 \), and let \( X_n \Rightarrow X \). Let \( \{\alpha_n\} \) be a sequence of numbers such that \( \alpha_n \downarrow 0 \). Then \( f(\alpha_n)(X_n) \Rightarrow f(X) \).

**Proof:** By Proposition 1.2.4, since \( X(C[0,1]) = 1 \), it suffices to show that for \( y \in C[0,1], \ y_n \to y \) in \( D[0,1] \) implies that \( f(\alpha_n)(y_n) \to f(y) \).

We first show that \( f(y) \leq \limsup_{n \to \infty} f(\alpha_n)(y_n) \), i.e., that given \( \epsilon > 0 \), for \( n \) sufficiently large,

\[ \sup_{F \in F} \frac{1}{q(F)+1} \int_{F} dy(t) \leq \sup_{\alpha_n} \frac{1}{q(F)+1} \int_{F} dy_n(t) + \epsilon, \]

This will be done by showing that for \( n \) sufficiently large depending only on \( \epsilon \) and on \( y \), then for each \( F \in F \), there exists a set \( G \in F(\alpha_n) \) such that

\[ \frac{1}{q(F)+1} \int_{F} dy(t) \leq \frac{1}{q(G)+1} \int_{G} dy_n(t) + \epsilon. \]

\( G \) will be constructed from \( F \) as in Proposition 1.3.10. I.e., if \( F = \bigcup_{i=1}^{k} [b_i, c_i] \) as above, then \( G \) is defined as the union of \( F \) with every interval \( [c_i, b_{i+1}] \) such that \( b_{i+1} - c_i \leq 2\alpha_n \) and with \( [0, b_1] \) if \( b_1 \leq \alpha_n \) and with \( [c_m, 1] \) if \( c_m \geq 1 - \alpha_n \). Then clearly \( G \in F(2\alpha_n) \).

**Remark 1:** \( q(F) \geq q(G) \). (Clearly \( G \in F \) so \( q(G) \) is defined.) Indeed, each component interval of \( G \) contains a point of some component interval of \( F \) and hence must contain the whole component interval. Each component interval of \( F \) intersects precisely one component interval of
G. If a component interval of $G$ contains a component interval of $F$ which has an endpoint at $0$ (resp. $1$), then that component interval of $G$ has an endpoint at $0$ (resp. $1$). Thus the endpoints in $(0,1)$ of component intervals of $G$ can be placed in 1-1 correspondence with the endpoints in $(0,1)$ of a subset of the component intervals of $F$. (Of course, a component interval of $G$ may contain several component intervals of $F$.)

Remark 2: $G - F$ consists of at most $q(F)$ intervals each of length $< 2\alpha_n$. (These will be referred to as component intervals although in general they are not closed intervals and $G - F \not\subset F$.) Indeed, it is evident that for each component interval of $G - F$, at least one of its endpoints must also be the endpoint of a component interval of $F$. If that endpoint is $0$ (resp. $1$), then we must have $b_1 = c_1 = 0$ (resp. $b_1 = c_1 = 1$), in which case the endpoint is counted once, as a right-hand (resp. left-hand) endpoint. Therefore the number of component intervals of $G - F$ is less than or equal to $q(F)$.

Choose $n$ sufficiently large that

$$\sup_{t_1, t_2 \in [0,1]} |y(t_1) - y(t_2)| < \frac{\varepsilon}{3} \quad \text{and} \quad \sup_{0 \leq t \leq 1} |y_n(t) - y(t)| < \frac{\varepsilon}{3}.$$  

$$|t_1 - t_2| < 2\alpha_n$$

This is possible because $y \in C[0,1]$ implies that $y$ is uniformly continuous and further that convergence to $y$ in $D[0,1]$ is equivalent to convergence in the uniform metric.

But

$$\left| \frac{1}{q(F)+1} \int_F dy(t) - \frac{1}{q(F)+1} \int_G dy_n(t) \right| \leq$$
\[
\frac{1}{q(F) + 1} \left| \int_G dy_n(t) - \int_G dy(t) \right| + \frac{1}{q(F) + 1} \left| \int_F dy(t) - \int_F dy(t) \right|.
\]

By remark 2,

\[
\left| \int_G dy - \int_F dy \right| \leq 2q(F) \sup_{t_1, t_2 \in [0, 1], |t_1 - t_2| < 2\alpha n} |y(t_1) - y(t_2)| \leq 2q(F)\frac{\varepsilon}{3}.
\]

\[
\left| \int_G dy_n - \int_G dy \right| \leq q(G) \sup_{0 \leq t \leq 1} |y_n(t) - y(t)| < q(G)\frac{\varepsilon}{3} \leq q(F)\frac{\varepsilon}{3}.
\]

Therefore,

\[
\left| \frac{1}{q(F) + 1} \int_F dy(t) - \frac{1}{q(F) + 1} \int_G dy_n(t) \right| \leq \varepsilon.
\]

Hence,

\[
(*) \quad \frac{1}{q(F) + 1} \int_F dy(t) \leq \frac{1}{q(F) + 1} \int_G dy_n(t) + \varepsilon.
\]

We have two cases: (i) \( \int_G y_n(t) \geq 0 \) and (ii) \( \int_G y_n(t) < 0 \). In case (ii), replace \( G \) by \( G' = \emptyset \). Then \( G' \in G^{(2\alpha_n)} \). (ii) and (*) imply that

\[
\frac{1}{q(F') + 1} \int_{G'} dy(t) \leq \frac{1}{q(G') + 1} \int_{G'} dy_n(t) + \varepsilon.
\]

In case (i), (*) and Remark 2 imply that

\[
\frac{1}{q(F) + 1} \int_F dy(t) \leq \frac{1}{q(G) + 1} \int_G dy_n(t) + \varepsilon.
\]

In either case we thus have the necessary inequality, showing that

\[
f(y) \leq \limsup_{n \to \infty} f^{(\alpha_n)}(y_n).
\]

We complete the proof by showing that \( \limsup_{n \to \infty} f^{(\alpha_n)}(y_n) \leq f(y) \).
As above, given \( \varepsilon \) and choosing \( n \) sufficiently large, then for any \( F \in F \),

\[
\left| \int_F \tilde{d}y_n(t) - \int_F \tilde{d}y(t) \right| \leq q(F)\varepsilon,
\]

so

\[
\int_F \tilde{d}y_n(t) \leq \int_F \tilde{d}y(t) + q(F)\varepsilon.
\]

Consequently,

\[
\sup_{F \in F} \frac{1}{(2\alpha_n)^q(F)+1} \int_F \tilde{d}y_n(t) \leq \sup_{F \in F} \frac{1}{(2\alpha_n)^q(F)+1} \int_F \tilde{d}y(t) + \varepsilon
\]

\[
\leq \sup_{F \in F} \frac{1}{q(F)+1} \int_F \tilde{d}y(t) + \varepsilon = f(y) + \varepsilon.
\]

Since \( \varepsilon \) is arbitrary, \( \limsup_{n \to \infty} f(\alpha_n) (y_n) \leq f(y) \).

Recall from Proposition 1.3.9 that

\[
\Pr\{d_F(P, Q) \leq c\} = \Pr\{ \sup_{F \in F(2c)} P(F) - Q(F^c) \leq c\}
\]

\[
= \Pr\{P(F) \leq Q(F^c) + c; \text{ for all } F \in F(2c)\}
\]

Letting \( \mu \) denote uniform measure on \([0,1]\), let \( \mu_m \) and \( \nu_n \) be the empirical measures of the sets \( \{r_1/N\} \) and \( \{s_1/N\} \), where \( r_1, \ldots, r_m \) and \( s_1, \ldots, s_n \) are the ranks among the pooled observations of two i.i.d. random samples from \( \mu \), and \( N = m+n \). Let \( \gamma_N \) be defined as \( \frac{m}{N^2} + \frac{n}{N^2} \). In other words, \( \gamma_N \) is the measure which places mass \( \frac{1}{N} \) at each of the points \( 1/N, 2/N, \ldots, 1 \).

Proposition 4.1.2: \( n^{1/2} d_F(\mu_n, \mu) \to f(W^*) \) as \( n \to \infty \).
Proof: Let \( c \) be a continuity point of the distribution of \( f(W^o) \).

\[
\Pr\left( d_F^n\left( \mu_n, \mu \right) \leq cn^{-\frac{k}{n}} \right) = \Pr\left( \mu_n(F) \leq \mu(F^{cn^{-\frac{k}{n}}}) + cn^{-\frac{k}{n}} \right) \text{ for all } F \in F(2cn^{-\frac{k}{n}}) \\
= \Pr\left( n^{-\frac{k}{n}}\left( \mu_n(F) - \mu(F) \right) \leq n^{-\frac{k}{n}}\mu(F^{cn^{-\frac{k}{n}}}) - F \right) + c; \text{ for all } F \in F(2cn^{-\frac{k}{n}}) .
\]

For \( F \in F(2cn^{-\frac{k}{n}}) \), \( \mu(F^{cn^{-\frac{k}{n}}}) - F = q(F)cn^{-\frac{k}{n}} \), since for such a set \( F \) there will be no overlap at the ends of the component interval. Hence the last expression is equal to

\[
\Pr\left( n^{-\frac{k}{n}}(\mu_n(F) - \mu(F)) \leq (q(F)+1)c; \text{ for all } F \in F(2cn^{-\frac{k}{n}}) \right)
\]

\[
= \Pr\left\{ \sup_{F \in F(2cn^{-\frac{k}{n}})} \frac{1}{q(F)+1} \left\{ dX_n(t) \leq c \right\} = \Pr\left\{ f(cn^{-\frac{k}{n}})(X_n) \leq c \right\} \right.
\]

where \( X_n(t) = n^{-\frac{k}{n}}(\mu_n[0,t] - \mu[0,t]) \). By 1.4.2, \( X_n \Rightarrow W^o \) and hence by 4.1.1 \( f(cn^{-\frac{k}{n}})(X_n) \Rightarrow f(W^o) \).

Hence

\[
\lim_{n \to \infty} \Pr\left( n^{-\frac{k}{n}}d_F^n(\mu_n, \mu) \leq c \right) = \lim_{n \to \infty} \Pr\left\{ f(cn^{-\frac{k}{n}})(X_n) \leq c \right\} = \Pr\{ f(W^o) \leq c \} .
\]

**Proposition 4.1.3:** \( \left( \frac{n}{mN} \right)^{-\frac{k}{n}} d_F^n(\mu_m, \gamma_N) \Rightarrow f(W^o) \) as \( m, n \to \infty \)

Proof: We treat separately the cases (i) \( 0 < a \leq \frac{m}{n} \leq b < \infty \), for some constants \( a \) and \( b \); (ii) \( \frac{m}{n} \to 0 \); (iii) \( \frac{m}{n} \to \infty \). This will suffice to show that \( \left( \frac{n}{mN} \right)^{-\frac{k}{n}} d_F^n(\mu_m, \gamma_N) \Rightarrow f(W^o) \) as \( m, n \to \infty \); indeed, any sub-sequence \( (m_i, n_i) \) must contain a sub-subsequence satisfying (i), (ii), or (iii), so weak convergence would hold for that sub-subsequence.
(i) Let $\delta(m,n) = \left(\frac{n}{mN}\right)^{\frac{1}{2}}$. Let $c$ be a continuity point of the distribution of $f(W^o)$. By 1.3.9,

$$\Pr\{p_{\mu_{m,n}}, \gamma_n \leq c\delta(m,n)\} = \Pr\{\mu_{m,F} \leq \gamma_n (F^{c\delta(m,n)})\}$$

$$+ c\delta(m,n); \text{ for all } F \in F(2c\delta(m,n))$$

$$= \Pr\{\mu_{m,F} - \gamma_n (F) \leq \gamma_n (F^{c\delta(m,n)}) - F + c\delta(m,n); \text{ for all } F \in F(2c\delta(m,n))\}$$

$$= \Pr\{\mu_{m,F} - \gamma_n (F) \leq q(F)(c\delta(m,n) + \eta(F))$$

$$+ c\delta(m,n); \text{ for all } F \in F(2c\delta(m,n))\} ,$$

where $\eta(F)$ satisfies $\frac{1}{N} \leq \eta(F) \leq \frac{1}{N}$, since for any interval $I \subseteq [0,1]$, $|\gamma_n (I) - \mu(I)| \leq \frac{1}{N}$. Thus $\eta(F) = O\left(\frac{1}{N}\right)$, uniformly in $F$.

Multiplying through by $\delta(m,n)^{-1}$, we get

$$\Pr\{\left(\frac{mN}{m}\right)^{\frac{1}{2}} (\mu_{m,F} - \gamma_n (F)) \leq (q(F) + 1)c$$

$$+ q(F)\left(\frac{m}{n}\right)^{\frac{1}{2}} 0(N^{-\frac{1}{2}}); \text{ for all } F \in F(2c\delta(m,n))\}$$

$$= \Pr\left\{\sup_{F \in F(2c\delta(m,n))} \frac{1}{q(F)+1} \int_{F}^{F} \frac{d\gamma_{m,n}(t)}{1} \leq c + \left(\frac{m}{n}\right)^{\frac{1}{2}} 0(N^{-\frac{1}{2}})$$

$$\text{for all } F \in F(2c\delta(m,n))\}$$

$$= \Pr\{f(c\delta(m,n))(Y_{m,n}) \leq c + \left(\frac{m}{n}\right)^{\frac{1}{2}} 0(N^{-\frac{1}{2}})\} ,$$

where $Y_{m,n}$ is defined in 1.4.2. By 1.4.2, as $m,n \to \infty$ in such a way that (i) is satisfied, $Y_{m,n} \Rightarrow W^o$, so by 4.1.1, $f(c\delta(m,n))(Y_{m,n}) \Rightarrow f(W^o)$.

Since we have assumed $0 < a \leq \frac{m}{n}$, $\left(\frac{m}{n}\right)^{\frac{1}{2}} 0(N^{-\frac{1}{2}}) = 0(N^{-\frac{1}{2}})$. Hence since $c$ is a continuity point of the distribution of $f(W^o)$,
\[
\lim_{m,n \to \infty} \Pr\{d_P(\nu_m, \gamma_N) \leq c\} = \Pr\{f(W^o) \leq c\}.
\]

(ii) In this case we consider

\[
\Pr\{d_P(\nu_m, \gamma_N) \leq cm^{-\frac{1}{2} \gamma}\} = \Pr\{\nu_m(F) - \mu(F) \leq (\gamma_N(F^{cm^{-\frac{1}{2} \gamma}}) - \mu(F)) + cm^{-\frac{1}{2} \gamma} \text{ for all } F \in F(2cm^{-\frac{1}{2} \gamma})\}
\]

\[
= \Pr\{\nu_m(F) - \mu(F) \leq (\gamma_N(F^{cm^{-\frac{1}{2} \gamma}}) - \mu(F^{cm^{-\frac{1}{2} \gamma}})) + \mu(F^{cm^{-\frac{1}{2} \gamma}})_F
\]

\[
+ cm^{-\frac{1}{2} \gamma} \forall F \in F(2cm^{-\frac{1}{2} \gamma})\}
\]

\[
= \Pr\{m^{\frac{1}{2} \gamma}(\nu_m(F) - \mu(F)) \in q(F)(c + m^{\frac{1}{2} \gamma}\eta(F)) + c \text{ for all } F \in F(2cm^{-\frac{1}{2} \gamma})\},
\]

where \(-\frac{1}{N} \leq \eta(F) \leq \frac{1}{N}\),

\[
= \Pr\{\sup_{F \in F(2cm^{-\frac{1}{2} \gamma})} \frac{1}{q(F)+1} \int_F \mathcal{X}_m(t) \leq c + m^{\frac{1}{2} \gamma}0(\frac{1}{N}) \}\,
\]

where \(\chi_m\) is as defined in 4.1.2. Since \(m^{\frac{1}{2} \gamma}0(N^{-1}) = 0(N^{-1})\).

