ASYMPTOTIC THEORY OF RANK TESTS FOR INDEPENDENCE

F.H. RUYMGAART

MATHEMATICAL CENTRE TRACTS 43

MATHEMATICAL CENTRE TRACTS 43

F.H. RUYMGAART

ASYMPTOTIC THEORY OF RANK TESTS FOR INDEPENDENCE

MATHEMATISCH CENTRUM AMSTERDAM 1973

AMS(MOS) subject classification scheme (1970): 62G10

PREFACE

The approach introduced by Chernoff and Savage to prove asymptotic normality under fixed alternatives of rank statistics for the two-sample problem has first been applied, in a suitably modified form, by Bhuchongkul to prove asymptotic normality of rank statistics for testing the hypothesis of independence. In the present tract this method is generalized so that weaker conditions on the limiting score function suffice for the asymptotic normality under fixed continuous bivariate distribution functions, belonging to the alternative. The condition frequently encountered in the literature that the limiting score function, which is a function on the unit square in the independence problem, be a product of functions on the unit interval is abandoned. Apart from this secondary generalization, the main results lie in an essential enlargement of the orders of magnitude of the limiting score function near the boundary of the unit square, and in the fact that this function may exhibit discontinuities on a finite number of lines in the unit square, parallel to the axes. Discontinuity of the limiting score function entails, however, a local differentiability condition on the underlying bivariate distribution function.

The presentation of the material is self-contained. The first three chapters are devoted to the above general problem. In Chapter 4 the results of the preceding chapters are applied to prove asymptotic normality under converging sequences of underlying distribution functions. Moreover, applications to consistency and asymptotic relative efficiency are given in this chapter. The verification of the conditions of the theorems in some important special cases and a comparison with earlier results is given in Chapter 5. Finally, in Chapter 6 attention is payed to discrete underlying distribution functions, entirely concentrated on a finite lattice of points in the plane.

Although this tract exclusively deals with the independence problem in its usual setting, where all random elements have the same bivariate distribution function, the validity of essentially the same techniques in similar situations may be conjectured, e.g. in the case where the random elements have (a fixed finite number of) possibly different distributions. The latter conjectured generalization opens the way to a treatment of k-sample and regression statistics as special cases of statistics for the independence problem.

The technique which is used depends on properties of the joint and marginal empirical distribution functions, also when discontinuous limiting score functions are considered. Hence it remains methodologically most closely related to the approach by Chernoff and Savage or Bhuchongkul. The approach is quite different from a technique used by Pyke and Shorack for k-sample rank statistics, although some of their lemmas are of essential importance. It also differs from the approach by Hájek or Dupač and Hájek, who derived similar results for regression rank statistics.

The problem, leading to this tract, was suggested by Professor G.R. Shorack, from the University of Washington, in the year 1969-1970 he was a visitor at the Mathematisch Centrum. Thanks to his unselfish help an important part of the present work has been finished during his stay at the statistical department. The author is very grateful to him and will not forget this pleasant and stimulating cooperation. The author's sincere thanks go to Professor W.R. van Zwet for his instructive and kind guidance, which has been essential for the realization of this work.

To the members of the statistical department and in particular to Professor J. Hemelrijk, head of the department, the author is obliged for the ideal working conditions. He is indebted to Mr. R. Helmers for introducing him to the subject of nonparametric statistics.

CONTENTS

Chapter 1

INTRODUCTION

1.1. FORMULATION OF THE PROBLEM

The earliest general theorem on asymptotic normality of rank statistics, used for testing the hypothesis of independence, under fixed alternatives is due to Bhuchongkul [4] in 1964. In [4] a class of rank statistics is determined, having a normal distribution in the limit for any underlying continuous bivariate distribution function. It is our purpose to generalize this result and similar ones of more recent date. The present work unifies and extends the papers [33] and [34] by Shorack, van Zwet and the author. In order to arrive at a detailed description of the problem, we shall first have to introduce some basic concepts and notation and exhibit some useful relations between the various concepts.

Let (Ω,A,P) be a probability space on which a pair (X,Y) of random variables (rvs) is defined, having the joint distribution function (df) $H(x,y) = P(\{X \leq x, Y \leq y\})$ and marginal dfs $F(x) = P(\{X \leq x\})$ and $G(y) = P(\{Y \leq y\})$ for all $-\infty < x < \infty$, $-\infty < y < \infty$. Let be given a sequence of mutually independent and identically distributed (iid) random vectors $(X_1,Y_1),(X_2,Y_2),\ldots$, all defined on the above probability space and all possessing the above bivariate df H. To display the underlying df H, the probability measure will occasionally be denoted by P_H rather than P. The problem of testing the hypothesis of independence is to decide on the basis of a random sample of size N from some df H, consisting e.g. of the first N elements $(X_1,Y_1),\ldots,(X_N,Y_N)$ of the above sequence, whether the null hypothesis that $H(x,y) = F(x)G(y)$ for all $-\infty < x < \infty$, $-\infty < y < \infty$ (corresponding to independence of X and Y) should be rejected or not.

Given a positive integer N, for all $-\infty < x < \infty$, $-\infty < y < \infty$ the bivariate empirical df H_N based on the first N random vectors is defined by $NH_N(x,y) =$ [number of (X_n,Y_n) among $(X_1,Y_1),\ldots,(X_N,Y_N)$: $X_n \leq x$, $Y_n \leq y$], and its marginal empirical dfs $F_N(x)$ and $G_N(y)$ by $NF_N(x) =$ [number of X_n among $X_1,\ldots,X_N$: $X_n \leq x$] and $NG_N(y) =$ [number of Y_n among $Y_1,\ldots,Y_N$: $Y_n \leq y$] respectively. For $1 \leq n \leq N$ the rank R_{nN} of X_n is defined as the number of X_m among $X_1,\ldots,X_N$ for which $X_m \leq X_n$, and the rank Q_{nN} of Y_n as the number of Y_m among $Y_1,\ldots,Y_N$ for which $Y_m \leq Y_n$. Before listing some relations between ranks and empirical dfs, let us introduce the function

(1.1.1) $c(z) = 0$ for $z < 0$, $c(z) = 1$ for $z \geq 0$.

Then we obviously have

(1.1.2) $F_N(X_n) = R_{nN}/N$, $G_N(Y_n) = Q_{nN}/N$,

(1.1.3) $H_N(X_n,Y_n) = [\sum_{m=1}^{N} c(R_{nN}-R_{mN})c(Q_{nN}-Q_{mN})]/N$,

$n = 1,\ldots,N$. Furthermore the set of ordered first and second coordinates will be denoted by $X_{1:N} \leq \ldots \leq X_{N:N}$ and $Y_{1:N} \leq \ldots \leq Y_{N:N}$ respectively. For any rv Z the df $\Psi(z) = \Pr(Z \leq z)$ is defined and right-continuous for $-\infty < z < \infty$. Let us define an inverse of this function by

(1.1.4) $\Psi^{-1}(u) = \inf \{z : \Psi(z) \geq u\}$,

for $0 < u \leq 1$. Here by way of exception a function is introduced which may assume an infinite value. According to (1.1.4) $\Psi^{-1}(u)$ is non-decreasing, left-continuous and satisfies $\Psi(\Psi^{-1}(u)) \geq u$ (with equality if and only if Ψ is continuous) for all $0 < u \leq 1$. Furthermore it has the property that $\Psi^{-1}(\Psi(z)) \leq z$ (with equality if and only if Ψ is strictly increasing) for all $-\infty < z < \infty$. Taking F_N and G_N for Ψ we find the relations

(1.1.5) $F_N^{-1}(n/N) = X_{n:N}$, $G_N^{-1}(n/N) = Y_{n:N}$,

$n = 1,\ldots,N$, between order statistics and inverses of empirical dfs.

Although all of the above makes sense for any underlying df H, we shall primarily be concerned with the case where H is restricted to the class $\mathcal{H}$, defined by

(1.1.6) $\mathcal{H} = \{H : H$ is a bivariate df, continuous on $(-\infty,\infty) \times (-\infty,\infty)\}$.

For H in $\mathcal{H}$ the $X_1,Y_1,X_2,Y_2,\ldots$ assume different values with probability 1 and so do the $R_{1N},\ldots,R_{NN}$ and the $Q_{1N},\ldots,Q_{NN}$ so that the notion of order statistic and rank has an unambiguous meaning. In this case the null hypothesis $\mathcal{H}_0$ and the alternative hypothesis $\mathcal{H}_1$ are the subclasses

(1.1.7) $\mathcal{H}_0 = \{H \in \mathcal{H} : H = F \times G\}$, $\mathcal{H}_1 = \mathcal{H} \cap \mathcal{H}_0^c$

of H. Clearly it will be impossible to obtain a uniformly most powerful test in the class of all tests with size not exceeding a given level of significance $0 < \alpha < 1$. In a case like this one may try to find an optimal procedure in a restricted class, and frequently a natural reduction is obtained by invariance. In our case the testing problem remains invariant under the group of transformations $(X_1', Y_1', \ldots, X_N', Y_N') = (f(X_1), g(Y_1), \ldots, f(X_N), g(Y_N))$ such that f and g are continuous and strictly increasing functions on $(-\infty, \infty)$, leaving as maximal invariant the vector of ranks $(R_{1N}, Q_{1N}, \ldots, R_{NN}, Q_{NN})$. Both for the general theory of invariance and for its application in non-parametric statistics the reader is referred to Lehmann [26] and e.g. to Schmetterer [35] and Witting and Nölle [42]. Henceforth let us restrict our attention to tests that are functions of the vector of ranks only, to be called rank tests.

Let us note meanwhile that such rank tests have the desirable property of being similar on the null hypothesis H_0. For let $(i_1, \ldots, i_N)$ and $(j_1, \ldots, j_N)$ be arbitrary permutations of $(1, \ldots, N)$, then

$$P_H(\{R_1 = i_1, Q_1 = j_1, \ldots, R_N = i_N, Q_N = j_N\}) = (N!)^{-2},$$

for any underlying df H in H_0. Because of this property each probability measure corresponding to an underlying df H in H_0 may - and will - be denoted by P_0 without risking ambiguity. In order to avoid randomization, only significance levels $\alpha = n/(N!)^2$, $n = 1, \ldots, (N!)^2 - 1$, will be considered. These levels of significance will be called *natural significance levels*. Note that if α is a natural significance level for a sample of size N, it remains a natural significance level for samples of size larger than N. The critical region of any rank test may be written in the form $\{T_N \geq z\}$ for some $-\infty < z < \infty$ and rank statistic $T_N = T_N(R_{1N}, Q_{1N}, \ldots, R_{NN}, Q_{NN})$. Because T_N may assume fewer than $(N!)^2$ different values, the set of natural significance levels that can actually be attained by this procedure for a fixed choice of T_N and varying z is usually a subset, called the set of *natural significance levels for T_N*. Given a natural significance level α for T_N, the corresponding critical regions are uniquely determined by the numbers $C_{\alpha,N}$, satisfying

$$(1.1.8) \qquad P_0(\{T_N \geq C_{\alpha,N}\}) = \alpha.$$

Even the restricted class of rank procedures does not contain a uniformly most powerful test. In Blum, Kiefer and Rosenblatt [6] a test has been proposed, having reasonable power properties under most of the alternative dfs. This test can be used if nothing is known about the type of dependence to be expected. To describe the statistic in [6], a few conventions concerning integration, maintained throughout, will be introduced. All integrals are Lebesgue-Stieltjes or Lebesgue integrals. In the notation for Lebesgue-Stieltjes integrals over a measurable subset of k-dimensional number space with respect to some (random) measure, we write (in upper case letters such as F, G, H or F_N, G_N, H_N) the symbol for a distribution function corresponding to this (random) measure, instead of the symbol for the (random) measure itself. Ordinary Lebesgue integrals over measurable subsets of k-dimensional number space are always written as repeated integrals, with lower case letters (such as s, t, u or v) at the place of a distribution function corresponding to Lebesgue measure over the real line. Integration should be extended over the entire k-dimensional number space if no domain is indicated. We are now in a position to give the expression for the statistic in [6], which is

$$\int\int [H_N(x,y)-F_N(x)G_N(y)]^2 dH_N(x,y).$$

Because for any function f defined on $(-\infty,\infty) \times (-\infty,\infty)$ we have

$$\int\int f(x,y)dH_N(x,y) = [\textstyle\sum_{n=1}^{N} f(X_n,Y_n)]/N,$$

it is immediate that this statistic is a rank statistic in view of formulas (1.1.2) and (1.1.3). The asymptotic behavior of this statistic is investigated in [6], both under the hypothesis and the alternative. See also Neuhaus [29] and Durbin [9].

One can usually do better if one knows a priori that the dependence is of a special type. For the moment let us mention a well known test, introduced by Kendall (see e.g. Kendall [20]), based on a statistic which is equivalent to

$$\int\int H_N(x,y)dH_N(x,y).$$

This test is intuitively seen to be suitable for testing against positive (or negative) dependence. For the latter concept see Konijn [23], Lehmann [27] and e.g. Yanagimoto [43]. Both statistics considered so far are special cases of statistics of the type

$$\iint \tilde{J}_N(F_N(x), G_N(y), H_N(x,y))\, dH_N(x,y),$$

where $\tilde{J}_N$ is a function defined on $(0,1] \times (0,1] \times (0,1]$. This class of statistics is too general to be dealt with here.

The statistics that will be considered have the form $T_N = [\sum_{n=1}^{N} A_N(R_{nN}, Q_{nN})]/N$, where the numbers $A_N(m,n)$ are defined and finite for $m = 1,\ldots,N$ and $n = 1,\ldots,N$. These numbers are called *scores*. A motivation to consider statistics of this type will be given in Section 1.2, where it is observed that such statistics arise in the search for locally most powerful rank tests against families of alternative dfs depending on a real valued parameter (see Hãjek and Šidãk [17]). As has been noticed in [4], an alternative expression for T_N is

(1.1.9) $$T_N = \iint J_N(F_N(x), G_N(y))\, dH_N(x,y),$$

where the function J_N is defined (and finite) on $(0,1] \times (0,1]$ by

(1.1.10) $$J_N(s,t) = A_N(m,n) \text{ for } (s,t) \in ((m-1)/N, m/N] \times ((n-1)/N, n/N],$$

$m = 1,\ldots,N$ and $n = 1,\ldots,N$. This function will be called the *score function*. Suppose that for each $N = 1,2,\ldots$ such a score function J_N is given, defining a sequence of statistics T_N. It is our aim to find general conditions under which the T_N, after suitable standardization, will converge in distribution to a normal law. For this purpose we introduce a function

(1.1.11) $$J(s,t),$$

defined (and finite) for (s,t) in $(0,1) \times (0,1)$, to be thought of as the limit of the sequence $J_N (N = 1,2,\ldots)$ in a sense to be explained in the next chapters. This function J will therefore be called the *limiting score function*. It is allowed to tend to infinity only when the argument tends to the boundary of the unit square. The rate of growth of J near the boundary of the unit square plays an important role among the conditions for asymptotic normality of the standardized T_N.

Besides this growth condition we consider two different smoothness conditions on the function J. In Chapter 2 asymptotic normality under fixed alternative dfs is obtained for limiting score functions that are continuous throughout the open unit square. A similar result is established in Chapter 3 for limiting score functions that are allowed to have discontinuities of a special kind. In Chapter 4 the results of the preceding chapters are applied to prove asymptotic normality under converging sequences of alternative dfs. Moreover applications to consistency and asymptotic relative efficiency are given in this chapter. The verification of the conditions of these theorems in some important special cases is dealt with in Chapter 5, where in addition our results are compared with earlier ones. Finally in Chapter 6 attention is payed to discrete underlying alternative dfs, entirely concentrated on a finite lattice of points in the plane. As far as the present chapter is concerned, in Section 1.2 some locally most powerful rank tests of the type (1.1.9) are introduced and Section 1.3 presents a survey of the basic machinery for non-parametric asymptotic theory.

To conclude this section we give an alternative expression for the statistic T_N in (1.1.9), which is analogous to the expression for the two-sample rank statistic given by Pyke and Shorack [31]. First let us introduce the bivariate empirical df $\bar{H}_N$, based on the vectors of ranks of $(X_1,Y_1),\ldots,(X_N,Y_N)$ and defined for all $0 < s \leq 1$, $0 < t \leq 1$ by $N\bar{H}_N(s,t)$ = [number of (R_{nN},Q_{nN}) among $(R_{1N},Q_{1N}),\ldots,(R_{NN},Q_{NN})$: $R_{nN}/N \leq s, Q_{nN}/N \leq t$]. Note that the corresponding random measure puts mass $1/N$ at each of the points $(R_{nN}/N, Q_{nN}/N)$, $n = 1,\ldots,N$. From (1.1.5) we see that

$$\bar{H}_N(m/N,n/N) = H_N(F_N^{-1}(m/N),G_N^{-1}(n/N)),$$

for $m = 1,\ldots,N$ and $n = 1,\ldots,N$. In Figure 1.1.1 the function F_N^{-1} and the points where $\bar{H}_N$ puts mass $1/N$ are drawn for $N = 5$. We may think of $\bar{H}_N$ as a

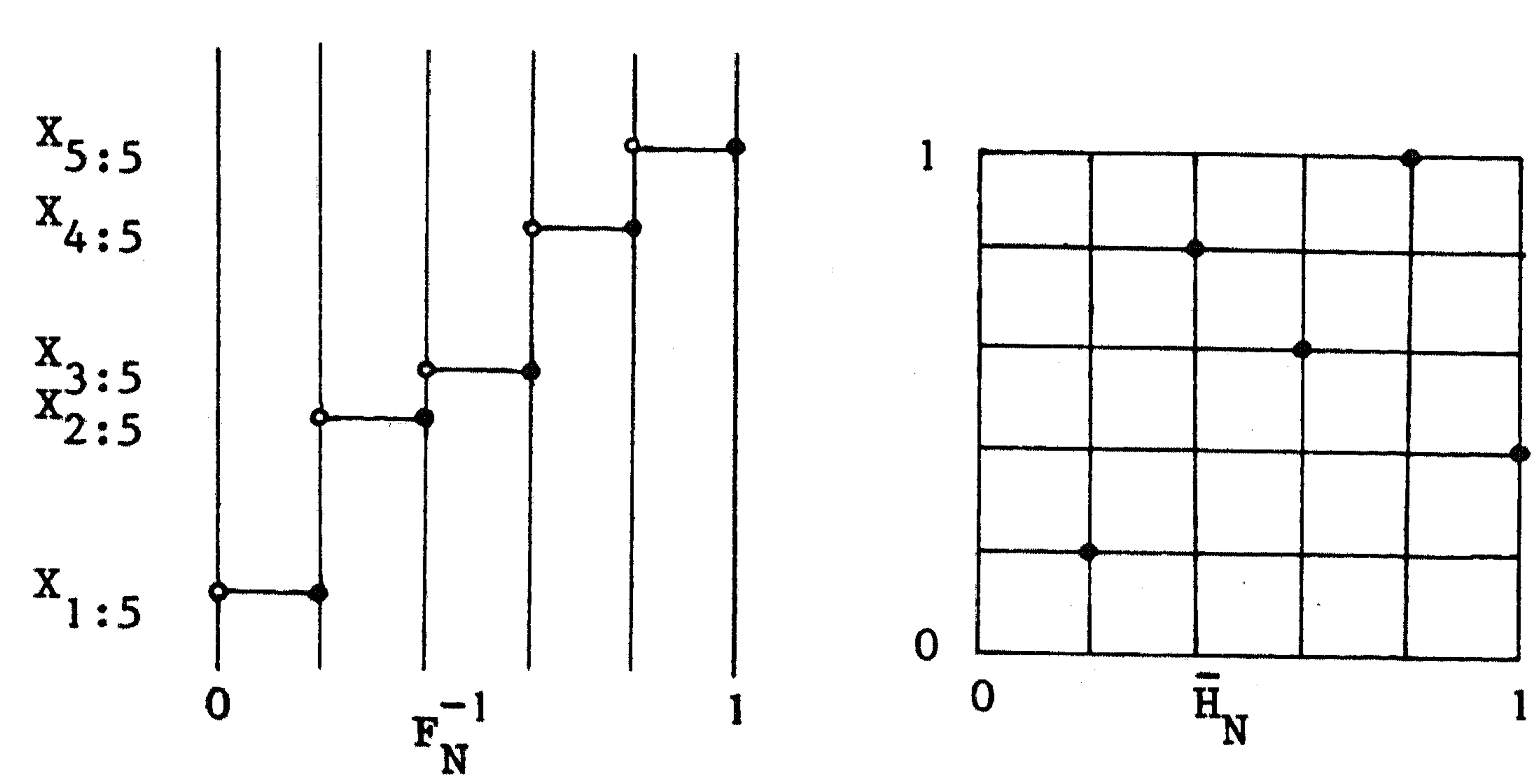

Figure 1.1.1

scalefree version of H_N. The corresponding random measure is concentrated on $(0,1] \times (0,1]$, and the measure of the interval $((m-1)/N,m/N] \times ((n-1)/N,n/N]$ equals $\bar{H}_N(m/N,n/N) - \bar{H}_N((m-1)/N,n/N) + \bar{H}_N((m-1)/N,(n-1)/N) - \bar{H}_N(m/N,(n-1)/N)$ for $m = 1,\ldots,N$ and $n = 1,\ldots,N$. Combination of the above with (1.1.9), (1.1.10) and the continuity of H yields that the statistic may be written as

$$T_N = \iint_{(0,1]\times(0,1]} J_N(s,t)d\bar{H}_N(s,t)$$

$$= \sum_{m=1}^{N-1} \sum_{n=1}^{N-1} \bar{H}_N(m/N,n/N)[J_N((m+1)/N,(n+1)/N) -$$

$$J_N(m/N,(n+1)/N) + J_N(m/N,n/N) - J_N((m+1)/N,n/N)] + c_N.$$

The constant c_N arises from terms involving only the function J_N and the non-random values $\bar{H}_N(m/N,0) = 0$, $\bar{H}_N(m/N,1) = m/N$, $\bar{H}_N(0,n/N) = 0$, $\bar{H}_N(1,n/N) = n/N$ for $m = 1,\ldots,N$ and $n = 1,\ldots,N$. Hence instead of T_N we may consider the equivalent statistic $T_N - c_N$. Let ν_N be the signed measure on $(0,1) \times (0,1)$ associated with the (left-continuous) function J_N, i.e. for $0 < s_1 < s_2 < 1$, $0 < t_1 < t_2 < 1$ we have $\nu_N([s_1,s_2) \times [t_1,t_2)) = J_N(s_2,t_2) - J_N(s_1,t_2) + J_N(s_1,t_1) - J_N(s_2,t_1)$. Writing dJ_N for integration with respect to ν_N, we find from the definition of $\bar{H}_N$ that

$$\text{(1.1.12)} \qquad T_N - c_n = \iint_{(0,1)\times(0,1)} \bar{H}_N(s,t)dJ_N(s,t).$$

It would be interesting to see whether results comparable to those of Chapters 2-4 can be obtained with (1.1.12) as a starting point. One would have to study the weak convergence of the suitably standardized processes $\bar{H}_N$ and the rate of growth of these processes near the boundary of the unit square.