Hence if \(c\) is a continuity point of the distribution of \(f(W^o)\), then

\[
\lim_{m,n \to \infty} \Pr\{\left(\frac{n}{mn}\right)^{\frac{1}{2} \gamma} d_P(\nu_m, \gamma_N) \leq c\}
\]

\[
= \lim_{m,n \to \infty} \Pr\{m^{\frac{1}{2} \gamma}d_P(\nu_m, \gamma_N) \leq c\} = \Pr\{f(W^o) \leq c\}.
\]

(iii) Let \(\delta(m,n) = n^{\frac{1}{2} \gamma}/m\).

\[
\Pr\{d_P(\nu_m, \gamma_N) \leq c\delta(m,n)\} = \Pr\{\nu_m(F) \leq \gamma_N(F^{c\delta(m,n)})\}
\]

\[
+ c\delta(m,n) \text{ for all } F \in F(2c\delta(m,n))\} =
\]
\[
\Pr(\mu_m(F) - \gamma_N(F) \leq (\gamma_N(F^{c\delta(m,n)}) - \mu(F^{c\delta(m,n)})) + (\mu(F) - \gamma_N(F)) + \mu(F^{c\delta(m,n)}) \cdot F) + c\delta(m,n)
\]
for all \(F \in F(2c\delta(m,n))\},
\]
\[
= \Pr(\max_{m}^{N}(\gamma_N(F) - \eta_n(F)) \leq q(F)\eta(F) + (\mu(F) - \gamma_N(F)) + c\delta(m,n)(q(F)+1);
\]
for all \(F \in F(2c\delta(m,n))\},
\]
where \(-\frac{1}{N} \leq \eta(F) \leq \frac{1}{N}\),
\[
= \Pr(\max_{m}^{N}(\mu(F) - \eta_n(F)) + \max_{m}^{N}(\gamma_N(F) - \mu(F)) \leq q(F)\eta(F) + (\mu(F) - \gamma_N(F)) + c\delta(m,n)(q(F)+1);
\]
for all \(F \in F(2c\delta(m,n))\},
\]
\[
= \Pr(\eta_n^{-\frac{1}{2}}(\mu(F) - \eta_n(F)) \leq q(F)\eta(F)(2+ \frac{n}{m})\delta^{-1}(m,n) + c(q(F)+1);
\]
for all \(F \in F(2c\delta(m,n))\},
\]
\[
= \Pr(\sup_{F \in F(2c\delta(m,n))} \frac{1}{q(F)+1} \int_{F} dX_n(t) \leq \eta(F)(2+ \frac{n}{m})\delta(m,n)^{-1} + c),
\]
\[
\eta(F)(2+ \frac{n}{m})\delta(m,n)^{-1} = O(n^{\frac{1}{2}}/N) = o(1). \text{ Hence in the case } \frac{m}{n} \to \infty,
\]
\[
\lim_{m,n \to \infty} \Pr(\{\frac{n}{mN})^{-\frac{1}{2}} \cdot d_p(u_m, \gamma_N) \leq c\}
\]
\[
= \lim_{m,n \to \infty} \Pr(d_p(u_m, \gamma_N) \leq c \cdot \frac{n}{m}) = \Pr(f(W^o) \leq c) . \quad \Box
\]

**Proposition 4.1.4:** \((\frac{N}{mn})^{-\frac{1}{2}} \cdot d_p(u_m, \gamma_n) \Rightarrow f(W^o)\) as \(m,n \to \infty\).

**Proof:** We treat separately the cases (i) \(0 < a \leq \frac{m}{n} \leq b < \infty\) and
(ii) \( \frac{m}{n} \to 0 \). The case \( \frac{m}{n} \to \infty \), i.e., \( \frac{n}{m} \to 0 \), follows due to the symmetric roles of \( m \) and \( n \).

Let \( \delta(m,n) = \left( \frac{N}{mn} \right)^{\frac{1}{2}} \).

(i) \[ \Pr\{d_P(\mu_m, \nu_n) \leq c\delta(m,n)\} = \Pr\{\mu_m(F) \leq \nu_n(F^{c\delta(m,n)})\} \]

+ \( c\delta(m,n) \); for all \( F \in \mathcal{F}(2c\delta(m,n)) \)

= \[ \Pr\{\mu_m(F) - \nu_n(F) \leq (\nu_n(F^{c\delta(m,n)}) - F) - \mu(F^{c\delta(m,n)}) - F) \]

+ \( \mu(F^{c\delta(m,n)}) - F) \) + \( c\delta(m,n) \);

for all \( F \in \mathcal{F}(2c\delta(m,n)) \)

= \[ \Pr\{\delta(m,n)^{-1}(\mu_m(F) - \nu_n(F)) \leq (\nu_n(F^{c\delta(m,n)}) - F) \]

- \( \mu(F^{c\delta(m,n)}) - F) \) + \( c(q(F) + 1) \);

for all \( F \in \mathcal{F}(2c\delta(m,n)) \)

= \[ \Pr\left\{ \sup_{F \in \mathcal{F}(2c\delta(m,n))} \frac{1}{q(F) + 1} (Y_{m,n}(F) - Z_{m,n}(F)) \leq c \right\} \]

where

\[ Y_{m,n}(F) = \delta(m,n)^{-1}(\mu_m(F) - \nu_n(F)) \]

and

\[ Z_{m,n}(F) = \delta(m,n)^{-1}(\nu_n(F^{c\delta(m,n)}) - F) - \mu(F^{c\delta(m,n)}) - F) \).

\[ Y_{m,n}(t) = \delta(m,n)^{-1}(\mu_m[0,t] - \nu_n[0,t]) \]

is a random process on \( D[0,1] \)

and \( Y_{m,n} \Rightarrow W^0 \) as \( m,n \to \infty \) where \( 0 < a \leq \frac{m}{n} \leq b < \infty \), by 1.4.2. Hence \( f(c\delta(m,n)) \Rightarrow f(W^0) \). We can thus complete the proof for case (i) by

showing that \( \sup_{F \in \mathcal{F}(2c\delta(m,n))} \frac{1}{q(F) + 1} Z_{m,n}(F) \Rightarrow h(W^0) \), where \( h(y) \equiv 0 \).
for \( y \in D[0,1] \). For each \( y \in D[0,1] \), let

\[
h_{m,n}(y) = \left( \frac{N}{mn} \right)^{-\frac{1}{2}} \sup_{F \in F(2c\delta(m,n))} (q(F)+1)^{-1} \int_{F \cap \delta(m,n)} dy.
\]

It suffices to show that \( h_{m,n}(\nu_n - \mu) \Rightarrow h(W^0) \). Since

\[
|\left( \frac{N}{mn} \right)^{-\frac{1}{2}}(\nu_n - \mu)| \leq (1+b)|n^{\frac{1}{2}}(\nu_n - \mu)|,
\]

and \( n^{\frac{1}{2}}(\nu_n - \mu) \Rightarrow W^0 \) by 1.2.5, it suffices (as in Lemma 4.1.1) to show that for any \( y \in C[0,1] \) and \( \{y_n\} \) in \( D[0,1] \) such that \( y_n \to y \),

\[
\frac{1}{q(F)+1} \int_{F_\delta(m,n)} dy_n \to 0 \text{, for each } F \in F(2c\delta(m,n)).
\]

Let \( m \) and \( n \) be sufficiently large that \( \sup_{0 \leq t \leq 1} |y_n[0,t] - y[0,t]| < \frac{\varepsilon}{3} \)

and

\[
\sup_{s-t \leq c\delta(m,n)} |y(s) - y(t)| < \frac{\varepsilon}{3}.
\]

For any \( F \in F(2c\delta(m,n)) \),

\[
\int_{F_\delta(m,n)} dy \leq q(F) \sup_{s-t < c\delta(m,n)} |y(s) - y(t)| < q(F)\frac{\varepsilon}{3},
\]

\[
\int_{F_\delta(m,n)} dy_n - \int_{F_\delta(m,n)} dy \leq 2q(F) \sup_{0 \leq t \leq 1} |y_n[0,t] - y[0,t]| \leq 2q(F)\frac{\varepsilon}{3},
\]

as in the proof of 4.1.1. Hence

\[
\frac{1}{q(F)+1} \int_{F_\delta(m,n)} dy_n < \varepsilon.
\]
(ii) \[ \Pr\{d_p(u_m, v_n) \leq cm^{-\frac{k}{2}}\} = \Pr\{u_m(F) \leq v_n(Fc^{-\frac{k}{2}})\} \]
\[ + cm^{-\frac{k}{2}}, \text{ for all } F \in F(2cm^{-\frac{k}{2}}) \]
\[ = \Pr\{u_m(F) - u(F) \leq (v_n(Fc^{-\frac{k}{2}}) - u(Fc^{-\frac{k}{2}})) + u(Fc^{-\frac{k}{2}})\} \]
\[ + cm^{-\frac{k}{2}}, \text{ for all } F \in F(2cm^{-\frac{k}{2}}) \]
\[ = \Pr\{X_{m,n}(F) \leq (q(F)+1)c; \text{ for all } F \in F(2cm^{-\frac{k}{2}})\}, \]

where
\[ X_{m,n}(F) = m^{\frac{k}{2}}(u_m(F) - u(F)) \text{ and } Z_{m,n}(F) = m^{\frac{k}{2}}(v_n(Fc^{-\frac{k}{2}}) - u(Fc^{-\frac{k}{2}})). \]

We wish to show that
\[ \sup_{F \in F(2cm^{-\frac{k}{2}})} \frac{1}{q(F)+1} m^{\frac{k}{2}}(v_n(Fc^{-\frac{k}{2}}) - u(Fc^{-\frac{k}{2}})) \]
converges in probability to 0. But
\[ v_n(Fc^{-\frac{k}{2}}) - u(Fc^{-\frac{k}{2}}) \leq q(F) \sup_{0 \leq t \leq 1} (v_n[0,t] - u[0,t]), \]
for any set \( F \in F(2cm^{-\frac{k}{2}}) \). So
\[ \sup_{F \in F} \frac{1}{q(F)+1} m^{\frac{k}{2}}(v_n(Fc^{-\frac{k}{2}}) - u(Fc^{-\frac{k}{2}})) \]
\[ \leq \left(\frac{m}{n}\right)^{\frac{k}{2}} \sup_{0 \leq t \leq 1} (v_n[0,t] - u[0,t]), \]
which converges in probability to 0, since \( \frac{m}{n} \to 0 \). \( \square \)

**Proposition 4.1.5:** For \( p = 1, 2, \ldots \) define \( T_p(W^o) \) as
\[ \sup_{0 \leq t_1 \leq \ldots \leq t_p \leq 1} |W^o(t_1) - W^o(t_2) + \ldots + (-1)^{p-1} W^o(t_p)|. \] For \( c > 0, \)
\[ \{T_1(W^o) \leq 2c, T_2(W^o) \leq 3c, \ldots, T_{p-1}(W^o) \leq pc, T_p(W^o) \leq pc\} \subseteq \{f(W^o) \leq c\} \subseteq \]
\{T_1(W^o) \leq 2c, T_2(W^o) \leq 3c, \ldots, T_{p-1}(W^o) \leq pc, T_p(W^o) \leq (p+1)c \}.

Proof: \{f(W^o) \leq c\} = \{T_1(W^o) \leq 2c, T_2(W^o) \leq 3c, \ldots\} \subseteq T_1(W^o) \leq 2c, T_2(W^o) \leq 3c, \ldots, T_{p-1} \leq pc, T_p \leq (p+1)c\}. Assume that T_1 \leq 2c, T_2 \leq 3c, \ldots, T_{p-1} \leq pc, T_p \leq pc. Then T_p \leq (p+1)c. We wish to show that this implies that T_i \leq (i+1)c for all i; assume for purposes of induction that for some j > p, T_i \leq (i+1)c for i = 1, 2, \ldots, j-1.

By the triangle inequality, T_j \leq T_p + T_{j-p}. Since T_p \leq pc and T_{j-p} \leq (j-p+1)c, it follows that T_j \leq (j+1)c. \)

Using Proposition 1.4.3, we can evaluate \Pr(T_1(W^o) \leq 2c, T_2(W^o) \leq 2c) = \Pr(T_2(W^o) \leq 2c), and \Pr(T_1(W^o) \leq 2c, T_2(W^o) \leq 3c) directly to give rough upper and lower bounds on \Pr(f(W^o) \leq c). The joint distribution of (T_1, T_2, T_3) seems intractable.

4.2 Asymptotic distribution for d_L.

Proposition 4.2.1:

a) \(n^{-\frac{k}{2}}d_L(\mu_n, \mu) \Rightarrow \sup_{0 \leq t \leq 1} |\frac{1}{2}W^o(t)|.\)

b) \(\left(\frac{n}{mn}\right)^{-\frac{k}{2}}d_L(\mu_m, \gamma_N) \Rightarrow \sup_{0 \leq t \leq 1} |\frac{1}{2}W^o(t)|.\)

c) \(\left(\frac{N}{mn}\right)^{-\frac{k}{2}}d_L(\mu_m, \nu_n) \Rightarrow \sup_{0 \leq t \leq 1} |\frac{1}{2}W^o(t)|.\)

Proof: We shall prove only (a), since the proofs are completely analogous to 4.1.3, 4.1.4, and 4.1.5, respectively.

a) \(\Pr(d_L(\mu_n, \mu) \leq cn^{-\frac{k}{2}}) = \)
\[
\Pr\{\mu(-\omega, x - cn^{1/k}) - cn^{-1/k} \leq \mu_n(x) \leq \mu(-\omega, x + cn^{1/k}) + cn^{-1/k};
\]

for all \( x \in [0, 1] \}

= \Pr\{-2cn^{-1/k} \leq \mu_n(x) - \mu(x) \leq 2cn^{-1/k}; \text{ for all } x \in [0, 1]\}

= \Pr\{ \sup_{0 \leq t \leq 1} n^{1/k} |\mu_n(x) - \mu(x)| \leq 2c \} \rightarrow \Pr\{ \sup_{0 \leq t \leq 1} |W^o(t)| \leq 2c \},

by 1.4.2.

Thus \( d_{\mu, \nu}(\mu_n, \mu) \) has the same asymptotic distribution as \( 2d_{\mu, \nu}(\mu_n, \mu) \).

4.3 Absolute continuity of \( f(W^o) \).

Definition 4.3.1: For \( k \geq 1 \) and \( i=1,\ldots,k \), define

\[
U_i = U_i^{(k)} = \sup_{\frac{i-1}{k} \leq t \leq \frac{i}{k}} W^o(t); \quad V_i = V_i^{(k)} = \inf_{\frac{i-1}{k} \leq t \leq \frac{i}{k}} W^o(t).
\]

For \( p \leq k \), define

\[
T_p^{(k)}(W^o) = \max \left( \max_{1 \leq i_1 < \ldots < i_p \leq k} U_{i_1} - V_{i_1} + U_{i_2} - \ldots + (-1)^{p+1} U_{i_p} \right) \],

max \left( \max_{1 \leq i_1 < \ldots < i_p \leq k} V_{i_1} - U_{i_1} + \ldots + (-1)^{p+1} V_{i_p} \right) \text{ if } p \text{ is odd,}

\[
T_p^{(k)}(W^o) = \max \left( \max_{1 \leq i_1 < \ldots < i_p \leq k} U_{i_1} - V_{i_1} + U_{i_2} - \ldots + (-1)^{p+1} U_{i_p} \right) \],

max \left( \max_{1 \leq i_1 < \ldots < i_p \leq k} V_{i_1} - U_{i_1} + \ldots + (-1)^{p+1} V_{i_p} \right) \text{ if } p \text{ is even.}
\]

Remark 4.3.2: For a continuous sample path \( W^o(t), t \in [0, 1] \), it is apparent that \( T_p^{(k)}(W^o) \) may be defined in the same way as \( T_p(W^o) \) with
the restriction that no two of the points \( t_1, t_2, \ldots, t_p \) (in the definition) are in the interval \([\frac{j}{k}, \frac{j+1}{k}]\), for any \( j \).

Definition 4.3.3: For \( c > 0 \), let \( A(k,c) \) denote the event that

\[
\max_{1 \leq i \leq k} \left( \sup_{\frac{i-1}{k} \leq s \leq \frac{i}{k}} |W^0(s) - W^0(t)| \right) \leq c. \]

We will establish the absolute continuity (with respect to Lebesgue measure) of \( f(W^0) \) by showing the following.

(I) \( T_p^{(k)}(W^0) \) is absolutely continuous for \( 1 \leq p \leq k \).