1.2. LOCALLY MOST POWERFUL RANK TESTS

In the previous section we have alluded to the fact that rank statistics which provide locally most powerful rank tests against families of alternative dfs depending on a single real valued parameter exist, and are of the type introduced in (1.1.9). Of course it may be hard to decide on such a special family of alternative dfs. Even when the dependence is known to be of a special kind such as positive or circular dependence, the restriction to a real valued one parameter family of alternative dfs displaying this particular dependence remains an arbitrary matter. We shall not dwell on this problem, since we merely wish to give examples of situations in which statistics of the type (1.1.9) are desirable.

It turns out that in general for such *locally most powerful rank tests* a limiting score function J can be used that is determined by this special family of alternative dfs and that, moreover, generates the score functions J_N in a natural way. Therefore we shall in this section speak of the *scores generating function* J rather than of the limiting score function. To be more explicit let us put, for brevity,

(1.2.1) $$b_{\nu_1,\nu_2}(u) = \Gamma(\nu_1+\nu_2)[\Gamma(\nu_1)\Gamma(\nu_2)]^{-1}u^{\nu_1-1}(1-u)^{\nu_2-1},$$

for $0 < u < 1$ and $b_{\nu_1,\nu_2}(u) = 0$ for u in $(-\infty,0] \cup [1,\infty)$. Here Γ denotes the gamma-function and $\nu_1 > 0$ and $\nu_2 > 0$ are given constants. This function is the beta-density with parameters ν_1 and ν_2, and $b_{n,N-n+1}$ is the density of the n-th order statistic of a sample of size N from the uniform distribution on (0,1). Then for measurable and square integrable functions J, the score functions of the locally most powerful rank test are in general obtained by

(1.2.2) $$J_N(s,t) = \int_0^1\int_0^1 J(u,v)b_{m,N-m+1}(u)b_{n,N-n+1}(v)du\, dv,$$

for (s,t) in $((m-1)/N,m/N] \times ((n-1)/N,n/N]$, $m = 1,\dots,N$ and $n = 1,\dots,N$. Score functions J_N obtained from J according to (1.2.2) are called *exact score functions derived from J (for the sample size N)*. In practice exact score functions are usually hard to calculate. For this reason one frequently uses the score functions

(1.2.3) $$J_N(s,t) = J(m/(N+1),n/(N+1)),$$

for (s,t) in $((m-1)/N,m/N] \times ((n-1)/N,n/N]$, $m = 1,\dots,N$ and $n = 1,\dots,N$, which are called *approximate score functions derived from J (for the sample size N)*. When using approximate score functions instead of exact ones, in most cases the corresponding test is no longer locally most powerful for the sample size N, although asymptotic optimality is maintained.

By way of examples some families of alternative dfs depending on a single real parameter will be considered. Throughout this section X_0 and Y_0

are understood to be mutually independent rvs with known joint df $H_0 = F_0 \times G_0$ in $\mathcal{H}_0$. We shall avoid a complete enumeration of the conditions that F_0 and G_0 have to satisfy and confine ourselves to mentioning that these marginal dfs should possess densities f_0 and g_0 with almost everywhere continuous derivatives f_0' and g_0' respectively.

Let us start with a family of dfs exhibiting *positive dependence* and given by

$$(1.2.4) \qquad \{H_\theta \in \mathcal{H} : H_\theta \text{ is the df of } (X_0+\theta Z, Y_0+\theta Z),\ 0 \leq \theta < \infty\}.$$

Here Z is an arbitrary rv with $0 < \mathrm{Var}(Z) < \infty$, which is independent of both X_0 and Y_0. This class is considered in Hājek and Šidāk [17] and is essentially the same as the one introduced by Bhuchongkul [4]. In [17] it is shown that the scores generating function associated with the family (1.2.4) is of product type and given by

$$(1.2.5) \qquad J(s,t) = [-f_0'(F_0^{-1}(s))/f_0(F_0^{-1}(s))][-g_0'(G_0^{-1}(t))/g_0(G_0^{-1}(t))]$$

for $0 < s < 1$, $0 < t < 1$. The following four examples refer to the model (1.2.4).

<u>EXAMPLE 1.2.1</u> (Double exponential distribution). Suppose $f_0(z) = g_0(z) = 2^{-1} e^{-|z|}$ for $-\infty < z < \infty$. Then the scores generating function is equal to

$$J(s,t) = \mathrm{sgn}(s-1/2)\mathrm{sgn}(t-1/2).$$

Here sgn(z) is the function defined by

$$(1.2.6) \qquad \mathrm{sgn}(z) = -1 \text{ for } z < 0,\ \mathrm{sgn}(0) = 0,\ \mathrm{sgn}(z) = 1 \text{ for } z > 0,$$

for all $-\infty < z < \infty$. Utilizing approximate score functions the so called quadrant test is obtained. Note that J is bounded but discontinuous on the lines $s = 1/2$ and $t = 1/2$.

<u>EXAMPLE 1.2.2</u> (Logistic distribution). For $f_0(z) = g_0(z) = e^{-z}(1+e^{-z})^{-2}$, $-\infty < z < \infty$, the scores generating function equals

$$J(s,t) = (2s-1)(2t-1).$$

For exact as well as approximate score functions we arrive at a statistic which is equivalent to Spearman's correlation coefficient. Note that J is bounded and continuous.

EXAMPLE 1.2.3 (Normal distribution). When $f_0(z) = g_0(z) = (2\pi)^{-1/2}e^{-z^2/2}$, $-\infty < z < \infty$, the scores generating function is

$$J(s,t) = \Phi^{-1}(s)\Phi^{-1}(t),$$

where Φ is the standard normal df. The corresponding test is called the Fisher-Yates or normal scores test. Note that J is continuous but no longer bounded.

EXAMPLE 1.2.4 (Beta-distribution). Suppose $f_0(z) = g_0(z) = b_{\nu,\nu}(z)$, $-\infty < z < \infty$, where $b_{\nu,\nu}$ is given by (1.2.1) with $\nu_1 = \nu_2 = \nu > 0$ and $\nu \neq 1$. We shall not derive the scores generating function explicitly but give a bound and specify its behavior near the boundary of the unit square. We have $J(s,t) = K(s)K(t)$, where for some finite constants $D > 0$ and $c > 0$

$$|J(s,t)| \leq D[s(1-s)t(1-t)]^{-1/\nu},$$

$$\lim_{s\downarrow 0}|K(s)|[s(1-s)]^{1/\nu} = \lim_{s\uparrow 1}|K(s)|[s(1-s)]^{1/\nu} = c.$$

Note that for any $\nu > 0$, $\nu \neq 1$ the growth of J near the boundary of the unit square is essentially larger than in Example 1.2.3. If $\nu > 2$ ($\nu > 4$) the function $|J|^{2+\delta}$ ($|J|^{4+\delta}$) is integrable over the unit square for some positive number δ.

Our next example is a family, also exhibiting *positive dependence*, determined by

(1.2.7) $\{H_\theta \in \mathcal{H} : H_\theta$ is the df of $((1-\theta)X_0+\theta Y_0, \theta X_0+(1-\theta)Y_0), 0 \leq \theta < 1/2\}$.

This family is a special case of a class considered by Konijn [22]. By means of a technique explained in Lehmann [26], page 237 (see also Lehmann [25], or [42]), one can show that the scores generating function associated with the family (1.2.7) is a sum of functions of product type $2 + [F_0^{-1}(s)+G_0^{-1}(t)-2] \times [-f_0'(F_0^{-1}(s))/f_0(F_0^{-1}(s)) - g_0'(G_0^{-1}(t))/g_0(G_0^{-1}(t))]$, for $0 < s < 1$, $0 < t < 1$.

Equivalent statistics may be obtained from the scores generating function

$$(1.2.8) \qquad J(s,t) = F_0^{-1}(s)[-g_0'(G_0^{-1}(t))/g_0(G_0^{-1}(t))] + [-f_0'(F_0^{-1}(s))/f_0(F_0^{-1}(s))]G_0^{-1}(t),$$

for $0 < s < 1$, $0 < t < 1$. The next three examples refer to the model (1.2.7).

EXAMPLE 1.2.5 (Double exponential distribution). Suppose F_0 and G_0 are as in Example 1.2.1. Then the scores generating function in (1.2.8) equals

$$J(s,t) = \lambda(s)[\mathrm{sgn}(t-1/2)] + [\mathrm{sgn}(s-1/2)]\lambda(t).$$

Here the function sgn(z) is defined in (1.2.6) and

$$\lambda(s) = \log 2s \text{ for } 0 < s < 1/2, \ \lambda(s) = -\log 2(1-s) \text{ for } 1/2 \le s < 1.$$

The function J is discontinuous on the lines $s = 1/2$ and $t = 1/2$ and it is unbounded.

EXAMPLE 1.2.6 (Logistic distribution). If F_0 and G_0 are as in Example 1.2.2 we find for the scores generating function in (1.2.8)

$$J(s,t) = [\log(s/(1-s))](2t-1)+(2s-1)[\log(t/(1-t))].$$

This function is unbounded but continuous.

EXAMPLE 1.2.7 (Normal distribution). For F_0 and G_0 as in Example 1.2.3 the scores generating function in (1.2.8) equals

$$J(s,t) = 2\Phi^{-1}(s)\Phi^{-1}(t),$$

where Φ is the standard normal df. In this case the scores generating function is of product type and equivalent to the one of Example 1.2.3.