(II) For \( p \geq 2 \) if \( T_p(W^0) - T_{p-2}(W^0) > c \) and \( A(k,c) \) occurs then

\[
T_p^{(k)}(W^0) = T_p(W^0), \text{ where } T_0 = 0. \text{ Of course, } T_1^{(k)} = T_1.
\]

(III) As \( c \downarrow 0 \), \( \Pr(T_p(W^0) - T_{p-2}(W^0) > c) \uparrow 1 \). For fixed \( c \), as \( k \to \infty \), \( \Pr(A(k,c)) \to 1 \).

(Note: It is not hard to show that \( T_p^{(k)}(W^0) \Rightarrow T_p(W^0) \) as \( k \to \infty \), in particular that \( |T_p(W^0) - T_p^{(k)}(W^0)| \) is arbitrarily small with probability arbitrarily close to 1, for \( k \) sufficiently large. However, even when combined with (I), this weak convergence does not imply the absolute continuity of the distribution of \( T_p(W^0) \).

(II) and (III) together imply

(IV) Given \( \varepsilon > 0 \), for \( k \) sufficiently large, \( \Pr(T_p^{(k)} = T_p) > 1 - \varepsilon \).

(Indeed, choose \( c \) first, then choose \( k \).)

Proposition 4.3.4: (I) and (IV) together imply that \( f(W^0) \) is absolutely continuous.

Proof: We need the following lemmas:
Lemma 4.3.5: Let $X$ be a random variable on some probability space. Let $X_1, X_2, \ldots$ be absolutely continuous random variables such that $\Pr\{X_k \neq X\} \to 0$ as $k \to \infty$. Then $X$ is absolutely continuous.

Proof: Let $A$ have Lebesgue measure zero. Then $\Pr\{X_k \in A\} = 0$, for all $k$. Choose $\varepsilon > 0$. $\Pr\{X \in A\} \leq \Pr\{X_k \in A\} + \Pr\{X_k \neq X\} < \varepsilon$, for $k$ sufficiently large. Since $\varepsilon$ is arbitrary, $\Pr\{X \in A\} = 0$.

It follows that if (I) and (IV) hold, then $T_p(W^o)$ is absolutely continuous for each $p$.

Lemma 4.3.6: Let $X_1, X_2, \ldots$ be a sequence of random variables on some probability space and $N$ a positive integer-valued random variable. If each random variable $X_i$ is absolutely continuous, then $X_N$ is absolutely continuous.

Proof: For any Lebesgue-measurable set $A$,

$$\Pr\{X_N \in A\} = \sum_{n=1}^{\infty} \Pr\{X_n \in A, N=n\}.$$ 

But if $A$ has Lebesgue measure 0, $\Pr\{X_n \in A, N=n\} = 0$ for each $n$.

Lemma 4.3.7: $(p+1)^{-1} T_p(W^o) \to 0$ a.s. as $p \to \infty$.

Before proving 4.3.7, we show that it will imply that $f(W^o)$ is absolutely continuous, assuming (I) and (IV) are true. Indeed, by 4.3.7, with probability one, $f(W^o) = \sup_{1 \leq p < \infty} (p+1)^{-1} T_p(W^o) = (n+1)^{-1} T_n(W^o)$ for some $n$, depending on $W^o$. 

Let \( N = N(W^o) \) be the least such value of \( n \) (define \( N \) in any manner on the probability zero set where there may not be such a value of \( n \)). Then by 4.3.5 and 4.3.6, \( f(W^o) = (N+1)^{-1} T_N(W^o) \) is absolutely continuous. We complete the proof of 4.3.4 by proving 4.3.7.

**Proof of 4.3.7:** Let \( M_k(W^o) = \sup\limits_{0 \leq s-t \leq 1/k} |W^o(s)-W^o(t)|, 0 \leq s, t \leq 1. \) Since \( W^o(t) \) is almost surely a uniformly continuous function on \([0,1]\), \( M_k(W^o) \to 0 \) a.s.

Let \( 0 \leq t_1 < t_2 \ldots < t_n \leq 1 \). The \( n \) points \( t_1, \ldots, t_n \) can be partitioned into pairs and singletons such that for each pair, \((t_i, t_{i+1})\), both points \( t_i \) and \( t_{i+1} \) are in one of the intervals \([\frac{1}{k}, \frac{j+1}{k}]\), \( j=0, \ldots, k-2 \), or \([\frac{k-1}{k}, 1]\), and such that there are at most \( k \) singletons. It easily follows that for any \( n \geq k \)

\[ T_n \leq T_k + \frac{n-k}{2} M_k(W^o) . \]

Hence \( (n+1)^{-1} T_n \leq \frac{k}{n+1} T_1(W^o) + \frac{1}{4} M_k(W^o) \). Since \( T_1(W^o) \) is a.s. finite and \( M_k(W^o) \to 0 \) a.s., it follows that \( (n+1)^{-1} T_n \to 0 \) a.s. as \( n \to \infty \).

We will prove (II) first and then (I) and (III).

**Proposition 4.3.8:** If \( T_p(W^o) - T_{p-2}(W^o) > c \) and \( A(k,c) \) occurs, then \( T_p^{(k)}(W^o) = T_p(W^o) \) a.s.

**Proof:** Since \( W^o(t) \) is a.s. a uniformly continuous function on \([0,1]\), assume that \( W^o(t) \) is uniformly continuous. Then we may choose points \( 0 \leq t_1 \leq \ldots \leq t_p \leq 1 \) such that

\[ T_p(W^o) = (-1)^S(W^o(t_1) - W^o(t_2) + \ldots + (-1)^{P+1} W^o(t_p)) \text{ for } s=0 \text{ or } 1. \]
Since by hypothesis, \( T_p(W^o) - T_{p-2}(W^o) > c \), we must have \( t_1 < \ldots < t_p \). From the triangle inequality applied to the expression for \( T_p(W^o) \), it follows that \( T_p(W^o) \leq T_{p-2}(W^o) + |W^o(t_{i_{-i_{1}}}) - W^o(t_{i+1})| \) for \( 1 \leq i \leq p-1 \). Thus if \( T_p(W^o) - T_{p-2}(W^o) > c \), then

\[
\min_{1 \leq i \leq p-1} |W^o(t_{i_{1}}) - W^o(t_{i+1})| > c.
\]

If in addition \( A(k,c) \) occurs, this implies that for no values of \( j \) and \( i \) are \( \frac{j-1}{k} \leq t_{i_{1}} < t_{i+1} \leq \frac{j}{k} \). Then by remark 4.3.2, \( T_p^{(k)}(W^o) = T_p(W^o) \).

Proposition 4.3.9: \( T_p^{(k)}(W^o) \) is absolutely continuous for each \( p \leq k \).

Proof: Since the maximum of finitely many absolutely continuous random variables is itself absolutely continuous, it suffices to fix the values \( 1 \leq i_1 < \ldots < i_p \leq k \) and (if \( p \) is even) show that \( U_{i_1} - V_{i_2} + U_{i_3} - \ldots \) \(+(-1)^{p+1}V_{i_p}\) is absolutely continuous. (It will be seen that the same proof applies for \( V_{i_1} - U_{i_2} + V_{i_3} - \ldots +(-1)^{p+1}U_{i_p} \), and for both cases when \( p \) is odd.)

Let \( 1 \leq i_1 < \ldots < i_p \leq k \) be fixed for the remainder of the proof.

We wish to consider the conditional joint distribution of \( U_{i_1}, V_{i_2}, U_{i_3}, \ldots, V_{i_p} \), given \( W^o(\frac{i_{1_{-1}}}{k}), W^o(\frac{i_1}{k}), W^o(\frac{i_{2_{-1}}}{k}), W^o(\frac{i_2}{k}), \ldots, W^o(\frac{i_{p_{-1}}}{k}), W^o(\frac{i_p}{k}) \). The values \( \frac{i_{j_{1}}}{k} \) and \( \frac{i_{j+1_{-1}}}{k} \) might not be distinct.

Let \( \ell_1, \ldots, \ell_r \) be the points in order of the set

\[ \{ \ell := \frac{i_j}{k} = \ell \text{ or } \frac{i_{j_{-1}}}{k} = \ell, j=1,\ldots,p \} \], where \( p \leq r \leq 2p \).

\( Y(W^o) = (W^o(\ell_1), \ldots, W^o(\ell_r)) \) has a (non-singular) \( r \)-variate normal distribution with density \( \phi(X_1, \ldots, X_r) \). For any Lebesgue-measurable set \( A \),
\[
\Pr \left( \sum_{1 \leq i \leq 3} (-1)^{i} U_{i,1} + U_{1,2} + U_{1,3} \in A \right) \\
= \int_{V} \Pr \left( \sum_{1 \leq i \leq 3} (-1)^{i} U_{i,1} + U_{1,2} + U_{1,3} \in A \right) \\
\quad / W^{\circ}(z_{1}) = b_{1}, \ldots, W^{\circ}(z_{r}) = b_{r} \} \phi(b_{1}, \ldots, b_{r}) \\
\quad \, db_{1}, \ldots, db_{r}.
\]

We complete the proof by showing that if \( A \) has Lebesgue measure 0, \( \Pr \left( \sum_{1 \leq i \leq 3} (-1)^{i} U_{i,1} + U_{1,2} + U_{1,3} \in A \right) / W^{\circ}(z_{1}) = b_{1}, \ldots, W^{\circ}(z_{r}) = b_{r} \} = 0 \)

for each \( b_{1}, \ldots, b_{r} \).

**Lemma 4.3.10:** Let \( 0 \leq c_{1} < c_{2} \leq 1 \) and let \( Z(t) = \\
W^{\circ}(c_{1}) + \frac{t-c_{1}}{c_{2}-c_{1}} [W^{\circ}(c_{2}) - W^{\circ}(c_{1})] = -\frac{t-c_{1}}{c_{2}-c_{1}} W^{\circ}(c_{2}) + \frac{t-c_{2}}{c_{2}-t} W^{\circ}(c_{1}) \).

Then for \( c_{1} < t < c_{2} \), \( W^{\circ}(t) - Z(t) \) is a normal random variable and is independent of \( W^{\circ}(s) \) for each \( s \in [0, c_{1}] \cup [c_{2}, 1] \).

(Note: this result bears out the intuitive notion that in the plane, the point \( (t, E(W^{\circ}(t)/W^{\circ}(c_{1}), W^{\circ}(c_{2}))) \), for \( c_{1} < t < c_{2} \), lies on the line segment joining the points \( (c_{1}, W^{\circ}(c_{1})) \) and \( (c_{2}, W^{\circ}(c_{2})) \).)

**Proof:**

For \( s \in [0, c_{1}] \) and \( t \in (c_{1}, c_{2}) \), \( E W^{\circ}(s) = 0 = E[W^{\circ}(t) - Z(t)] \), and it is easy to check (using the relationship \( E W^{\circ}(t_{1}) W^{\circ}(t_{2}) = t_{1}(1-t_{2}) \) for \( 0 \leq t_{1} \leq t_{2} \leq 1 \)) that \( E W^{\circ}(s)(W^{\circ}(t) - Z(t)) = 0 \). \( W^{\circ}(t) - Z(t) \) is a linear combination of normal random variables and therefore normal, so \( W^{\circ}(t) - Z(t) \) is independent of \( W^{\circ}(s) \). Similarly for \( s \in [c_{1}, 1] \). \( \square \)
It follows that given $W^o(\hat{Z}_1), \ldots, W^o(\hat{Z}_r), U_{i_1}, V_{i_2}, U_{i_3}, \ldots, V_{i_p}$ are mutually independent. If we can show that the conditional distribution of $U_{i_1}$ is absolutely continuous, the absolute continuity of $T^{(k)}_{(p)}(W^o)$ will follow.

**Lemma 4.3.11:** $\Pr\{ \sup_{c_1 \leq t \leq c_2} W^o(t) \leq \lambda / W^o(c_1) = b_1, W^o(c_2) = b_2 \} = 1 - e^{-2\lambda' \left( \frac{\lambda' - b}{c_2^2} \right)}$, where $b = b_2 - b_1$, $\lambda' = \lambda - b_1$, $c = (c_2 - c_1)^{\frac{1}{2}}$.

**Proof:** Conditional on the event $\{ W^o(c_1) = b_1 \text{ and } W^o(c_2) = b_2 \}$, for $c_1 < t < c_2$, $W^o(t)$ is the sum of the normal random variable $W^o(t) - Z(t)$ and a constant $\frac{c_2 - t}{c_2 - c_1} b_2 - \frac{t - c_1}{c_2 - c_1} b_1$.

It is easy to compute that for $c_1 \leq s \leq t \leq c_2$ $E(W^o(s) - Z(s))(W^o(t) - Z(t)) = \frac{(s - c_1)(c_2 - t)}{(c_2 - c_1)}$. Making the change of variables $s' = \frac{s - c_1}{c_2 - c_1}$ and $t' = \frac{t - c_1}{c_2 - c_1}$, $0 \leq s' \leq t' \leq 1$, and writing $X(s') = W^o((c_2 - c_1)s' + c_1) - Z((c_2 - c_1)s' + c_1)$, we have $EX(s')X(t') = (c_2 - c_1)s'(1 - t')$, for $0 \leq s' \leq t' \leq 1$. Hence $X(t')$, $0 \leq t' \leq 1$ in $(c_2 - c_1)^{\frac{1}{2}}$ times a "Brownian Bridge" random process as defined in 1.4.1.

$\sup_{c_1 \leq t \leq c_2} W^o(t) - Z(t) = \sup_{0 \leq t' \leq 1} X(t')$.

Hence

$\Pr\{ \sup_{c_1 \leq t \leq c_2} W^o(t) \leq \lambda / W^o(c_1) = b_1, W^o(c_2) = b_2 \}$
\[ = \Pr\{ \sup_{0 \leq t \leq 1} (X(t') + Z((c_2-c_1)t'+c_1)) \leq \lambda / W'(c_1) = b_1, W'(c_2) = b_2 \} \]

\[ = \Pr\{ \sup_{0 \leq t' \leq 1} (X(t') + (b_2-b_1)t') + b_1 \leq \lambda \} \]

\[ = \Pr\{ \sup_{0 \leq t' \leq 1} (cW'(t') + t'b) \leq \lambda' \}, \text{ where } c = (c_2-c_1)^{\frac{1}{2}} \]

\[ \lambda' = \lambda - b_1 \]

\[ = \Pr\{ \sup_{0 \leq t < \infty} (cW'(\frac{t}{t+1}) + \frac{t}{t+1} b) \leq \lambda' \} \]

\[ = \Pr\{ \sup_{0 \leq t < \infty} (W(t) - \frac{\lambda' - b}{c} t - \frac{\lambda'}{c} \leq 0 \}, \text{ where } W \text{ is the Wiener process} \]

(see 1.4.1). By a result in [10], this last probability is equal to

\[ 1 - e^{-\frac{\lambda'}{c} \cdot \left( \frac{\lambda' - b}{c} \right)} . \]

We next prove III, in two parts.

**Proposition 4.3.12.** For each \( c > 0 \), \( \Pr(A(k,c)) \to 1 \) as \( k \to \infty \), where \( A(k,c) \) is as defined in 4.3.3.

**Proof:** We will consider \( W'(t) = W(t) - tW(1), 0 \leq t \leq 1 \). Then

\[ Z_i^k = \sup_{\frac{i-1}{k} < s \leq \frac{i}{k}} |W'(s) - W'(t)| \]

\[ = \sup_{\frac{i-1}{k} < s \leq \frac{i}{k}} (W'(s) - W'(\frac{i-1}{k})) - \inf_{\frac{i-1}{k} < s \leq \frac{i}{k}} (W'(t) - W'(\frac{i-1}{k})) \]

\[ = \sup_{\frac{i-1}{k} < s \leq \frac{i}{k}} [(W(s) - W(\frac{i-1}{k})) + (s - \frac{i-1}{k})W(1)] \]

\[ - \inf_{\frac{i-1}{k} < s \leq \frac{i}{k}} [(W(t) - W(\frac{i-1}{k})) + (t - \frac{i-1}{k})W(1)] . \]
If \( Z_i = \frac{1}{k} \sup_{\frac{i-1}{k} \leq s \leq \frac{i}{k}} |W(s) - W(t)| \), then it is clear that
\[
|Z_i - Z_0| \leq \frac{2}{k} W(1) \rightarrow 0 \quad \text{as} \quad k \rightarrow \infty.
\]
Hence it will suffice to show that \( \Pr\{ \max_{1 \leq i \leq k} Z_i \leq c\} \rightarrow 1 \quad \text{as} \quad k \rightarrow \infty. \)

By the reflection principle, \( \sup_{\frac{i-1}{k} \leq s \leq \frac{i}{k}} (W(s) - W(\frac{i-1}{k})) \) and \( \inf_{\frac{i-1}{k} \leq t \leq \frac{i}{k}} (W(t) - W(\frac{i-1}{k})) \) are each distributed as twice a normal random variable with mean zero and variance \( \frac{1}{k} \). Then
\[
Z_i \leq 2 \left| \sup_{\frac{i-1}{k} \leq s \leq \frac{i}{k}} (W(s) - W(\frac{i-1}{k})) \right| + 2 \left| \inf_{\frac{i-1}{k} \leq t \leq \frac{i}{k}} (W(t) - W(\frac{i-1}{k})) \right|
\]
is stochastically smaller than 8 times the absolute value of a normal random variable with mean zero and variance \( \frac{1}{k} \). Since \( Z_1, \ldots, Z_k \) are mutually independent,
\[
\Pr\{ \max_{1 \leq i \leq k} Z_i \leq c\} = (\Pr\{Z_i \leq c\})^k \geq \left( \Phi(\frac{\sqrt{c}}{\sqrt{8}}) \right)^k.
\]
We claim that \( (\Phi(\alpha k^{\frac{1}{2}}))^k \rightarrow 1 \), which will complete the proof. Indeed, it is well known that
\[
\Phi(\alpha k^{\frac{1}{2}}) > 1 - \frac{1}{(2\pi)^{\frac{1}{2}}} e^{-\frac{1}{2} \left( \frac{1}{\alpha k^{\frac{1}{2}}} \right)^2},
\]
so
\[
(\Phi(\alpha k^{\frac{1}{2}}))^k > 1 - \frac{k^{\frac{1}{2}}}{(2\pi)^{\frac{1}{2}}} e^{-\frac{1}{2} \frac{2k^{\frac{1}{2}}}{\alpha}}.
\]
for \( k \) sufficiently large. Hence \( (\Phi(\alpha k^{\frac{1}{2}}))^k \rightarrow 1 \) as \( k \rightarrow \infty. \)
Proposition 4.3.13. For $p \geq 2$, $\Pr(T_p(W^o) - T_{p-1}(W^o) > c) \to 1$ as $c \to 0$.

Proof: Let $A(K) = \{ \max_{1 \leq i \leq 2p} |W^o(\frac{i-1}{2p}) - W^o(\frac{i}{2p})| \leq K \}$. For fixed $\epsilon > 0$, choose $K$ sufficiently large that $\Pr(A) > 1 - \frac{\epsilon}{2}$.

Let $0 \leq t_1 \leq \ldots \leq t_{p-2} \leq 1$ be chosen so that for $s = 0$ or $s = 1, T_{p-2}(W^o) = (-1)^S(W^o(t_1) - W^o(t_2) + \ldots + (-1)^{p-1}W^o(t_{p-2}))$; this can be done whenever $W^o(t)$ is a continuous function of $t$, i.e., almost surely.

For some $i'$ and $j'$, $1 \leq i' \leq p-1$ and $1 \leq j' \leq 2p$, $t_{i'-1} \leq \frac{j'-1}{2p} < \frac{j'}{2p} \leq t_i'$, where we define $t_0 = 0, t_{p-1} = 1$. Indeed, if $p \geq 2$, for some $i = 1, \ldots, p-1, t_i - t_{i-1} \geq \frac{1}{p-1}$.

We wish to show that when $A$ occurs, then for $c$ sufficiently small, with probability greater than $1 - \frac{\epsilon}{2}$ each interval $[\frac{j-1}{2p}, \frac{j}{2p}]$, $j = 1, \ldots, 2p$, contains points $u$ and $v$ such that $u < v$ and $W^o(u) - W^o(v) > c$ and points $u'$ and $v'$ such that $u' < v'$ and $W^o(v') - W^o(u') > c$. Since such points therefore can be found in the interval $[\frac{j'-1}{2p}, \frac{j'}{2p}]$ defined in the last paragraph, then either $(-1)^S[(-1)^{i'}(W^o(u) - W^o(v))] > c$ or $(-1)^S[(-1)^{i'}(W^o(u') - W^o(v'))] > c$, so $T_p(W^o) - T_{p-2}(W^o) > c$.

Assume for definiteness that $W^o(\frac{j-1}{2p}) = b_1$ and $W^o(\frac{j}{2p}) = b_2$ with $b_1 \leq b_2$. As in 4.3.11, for $\lambda \geq b_2$, $\Pr(\sup_{\frac{j-1}{2p} \leq t \leq \frac{j}{2p}} W^o(t) \leq \lambda/W^o(\frac{j-1}{2p}) = b_1, W^o(\frac{j}{2p}) = b_2) = 1 - e^{-4p(\lambda - b_2)(\lambda - b_1)}$. 

Under the assumptions $W^o\left(\frac{j'_{i-1}}{2p}\right) = b_1 \leq b_2 = W^o\left(\frac{j'_i}{2p}\right)$, if
\[
\sup_{\frac{j'_{i-1}}{2p} \leq s \leq \frac{j'_i}{2p}} W^o(t) > b_2 + c,
\]
then points $u$ and $v$, $u'$ and $v'$ as described may be found; indeed in the interval $[\frac{j'_{i-1}}{2p}, \frac{j'_i}{2p}]$, the process starts at $b_1$, goes up to $b_2 + c$ and back down to $b_2$. Then
\[
\Pr\left\{ \sup_{\frac{j'_{i-1}}{2p} \leq s \leq \frac{j'_i}{2p}} W^o(t) \leq b_2 + c/W^o\left(\frac{j'_{i-1}}{2p}\right) = b_1, W^o\left(\frac{j'_i}{2p}\right) = b_2 \right\} = 1 - e^{-4p c(c+b_2-b_1)} \leq 1 - e^{-4p c(c+K)},
\]
under the assumption that $A(K)$ occurs. If $c$ is chosen sufficiently small, this value may be made smaller than $\frac{\varepsilon}{2}$. \[\square\]

We have now proved (I), (II), and (III), showing by 4.3.4 that $f(W^o)$ is absolutely continuous.
CHAPTER 5

Statistical Properties

5.1 Related test statistics.

From among the metrics which generate the topology of weak convergence on $P(S)$, we have elected to examine $d_P$ and $d_L$ because they are the ones which have appeared most often in the literature. As an alternative, we may consider, for any $b > 0$, $d_{P}^{(b)}(P,Q) = \inf\{\varepsilon > 0 : P(F) \leq Q(F^c) + b\varepsilon ; \text{ all closed } F \subseteq S\}$ or $d_{L}^{(b)}(P,Q) = \inf\{\varepsilon > 0 : P(-\infty,x-\varepsilon] - b\varepsilon \leq Q(-\infty,x] \leq P(-\infty,x+\varepsilon] + b\varepsilon ; \text{ for all } x \in E^k\}$, which are topologically equivalent to $d_P$ and $d_L$ and have respectively the same uniformities. (Indeed, if $b > 1$, $d_{P}^{(b)}(P,Q)/b \leq d_{P}(P,Q) \leq bd_{P}^{(b)}(P,Q)$, and $d_{L}^{(b)}(P,Q)/b \leq d_{L}(P,Q) \leq bd_{L}^{(b)}(P,Q) ;$ if $b < 1$, reverse the inequalities.)

The two-sample test (for equal sample sizes) based on $d_{P}^{(2)}(\mu_n,\nu_n)$ seems to have advantages over that based on $d_P(\mu_n,\nu_n)$ because of differences in the null distribution for finite sample sizes. $Pr\{d_L(\mu_n,\nu_n) \leq c\}$ is constant in the intervals $[i/2n, (i+1)/2n)$, $1 \leq i \leq 2n$, because $\mu_n(F) - \nu_n(F^c)$ is always a multiple of $1/n$, and the points at which the atoms of $\mu_n$ and $\nu_n$ are located are separated by multiples of $1/2n$. Thus, intuitively, there is more "going on" at the points $2/2n, 4/2n, \ldots, 1$ than at the points $1/2n, 3/2n, \ldots, (2n-1)/2n.$
For $d_P^{(2)}(\mu_n, \nu_n)$ the interval between atoms is the same as the common divisor of the possible values of the test statistic and so there is no such distinction between points of the form $2i/2n$ and those of the form $(2i+1)/2n$. This difference between $d_P$ and $d_P^{(2)}$ can be seen in the tables in the appendix. The result seems to be a faster, or at least more uniform convergence of $d_P^{(2)}(\mu_n, \nu_n)$ to its asymptotic distribution. A further difference is that given our restriction on the values of $z$, as described in Chapter 3, we can construct more extensive tables of the distribution of $d_P^{(2)}$, using the same methods as for $d_P$. Indeed, in place of 3.1.1, we have

$$\Pr(d_P^{(2)}(\mu_n, \nu_n) \leq c) = \left(\frac{2n}{n}\right)^{-1} \sum_{0 \leq p \leq z} A(p, n, z).$$

The rate of convergence to the asymptotic distribution for the one-sample $d_P$ test statistic seems roughly on the order of that of the Kolmogorov test judging from empirical distribution functions based on actual randomly-generated samples. Our exact distributions show that the convergence for the two-sample $d_P$ test is slower than for the two-sample Kolmogorov-Smirnov test. The proof of Proposition 4.1.1 may give some explanation for this, but we have not been able to give a precise explanation.

The Monte Carlo one-sample distribution functions indicate that the upper bound given by $\Pr(T_1(W^\circ) \leq 2c, T_2(W^\circ) \leq 3c)$ is fairly close to the true asymptotic distribution and the lower bound, $\Pr(T_2(W^\circ) \leq 2c)$ is not as good.

In [26], Rosenblatt suggests a test based on the metric

$$d_2(P, Q) = \sup_{I \text{ an interval}} |P(I) - Q(I)|,$$

stating that it has somewhat better
power than \( d_K \) against certain alternatives. The asymptotic distribution for this test statistic, multiplied by \( n^{\frac{1}{2}} \), is \( \Pr(T_2(W^o) \leq c) \). As a generalization of \( d_2 \), we suggest:

\[
(1) \quad d_k(P,Q) = \sup_{F \in \bigcup_{i=1}^k F_i} |P(F) - Q(F)|, \quad \text{the supremum being taken over all}
\]
\[F \in \bigcup_{i=1}^k F_i, \text{ where } F_i = \{F \in F: q(F) = i\}, \quad q(F) \text{ being defined as in Chapter 4, and}
\]
\[
(2) \quad d^*_P(P,Q) = \inf\{\epsilon > 0: \text{ for } k=1,2,\ldots, |P(F) - Q(F)| \leq (k+1)\epsilon \text{ for all } F \in F\}.
\]

\( d^*_P \) is thus the same as \( d_P \) except that the "end effects" for each interval are ignored; \( d^*_P \) and \( d_P \) put a "penalty" of a factor of \((k+1)\) on the number of component intervals (more precisely on \( q(F) \)) allowed for the set \( F \) under consideration. Some such penalty is necessary if the metric is to take into account all \( F \in F \); indeed,

\[
\Pr(T_k(W^o) \leq c) \to 0 \text{ as } k^{\infty} \text{ (because } W^o(t) \text{ a.s. has infinite total variation). Intuition does not indicate why the factor should be } k+1. \text{ However, a factor of } k \text{ would not be appropriate, since by the triangle inequality, } |P(F) - Q(F)| \leq \epsilon \text{ for all } F \in F_1 \text{ implies that for } k=1,2,\ldots, |P(F) - Q(F)| \leq k\epsilon \text{ for all } F \in F_k.
\]

It should be noted that \( d^*_P \) does not generate the same topology on \( P(S) \) as \( d_P \) and \( d_L \). However, we would expect that the power properties of \( d_P \) (relative to \( d_K \)) which will be described in the next section, would be shared by \( d^*_P \).

5.2 Power of the test based on \( d_P \).

We first show that the tests based on \( d_P \) and \( d_L \) are consistent.

Let \( U \) be the uniform measure on \([0,1] \). Let \( V \in P([0,1]) \), \( V \neq U \), and
let \( V \) be the empirical measure of a sample of size \( n \) from \( V \). We test the hypothesis that \( V = U \) by using one of the test statistics \( d_L(V_n, U) \) or \( d_P(V_n, U) \). By Proposition 1.3.5, \( d_P(V_n, U) + d_P(V, U) > 0 \) a.s. and \( d_L(V_n, U) + d_L(V, U) \) a.s. Since, if \( F \) is a continuous distribution function, \( \{F(X_i)\} \) will be a sample from \( U \) if and only if the sample \( \{X_i\} \) actually comes from \( F \), our one-sample test is consistent against all alternatives to a continuous null distribution.

To show consistency for the two-sample test, we use the following result related to Proposition 1.3.12.

**Proposition 5.2.1**: Let \( P_n \) put mass \( 1/n \) on \( n \) of the points \( \{1/2n, 2/2n, \ldots, 1\} \) and let \( Q_n \) put mass \( 1/n \) on the other \( n \) points. Then \( d_P(P_n, Q_n) \geq d_K(P_n, Q_n) / 3 - 1/n \).

**Proof**: If for some \( x \in [0,1] \), \( |P_n([0,x]) - Q_n([0,x])| \geq 3\varepsilon > 0 \), then either

\[
P_n([0,x]) \geq Q_n([0,x]^{\varepsilon}) + \varepsilon - 1/n, \quad \text{or} \quad P_n([0,x]) \geq Q_n([x,1]^{\varepsilon}) + \varepsilon - 1/n.
\]

Indeed, \( Q_n([x,x+\varepsilon]) \leq (2n\varepsilon + 1)/n \), and \( Q_n([x-\varepsilon,x]) \leq (2n\varepsilon + 1)/n \). Hence

\[d_P(P_n, Q_n) \geq \varepsilon - 1/n.\]

(By using a more laborious argument along the lines of 4.1.5, \( d_K(P_n, Q_n)/3 \) can be replaced by \( d_K(P_n, Q_n)/2 \).) The consistency of the \( d_P \) and \( d_L \) two-sample tests now follows from the consistency of the Kolmogorov-Smirnov test.

Massey [20] has given an example to show that the Kolmogorov-Smirnov test is not unbiased. Similarly, we can show that the one-sample test based on \( d_P \) is not unbiased. Suppose we take a sample \( \{X_i\} \) of
size \( n \) from the probability measure \( \nu \) which equals Lebesgue measure on \((c_{\alpha} n^{-\frac{1}{\alpha}}, 1 - c_{\alpha} n^{-\frac{1}{\alpha}})\), where \( c_{\alpha} \) is the critical value for the size \( \alpha \) test, and which places mass \( c_{\alpha}^{-\frac{1}{\alpha}} \) at the points \( c_{\alpha}^{-\frac{1}{\alpha}} \) and \( 1 - c_{\alpha}^{-\frac{1}{\alpha}} \).

(Assume that \( n \) is sufficiently large that \( c_{\alpha}^{-\frac{1}{\alpha}} < 1 - c_{\alpha}^{-\frac{1}{\alpha}} \).) We may generate such a sample \( \{X_i\} \) by "trimming" a sample \( \{X'_i\} \) from \( U[0,1] \), i.e., by letting \( X_i = X'_i \) if \( X'_i \in \left(c_{\alpha}^{-\frac{1}{\alpha}}, 1 - c_{\alpha}^{-\frac{1}{\alpha}}\right)\), and letting \( X_i = c_{\alpha}^{-\frac{1}{\alpha}} \) if \( X'_i \leq c_{\alpha}^{-\frac{1}{\alpha}} \), and \( X_i = 1 - c_{\alpha}^{-\frac{1}{\alpha}} \) if \( X'_i \geq 1 - c_{\alpha}^{-\frac{1}{\alpha}} \). Let \( U_n^* \) be the empirical measure of such a trimmed sample from \( U \) and let \( U_n \) the the empirical measure of the original sample from \( U \).