Let us further consider the family of dfs

$$(1.2.9)\qquad \{H_\theta \in \mathcal{H} : H_\theta(x,y) = c(\theta)\int_{-\infty}^{x}\int_{-\infty}^{y}\exp(-[(\xi^2+\eta^2)^{1/2}-\theta]^2/2)\,d\xi d\eta,\quad 0 \leq \theta < \infty\}.$$

Here $c(\theta) = \{2\pi[\exp(-\theta^2/2)+(2\pi)^{1/2}\theta\Phi(\theta)]\}^{-1}$, where Φ is the standard normal df. This family displays a *circular dependence*. Note that H_0 is the bivariate normal df with standard normal marginal dfs and zero correlation coefficient. With the aid of the same technique as was used for the family (1.2.7) we find that an equivalent limiting scores generating function, associated with the model (1.2.9), equals

$$(1.2.10)\qquad J(s,t) = [(\Phi^{-1}(s))^2 + (\Phi^{-1}(t))^2]^{1/2},$$

for $0 < s < 1$, $0 < t < 1$. Here J is no longer a finite sum of functions of product type.

All families of alternative dfs considered so far are families of *parametric alternatives*. Let us now indicate how a family of so called *Lehmann* - or *non-parametric alternatives* may be constructed from each family of parametric alternatives. To each df H in $\mathcal{H}$ there corresponds the df $\bar{H}$, defined by

$$(1.2.11)\qquad \bar{H}(s,t) = H(F^{-1}(s),G^{-1}(t)) = P(\{F(X) \leq s, G(Y) \leq t\}),$$

$0 < s < 1$, $0 < t < 1$ (for X and Y see the beginning of Section 1.1). This df concentrates mass 1 on $(0,1) \times (0,1)$ where, moreover, it is continuous. Furthermore $\bar{H}(s,1) = s$, $0 < s < 1$, and $\bar{H}(1,t) = t$, $0 < t < 1$, i.e. the marginals are uniform distributions over (0,1). Under suitable regularity conditions on H, the transformed df $\bar{H}$ possesses a density equal to

$$(1.2.12)\qquad \bar{h}(s,t) = h(F^{-1}(s),G^{-1}(t))[f(F^{-1}(s))g(G^{-1}(t))]^{-1},$$

for $0 < s < 1$, $0 < t < 1$. Here h is the density of the bivariate df H and f and g are the densities of the marginal dfs F and G respectively (see also [43]). Note that for any df $H \in \mathcal{H}_0$ the transformed df equals

$$(1.2.13)\qquad \bar{H}(s,t) = st,$$

for $0 < s < 1$, $0 < t < 1$. By way of a further example, let

$$\Phi_\theta(x,y) = [2\pi(1-\theta^2)^{1/2}]^{-1}\int_{-\infty}^{x}\int_{-\infty}^{y} \exp(-[2(1-\theta^2)]^{-1}(\xi^2-2\theta\xi\eta+\eta^2))d\xi d\eta,$$

$-\infty < x < \infty$, $-\infty < y < \infty$, $-1 < \theta < 1$, be the bivariate normal df with standard normal marginal dfs Φ and correlation coefficient θ. Then we have

$$(1.2.14) \qquad \bar{\Phi}_\theta = \Phi_\theta(\Phi^{-1},\Phi^{-1})$$

on $(0,1) \times (0,1)$.

Suppose that $\{H_\theta, 0 \leq \theta < \theta_0\}$ is a family of parametric alternatives, such as (1.2.4), (1.2.7) or (1.2.9). With the aid of the dfs $\bar{H}_\theta$ defined in (1.2.11) the family of non-parametric alternatives

$$\{H_\theta^* \in \mathcal{H} : H_\theta^* = \bar{H}_\theta(F^*,G^*), F^* \times G^* \in \mathcal{H}_0, 0 \leq \theta < \theta_0\}$$

may be constructed. The marginal dfs of H_θ^* are F^* and G^* for all $0 \leq \theta < \theta_0$ and, given a fixed parameter value θ, the probability distribution of the vector of ranks $(R_{1N},Q_{1N},\dots,R_{NN},Q_{NN})$ is constant on the orbit of all dfs $\bar{H}_\theta(F^*,G^*)$, where $F^* \times G^*$ runs through $\mathcal{H}_0$. This property entails the existence of a locally most powerful rank test, to be derived from the same scores generating function which is associated with the original parametric family. We conclude this section with two examples of such non-parametric alternatives. For further information on these alternatives the reader is referred to Lehmann [25] or Witting and Nölle [42] and Behmen [2,3].

In the first place let us consider the family

$$(1.2.15) \qquad \{H_\theta^* \in \mathcal{H} : H_\theta^* = \bar{\Phi}_\theta(F^*,G^*), F^* \times G^* \in \mathcal{H}_0 \text{ and } -1 < \theta < 1\},$$

where $\bar{\Phi}_\theta$ is given in (1.2.14). When restricting the parameter θ to $0 \leq \theta < 1$ (positive dependence), the associated scores generating function equals

$$(1.2.16) \qquad J(s,t) = \Phi^{-1}(s)\Phi^{-1}(t).$$

This function is the same as found in Example 1.2.3 and equivalent to the one of Example 1.2.7.

Secondly consider the family

$$(1.2.17) \qquad \{H_\theta^* \in \mathcal{H} : H_\theta^* = F^*G^*[1+\theta(1-F^*)(1-G^*)], F^* \times G^* \in \mathcal{H}_0 \text{ and } -1 < \theta < 1\},$$

which has been introduced by Gumbel [13]. When restricting the parameter θ to $0 \leq \theta < 1$ (positive dependence), the associated scores generating function is

(1.2.18) $\quad J(s,t) = (2s-1)(2t-1),$

which happens to be the same as in Example 1.2.2. This fact permits an other way of looking at Spearman's rank statistic.

The above examples illustrate the variety of scores generating functions that may occur. Rapid growth near the boundary of the unit square can in particular be found in Example 1.2.4, whereas Examples 1.2.1 and 1.2.5 are typical for the kind of discontinuities that will be allowed for J. In principle asymptotic normality can be shown in each of the above examples when approximate scores are used, provided the growth of J is not too fast. When using exact scores we have to restrict ourselves to continuous J. In Section 5.3 we shall briefly return to these examples.

1.3. SOME FUNDAMENTAL TOOLS

In this section we collect some basic tools for non-parametric asymptotic theory. At the same time we shall make some further notational conventions and introduce two important classes of functions. Throughout, the notation

(1.3.1) $\quad \mathcal{U}n(0,1),\ \mathcal{B}i(N,\pi),\ \mathcal{N}(\mu,\sigma^2)$

will be used for the uniform distribution over (0,1), the binomial distribution based on N trials with success probability $0 \leq \pi \leq 1$ and the normal distribution with mean $-\infty < \mu < \infty$ and variance $0 \leq \sigma^2 < \infty$ respectively. Here $\mathcal{N}(\mu,0)$ should be interpreted as a distribution, degenerated at the point μ.

The order symbols $\mathcal{O}_P$ and o_P have the following meaning. Let T be an arbitrary set and suppose $Z_{N,\theta}$ $(N = 1,2,\ldots)$ is a sequence of rvs depending on $\theta \in T$ for all N, and defined on (Ω,A,P) where P may be a member of a class P of probability measures on (Ω,A). Furthermore let $\psi(N,\theta)$ be a real valued function defined for all positive integers N and all $\theta \in T$. Then the statements

$$Z_{N,\theta} = O_P(\psi(N,\theta)) \text{ as } N \to \infty, \text{ uniformly on } T \times P,$$

(1.3.2) or

$$Z_{N,\theta} = o_P(\psi(N,\theta)) \text{ as } N \to \infty, \text{ uniformly on } T \times P$$

mean that for each $\varepsilon > 0$ there exist an index N_ε and a finite number M_ε, both independent of θ and P, such that

$$P(\{|Z_{N,\theta}| \leq M_\varepsilon \psi(N,\theta)\}) \geq 1-\varepsilon,$$

(1.3.3) or

$$P(\{|Z_{N,\theta}| \leq \varepsilon\psi(N,\theta)\}) \geq 1-\varepsilon$$

for all $N \geq N_\varepsilon$ and all $(\theta,P) \in T \times P$, respectively. If θ or P are fixed, i.e. $T = \{\theta\}$ or $P = \{P\}$, there is of course no need to mention uniformity on T or P nor to require independence of θ or P in the above definitions. To indicate convergence in distribution, convergence in probability, and almost sure convergence, the symbols

$$\longrightarrow_d \quad ,\longrightarrow_P \quad ,\longrightarrow_{a.s.}$$

will be used. Clearly the statement $Z_{N,\theta} = o_P(1)$ as $N \to \infty$ (uniformly on $T \times P$) is equivalent to $Z_{N,\theta} \to_P 0$ as $N \to \infty$ (uniformly on $T \times P$).

Furthermore let us make some remarks on conditional expectations. Most conditional expectations we deal with involve conditioning on F(X) or G(Y) (for X and Y see the beginning of Section 1.1). Let ϕ and ψ be functions defined and measurable on (0,1) and such that all integrals below are finite. Then the conditional expectation of $\psi(G(Y))$ given F(X) and of $\phi(F(X))$ given G(Y) will, as usual, be denoted by

$$E(\psi(G(Y))|F(X)),\ E(\phi(F(X))|G(Y))$$

respectively. Frequently conditional expectations will be used to write a double integral as a repeated integral (see e.g. Lehmann [26], page 47). In connection with this let us recall the definition of $\bar{H}$ given in (1.2.11) and note that we have

$$(1.3.4)\qquad \iint \phi(F(x))\psi(G(y))dH(x,y) = \int_0^1\int_0^1 \phi(s)\psi(t)d\bar{H}(s,t)$$

$$= \int_0^1 \phi(s)E(\psi(G(Y))|F(X) = s)ds = \int_0^1 E(\phi(F(X))|G(Y) = t)\psi(t)dt.$$

A similar notation will be used for conditional probabilities.

Let us now return to the sequence of random vectors $(X_1,Y_1),(X_2,Y_2),\ldots$ and the basic concepts defined with the aid of this sequence in the beginning of Section 1.1. It is of general importance to note that the set

$$(1.3.5)\qquad \Omega_0 = \{\omega \in \Omega : F_N(F^{-1}(F(x))) = F_N(x), G_N(G^{-1}(G(y))) = G_N(y),$$
$$\text{for all } -\infty < x < \infty,\ -\infty < y < \infty \text{ and } N = 1,2,\ldots\}$$

has probability $P_H(\Omega_0) = 1$ for all $H \in \mathcal{H}$. It follows from this property of the set Ω_0 that the random functions $F_N(F^{-1})$ and $G_N(G^{-1})$ are with probability 1 the empirical dfs of the sets of independent $\mathcal{U}n(0,1)$ distributed rvs $F(X_1),\ldots,F(X_N)$ and $G(Y_1),\ldots,G(Y_N)$ respectively, for each H in $\mathcal{H}$. The processes $U_N(s)$ and $V_N(t)$, defined for $0 \leq s \leq 1$ and $0 \leq t \leq 1$ by

$$(1.3.6)\qquad U_N(s) = N^{1/2}[F_N(F^{-1}(s))-s],\ V_N(t) = N^{1/2}[G_N(G^{-1}(t))-t],$$

ill be called *(the marginal) empirical processes*. These processes satisfy

$$(1.3.7)\qquad U_N(F) = N^{1/2}(F_N-F),\ V_N(G) = N^{1/2}(G_N-G)$$

on $(-\infty,\infty)$ for all ω in Ω_0, i.e. with probability 1. Let W be a Brownian bridge, i.e. a normal process $\{W(u) : 0 \leq u \leq 1\}$ with all sample pathes continuous, $E(W(u)) = 0$ for $0 \leq u \leq 1$ and covariance function $E(W(u)W(v)) = \min\{u,v\} - uv$ for $0 \leq u \leq 1$, $0 \leq v \leq 1$. It is well known that

$$U_N \to_d W,\ V_N \to_d W,$$

as $N \to \infty$. A proof can be found e.g. in Billingsley [5].