For any sample \( \{X'_i\} \) from \( U \), \( d_P(U_n^*, U) \geq c_{\alpha}^{-\frac{1}{\alpha}} \) implies that \( d_P(U_n, U) \geq c_{\alpha}^{-\frac{1}{\alpha}} \). Indeed, for any closed set \( F \subseteq [0,1] \),

\[
U_n^*(F \cap \left[c_{\alpha}^{-\frac{1}{\alpha}}, \infty\right)) \geq U_n(F \cap \left[c_{\alpha}^{-\frac{1}{\alpha}}, \infty\right)),
\]

so

\[
\sup_{\text{closed } F \subseteq [0,1]} U(F) - U_n(F \cap \left[c_{\alpha}^{-\frac{1}{\alpha}}, \infty\right)) \geq \sup_{\text{closed } F \subseteq [0,1]} U(F) - U_n^*(F \cap \left[c_{\alpha}^{-\frac{1}{\alpha}}, \infty\right)).
\]

With positive probability, \( d_P(U_n, U) > c_{\alpha}^{-\frac{1}{\alpha}} \) but \( d_P(U_n^*, U) < c_{\alpha}^{-\frac{1}{\alpha}} \). For example, if \( U_n \) has about \( n - 4c_{\alpha} n^{1/\alpha} \) points close to being uniformly distributed on the interior of \((2c_{\alpha}^{-\frac{1}{\alpha}}, 1 - 2c_{\alpha}^{-\frac{1}{\alpha}})\) and the remainder evenly divided between small intervals about 0 and 1, then \( d_P(U_n, U) \) is close to \( \frac{4}{3} c_{\alpha}^{-\frac{1}{\alpha}} \) (worst set is the union of the small intervals about 0 and 1) and \( d_P(U_n^*, U) \) is close to \( \frac{4}{5} c_{\alpha}^{-\frac{1}{\alpha}} \) (worst set \( \{c_{\alpha}^{-\frac{1}{\alpha}}, 1-c_{\alpha}^{-\frac{1}{\alpha}}\} \)). Hence the one-sample test based on \( d_P \) is not unbiased.

With additional complications, a similar example can be constructed for the two-sample case.
To attempt to gain some insight into the power of the $d_P$ test against various kinds of alternatives we have made some empirical comparisons of power of the one- and two-sample tests based on $d_P$ with that of the test of the same size based on $d_K$. For the one-sample test the $d_P$ critical value was taken from the distribution functions which were estimated in a separate Monte Carlo simulation. For the sake of convenience, for the two sample tests the size was chosen so that a non-randomized test could be used for both $d_K$ and $d_P$. Random samples were generated from several alternatives to the null distribution. The results of these comparisons are given in the following table.
One-sample tests $\alpha = .05$

<table>
<thead>
<tr>
<th>Description</th>
<th>(sample size)</th>
<th>Number Trials</th>
<th>Number Rejected</th>
<th>Proportion Rejected</th>
<th>$d_K$</th>
<th>$d_P$</th>
<th>$d_K$</th>
<th>$d_P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$U[0,1]$ vs. $U[0,1]$</td>
<td>(40)</td>
<td>4000</td>
<td>189</td>
<td>194</td>
<td>.047</td>
<td>.048</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U[0,1]$ vs. $U^2$</td>
<td>(40)</td>
<td>2000</td>
<td>1899</td>
<td>1900</td>
<td>.950</td>
<td>.950</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U[0,1]$ vs. 2-sided $U^2$</td>
<td>(40)</td>
<td>2000</td>
<td>743</td>
<td>1445</td>
<td>.371</td>
<td>.722</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U[0,1]$ vs. bimodal</td>
<td>(40)</td>
<td>2000</td>
<td>287</td>
<td>358</td>
<td>.143</td>
<td>.179</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U[0,1]$ vs. trimodal</td>
<td>(40)</td>
<td>2000</td>
<td>173</td>
<td>195</td>
<td>.086</td>
<td>.097</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U[0,1]$ vs. $U[0,1]$</td>
<td>(100)</td>
<td>1000</td>
<td>47</td>
<td>53</td>
<td>.047</td>
<td>.053</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U[0,1]$ vs. 2-sided $U^2$</td>
<td>(100)</td>
<td>1000</td>
<td>923</td>
<td>999</td>
<td>.923</td>
<td>.999</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U[0,1]$ vs. bimodal</td>
<td>(100)</td>
<td>1000</td>
<td>293</td>
<td>672</td>
<td>.293</td>
<td>.672</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U[0,1]$ vs. trimodal</td>
<td>(100)</td>
<td>1000</td>
<td>150</td>
<td>239</td>
<td>.150</td>
<td>.239</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Two-sample tests $\alpha = .25$

<table>
<thead>
<tr>
<th>Description</th>
<th>(sample size)</th>
<th>Number Trials</th>
<th>Number Rejected</th>
<th>Proportion Rejected</th>
<th>$d_K$</th>
<th>$d_P$</th>
<th>$d_K$</th>
<th>$d_P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N(0,1)$ vs. $N(0,1)$</td>
<td>(30)</td>
<td>200</td>
<td>46</td>
<td>48</td>
<td>.230</td>
<td>.240</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N(0,1)$ vs. $N(.5,1)$</td>
<td>(30)</td>
<td>1000</td>
<td>703</td>
<td>694</td>
<td>.703</td>
<td>.694</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N(0,1)$ vs. $N(.75,1)$</td>
<td>(30)</td>
<td>1000</td>
<td>914</td>
<td>913</td>
<td>.914</td>
<td>.913</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N(0,1)$ vs. $N(1,1)$</td>
<td>(30)</td>
<td>1000</td>
<td>987</td>
<td>983</td>
<td>.987</td>
<td>.983</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N(0,1)$ vs. $N(0,2)$</td>
<td>(30)</td>
<td>1000</td>
<td>369</td>
<td>438</td>
<td>.369</td>
<td>.438</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N(0,1)$ vs. $N(0,4)$</td>
<td>(30)</td>
<td>1000</td>
<td>693</td>
<td>825</td>
<td>.693</td>
<td>.825</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U[0,1]$ vs. $U^2$</td>
<td>(30)</td>
<td>1000</td>
<td>808</td>
<td>798</td>
<td>.808</td>
<td>.798</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U[0,1]$ vs. $U^3$</td>
<td>(30)</td>
<td>1000</td>
<td>988</td>
<td>988</td>
<td>.988</td>
<td>.988</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$U[0,1]$ vs. 2-sided $U^2$</td>
<td>(30)</td>
<td>1000</td>
<td>562</td>
<td>657</td>
<td>.562</td>
<td>.657</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Explanation: $U[0,1]$ denotes the uniform distribution on $[0,1]$. $N(\mu, \sigma^2)$ denotes the normal distribution with mean $\mu$ and variance $\sigma^2$. $U^k$ is the "Lehmann" alternative generated as $\max(U_1, \ldots, U_k)$, where $U_1, \ldots, U_k$
are uniform random variables. "2-sided $U^2$", "bimodal" and "trimodal" are random variables with respectively unimodal, bimodal, and trimodal distribution functions generated by taking appropriate mixtures of random variables with distribution $U^2$; specifically if $B_1$ and $B_2$ are independent Bernoulli random variables mutually independent of $V$, a random variable with distribution $U^2$, then $V_1 = B_1 V/2 + (1-B_1)(1-V/2)$ is our "2-sided" random variable, and $V_2 = B_2 V_1/2 + (1-B_2)(.5+V_1/2)$ is our "bimodal" random variable. The "trimodal" random deviates were generated as follows. Let $B$ take the values 0, 1, and 2 with equal probability. Then $V_3 = \frac{B}{3} + V_1/3$. Only uniform random numbers were used in the one-sample tests, because the time required to evaluate $\phi(X_1)$ for the normal distribution function $\phi$ threatened to become excessive for the large samples and large number of trials involved. The "2-sided", "bimodal" and "trimodal" alternatives were suggested by the examples given below, as being potentially favorable to $d_P$, because $d_P$ can "look at" several intervals simultaneously, not just at one as does $d_2$. The evidence is not sufficient to substantiate or refute such a conjecture.

It can be seen that in general $d_P$ performed better than $d_K$, except in the case of alternatives which were stochastically greater than the null distribution, in which case the differences are not significant.

Rosenblatt [26] gives examples comparing $d_2$ and $d_K$ which are useful for our purposes, shedding some light on the above results. An example in which $d_2$ performs better is an alternative $V$ to $U[0,1]$
which has distribution function

\[ F(x) = 0.03, \quad 0 \leq x \leq 0.06 \]
\[ F(x) = x, \quad x \geq 0.06, \]

where the sample size \( n = 1600 \) and \( \alpha = 0.05 \). In this case the acceptance region for the \( d_K \) test is (using [29]) \( d_K(V_n, U) > 1.36/40 = 0.034 \), and for \( d_2 \) is (from Rosenblatt's results) \( d_2(V_n, U) > 1.75/40 = 0.044 \).

A sample from \( V \) will always satisfy \( d_2(V_n, U) \geq 0.06 \), since there are no observations in (0, 0.06), and hence the \( d_2 \) test rejects with probability 1. On the other hand, the \( d_K \) test has power close to 0.05, since it can be shown (for example as Rosenblatt suggests, using a result of Uspensky cited in [25]) that \( \Pr\{ \sup_{0 \leq t \leq 0.06} |U[0,t] - V_n[0,t]| > 0.034 \} \)

is small for this value of \( n \).

The dramatic difference between \( d_2 \) and \( d_K \) in this case is due to the proper choice of \( n \) and, of course, \( V \). \( d_P \) is also similarly superior to \( d_2 \) in this case; the critical value for the \( d_P \) test is at most \( 0.7/40 = 0.0175 \), based on our Monte Carlo tables, but \( d_P(V_n, U) \geq 0.02 \) (consider the set (0, 0.06)), so the \( d_P \) test rejects with probability 1. However, if \( F = 0.0225 \) on [0, 0.045], then \( d_P \) also has poor power, while \( d_2 \) rejects a.s. (Rosenblatt also gives an example,

\[ F(x) = x; \quad 0 \leq x < 0.5 - \delta \]
\[ = 0.5 - \delta; \quad 0.5 - \delta \leq x \leq 0.5 + \delta \]
\[ = x; \quad 0.5 + \delta < x \leq 1, \]

for which \( d_K \) is superior to \( d_2 \) and also to \( d_P \), for proper choice of \( \delta \) and \( n \).

As an example of the same type in which \( d_P \) performs better than
either \( d_2 \) or \( d_K \), let \( n = 1600, \alpha = .05 \), and let \( V \) place mass .01 at the points 0 and .04 and mass .02 at the point .02, and be equal to Lebesgue measure on (.04,1]. Then if \( V_n \) is the empirical measure of a sample of size \( n \) from \( V \), \( d_P(V, U) \geq .02 \) (consider the set \( \{0, .02\} \)), so \( d_P \) rejects with probability 1. However, \( d_K(V, U) = .02 = d_2(V, U) \), while their critical values are .034 and .044 respectively, so as in the previous examples, it can be shown that the \( d_K \) and \( d_2 \) tests have low power for the values of \( n \) and \( \alpha \) chosen.

Massey [20] gives a lower bound for the power of the Kolmogorov-Smirnov test against a class of alternatives to the null distribution \( P \) of the form \( \{P' \in P(S): d_K(P, P') = c\} \). The result provides an alternative proof of consistency of the test. We give a similar result for \( d_P \).

**Proposition 5.3.2:** Let \( P \) and \( Q \) be elements of \( P(S) \) and \( P_n \) and \( Q_n \) be the empirical measures of samples of size \( n \) from \( P \) and \( Q \) respectively. Let \( \Pr\{d_P(P_n, P) \geq c_\alpha\} = \alpha \). If \( Q \in \{P' \in P(S): d_P(P, P') = c\} \) for some \( c > c_\alpha \), then \( \Pr\{d_P(Q_n, P) \geq c_\alpha\} \geq 1 - \Pr\{Y - \frac{n}{2} \leq -n(c - c_\alpha)\} \), where \( Y \) is a \( B(n, \frac{1}{2}) \) (i.e., binomial based on \( n \) trials with probability \( \frac{1}{2} \) of success on each trial) random variable.

**Proof:** Since \( d_P(Q, P) = c \), there exists a closed set \( F_0 \subseteq S \) such that \( Q(F_0) \geq P(F_0^C) + c \).

\[
\Pr\{d_P(Q_n, P) > c_\alpha\} = 1 - \Pr\{Q_n(F) \leq P(F_c^\alpha) + c_\alpha \; \text{all closed } F \subseteq S\}
\]

\[
\geq 1 - \Pr\{Q_n(F_0) \leq P(F_0^C) + c_\alpha\}
\]
\[ \geq 1 - \Pr \{ Q_n(F_0) \leq Q(F_0) - c + c_\alpha \} \]

\[ = 1 - \Pr \{ n(Q_n(F_0) - Q(F_0)) \leq -n(c-c_\alpha) \} \cdot \]

\( nQ_n(F_0) \) is a \( B(n,Q(F_0)) \) random variable. \( \Pr \{ n(Q_n(F_0) - Q(F_0)) \leq -n(c-c_\alpha) \} \) is maximized for \( Q(F_0) = \frac{1}{2} \), since the binomial random variable has maximum variance in this case. \( \square \)
APPENDIX

A.1 Tables of the one-sample distribution function.

The distribution function of $n^{1/2}d_p(u_n, \mu)$ has been estimated for $n = 5, 10, 20, 40, 60$ and 80, where $\mu$ is uniform measure on $[0, 1]$ and $\mu_n$ is the empirical measure of a sample of size $n$ from $\mu$. Our procedure was to generate 4000 random samples of size $n$, to compute for each sample the value of $d_p(u_n, \mu)$ and record the number of samples for which the value fell in the interval $[.001i, .001(i+1))$ for $i = 0, \ldots, 199$ and $[.2, \infty)$. (For $n = 5$ and $n = 10$, we used $[.002i, .002(i+1))$.) To get the tables which give $\text{Pr}(n^{1/2}d_p(u_n, \mu) \leq c)$ for values of $c$ which are multiples of .01, we interpolated linearly between the values estimated for $\text{Pr}(n^{1/2}d_p(u_n, \mu) \leq .001i \cdot n^{1/2})$ and $\text{Pr}(n^{1/2}d_p(u_n, \mu) \leq .001(i+1)n^{1/2})$ for $i = 0, \ldots, 199$.

The random samples from $\mu$ were generated using a program from the UNC Computation Center's "Scientific Subroutine Package." The program uses the "congruential" method, as explained in [Maclaren and Marsaglia, J.A.C.M. 12, 83-84]. It is noted in this article that the random numbers produced by such a method may not be mutually independent, and a correction suggested in the article has been used: 100 numbers generated by this method are placed in a table and labeled 00, \ldots, 99; two digits from an additional random number are used to determine which location in the table to take the next "observation" from, and the observation used is replaced by a new random number.
The same procedure was used to generate the $U[0,1]$ random sample for the power comparisons of Chapter 5. The Box-Mueller method (see Bell *C.A.C.M.* 11, 498 and Knop *C.A.C.M.* 12, 281) was used to generate the normally distributed random numbers for the power comparisons.

Several computer programs were written to compute $d_p$ using the methods of Chapter 2: one to use the Hungarian algorithm in the general two-sample case; one for the case of two discrete measures on $E^1$, and one for the one-sample test statistic on $E^1$. The programs have been checked by using them to find $d_p$ in a number of cases which can be computed by hand, involving various arrangements of the points of the "worst set." The two-sample programs provide a check on each other.

The programs for $E^1$ take on the order of 75 seconds to perform 4000 evaluations of $d_p(\mu_n, \mu)$ for $n = 80$. The program written for the Hungarian algorithm is considerably slower, taking on the order of 30 seconds to perform a single evaluation for $n = 30$, and the time required increases rapidly with $n$. The Hopcroft-Karp algorithm presumably will permit the evaluation for general metric spaces for larger values of $n$. 
One-sample distribution: $\Pr(\sqrt{n} d_p(\mu_n, \mu) \leq x)$

<table>
<thead>
<tr>
<th>$x$</th>
<th>n=5</th>
<th>10</th>
<th>20</th>
<th>40</th>
<th>60</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>.21</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>.22</td>
<td>.000</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>.23</td>
<td>.002</td>
<td>.002</td>
<td>.003</td>
<td>.001</td>
<td>.000</td>
<td>.001</td>
</tr>
<tr>
<td>.24</td>
<td>.007</td>
<td>.005</td>
<td>.004</td>
<td>.003</td>
<td>.002</td>
<td>.001</td>
</tr>
<tr>
<td>.25</td>
<td>.012</td>
<td>.010</td>
<td>.008</td>
<td>.006</td>
<td>.003</td>
<td>.002</td>
</tr>
<tr>
<td>.26</td>
<td>.021</td>
<td>.018</td>
<td>.015</td>
<td>.011</td>
<td>.008</td>
<td>.005</td>
</tr>
<tr>
<td>.27</td>
<td>.035</td>
<td>.028</td>
<td>.022</td>
<td>.015</td>
<td>.012</td>
<td>.012</td>
</tr>
<tr>
<td>.28</td>
<td>.052</td>
<td>.038</td>
<td>.033</td>
<td>.021</td>
<td>.017</td>
<td>.017</td>
</tr>
<tr>
<td>.29</td>
<td>.070</td>
<td>.056</td>
<td>.052</td>
<td>.031</td>
<td>.026</td>
<td>.029</td>
</tr>
<tr>
<td>.30</td>
<td>.095</td>
<td>.081</td>
<td>.069</td>
<td>.054</td>
<td>.037</td>
<td>.046</td>
</tr>
<tr>
<td>.31</td>
<td>.124</td>
<td>.103</td>
<td>.096</td>
<td>.071</td>
<td>.063</td>
<td>.060</td>
</tr>
<tr>
<td>.32</td>
<td>.156</td>
<td>.131</td>
<td>.120</td>
<td>.098</td>
<td>.083</td>
<td>.078</td>
</tr>
<tr>
<td>.33</td>
<td>.194</td>
<td>.158</td>
<td>.145</td>
<td>.119</td>
<td>.103</td>
<td>.109</td>
</tr>
<tr>
<td>.34</td>
<td>.236</td>
<td>.194</td>
<td>.184</td>
<td>.140</td>
<td>.141</td>
<td>.131</td>
</tr>
<tr>
<td>.35</td>
<td>.279</td>
<td>.235</td>
<td>.219</td>
<td>.180</td>
<td>.165</td>
<td>.157</td>
</tr>
<tr>
<td>.36</td>
<td>.322</td>
<td>.282</td>
<td>.262</td>
<td>.208</td>
<td>.199</td>
<td>.209</td>
</tr>
<tr>
<td>.37</td>
<td>.358</td>
<td>.318</td>
<td>.301</td>
<td>.253</td>
<td>.236</td>
<td>.240</td>
</tr>
<tr>
<td>.38</td>
<td>.400</td>
<td>.356</td>
<td>.336</td>
<td>.286</td>
<td>.264</td>
<td>.271</td>
</tr>
<tr>
<td>.40</td>
<td>.487</td>
<td>.440</td>
<td>.421</td>
<td>.368</td>
<td>.360</td>
<td>.362</td>
</tr>
<tr>
<td>.41</td>
<td>.527</td>
<td>.476</td>
<td>.457</td>
<td>.416</td>
<td>.392</td>
<td>.392</td>
</tr>
<tr>
<td>.42</td>
<td>.560</td>
<td>.513</td>
<td>.501</td>
<td>.448</td>
<td>.431</td>
<td>.444</td>
</tr>
<tr>
<td>.43</td>
<td>.592</td>
<td>.552</td>
<td>.530</td>
<td>.481</td>
<td>.482</td>
<td>.472</td>
</tr>
<tr>
<td>.44</td>
<td>.628</td>
<td>.592</td>
<td>.568</td>
<td>.526</td>
<td>.509</td>
<td>.501</td>
</tr>
<tr>
<td>.45</td>
<td>.655</td>
<td>.627</td>
<td>.604</td>
<td>.555</td>
<td>.561</td>
<td>.541</td>
</tr>
<tr>
<td>.46</td>
<td>.684</td>
<td>.660</td>
<td>.630</td>
<td>.598</td>
<td>.589</td>
<td>.590</td>
</tr>
<tr>
<td>.47</td>
<td>.709</td>
<td>.688</td>
<td>.659</td>
<td>.622</td>
<td>.613</td>
<td>.618</td>
</tr>
<tr>
<td>.48</td>
<td>.731</td>
<td>.712</td>
<td>.687</td>
<td>.660</td>
<td>.655</td>
<td>.641</td>
</tr>
<tr>
<td>.49</td>
<td>.753</td>
<td>.738</td>
<td>.712</td>
<td>.685</td>
<td>.676</td>
<td>.678</td>
</tr>
<tr>
<td>.50</td>
<td>.774</td>
<td>.764</td>
<td>.739</td>
<td>.706</td>
<td>.710</td>
<td>.701</td>
</tr>
<tr>
<td>.51</td>
<td>.794</td>
<td>.784</td>
<td>.760</td>
<td>.737</td>
<td>.736</td>
<td>.720</td>
</tr>
<tr>
<td>.52</td>
<td>.815</td>
<td>.804</td>
<td>.784</td>
<td>.756</td>
<td>.754</td>
<td>.758</td>
</tr>
<tr>
<td>.53</td>
<td>.833</td>
<td>.828</td>
<td>.801</td>
<td>.776</td>
<td>.776</td>
<td>.770</td>
</tr>
<tr>
<td>.54</td>
<td>.848</td>
<td>.841</td>
<td>.818</td>
<td>.795</td>
<td>.804</td>
<td>.785</td>
</tr>
<tr>
<td>.55</td>
<td>.863</td>
<td>.855</td>
<td>.837</td>
<td>.820</td>
<td>.818</td>
<td>.807</td>
</tr>
<tr>
<td>.56</td>
<td>.877</td>
<td>.869</td>
<td>.850</td>
<td>.834</td>
<td>.843</td>
<td>.828</td>
</tr>
<tr>
<td>.57</td>
<td>.889</td>
<td>.881</td>
<td>.863</td>
<td>.855</td>
<td>.854</td>
<td>.842</td>
</tr>
<tr>
<td>.58</td>
<td>.901</td>
<td>.894</td>
<td>.877</td>
<td>.867</td>
<td>.866</td>
<td>.863</td>
</tr>
<tr>
<td>.59</td>
<td>.913</td>
<td>.905</td>
<td>.888</td>
<td>.878</td>
<td>.884</td>
<td>.877</td>
</tr>
<tr>
<td>.60</td>
<td>.923</td>
<td>.915</td>
<td>.899</td>
<td>.893</td>
<td>.892</td>
<td>.884</td>
</tr>
<tr>
<td>x</td>
<td>n=5</td>
<td>10</td>
<td>20</td>
<td>40</td>
<td>60</td>
<td>80</td>
</tr>
<tr>
<td>-----</td>
<td>------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>.61</td>
<td>.933</td>
<td>.922</td>
<td>.907</td>
<td>.902</td>
<td>.899</td>
<td>.894</td>
</tr>
<tr>
<td>.62</td>
<td>.941</td>
<td>.930</td>
<td>.915</td>
<td>.912</td>
<td>.911</td>
<td>.909</td>
</tr>
<tr>
<td>.63</td>
<td>.947</td>
<td>.938</td>
<td>.925</td>
<td>.919</td>
<td>.920</td>
<td>.915</td>
</tr>
<tr>
<td>.64</td>
<td>.952</td>
<td>.944</td>
<td>.932</td>
<td>.927</td>
<td>.928</td>
<td>.926</td>
</tr>
<tr>
<td>.65</td>
<td>.958</td>
<td>.951</td>
<td>.940</td>
<td>.934</td>
<td>.935</td>
<td>.936</td>
</tr>
<tr>
<td>.66</td>
<td>.962</td>
<td>.958</td>
<td>.950</td>
<td>.941</td>
<td>.941</td>
<td>.942</td>
</tr>
<tr>
<td>.67</td>
<td>.967</td>
<td>.963</td>
<td>.954</td>
<td>.950</td>
<td>.948</td>
<td>.917</td>
</tr>
<tr>
<td>.68</td>
<td>.970</td>
<td>.967</td>
<td>.959</td>
<td>.956</td>
<td>.955</td>
<td>.954</td>
</tr>
<tr>
<td>.69</td>
<td>.973</td>
<td>.970</td>
<td>.963</td>
<td>.959</td>
<td>.960</td>
<td>.960</td>
</tr>
<tr>
<td>.70</td>
<td>.974</td>
<td>.973</td>
<td>.968</td>
<td>.963</td>
<td>.966</td>
<td>.964</td>
</tr>
<tr>
<td>.71</td>
<td>.978</td>
<td>.978</td>
<td>.972</td>
<td>.968</td>
<td>.968</td>
<td>.969</td>
</tr>
<tr>
<td>.72</td>
<td>.981</td>
<td>.980</td>
<td>.974</td>
<td>.970</td>
<td>.971</td>
<td>.973</td>
</tr>
<tr>
<td>.73</td>
<td>.983</td>
<td>.983</td>
<td>.976</td>
<td>.975</td>
<td>.974</td>
<td>.975</td>
</tr>
<tr>
<td>.74</td>
<td>.986</td>
<td>.984</td>
<td>.978</td>
<td>.978</td>
<td>.977</td>
<td>.979</td>
</tr>
<tr>
<td>.75</td>
<td>.988</td>
<td>.985</td>
<td>.980</td>
<td>.982</td>
<td>.980</td>
<td>.982</td>
</tr>
<tr>
<td>.76</td>
<td>.990</td>
<td>.988</td>
<td>.983</td>
<td>.984</td>
<td>.982</td>
<td>.984</td>
</tr>
<tr>
<td>.77</td>
<td>.991</td>
<td>.990</td>
<td>.986</td>
<td>.986</td>
<td>.983</td>
<td>.985</td>
</tr>
<tr>
<td>.78</td>
<td>.992</td>
<td>.991</td>
<td>.988</td>
<td>.988</td>
<td>.986</td>
<td>.987</td>
</tr>
<tr>
<td>.79</td>
<td>.994</td>
<td>.992</td>
<td>.989</td>
<td>.988</td>
<td>.987</td>
<td>.988</td>
</tr>
<tr>
<td>.80</td>
<td>.995</td>
<td>.994</td>
<td>.990</td>
<td>.989</td>
<td>.989</td>
<td>.990</td>
</tr>
<tr>
<td>.81</td>
<td>.996</td>
<td>.995</td>
<td>.992</td>
<td>.990</td>
<td>.990</td>
<td>.992</td>
</tr>
<tr>
<td>.82</td>
<td>.996</td>
<td>.996</td>
<td>.993</td>
<td>.991</td>
<td>.991</td>
<td>.992</td>
</tr>
<tr>
<td>.83</td>
<td>.996</td>
<td>.996</td>
<td>.994</td>
<td>.992</td>
<td>.993</td>
<td>.994</td>
</tr>
<tr>
<td>.84</td>
<td>.997</td>
<td>.997</td>
<td>.995</td>
<td>.993</td>
<td>.993</td>
<td>.995</td>
</tr>
<tr>
<td>.85</td>
<td>.998</td>
<td>.997</td>
<td>.996</td>
<td>.994</td>
<td>.994</td>
<td>.996</td>
</tr>
<tr>
<td>.86</td>
<td>.999</td>
<td>.997</td>
<td>.996</td>
<td>.994</td>
<td>.995</td>
<td>.996</td>
</tr>
<tr>
<td>.87</td>
<td>.999</td>
<td>.998</td>
<td>.997</td>
<td>.995</td>
<td>.996</td>
<td>.996</td>
</tr>
<tr>
<td>.88</td>
<td>.999</td>
<td>.999</td>
<td>.997</td>
<td>.996</td>
<td>.997</td>
<td>.997</td>
</tr>
<tr>
<td>.89</td>
<td>.999</td>
<td>.999</td>
<td>.998</td>
<td>.996</td>
<td>.997</td>
<td>.997</td>
</tr>
<tr>
<td>.90</td>
<td>.999</td>
<td>.999</td>
<td>.999</td>
<td>.996</td>
<td>.998</td>
<td>.998</td>
</tr>
<tr>
<td>.91</td>
<td>.999</td>
<td>.999</td>
<td>.999</td>
<td>.996</td>
<td>.998</td>
<td>.998</td>
</tr>
<tr>
<td>.92</td>
<td>.999</td>
<td>.999</td>
<td>.999</td>
<td>.997</td>
<td>.998</td>
<td>.998</td>
</tr>
<tr>
<td>.93</td>
<td>.999</td>
<td>.999</td>
<td>.999</td>
<td>.997</td>
<td>.998</td>
<td>.999</td>
</tr>
<tr>
<td>.94</td>
<td>.999</td>
<td>.999</td>
<td>.999</td>
<td>.998</td>
<td>.998</td>
<td>.999</td>
</tr>
<tr>
<td>.95</td>
<td>.999</td>
<td>.999</td>
<td>.999</td>
<td>.998</td>
<td>.999</td>
<td>.999</td>
</tr>
</tbody>
</table>
A.2 Tables of the two-sample distribution function.

The values of \( K(i, n, p, s, t) \) were computed for \( i = 0, 1, 2; \)
\( z = 0, 1, \ldots, 12; \) and \( n = 1, 2, \ldots, 100, \) for each value of \( d, s, \) and \( t, \)
\( 0 \leq p \leq z, \) \( 0 \leq s \leq \min(z, p), \) using equation 3.5.2 with the simplifications described in Section 3.6. \( A(p, n, z) \) as defined in 3.1.1 is then computed as
\[
\frac{2}{n} \sum_{i=0}^{2} \sum_{s=0}^{z} \sum_{t=0}^{p} K(i, n, p, s, t),
\]
where as described in Section 3.1, \( K(i, n, p, s, t) \) depends on \( z. \)

In the following tables we have listed \( \left( \frac{2n}{n} \right)^{-1} \sum_{0 \leq p \leq z/2} A(p, n, z), \)
which is \( \Pr\{d_1(\mu_n, \nu_n) \leq z/2n\} \), and also \( \left( \frac{2n}{n} \right)^{-1} \sum_{0 \leq p \leq z} A(p, n, z), \)
which is \( \Pr\{d(2)_p(\mu_n, \nu_n) \leq z/2n\}, \) where \( d(2)_p \) is defined as in Section 5.1.

Because of the recursiveness in equations 3.5.2, there is reason to suspect that rounding errors will affect the results. To check for this, the computations were repeated with values of \( n \) up to 50 for larger values of \( z, \) using "double precision" storage in the computer memory. Storing the \( K(\cdot, \cdot, \cdot, \cdot, \cdot) \) values in double precision did not affect the result while storing the \( D(\cdot, \cdot) \) values in double precision made a difference in the fourth decimal place of the computed probabilities. (There was no marked effect for smaller values of \( z. \)) The tables presented here were computed with the \( D_i(\cdot, \cdot) \) values in double precision. Storage limitations preclude the use of double precision for all the numbers for the full range of parameters.