The next lemmas concerning the marginal empirical processes or dfs will prove essential for controling the growth of the limiting score function J near the boundary of the unit square. They can be found either in Pyke and Shorack [31] or in Shorack [39], Appendix. To formulate the first lemma we have to introduce a special class of functions.

DEFINITION 1.3.1 (Shorack [39]). Let $\mathcal{Q}$ denote the class of all functions q defined and continuous on [0,1], which are positive on (0,1), symmetric about 1/2, increasing on (0,1/2] and for which $\int_0^1 [q(u)]^{-2} du < \infty$. The members of $\mathcal{Q}$ will be called *q-functions*.

Important examples of elements of $\mathcal{Q}$ can be obtained from the function

$$(1.3.8) \qquad R(u) = [u(1-u)]^{-1},$$

for $0 < u < 1$. For any $0 < \delta < 1/2$ the function $R^{-1/2+\delta}$ is an element of $\mathcal{Q}$.

LEMMA 1.3.1 (Pyke and Shorack [31]). Let U_N and V_N be the marginal empirical processes defined in (1.3.7) and let $q \in \mathcal{Q}$. Then, as $N \to \infty$, $\sup_{(-\infty,\infty)} |U_N(F)[q(F)]^{-1}| = \mathcal{O}_P(1)$ and $\sup_{(-\infty,\infty)} |V_N(G)[q(G)]^{-1}| = \mathcal{O}_P(1)$, uniformly for $H \in \mathcal{H}$.

PROOF. See Pyke and Shorack [31], Lemma 2.2. □

As has been remarked in Shorack [39], the next lemma is especially useful in connection with the following class of functions (see also Lemma 2.3.2 and the beginning of Section 2.4).

DEFINITION 1.3.2 (Shorack [39]). A function r, defined and positive on (0,1), which is symmetric about 1/2, will be called *u-shaped* if it is decreasing on (0,1/2]. If $0 < \beta < 1$ we introduce the notation r_β for the function satisfying $r_\beta(u) = r(\beta u)$ for $0 < u \le 1/2$ and $r_\beta(u) = r(1-\beta(1-u))$ for $1/2 < u < 1$. If for all $\beta > 0$ in a neighborhood of 0 there exists a constant M_β such that $r_\beta \le M_\beta r$ on (0,1), then r will be called a *reproducing u-shaped function*. The class of all reproducing u-shaped functions will be denoted by $\mathcal{R}$.

Typical examples of reproducing u-shaped functions are DR^τ, where $D \ge 0$ and $\tau \ge 0$ are arbitrary constants and where R is the function defined in (1.3.8). Obviously these functions are u-shaped, and for any $0 < \beta < 1$ and for $0 < u \le 1/2$ we have $r_\beta(u) = D[R(\beta u)]^\tau = D(\beta u)^{-\tau}(1-\beta u)^{-\tau} \le D\beta^{-\tau}[R(u)]^\tau = D\beta^{-\tau} r(u)$.

LEMMA 1.3.2 (Shorack [39]). Given an arbitrary $\varepsilon > 0$ there exists a $0 < \beta = \beta_\varepsilon < 1$ such that $P(\{\beta F(x) \leq F_N(x) \leq 1-\beta(1-F(x))$, for $x \in \{0 < F_N < 1\}\}) \geq 1-\varepsilon$ and $P(\{\beta G(y) \leq G_N(y) \leq 1-\beta(1-G(y))$, for $y \in \{0 < G_N < 1\}\}) \geq 1-\varepsilon$, for all $N = 1,2,\ldots$ and independently of $H \in \mathcal{H}$.

PROOF. See Shorack [38]. □

Roughly speaking the following two lemmas are needed to justify replacing certain integrals with respect to dH_N by integrals with respect to dH. It will turn out that this transition is a rather delicate one if the limiting score function J is allowed to possess a certain type of discontinuities. Let Z be a $\mathcal{B}i(N,\pi)$ distributed rv. In Hoeffding [18] it has been shown that for any N, $0 \leq \pi \leq 1$ and $\rho \geq 0$

$$(1.3.9) \qquad \Pr(|Z-N\pi| \geq N\rho) \leq 2\exp(-2N\rho^2).$$

We shall use (1.3.9) in the proof of Lemma 3.3.1. We shall also need a more sophisticated result of a similar type of Kiefer [21].

LEMMA 1.3.3 (Kiefer [21]). For any $\zeta > 0$ there is a finite constant M_ζ such that $P_H(\{\sup_{-\infty<x<\infty,-\infty<y<\infty}|H_N(x,y)-H(x,y)| \geq \rho\}) \leq M_\zeta \exp(-(2-\zeta)N\rho^2))$, for all $N = 1,2,\ldots$, all $\rho \geq 0$ and all bivariate dfs H (continuous or not).

PROOF. See Kiefer [21], Theorem 1-m. □

The next result is a corollary to the preceding lemma and is due to van Zwet [44]. Like Kiefer's theorem the result can be formulated for m-dimensional empirical dfs. To avoid additional notational conventions we shall restrict attention to the case where m=2. For a comparison between Lemma 1.3.4 and related results of Bahadur [1], Sen [36] and Ghosh [12] see Section 5.2. For any Borel set D in the plane we shall write $\iint_D dH = H\{D\}$ and $\iint_D dH_N = H_N\{D\}$. By an interval I in the plane the product set of two intervals on the line will be understood.

LEMMA 1.3.4 (van Zwet [44]). Let $I_1, I_2, \ldots$ be a sequence of intervals in the plane and let $\mathcal{I}_N = \{I_N^* : I_N^*$ is an interval contained in $I_N\}$, $N = 1,2,\ldots$ Then, as $N \to \infty$, $\sup_{I_N^* \in \mathcal{I}_N}|H_N\{I_N^*\}-H\{I_N^*\}| = O_P([H\{I_N\}/N]^{1/2})$, uniformly in all sequences of intervals $I_1, I_2, \ldots$ and all bivariate dfs H (continuous or not).

PROOF. Given any $0 < \varepsilon < 1$, the existence of a number $M = M_\varepsilon$, independent of the df H and the particular sequence $I_1, I_2, \ldots$, must be established such that

$$(1.3.10) \qquad P_H(\{\sup_{I_N^* \in \mathcal{I}_N} |H_N\{I_N^*\} - H\{I_N^*\}| \geq M[H\{I_N\}/N]^{1/2}\}) \leq \varepsilon,$$

for all N, all sequences $I_1, I_2, \ldots$ and all bivariate dfs H. If $H\{I_N\} = 0$ the lemma follows immediately. It proves to be convenient to consider the cases $0 < H\{I_N\} \leq 8/(\varepsilon N)$ and $H\{I_N\} > 8/(\varepsilon N)$ separately.

First suppose that $0 < H\{I_N\} \leq 8/(\varepsilon N)$ and choose $M = M_\varepsilon = (2/\varepsilon)^{3/2}$. It is always true that $\sup_{I_N^* \in \mathcal{I}_N} |H_N\{I_N^*\} - H\{I_N^*\}| \leq \max\{H_N\{I_N\}, H\{I_N\}\}$. By our choice of M we have $M[H\{I_N\}/N]^{1/2} \geq H\{I_N\}/\varepsilon$. Since $NH_N\{I_N\}$ is a $\mathcal{Bi}(N, H\{I_N\})$ distributed rv, application of Markov's inequality shows that the left side in (1.3.10) is bounded above by $P(\{\max\{H_N\{I_N\}, H\{I_N\}\} \geq H\{I_N\}/\varepsilon\}) = P(\{H_N\{I_N\} \geq H\{I_N\}/\varepsilon\}) \leq \varepsilon$.

Next we suppose that $H\{I_N\} > 8/(\varepsilon N)$. Then for $n = 0, 1, \ldots, N$ we may define the conditional probabilities

$$\pi(n) = P(\{\sup_{I_N^* \in \mathcal{I}_N} |H_N\{I_N^*\} - H\{I_N^*\}| \geq M[H\{I_N\}/N]^{1/2}\} \mid \{H_N\{I_N\} = n/N\}).$$

The probability on the left in (1.3.10) can now be written as

$$(1.3.11) \qquad \sum_{n < NH\{I_N\}/2} \pi(n) P(\{H_N\{I_N\} = n/N\}) + \sum_{n \geq NH\{I_N\}/2} \pi(n) P(\{H_N\{I_N\} = n/N\}).$$

By the Bienaymé-Chebyshev inequality we have

$$(1.3.12) \qquad \sum_{n < NH\{I_N\}/2} \pi(n) P(\{H_N\{I_N\} = n/N\}) \leq P(\{H_N\{I_N\} < H\{I_N\}/2\}) \leq P(\{|H_N\{I_N\} - H\{I_N\}| > H\{I_N\}/2\}) \leq 4/(NH\{I_N\}) < \varepsilon/2,$$

since by assumption $H\{I_N\} > 8/(\varepsilon N)$. In the second term in (1.3.11) only values $n \neq 0$ are involved. As $H\{I_N\} > 0$, we find that for any $n \neq 0$, we have, conditional on $H_N\{I_N\} = n/N$,

$$\sup_{I_N^*\in \mathcal{I}_N} |H_N\{I_N^*\}-H\{I_N^*\}|$$

$$\leq H\{I_N\} \left[\sup_{I_N^*\in \mathcal{I}_N} \left| \frac{H_N\{I_N^*\}}{H\{I_N\}} - \frac{H_N\{I_N^*\}}{H_N\{I_N\}} \right| + \sup_{I_N^*\in \mathcal{I}_N} \left| \frac{H_N\{I_N^*\}}{H_N\{I_N\}} - \frac{H\{I_N^*\}}{H\{I_N\}} \right| \right]$$

$$= |H_N\{I_N\}-H\{I_N\}| + H\{I_N\}\sup_{I_N^*\in \mathcal{I}_N} |\widetilde{H}_n\{I_N^*\}-\widetilde{H}\{I_N^*\}| .$$

Here $\widetilde{H}\{I_N^*\} = H\{I_N^*\}/H\{I_N\}$ is the conditional probability that the random vector (X,Y) is an element of $I_N^* \subset I_N$ under the hypothesis that it is an element of I_N. Given $H_N\{I_N\} = n/N$ with $n \neq 0$, the ratio $\widetilde{H}_n\{I_N^*\} = H_N\{I_N^*\}/H_N\{I_N\}$ is distributed as the empirical df corresponding to $\widetilde{H}$, based $n \neq 0$ observations. For $n \neq 0$ we have $\pi(n) \leq \pi_1(n) + \pi_2(n)$, where

$$\pi_1(n) = P(\{|H_N\{I_N\}-H\{I_N\}| \geq M[H\{I_N\}/4N]^{1/2}\}|\{H_N\{I_N\} = n/N\}),$$

$$\pi_2(n) = P(\{\sup_{I_N^*\in \mathcal{I}_N} |\widetilde{H}_n\{I_N^*\}-\widetilde{H}\{I_N^*\}| \geq M[4NH\{I_N\}]^{-1/2}\}).$$