The \( D_i(\ell, s) \) values can be checked by hand for moderately small values of \( \ell, s \) and \( z, \) sufficient to test the operation of all parts of this segment of the program. The probabilities in the table have
been checked for $n = 6, 7$ and $8$ for several values of $z$, by generating all permutations of the two samples and using the two-sample program for $t^1$ to find $d_p(\mu_n, \nu_n)$ for each permutation. An additional check on the correctness of the programming (although not necessarily of the theory) is that for $z$ large enough $\sum_{p=0}^{z} A(p, n, z) = \binom{2n}{n}$. 
\[
\Pr\{d_P(\nu_n, \nu_n) \leq z/2n\}
\]

<table>
<thead>
<tr>
<th>n</th>
<th>z=1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.6667</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.4000</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.2286</td>
<td>0.9429</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.1270</td>
<td>0.8651</td>
<td>0.9683</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.0693</td>
<td>0.7705</td>
<td>0.9199</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.0373</td>
<td>0.6719</td>
<td>0.8631</td>
<td>0.9953</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>0.0199</td>
<td>0.5747</td>
<td>0.7995</td>
<td>0.9849</td>
<td>0.9975</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>0.0105</td>
<td>0.4835</td>
<td>0.7329</td>
<td>0.9692</td>
<td>0.9914</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.0055</td>
<td>0.4010</td>
<td>0.6660</td>
<td>0.9490</td>
<td>0.9815</td>
<td>0.9997</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.0029</td>
<td>0.3283</td>
<td>0.6004</td>
<td>0.9248</td>
<td>0.9682</td>
<td>0.9986</td>
<td>0.9998</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.0015</td>
<td>0.2658</td>
<td>0.5376</td>
<td>0.8971</td>
<td>0.9519</td>
<td>0.9964</td>
<td>0.9992</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.0008</td>
<td>0.2131</td>
<td>0.4783</td>
<td>0.8666</td>
<td>0.9330</td>
<td>0.9929</td>
<td>0.9979</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.0004</td>
<td>0.1692</td>
<td>0.4232</td>
<td>0.8338</td>
<td>0.9117</td>
<td>0.9880</td>
<td>0.9957</td>
<td>0.9999</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.0002</td>
<td>0.1333</td>
<td>0.3725</td>
<td>0.7993</td>
<td>0.8886</td>
<td>0.9818</td>
<td>0.9926</td>
<td>0.9996</td>
<td>0.9999</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>0.0001</td>
<td>0.1043</td>
<td>0.3262</td>
<td>0.7635</td>
<td>0.8637</td>
<td>0.9742</td>
<td>0.9884</td>
<td>0.9991</td>
<td>0.9998</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.0001</td>
<td>0.0810</td>
<td>0.2846</td>
<td>0.7269</td>
<td>0.8375</td>
<td>0.9654</td>
<td>0.9832</td>
<td>0.9984</td>
<td>0.9995</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>0.0000</td>
<td>0.0626</td>
<td>0.2472</td>
<td>0.6900</td>
<td>0.8103</td>
<td>0.9552</td>
<td>0.9770</td>
<td>0.9972</td>
<td>0.9990</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>0.0000</td>
<td>0.0481</td>
<td>0.2139</td>
<td>0.6530</td>
<td>0.7823</td>
<td>0.9439</td>
<td>0.9698</td>
<td>0.9956</td>
<td>0.9983</td>
<td>0.9999</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.0000</td>
<td>0.0367</td>
<td>0.1845</td>
<td>0.6162</td>
<td>0.7537</td>
<td>0.9315</td>
<td>0.9616</td>
<td>0.9936</td>
<td>0.9972</td>
<td>0.9998</td>
<td>0.9999</td>
<td>1.0000</td>
</tr>
<tr>
<td>21</td>
<td>0.0000</td>
<td>0.0279</td>
<td>0.1585</td>
<td>0.5800</td>
<td>0.7248</td>
<td>0.9181</td>
<td>0.9526</td>
<td>0.9911</td>
<td>0.9959</td>
<td>0.9996</td>
<td>0.9999</td>
<td>1.0000</td>
</tr>
<tr>
<td>22</td>
<td>0.0000</td>
<td>0.0217</td>
<td>0.1359</td>
<td>0.5445</td>
<td>0.6957</td>
<td>0.9037</td>
<td>0.9427</td>
<td>0.9881</td>
<td>0.9941</td>
<td>0.9993</td>
<td>0.9998</td>
<td>1.0000</td>
</tr>
<tr>
<td>23</td>
<td>0.0000</td>
<td>0.0160</td>
<td>0.1161</td>
<td>0.5100</td>
<td>0.6666</td>
<td>0.8885</td>
<td>0.9320</td>
<td>0.9845</td>
<td>0.9920</td>
<td>0.9990</td>
<td>0.9996</td>
<td>1.0000</td>
</tr>
<tr>
<td>24</td>
<td>0.0000</td>
<td>0.0120</td>
<td>0.0989</td>
<td>0.4765</td>
<td>0.6377</td>
<td>0.8725</td>
<td>0.9205</td>
<td>0.9805</td>
<td>0.9895</td>
<td>0.9984</td>
<td>0.9993</td>
<td>1.0000</td>
</tr>
<tr>
<td>25</td>
<td>0.0000</td>
<td>0.0090</td>
<td>0.0841</td>
<td>0.4442</td>
<td>0.6091</td>
<td>0.8558</td>
<td>0.9084</td>
<td>0.9759</td>
<td>0.9866</td>
<td>0.9977</td>
<td>0.9990</td>
<td>0.9999</td>
</tr>
<tr>
<td>26</td>
<td>0.0000</td>
<td>0.0067</td>
<td>0.0713</td>
<td>0.4133</td>
<td>0.5809</td>
<td>0.8386</td>
<td>0.8957</td>
<td>0.9709</td>
<td>0.9833</td>
<td>0.9969</td>
<td>0.9985</td>
<td>0.9998</td>
</tr>
<tr>
<td>27</td>
<td>0.0000</td>
<td>0.0050</td>
<td>0.0603</td>
<td>0.3837</td>
<td>0.5531</td>
<td>0.8208</td>
<td>0.8824</td>
<td>0.9654</td>
<td>0.9796</td>
<td>0.9958</td>
<td>0.9979</td>
<td>0.9998</td>
</tr>
<tr>
<td>28</td>
<td>0.0000</td>
<td>0.0037</td>
<td>0.0509</td>
<td>0.3555</td>
<td>0.5260</td>
<td>0.8025</td>
<td>0.8686</td>
<td>0.9593</td>
<td>0.9754</td>
<td>0.9946</td>
<td>0.9972</td>
<td>0.9996</td>
</tr>
<tr>
<td>29</td>
<td>0.0000</td>
<td>0.0027</td>
<td>0.0429</td>
<td>0.3288</td>
<td>0.4996</td>
<td>0.7839</td>
<td>0.8543</td>
<td>0.9529</td>
<td>0.9709</td>
<td>0.9931</td>
<td>0.9963</td>
<td>0.9994</td>
</tr>
<tr>
<td>30</td>
<td>0.0000</td>
<td>0.0020</td>
<td>0.0361</td>
<td>0.3035</td>
<td>0.4738</td>
<td>0.7649</td>
<td>0.8396</td>
<td>0.9459</td>
<td>0.9660</td>
<td>0.9915</td>
<td>0.9953</td>
<td>0.9992</td>
</tr>
<tr>
<td>n</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>.0000</td>
<td>.0095</td>
<td>.0458</td>
<td>.2153</td>
<td>.3319</td>
<td>.5722</td>
<td>.6592</td>
<td>.8135</td>
<td>.8560</td>
<td>.9297</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>.0000</td>
<td>.0085</td>
<td>.0423</td>
<td>.2054</td>
<td>.3205</td>
<td>.5604</td>
<td>.6486</td>
<td>.8057</td>
<td>.8495</td>
<td>.9257</td>
<td></td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>.0000</td>
<td>.0076</td>
<td>.0390</td>
<td>.1958</td>
<td>.3094</td>
<td>.5487</td>
<td>.6380</td>
<td>.7979</td>
<td>.8429</td>
<td>.9216</td>
<td></td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>.0000</td>
<td>.0067</td>
<td>.0360</td>
<td>.1867</td>
<td>.2986</td>
<td>.5371</td>
<td>.6275</td>
<td>.7900</td>
<td>.8362</td>
<td>.9174</td>
<td></td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>.0000</td>
<td>.0060</td>
<td>.0332</td>
<td>.1778</td>
<td>.2880</td>
<td>.5256</td>
<td>.6169</td>
<td>.7820</td>
<td>.8295</td>
<td>.9131</td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>.0000</td>
<td>.0053</td>
<td>.0306</td>
<td>.1694</td>
<td>.2778</td>
<td>.5141</td>
<td>.6065</td>
<td>.7740</td>
<td>.8227</td>
<td>.9087</td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>.0000</td>
<td>.0047</td>
<td>.0281</td>
<td>.1612</td>
<td>.2678</td>
<td>.5028</td>
<td>.5960</td>
<td>.7660</td>
<td>.8158</td>
<td>.9042</td>
<td></td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>.0000</td>
<td>.0042</td>
<td>.0259</td>
<td>.1534</td>
<td>.2581</td>
<td>.4916</td>
<td>.5857</td>
<td>.7577</td>
<td>.8088</td>
<td>.8997</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>.0000</td>
<td>.0037</td>
<td>.0238</td>
<td>.1459</td>
<td>.2487</td>
<td>.4806</td>
<td>.5753</td>
<td>.7495</td>
<td>.8018</td>
<td>.8950</td>
<td></td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>.0000</td>
<td>.0033</td>
<td>.0219</td>
<td>.1388</td>
<td>.2396</td>
<td>.4696</td>
<td>.5651</td>
<td>.7413</td>
<td>.7948</td>
<td>.8903</td>
<td></td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>.0000</td>
<td>.0029</td>
<td>.0202</td>
<td>.1319</td>
<td>.2307</td>
<td>.4588</td>
<td>.5549</td>
<td>.7331</td>
<td>.7877</td>
<td>.8855</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>.0000</td>
<td>.0026</td>
<td>.0185</td>
<td>.1253</td>
<td>.2221</td>
<td>.4481</td>
<td>.5448</td>
<td>.7248</td>
<td>.7805</td>
<td>.8807</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>.0000</td>
<td>.0023</td>
<td>.0170</td>
<td>.1190</td>
<td>.2137</td>
<td>.4376</td>
<td>.5348</td>
<td>.7165</td>
<td>.7733</td>
<td>.8757</td>
<td></td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>.0000</td>
<td>.0020</td>
<td>.0156</td>
<td>.1130</td>
<td>.2056</td>
<td>.4272</td>
<td>.5248</td>
<td>.7081</td>
<td>.7660</td>
<td>.8707</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>.0000</td>
<td>.0018</td>
<td>.0143</td>
<td>.1072</td>
<td>.1978</td>
<td>.4170</td>
<td>.5149</td>
<td>.6998</td>
<td>.7588</td>
<td>.8657</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>.0000</td>
<td>.0016</td>
<td>.0132</td>
<td>.1017</td>
<td>.1902</td>
<td>.4069</td>
<td>.5051</td>
<td>.6914</td>
<td>.7514</td>
<td>.8605</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>.0000</td>
<td>.0014</td>
<td>.0121</td>
<td>.0965</td>
<td>.1828</td>
<td>.3969</td>
<td>.4955</td>
<td>.6831</td>
<td>.7441</td>
<td>.8553</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>.0000</td>
<td>.0012</td>
<td>.0111</td>
<td>.0915</td>
<td>.1757</td>
<td>.3871</td>
<td>.4859</td>
<td>.6747</td>
<td>.7367</td>
<td>.8501</td>
<td></td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>.0000</td>
<td>.0011</td>
<td>.0101</td>
<td>.0867</td>
<td>.1688</td>
<td>.3775</td>
<td>.4764</td>
<td>.6663</td>
<td>.7293</td>
<td>.8448</td>
<td></td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>.0000</td>
<td>.0009</td>
<td>.0093</td>
<td>.0814</td>
<td>.1621</td>
<td>.3680</td>
<td>.4670</td>
<td>.6579</td>
<td>.7219</td>
<td>.8394</td>
<td></td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>.0000</td>
<td>.0008</td>
<td>.0085</td>
<td>.0778</td>
<td>.1557</td>
<td>.3587</td>
<td>.4577</td>
<td>.6496</td>
<td>.7145</td>
<td>.8340</td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>.0000</td>
<td>.0007</td>
<td>.0078</td>
<td>.0737</td>
<td>.1495</td>
<td>.3495</td>
<td>.4485</td>
<td>.6412</td>
<td>.7070</td>
<td>.8285</td>
<td></td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>.0000</td>
<td>.0006</td>
<td>.0071</td>
<td>.0697</td>
<td>.1435</td>
<td>.3405</td>
<td>.4394</td>
<td>.6329</td>
<td>.6996</td>
<td>.8229</td>
<td></td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>.0000</td>
<td>.0006</td>
<td>.0065</td>
<td>.0660</td>
<td>.1377</td>
<td>.3317</td>
<td>.4304</td>
<td>.6245</td>
<td>.6921</td>
<td>.8174</td>
<td></td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>.0000</td>
<td>.0005</td>
<td>.0060</td>
<td>.0624</td>
<td>.1321</td>
<td>.3230</td>
<td>.4215</td>
<td>.6162</td>
<td>.6847</td>
<td>.8117</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>.0000</td>
<td>.0004</td>
<td>.0055</td>
<td>.0590</td>
<td>.1267</td>
<td>.3144</td>
<td>.4128</td>
<td>.6079</td>
<td>.6772</td>
<td>.8061</td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>.0000</td>
<td>.0004</td>
<td>.0050</td>
<td>.0558</td>
<td>.1215</td>
<td>.3061</td>
<td>.4041</td>
<td>.5997</td>
<td>.6697</td>
<td>.8004</td>
<td></td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>.0000</td>
<td>.0003</td>
<td>.0046</td>
<td>.0527</td>
<td>.1164</td>
<td>.2979</td>
<td>.3956</td>
<td>.5914</td>
<td>.6623</td>
<td>.7946</td>
<td></td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>.0000</td>
<td>.0003</td>
<td>.0042</td>
<td>.0498</td>
<td>.1116</td>
<td>.2898</td>
<td>.3872</td>
<td>.5832</td>
<td>.6548</td>
<td>.7888</td>
<td></td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>.0000</td>
<td>.0003</td>
<td>.0038</td>
<td>.0471</td>
<td>.1069</td>
<td>.2820</td>
<td>.3789</td>
<td>.5751</td>
<td>.6473</td>
<td>.7830</td>
<td></td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>.0000</td>
<td>.0002</td>
<td>.0035</td>
<td>.0444</td>
<td>.1024</td>
<td>.2742</td>
<td>.3707</td>
<td>.5669</td>
<td>.6399</td>
<td>.7772</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>.0000</td>
<td>.0002</td>
<td>.0032</td>
<td>.0420</td>
<td>.0981</td>
<td>.2667</td>
<td>.3627</td>
<td>.5588</td>
<td>.6325</td>
<td>.7713</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>.0000</td>
<td>.0002</td>
<td>.0029</td>
<td>.0396</td>
<td>.0940</td>
<td>.2593</td>
<td>.3547</td>
<td>.5508</td>
<td>.6251</td>
<td>.7654</td>
<td></td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>.0000</td>
<td>.0002</td>
<td>.0026</td>
<td>.0374</td>
<td>.0900</td>
<td>.2520</td>
<td>.3469</td>
<td>.5428</td>
<td>.6176</td>
<td>.7595</td>
<td></td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>.0000</td>
<td>.0001</td>
<td>.0024</td>
<td>.0353</td>
<td>.0861</td>
<td>.2449</td>
<td>.3392</td>
<td>.5348</td>
<td>.6103</td>
<td>.7535</td>
<td></td>
<td></td>
</tr>
<tr>
<td>99</td>
<td>.0000</td>
<td>.0001</td>
<td>.0022</td>
<td>.0332</td>
<td>.0824</td>
<td>.2380</td>
<td>.3316</td>
<td>.5269</td>
<td>.6029</td>
<td>.7475</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>.0000</td>
<td>.0001</td>
<td>.0020</td>
<td>.0313</td>
<td>.0788</td>
<td>.2312</td>
<td>.3242</td>
<td>.5190</td>
<td>.5956</td>
<td>.7415</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
\[
\text{Pr}\{q^{(2)}_p (\mu_n, \nu_n) \leq z/2n\}
\]