Applying the Bienaymé-Chebyshev inequality once more we obtain

$$\text{(1.3.13)} \qquad \sum_{n\geq NH\{I_N\}/2}\pi_1(n)P(\{H_N\{I_N\} = n/N\})$$

$$\leq P(\{|H_N\{I_N\}-H\{I_N\}| \geq M[H\{I_N\}/4N]^{1/2}\}) \leq 4/M^2.$$

Finally we have to consider the summation involving the $\pi_2(n)$. For any interval I in the plane we have
$|\widetilde{H}_n\{I\}-\widetilde{H}\{I\}| \leq 4 \sup_{-\infty<x<\infty,-\infty<y<\infty}|\widetilde{H}_n(x,y)-\widetilde{H}(x,y)|$. According to Lemma 1.3.3 applied to $\widetilde{H}_n$ and $\widetilde{H}$ with e.g. $\zeta = 1$, there exists a constant M_1 such that

$$\pi_2(n) \leq M_1 \exp(-nM^2/(64NH\{I_N\})),$$

and hence

(1.3.14) $\sum_{n\geq NH\{I_N\}/2}\pi_2(n)P(\{H_N\{I_N\} = n/N\})$

$$\leq M_1 \exp(-NH\{I_N\}M^2/(128NH\{I_N\})) = M_1 \exp(-M^2/128).$$

Combining (1.3.12), (1.3.13) and (1.3.14) we see that for $H\{I_N\} > 8/(\varepsilon N)$ inequality (1.3.10) holds, provided M is chosen so large that both (1.3.13) and (1.3.14) are smaller than $\varepsilon/4$. Let us finally note that the argument is independent of the sequence $I_1, I_2, \ldots$ and the bivariate df H. □

Our final lemma will be used when the limiting score function J has discontinuities.

<u>LEMMA 1.3.5</u>. As $N \to \infty$, $\sup_{m=1,\ldots,N} N^{1/2}|F(X_{m:N})-m/N| = \mathcal{O}_P(1)$ and $\sup_{n=1,\ldots,N} N^{1/2}|G(Y_{n:N})-n/N| = \mathcal{O}_P(1)$, uniformly for $H \in \mathcal{H}$.

<u>PROOF</u>. We only need consider the first expression. Because $m/N = F_N(X_{m:N})$ we have $\sup_{m=1,\ldots,N} N^{1/2}|F(X_{m:N})-m/N| \leq \sup_{(-\infty,\infty)}|U_N(F)| = \mathcal{O}_P(1)$, uniformly for H in $\mathcal{H}$. □

Chapter 2

ASYMPTOTIC NORMALITY WHEN THE LIMITING SCORE FUNCTION IS CONTINUOUS

2.1. STATEMENT OF THE MAIN RESULT

For asymptotic normality of the rank statistics $T_N = \int\int J_N(F_N, G_N)dH_N$ (see (1.1.9)) a suitable standardization will be $N^{1/2}(T_N - \mu)$. Here the asymptotic mean depends on the underlying df H in the class $\mathcal{H}$ (see (1.1.6)) and, if (X,Y) has df H, it equals

$$(2.1.1) \qquad \mu = \mu(H) = E[J(F(X), G(Y))].$$

The asymptotic variance, also depending on H, is given by

$$(2.1.2) \qquad \sigma^2 = \sigma^2(H) = \mathrm{Var}\{J(F(X),G(Y)) + \int\int [c(F(x)-F(X))-F(x)]\, J^{(1,0)}(F(x),G(y))dH(x,y) + \int\int [c(G(y)-G(Y))-G(y)]\, J^{(0,1)}(F(x),G(y))dH(x,y)\}.$$

In the above formulas J is the limiting score function, introduced in (1.1.11), and $J^{(1,0)}$ ($J^{(0,1)}$) is its first partial derivative with respect to the first (second) variable. Finally the function c(z) is defined in (1.1.1). Before listing the assumptions needed for the theorem, let us introduce some more notation. Define

$$(2.1.3) \qquad F_N^* = [N/(N+1)]F_N, \; G_N^* = [N/(N+1)]G_N,$$

$$(2.1.4) \qquad B_{0N}^* = N^{1/2}\int\int [J_N(F_N,G_N) - J(F_N^*,G_N^*)]dH_N.$$

<u>REMARK</u>. Throughout Assumptions 2.1.1-2.1.3 the functions $r_1, \tilde{r}_1, r_2, \tilde{r}_2$ are the same fixed members of $\mathcal{R}$ (see Definition 1.3.2), the points $0 = s_0 < s_1 < \dots < s_k < s_{k+1} = 1$ and $0 = t_0 < t_1 < \dots < t_l < t_{l+1} = 1$ are the same fixed elements of the unit interval, and $\mathcal{H}'$ is the same subclass of $\mathcal{H}$. For $\mathcal{Q}$ see Definition 1.3.1.

The first two assumptions concern the smoothness and integrability conditions to be imposed on the limiting score function.

ASSUMPTION 2.1.1. The limiting score function $J(s,t)$ is continuous for $(s,t) \in (0,1) \times (0,1)$. The first partial derivatives $J^{(1,0)}(s,t) = \partial J(s,t)/\partial s$ and $J^{(0,1)}(s,t) = \partial J(s,t)/\partial t$ exist and are continuous for $(s,t) \in (s_{i-1},s_i) \times (t_{j-1},t_j)$, $i = 1,\ldots, k+1$ and $j = 1,\ldots, l+1$.

The above functions satisfy

$$|J(s,t)| \leq r_1(s)r_2(t), \quad |J^{(1,0)}(s,t)| \leq \tilde{r}_1(s)r_2(t),$$

$$|J^{(0,1)}(s,t)| \leq r_1(s)\tilde{r}_2(t),$$

at those $(s,t) \in (0,1) \times (0,1)$ where they are defined.

ASSUMPTION 2.1.2. The integrals

$$\int\int [r_1(F)r_2(G)]^{2+\delta} dH,$$

$$\int\int [q_1(F)\tilde{r}_1(F)r_2(G)]^{1+\delta} dH, \quad \int\int [q_2(G)r_1(F)\tilde{r}_2(G)]^{1+\delta} dH,$$

are bounded on the subclass of dfs $\mathcal{H}' \subset \mathcal{H}$ for some constant $\delta \geq 0$ and fixed $q_1, q_2 \in \mathcal{Q}$, satisfying $\int_0^1 [q_i(u)]^{-2-\delta} du < \infty$, $i = 1,2$, for the above δ.

The last assumption concerns the limiting behavior of the score functions with respect to the limiting score function.

ASSUMPTION 2.1.3. As $N \to \infty$, $B^*_{0N} \to_P 0$, uniformly on the subclass of dfs $\mathcal{H}' \subset \mathcal{H}$.

The main theorem of this chapter not only establishes asymptotic normality for a fixed df H in $\mathcal{H}$, but also uniformity of this convergence in distribution on an appropriate subclass $\mathcal{H}'$ of $\mathcal{H}$.

THEOREM 2.1.1. If for the score functions J_N and the limiting score function J Assumptions 2.1.1-2.1.3 are satisfied with $\mathcal{H}' = \{H\}$ for some fixed df $H \in \mathcal{H}$ and with $\delta = 0$, then $N^{1/2}(T_N - \mu) \to_d N(0,\sigma^2)$ as $N \to \infty$. Here $\mu = \mu(H)$ and $\sigma^2 = \sigma^2(H)$ are finite and given by (2.1.1) and (2.1.2) respectively.

Suppose that for the score functions J_N and the limiting score function J Assumptions 2.1.1-2.1.3 are satisfied for some fixed subclass of dfs

$H' \subset H$ and with $\delta > 0$. Provided $\sigma^2 = \sigma^2(H)$ is bounded away from zero on H', the above mentioned convergence in distribution is uniform on H'.

In spite of their rather formidable appearance, the assumptions are satisfied in many interesting situations as will be pointed out in Section 5.1. Under the null hypothesis the asymptotic mean and variance will be denoted by μ_0 and σ_0^2 respectively. These numbers, of course, do no longer depend on the particular underlying null hypothesis df $H = F \times G$ and are equal to (see also Bhuchongkul [4])

$$(2.1.5) \qquad \mu_0 = \int_0^1\int_0^1 J(u,v)du\, dv,$$

$$(2.1.6) \qquad \sigma_0^2 = \int_0^1\int_0^1 [\mu_0+J(u,v) - \int_0^1 J(u,t)dt - \int_0^1 J(s,v)ds]^2 du\, dv.$$

The expression for μ_0 is immediate. As to σ_0^2 note that $\int\int[c(F(x)-F(X))-F(x)]\, J^{(1,0)}(F(x),G(y))dF(x)dG(y) = \int_0^1[-\int_0^{F(X)} s\, J^{(1,0)}(s,t)ds + \int_{F(X)}^1 (1-s)\, J^{(1,0)}(s,t)ds]dt = \int_0^1\int_0^1 J(s,t)ds\, dt - \int_0^1 J(F(X),t)dt$, which follows from partial integration. To justify this we note that $\lim_{s\downarrow 0} sJ(s,t) = \lim_{s\uparrow 1} (1-s)\, J(s,t) = 0$, for each $0 < t < 1$. Actually, Assumption 2.1.1 says that $|J(s,t)| \leq r_1(s)r_2(t)$ for all (s,t) in $(0,1) \times (0,1)$ and from Assumption 2.1.2 it follows that $J(s,t)$ is a square integrable function of s for each fixed $0 < t < 1$. A similar expression may be found for $\int\int[c(G(y)-G(Y))-G(y)]\, J^{(0,1)}(F(x),G(y))dF(x)dG(y)$. The expression for the variance will be studied more generally in Section 4.2. The proof of Theorem 2.1.1 is deferred to Sections 2.2-2.4.