<table>
<thead>
<tr>
<th>n</th>
<th>z=1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.9000</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.7429</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>.5714</td>
<td>.9841</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.4156</td>
<td>.9546</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>.2890</td>
<td>.9103</td>
<td>.9977</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>.1939</td>
<td>.8546</td>
<td>.9918</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>.1264</td>
<td>.7899</td>
<td>.9819</td>
<td>.997</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>.0804</td>
<td>.7194</td>
<td>.9678</td>
<td>.9986</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>.0501</td>
<td>.6462</td>
<td>.9493</td>
<td>.9965</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>.0307</td>
<td>.5728</td>
<td>.9266</td>
<td>.9930</td>
<td>.998</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>.0185</td>
<td>.5016</td>
<td>.9000</td>
<td>.9881</td>
<td>.9994</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>.0110</td>
<td>.4343</td>
<td>.8698</td>
<td>.9816</td>
<td>.9986</td>
<td>1.0000</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>.0065</td>
<td>.3721</td>
<td>.8366</td>
<td>.9733</td>
<td>.9973</td>
<td>.9999</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>.0038</td>
<td>.3157</td>
<td>.8007</td>
<td>.9633</td>
<td>.9955</td>
<td>.9997</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>.0022</td>
<td>.2654</td>
<td>.7627</td>
<td>.9516</td>
<td>.9931</td>
<td>.9994</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>.0013</td>
<td>.2212</td>
<td>.7232</td>
<td>.9381</td>
<td>.9900</td>
<td>.9990</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>.0007</td>
<td>.1830</td>
<td>.6826</td>
<td>.9230</td>
<td>.9862</td>
<td>.9983</td>
<td>.9999</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>.0004</td>
<td>.1502</td>
<td>.6414</td>
<td>.9063</td>
<td>.9816</td>
<td>.9974</td>
<td>.9998</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>.0002</td>
<td>.1225</td>
<td>.6002</td>
<td>.8881</td>
<td>.9761</td>
<td>.9962</td>
<td>.9996</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>.0001</td>
<td>.0993</td>
<td>.5593</td>
<td>.8684</td>
<td>.9699</td>
<td>.9947</td>
<td>.9994</td>
<td>.9999</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>.0001</td>
<td>.0800</td>
<td>.5191</td>
<td>.8475</td>
<td>.9628</td>
<td>.9929</td>
<td>.9990</td>
<td>.9999</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>.0000</td>
<td>.0641</td>
<td>.4799</td>
<td>.8254</td>
<td>.9549</td>
<td>.9907</td>
<td>.9986</td>
<td>.9998</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>.0000</td>
<td>.0511</td>
<td>.4420</td>
<td>.8022</td>
<td>.9461</td>
<td>.9881</td>
<td>.9980</td>
<td>.9997</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>.0000</td>
<td>.0405</td>
<td>.4057</td>
<td>.7782</td>
<td>.9365</td>
<td>.9851</td>
<td>.9972</td>
<td>.9996</td>
<td>.9999</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>.0000</td>
<td>.0320</td>
<td>.3711</td>
<td>.7533</td>
<td>.9261</td>
<td>.9817</td>
<td>.9963</td>
<td>.9994</td>
<td>.9999</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>.0000</td>
<td>.0251</td>
<td>.3383</td>
<td>.7278</td>
<td>.9149</td>
<td>.9778</td>
<td>.9952</td>
<td>.9992</td>
<td>.9999</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>.0000</td>
<td>.0197</td>
<td>.3074</td>
<td>.7019</td>
<td>.9029</td>
<td>.9735</td>
<td>.9940</td>
<td>.9989</td>
<td>.9998</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>61</td>
<td>.0000</td>
<td>.0044</td>
<td>.0967</td>
<td>.3588</td>
<td>.6375</td>
<td>.8226</td>
<td>.9207</td>
<td>.9669</td>
<td>.9872</td>
<td>.9954</td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>.0000</td>
<td>.0038</td>
<td>.0891</td>
<td>.3438</td>
<td>.6238</td>
<td>.8134</td>
<td>.9156</td>
<td>.9643</td>
<td>.9860</td>
<td>.9949</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>.0000</td>
<td>.0032</td>
<td>.0820</td>
<td>.3292</td>
<td>.6097</td>
<td>.8040</td>
<td>.9013</td>
<td>.9617</td>
<td>.9847</td>
<td>.9943</td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>.0000</td>
<td>.0028</td>
<td>.0755</td>
<td>.3150</td>
<td>.5958</td>
<td>.7945</td>
<td>.9048</td>
<td>.9588</td>
<td>.9834</td>
<td>.9937</td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>.0000</td>
<td>.0023</td>
<td>.0693</td>
<td>.3012</td>
<td>.5819</td>
<td>.7849</td>
<td>.8992</td>
<td>.9559</td>
<td>.9819</td>
<td>.9931</td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>.0000</td>
<td>.0020</td>
<td>.0636</td>
<td>.2878</td>
<td>.5681</td>
<td>.7751</td>
<td>.8935</td>
<td>.9529</td>
<td>.9805</td>
<td>.9924</td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>.0000</td>
<td>.0017</td>
<td>.0584</td>
<td>.2748</td>
<td>.5543</td>
<td>.7652</td>
<td>.8876</td>
<td>.9497</td>
<td>.9789</td>
<td>.9917</td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>.0000</td>
<td>.0014</td>
<td>.0535</td>
<td>.2622</td>
<td>.5406</td>
<td>.7551</td>
<td>.8816</td>
<td>.9465</td>
<td>.9773</td>
<td>.9909</td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>.0000</td>
<td>.0012</td>
<td>.0490</td>
<td>.2501</td>
<td>.5270</td>
<td>.7450</td>
<td>.8754</td>
<td>.9431</td>
<td>.9756</td>
<td>.9901</td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>.0000</td>
<td>.0010</td>
<td>.0448</td>
<td>.2383</td>
<td>.5134</td>
<td>.7347</td>
<td>.8690</td>
<td>.9396</td>
<td>.9738</td>
<td>.9892</td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>.0000</td>
<td>.0009</td>
<td>.0410</td>
<td>.2270</td>
<td>.5000</td>
<td>.7244</td>
<td>.8626</td>
<td>.9360</td>
<td>.9719</td>
<td>.9884</td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>.0000</td>
<td>.0007</td>
<td>.0374</td>
<td>.2161</td>
<td>.4868</td>
<td>.7140</td>
<td>.8560</td>
<td>.9323</td>
<td>.9700</td>
<td>.9874</td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>.0000</td>
<td>.0006</td>
<td>.0342</td>
<td>.2055</td>
<td>.4736</td>
<td>.7035</td>
<td>.8492</td>
<td>.9284</td>
<td>.9680</td>
<td>.9864</td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>.0000</td>
<td>.0005</td>
<td>.0312</td>
<td>.1954</td>
<td>.4606</td>
<td>.6929</td>
<td>.8424</td>
<td>.9245</td>
<td>.9659</td>
<td>.9854</td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>.0000</td>
<td>.0004</td>
<td>.0284</td>
<td>.1856</td>
<td>.4478</td>
<td>.6823</td>
<td>.8354</td>
<td>.9204</td>
<td>.9637</td>
<td>.9843</td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>.0000</td>
<td>.0004</td>
<td>.0289</td>
<td>.1763</td>
<td>.4351</td>
<td>.6716</td>
<td>.8283</td>
<td>.9163</td>
<td>.9615</td>
<td>.9832</td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>.0000</td>
<td>.0003</td>
<td>.0236</td>
<td>.1673</td>
<td>.4226</td>
<td>.6609</td>
<td>.8211</td>
<td>.9120</td>
<td>.9591</td>
<td>.9820</td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>.0000</td>
<td>.0003</td>
<td>.0214</td>
<td>.1587</td>
<td>.4103</td>
<td>.6501</td>
<td>.8137</td>
<td>.9076</td>
<td>.9567</td>
<td>.9807</td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>.0000</td>
<td>.0002</td>
<td>.0195</td>
<td>.1502</td>
<td>.3982</td>
<td>.6394</td>
<td>.8063</td>
<td>.9032</td>
<td>.9542</td>
<td>.9794</td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>.0000</td>
<td>.0002</td>
<td>.0177</td>
<td>.1425</td>
<td>.3863</td>
<td>.6286</td>
<td>.7988</td>
<td>.8986</td>
<td>.9517</td>
<td>.9781</td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>.0000</td>
<td>.0002</td>
<td>.0161</td>
<td>.1350</td>
<td>.3745</td>
<td>.6178</td>
<td>.7912</td>
<td>.8939</td>
<td>.9490</td>
<td>.9767</td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>.0000</td>
<td>.0001</td>
<td>.0146</td>
<td>.1277</td>
<td>.3630</td>
<td>.6070</td>
<td>.7834</td>
<td>.8891</td>
<td>.9463</td>
<td>.9752</td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>.0000</td>
<td>.0001</td>
<td>.0132</td>
<td>.1208</td>
<td>.3517</td>
<td>.5962</td>
<td>.7756</td>
<td>.8842</td>
<td>.9435</td>
<td>.9740</td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>.0000</td>
<td>.0001</td>
<td>.0120</td>
<td>.1142</td>
<td>.3406</td>
<td>.5854</td>
<td>.7677</td>
<td>.8792</td>
<td>.9406</td>
<td>.9721</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>.0000</td>
<td>.0001</td>
<td>.0108</td>
<td>.1079</td>
<td>.3297</td>
<td>.5747</td>
<td>.7598</td>
<td>.8741</td>
<td>.9377</td>
<td>.9705</td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>.0000</td>
<td>.0001</td>
<td>.0098</td>
<td>.1019</td>
<td>.3191</td>
<td>.5640</td>
<td>.7517</td>
<td>.8690</td>
<td>.9346</td>
<td>.9688</td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>.0000</td>
<td>.0001</td>
<td>.0089</td>
<td>.0962</td>
<td>.3086</td>
<td>.5533</td>
<td>.7436</td>
<td>.8637</td>
<td>.9315</td>
<td>.9671</td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>.0000</td>
<td>.0000</td>
<td>.0080</td>
<td>.0908</td>
<td>.2984</td>
<td>.5427</td>
<td>.7354</td>
<td>.8583</td>
<td>.9283</td>
<td>.9653</td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>.0000</td>
<td>.0000</td>
<td>.0072</td>
<td>.0856</td>
<td>.2884</td>
<td>.5321</td>
<td>.7272</td>
<td>.8529</td>
<td>.9250</td>
<td>.9635</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>90</td>
<td>0.02</td>
<td>0.04</td>
<td>0.05</td>
<td>0.06</td>
<td>0.07</td>
<td>0.08</td>
<td>0.09</td>
<td>0.10</td>
<td>0.11</td>
<td>0.12</td>
<td>0.13</td>
</tr>
<tr>
<td>91</td>
<td>0.14</td>
<td>0.15</td>
<td>0.16</td>
<td>0.17</td>
<td>0.18</td>
<td>0.19</td>
<td>0.20</td>
<td>0.21</td>
<td>0.22</td>
<td>0.23</td>
<td>0.24</td>
</tr>
<tr>
<td>92</td>
<td>0.25</td>
<td>0.26</td>
<td>0.27</td>
<td>0.28</td>
<td>0.29</td>
<td>0.30</td>
<td>0.31</td>
<td>0.32</td>
<td>0.33</td>
<td>0.34</td>
<td>0.35</td>
</tr>
<tr>
<td>93</td>
<td>0.36</td>
<td>0.37</td>
<td>0.38</td>
<td>0.39</td>
<td>0.40</td>
<td>0.41</td>
<td>0.42</td>
<td>0.43</td>
<td>0.44</td>
<td>0.45</td>
<td>0.46</td>
</tr>
<tr>
<td>94</td>
<td>0.47</td>
<td>0.48</td>
<td>0.49</td>
<td>0.50</td>
<td>0.51</td>
<td>0.52</td>
<td>0.53</td>
<td>0.54</td>
<td>0.55</td>
<td>0.56</td>
<td>0.57</td>
</tr>
<tr>
<td>95</td>
<td>0.58</td>
<td>0.59</td>
<td>0.60</td>
<td>0.61</td>
<td>0.62</td>
<td>0.63</td>
<td>0.64</td>
<td>0.65</td>
<td>0.66</td>
<td>0.67</td>
<td>0.68</td>
</tr>
<tr>
<td>96</td>
<td>0.69</td>
<td>0.70</td>
<td>0.71</td>
<td>0.72</td>
<td>0.73</td>
<td>0.74</td>
<td>0.75</td>
<td>0.76</td>
<td>0.77</td>
<td>0.78</td>
<td>0.79</td>
</tr>
<tr>
<td>97</td>
<td>0.80</td>
<td>0.81</td>
<td>0.82</td>
<td>0.83</td>
<td>0.84</td>
<td>0.85</td>
<td>0.86</td>
<td>0.87</td>
<td>0.88</td>
<td>0.89</td>
<td>0.90</td>
</tr>
<tr>
<td>98</td>
<td>0.91</td>
<td>0.92</td>
<td>0.93</td>
<td>0.94</td>
<td>0.95</td>
<td>0.96</td>
<td>0.97</td>
<td>0.98</td>
<td>0.99</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>99</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

109
A.3 Bounds on $\Pr(f(W^o) \leq c)$.  

Proposition 4.1.6 gives as upper and lower bounds on $\Pr(f(W^o) \leq c)$ the values $\Pr(T_1(W^o) \leq 2c, T_2(W^o) \leq 3c)$ and $\Pr(T_2(W^o) \leq 2c)$. Equivalently, we may write $\Pr((m^o, M^o) \in B)$ and $\Pr((m^o, M^o) \in A)$ where

$$A = \{(x, y): x \leq 0, y \geq 0 \text{ and } y-x \leq 2c\} \text{ and}$$
$$B = \{(x, y): -2c \leq x \leq 0, 0 \leq y \leq 2c \text{ and } y-x \leq 3c\},$$

where $m^o$ and $M^o$ are as in 1.4.3. 1.4.3(c) permits the evaluation of $\Pr((m^o, M^o) \in \{(x, y): 0 \geq x \geq a, 0 \leq y \leq b\})$. (The conditions $0 \geq x$ and $0 \leq y$ are unnecessary since $m^o \leq 0$ and $M^o \geq 0$.)

$A$ and $B$ can be approximated arbitrarily closely by sets of the form $\{(x, y) = x \geq a, y \leq b\}$. For example

$$A \leq \bigcup_{i=1}^{2n} \{(x, y): -\frac{ic}{n} \leq x \leq -\frac{(i-1)c}{n}; -\infty < y \leq 2c - \frac{(i-1)c}{n}\}$$
$$\cap \{-2c \leq x \leq 0, -\infty < y \leq 0\}.$$ 

$$A \geq \bigcup_{i=1}^{2n-1} \{(x, y): -\frac{ic}{n} \leq x \leq -\frac{(i-1)c}{n}; -\infty < y < 2c - \frac{(i+1)c}{n}\}$$
$$\cap \{-2c \leq x \leq 0, -\infty < y \leq 0\}.$$ 

Each set of the form $\{-\frac{i}{n} c \leq x \leq -\frac{(i-1)}{n} c; -\infty < y < 2c - \frac{(i-1)}{n} c\}$ may be expressed as

$$\{-\frac{i}{n} c \leq x; -\infty < y < 2c - \frac{(i-1)}{n} c\} - \{-\frac{(i-1)}{n} c \leq x \leq 0;$$
$$\cap \{-\infty < y < 2c - \frac{(i-1)c}{n}\},$$

so that 1.4.3(c) can be used to compute its probability. The set $B$ can be approximated in a similar manner.
We give the upper and lower bounds computed in this way for selected values of $c$. The intervals between values of $c$ are chosen to be smaller where the distribution functions are increasing most rapidly. Compare the upper bound for $\Pr(f(W^0) \leq c)$ with table A1 for $n = 40, 60$ and 80.
<table>
<thead>
<tr>
<th>c</th>
<th>Pr(T₂(W₀) ≤ 2c) (lower bound)</th>
<th>Pr(T₁(W₀) ≤ 2c, T₂(W₀) ≤ 3c) (upper bound)</th>
</tr>
</thead>
<tbody>
<tr>
<td>.225</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>.25</td>
<td>.006</td>
<td></td>
</tr>
<tr>
<td>.275</td>
<td>.022</td>
<td></td>
</tr>
<tr>
<td>.3</td>
<td>.054</td>
<td></td>
</tr>
<tr>
<td>.325</td>
<td>.000</td>
<td>.107</td>
</tr>
<tr>
<td>.35</td>
<td>.002</td>
<td>.178</td>
</tr>
<tr>
<td>.375</td>
<td>.009</td>
<td>.263</td>
</tr>
<tr>
<td>.4</td>
<td>.021</td>
<td>.356</td>
</tr>
<tr>
<td>.425</td>
<td>.043</td>
<td>.447</td>
</tr>
<tr>
<td>.45</td>
<td>.075</td>
<td>.536</td>
</tr>
<tr>
<td>.5</td>
<td>.178</td>
<td>.690</td>
</tr>
<tr>
<td>.55</td>
<td>.314</td>
<td>.802</td>
</tr>
<tr>
<td>.6</td>
<td>.465</td>
<td>.879</td>
</tr>
<tr>
<td>.65</td>
<td>.609</td>
<td>.929</td>
</tr>
<tr>
<td>.7</td>
<td>.728</td>
<td>.959</td>
</tr>
<tr>
<td>.75</td>
<td>.821</td>
<td>.977</td>
</tr>
<tr>
<td>.8</td>
<td>.889</td>
<td>.988</td>
</tr>
<tr>
<td>.85</td>
<td>.934</td>
<td>.994</td>
</tr>
<tr>
<td>.9</td>
<td>.963</td>
<td>.997</td>
</tr>
<tr>
<td>.95</td>
<td>.979</td>
<td>.999</td>
</tr>
<tr>
<td>1.00</td>
<td>.989</td>
<td>.999</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