2.2. PROOF OF THEOREM 2.1.1: ASYMPTOTIC NORMALITY OF THE LEADING TERMS

We start this section with some basic notation. Let

$$(2.2.1) \qquad \begin{aligned} \Delta_{N1}(\omega) &= \Delta_{N1} = [X_{1:N},X_{N:N}),\ \bar{\Delta}_{N1}(\omega) = \bar{\Delta}_{N1} = [X_{1:N},X_{N:N}], \\ \Delta_{N2}(\omega) &= \Delta_{N2} = [Y_{1:N},Y_{N:N}),\ \bar{\Delta}_{N2}(\omega) = \bar{\Delta}_{N2} = [Y_{1:N},Y_{N:N}], \end{aligned}$$

and let $\Delta_N(\omega) = \Delta_N = \Delta_{N1} \times \Delta_{N2}$ and $\bar{\Delta}_N(\omega) = \bar{\Delta}_N = \bar{\Delta}_{N1} \times \bar{\Delta}_{N2}$ be the random product sets in the plane (see Figure 2.2.1). Note that the

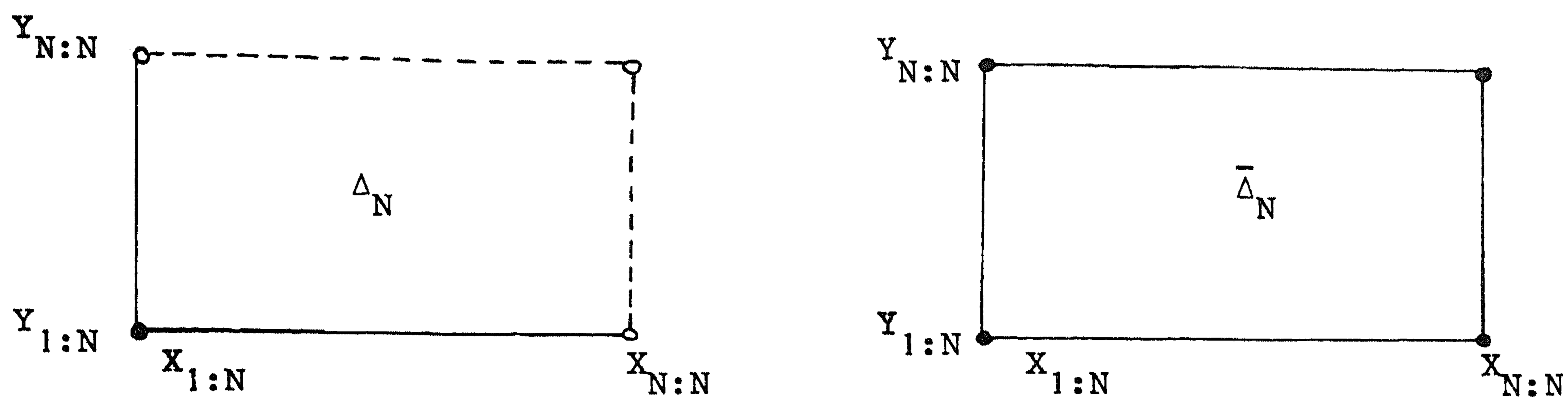

Figure 2.2.1

bivariate empirical df H_N assigns mass 1 to $\bar{\Delta}_N$. Without loss of generality we may prove Theorem 2.1.1 in the case where k = 1 = 1 in Assumption 2.1.1. For small positive γ define the sets

$$S_{\gamma 1} = \{x : F(x) \in [\gamma, s_1-\gamma] \cup [s_1+\gamma, 1-\gamma]\},$$

(2.2.2)

$$S_{\gamma 2} = \{y : G(y) \in [\gamma, t_1-\gamma] \cup [t_1+\gamma, 1-\gamma]\}.$$

Let $S_\gamma = S_{\gamma 1} \times S_{\gamma 2}$ be the product set in the plane. For γ as above and the set Ω_0 as in (1.3.5) we adopt furthermore the notation

$$(2.2.3) \qquad \Omega^*_{\gamma N} = \{\omega : \sup|F^*_N-F| < \gamma/2, \sup|G^*_N-G| < \gamma/2\} \cap \Omega_0.$$

Given any set D, D^c will denote its complement, $\chi(D)$ its indicator function and $\chi(D;z)$ the value of this function at the point z. Both the empirical processes U_N, V_N defined in (1.3.6) and their modifications $U^*_N(s) = N^{1/2}[F^*_N(F^{-1}(s))-s]$, $V^*_N(t) = N^{1/2}[G^*_N(G^{-1}(t))-t]$ defined for s,t in [0,1] will be needed. Like U_N, V_N, the latter processes satisfy $U^*_N(F) = N^{1/2}(F^*_N-F)$, $V^*_N(G) = N^{1/2}(G^*_N-G)$ on $(-\infty,\infty)$ for ω in Ω_0 (i.e. with probability 1).

For any ω in $\Omega^*_{\gamma N}$ the mean value theorem yields

$$(2.2.4) \qquad N^{1/2}J(F^*_N,G^*_N) = N^{1/2}J(F,G) + U^*_N(F)J^{(1,0)}(\Phi_N,\Psi_N) + V^*_N(G)J^{(0,1)}(\Phi_N,\Psi_N),$$

for all (x,y) in $\bar{\Delta}_N \cap S_\gamma$. In (2.2.4) the random point (Φ_N,Ψ_N) lies in the open random line segment joining (F,G) and (F^*_N,G^*_N) (see Figure 2.2.2). It

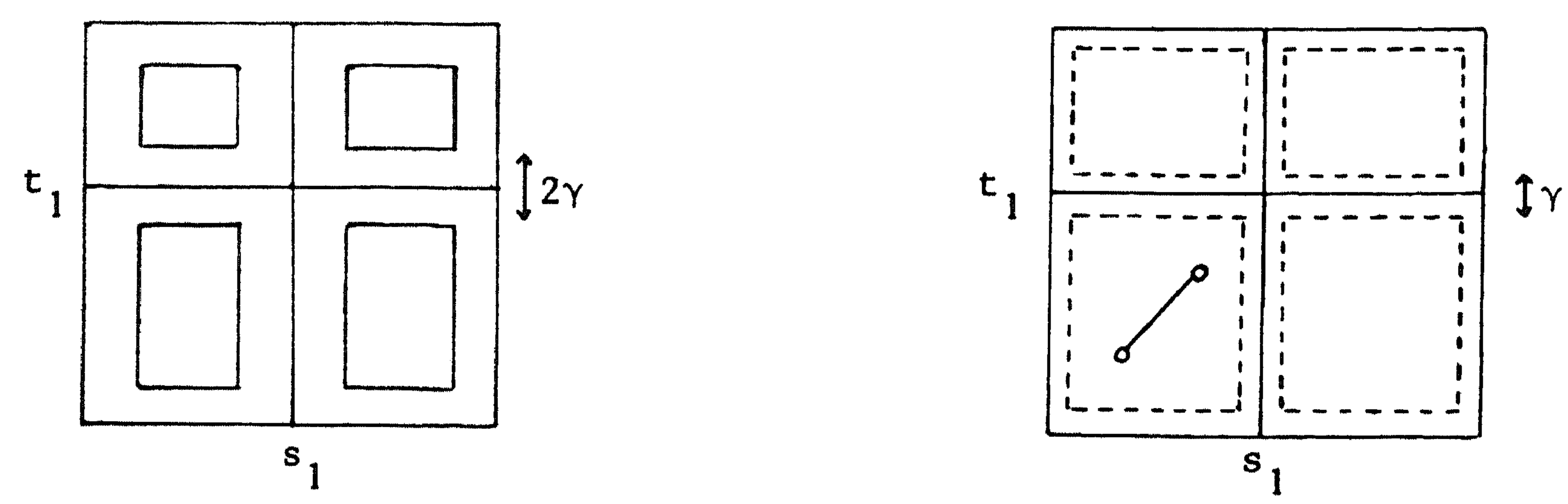

Figure 2.2.2

is defined by $(\Phi_N,\Psi_N) = (F,G) + \theta \times (F_N^*-F,G_N^*-G)$, where $\theta = \theta(\omega,x,y,N)$ satisfies $0 < \theta < 1$. In this context it is useful to define zero times a non-defined quantity as zero. Thus formula (2.2.4) leads to the fundamental decomposition

(2.2.5) $$N^{1/2}(T_N-\mu) = \sum_{i=0}^{2} A_{iN} + B_{ON}^* + \sum_{i=1}^{5} B_{iN},$$

where B_{ON}^* is defined in (2.1.4) and where

$$A_{ON} = N^{1/2}\iint J(F,G)d(H_N-H),$$

$$A_{1N} = \iint U_N(F)J^{(1,0)}(F,G)dH,$$

$$A_{2N} = \iint V_N(G)J^{(0,1)}(F,G)dH,$$

$$B_{1N} = \chi(\Omega_{\gamma N}^*)\iint_{S_\gamma} U_N^*(F)J^{(1,0)}(\Phi_N,\Psi_N)dH_N - A_{1N},$$

$$B_{2N} = \chi(\Omega_{\gamma N}^*)\iint_{S_\gamma} V_N^*(G)J^{(0,1)}(\Phi_N,\Psi_N)dH_N - A_{2N},$$

$$B_{3N} = \chi(\Omega_{\gamma N}^*)N^{1/2}\iint_{S_\gamma^c}[J(F_N^*,G_N^*)-J(F,G)]dH_N,$$

$$B_{4N} = \chi(\Omega_{\gamma N}^{*c})N^{1/2}[\iint J(F_N^*,G_N^*)dH_N-\iint J(F,G)dH],$$

$$B_{5N} = -\chi(\Omega_{\gamma N}^{*c})N^{1/2}\iint J(F,G)d(H_N-H).$$

In this section attention will be restricted to the A-terms.

The rv A_{ON} may be written in the form

(2.2.6) $$A_{ON} = N^{-1/2}\sum_{n=1}^{N} A_{OnN},$$

where the $A_{0nN} = J(F(X_n),G(Y_n)) - \mu$ are idd with mean zero. For the fixed df H (the fixed subclass of dfs H') the rv A_{0nN} has a finite moment of order 2 (a finite absolute moment of order larger than 2, bounded on H') by Assumption 2.1.2.

Note that for c as defined in (1.1.1) the empirical dfs F_N and G_N satisfy

$$(2.2.7) \qquad F_N(x) = N^{-1} \sum_{n=1}^{N} c(x-X_n), \quad G_N(y) = N^{-1} \sum_{n=1}^{N} c(y-Y_n),$$

for all x, y in $(-\infty,\infty)$. Since for ω in Ω_0 (i.e. with probability 1) we have $c(x-X_n) = c(F(x)-F(X_n))$ we may write $U_N(F) = N^{-1/2} \sum_{n=1}^{N}[c(F-F(X_n))-F]$. By this and a similar expression for $V_N(G)$ we obtain

$$(2.2.8) \qquad A_{1N} = N^{-1/2} \sum_{n=1}^{N} A_{1nN}, \quad A_{2N} = N^{-1/2} \sum_{n=1}^{N} A_{2nN},$$

where $A_{1nN} = \iint[c(F-F(X_n))-F]J^{(1,0)}(F,G)dH$ and $A_{2nN} = \iint[c(G-G(Y_n))-G] \times J^{(0,1)}(F,G)dH$ for $n = 1,..,N$. Each of these two sets of rvs consists of N iid elements with mean zero. To see the existence of higher order moments of A_{1nN} and A_{2nN} we need the following property of q-functions.

<u>LEMMA 2.2.1</u>. Let for arbitrary $u,v \in (0,1)$ the function $c(v-u)$ be defined as in (1.1.1), and let $q \in \mathcal{Q}$. Then there exists a constant $M = M_q$ (depending on q only) such that $|c(v-u)-v| \leq Mq(v)[q(u)]^{-1}$ for $u,v \in (0,1)$.

<u>PROOF</u>. Because of the properties of q-functions, there exists a number $\varepsilon = \varepsilon_q$ satisfying $0 < \varepsilon < \frac{1}{2}$, such that $u \leq q(u)$ for $0 \leq u \leq \varepsilon$. For suppose such ε does not exist. Then there is a sequence $u_n \downarrow 0$ satisfying $q(u_n) < u_n$, and hence $[q(u_n)]^{-2} > u_n^{-2}$. The reciprocal of q is square integrable on the unit interval; on the other hand $\int_0^1[q(u)]^{-2}du \geq u_n^{-1} \to \infty$ as $n \to \infty$, which yields a contradiction. (Similarly sharper bounds for q in the neighborhood of zero may be obtained.)

Let us first consider pairs $v < u$. Then $|c(v-u)-v|[q(v)]^{-1} = v[q(v)]^{-1}$. For $0 < v \leq \varepsilon$, with ε as above, we find $v[q(v)]^{-1} \leq 1 \leq M_1[q(u)]^{-1}$, if $M_1 = \max_{0<u<1} q(u)$. For $\varepsilon < v \leq \frac{1}{2}$ and M_1 as above we have $v[q(v)]^{-1} \leq [q(\varepsilon)]^{-1} \leq M_1[q(\varepsilon)]^{-1}[q(u)]^{-1}$. Finally for $\frac{1}{2} < v < 1$ we simply have $v[q(v)]^{-1} \leq [q(u)]^{-1}$. Evidently for $v < u$ the lemma holds with $M = \max \{M_1, M_1[q(\varepsilon)]^{-1}, 1\}$. For pairs $v \geq u$ the proof can be given in the same way. □

Lemma 2.2.1 applied with $q = q_1$, where q_1 is the function introduced in Assumption 2.1.2, guarantees the existence of a constant $M_1 = M_{q_1}$ such that for each ω

$$|A_{1nN}| \leq M_1[q_1(F(X_n))]^{-1}\iint q_1(F)\tilde{r}_1(F)r_2(G)dH,$$

$n = 1,\ldots,N$. By Assumption 2.1.2 for the fixed df H (the fixed subclass of dfs H') the random part $[q_1(F(X_n))]^{-1}$ of this upper bound possesses a finite moment of order 2 (a finite absolute moment of order larger than 2, bounded on H'). It is due to the same assumption that for the fixed df H (the fixed subclass of dfs H') the non-random integral is finite (bounded on H'). A similar argument deals with A_{2nN}.

Combining (2.2.6) and (2.2.8) we see that $\sum_{i=0}^{2} A_{inN}$ depends on $(X_1,Y_1),\ldots,(X_N,Y_N)$ through (X_n,Y_n) only, $n = 1,\ldots,N$. Moreover they are iid with mean zero and variance equal to $\mathrm{Var}(\sum_{i=0}^{2} A_{inN}) = \sigma^2(H)$ as given in (2.1.2). Hence application of the central limit theorem yields

$$(2.2.9) \qquad N^{-1/2} \sum_{n=1}^{N} \sum_{i=0}^{2} A_{inN} \to_d N(0,\sigma^2(H)),$$

for the fixed df H, as $N \to \infty$. Since, given the fixed subclass of dfs H', the absolute moment of an order larger than 2 is bounded on H' and since moreover the variance is given to be bounded away from zero on H' by Esseen's theorem (see Chernoff and Savage [7]) the above convergence in distribution is uniform on H'.

2.3. PROOF OF THEOREM 2.1.1: SOME LEMMAS

The first Lemma contains a modification of the basic Lemma 1.3.1 by Pyke and Shorack [31]. For notation see also (2.2.1).

<u>LEMMA 2.3.1</u>. For each $q \in Q$ we have as $N \to \infty$, uniformly for $H \in H$,

(i) $\sup_{\bar{\Delta}_{N1}} |U_N^*(F)|[q(F)]^{-1} = O_P(1)$, $\sup_{\bar{\Delta}_{N2}} |V_N^*(G)|[q(G)]^{-1} = O_P(1)$;

(ii) $\sup_{\bar{\Delta}_{N1}} |U_N^*(F)-U_N(F)|[q(F)]^{-1} \to_P 0$, $\sup_{\bar{\Delta}_{N2}} |V_N^*(G)-V_N(G)|[q(G)]^{-1} \to_P 0$.

<u>PROOF</u>. Note that $U_N^*(F) = U_N(F) + N^{1/2}(F_N^*-F_N)$, where $|N^{1/2}(F_N^*-F_N)| \leq N^{-1/2}$ on $(-\infty,\infty)$. The integrability of $[q]^{-2}$ entails that for any fixed $0 < \beta < 1$ we have $N^{-1/2}[q(\beta/N)]^{-1} = o(1)$ and $N^{-1/2}[q(1-\beta/N)]^{-1} = o(1)$ as $N \to \infty$. Because the $F(X_n)$ are independent $Un(0,1)$ rvs, given an arbitrary $\varepsilon > 0$ we can choose a number $0 < \beta_\varepsilon = \beta < 1$ such that

$P(\{\beta/N \le F(X_{1:N}) \le F(X_{N:N}) \le 1-\beta/N\}) \ge 1 - \varepsilon$ for all N and uniformly for all continuous F. The same remark holds true for $V_N^*(G)$.

From the above, part (ii) follows at once. For part (i) we also need Lemma 1.3.1. □

The next lemma is an application of Lemma 1.3.2 and displays a feature of reproducing u-shaped functions.

LEMMA 2.3.2. For each $r \in \mathcal{R}$ we have as $N \to \infty$, uniformly for $H \in \mathcal{H}$,

(i) $\sup_{\Delta_{N1}} r(F_N)[r(F)]^{-1} = \mathcal{O}_P(1)$, $\sup_{\Delta_{N2}} r(G_N)[r(G)]^{-1} = \mathcal{O}_P(1)$;

(ii) $\sup_{\bar{\Delta}_{N1}} r(F_N^*)[r(F)]^{-1} = \mathcal{O}_P(1)$, $\sup_{\bar{\Delta}_{N2}} r(G_N^*)[r(G)]^{-1} = \mathcal{O}_P(1)$.

PROOF. It suffices to prove the lemma for F_N and F_N^*.
(i) Because $\{0 < F_N < 1\} = \Delta_{N1}$, it follows from Lemma 1.3.2 that for each $\varepsilon > 0$ there exists a constant $0 < \beta_\varepsilon = \beta < 1$ such that

(2.3.1) $$\Omega_N = \{\beta F \le F_N \le 1-\beta(1-F) \text{ on } \Delta_{N1}\}$$

has probability $P(\Omega_N) \ge 1 - \varepsilon$ for all N and all continuous F. From the properties of reproducing u-shaped functions, in particular because these functions are symmetric about 1/2, it follows that $r(\beta s) \ge r(1-\beta(1-s))$ for $0 < s \le 1/2$ and that $r(1-\beta(1-s)) \ge r(\beta s)$ for $1/2 < s < 1$. Hence for each ω in Ω_N we have that $r(F_N) \le r(\beta F) \le Mr(F)$ for $0 < F \le 1/2$ and that $r(F_N) \le r(1-\beta(1-F)) \le Mr(F)$ for $1/2 < F < 1$ on Δ_{N1}, where $M = M_\beta$ is some finite positive constant. Consequently the inequality $r(F_N) \le Mr(F)$ holds true on Δ_{N1} for each ω in Ω_N.
(ii) The second part follows in the same way if it can be shown that for each $\varepsilon > 0$ there exists a constant $0 < \beta_\varepsilon = \beta < 1$ such that

(2.3.2) $$\Omega_N^* = \{\beta F \le F_N^* \le 1-\beta(1-F) \text{ on } \bar{\Delta}_{N1}\}$$

has probability $P(\Omega_N^*) \ge 1 - \varepsilon$ for all N and all continuous F. Since Ω_N (see (2.3.1)) has that property, it remains to show that
$P(\{N/(N+1) \le 1-\beta[1-F(X_{N:N})]\}) \ge 1 - \varepsilon$ for sufficiently small positive β. Because the $F(X_n)$ are independent $Un(0,1)$ rvs, the latter probability equals $1 - (1-[\beta(N+1)]^{-1})^N \ge 1 - \varepsilon$ for all N and all continuous F, provided $\beta = \beta_\varepsilon$ is chosen sufficiently small. □

To formulate the last lemma, let us first introduce a useful notation. For each positive integer m we define the function I_m on [0,1] by

(2.3.3) $\quad I_m(u) = (i-1)/m$ for $(i-1)/m \leq u < i/m$, $i = 1,\dots,m$ and $I_m(1) = 1$.

LEMMA 2.3.3. As m, $N \to \infty$, $\sup_{(-\infty,\infty)} |U_N(I_m(F))-U_N(F)| \to_P 0$ and $\sup_{(-\infty,\infty)} |V_N(I_m(G))-V_N(G)| \to_P 0$, uniformly for $H \in \mathcal{H}$.

PROOF. Note that for continuous F we have $\sup_{-\infty<x<\infty} |U_N(I_m(F(x)))-U_N(F(x))| = \sup_{0\leq s\leq 1} |U_N(I_m(s))-U_N(s)|$, which no longer depends on F. The U_N-processes converge weakly to a Brownian bridge W (see Section 1.3). In Pyke and Shorack [31] these U_N- and W-processes are replaced by $\tilde{U}_N$- and $\tilde{W}$-processes defined on a single new probability space $(\tilde{\Omega},\tilde{A},\tilde{P})$ and having the same finite dimensional distributions as the original processes (see also Skorokhod [40]). These new processes satisfy $\sup |\tilde{U}_N-\tilde{W}| \to_{a.s.} 0$ and hence also $\sup |\tilde{U}_N(I_m)-\tilde{W}(I_m)| \to_{a.s.} 0$ uniformly in m, as $N \to \infty$. Now $\sup |\tilde{U}_N(I_m)-\tilde{U}_N| \leq \sup |\tilde{U}_N-\tilde{W}| + \sup |\tilde{W}-\tilde{W}(I_m)| + \sup |\tilde{W}(I_m)-\tilde{U}_N(I_m)|$. For almost every $\tilde{\omega}$ in $\tilde{\Omega}$ the function $\tilde{W}$ is uniformly continuous on [0,1] so that $\sup |\tilde{W}-\tilde{W}(I_m)| \to_{a.s.} 0$ as $m \to \infty$. This proves that $\sup |\tilde{U}_N(I_m)-\tilde{U}_N| \to_{a.s.} 0$ as $m,N \to \infty$. This last result implies the first convergence in probability of the lemma. The second convergence is established in a similar fashion. □

2.4. PROOF OF THEOREM 2.1.1: ASYMPTOTIC NEGLIGIBILITY OF THE REMAINDER TERMS

The asymptotic negligibility of the remainder terms will be given as corollaries to the lemmas of the previous section. It will be convenient to introduce the following subsets of Ω:

$$\Omega_{1N} = \{|U_N(F)| \leq Mq_1(F) \text{ on } (-\infty,\infty); |V_N(G)| \leq Mq_2(G) \text{ on } (-\infty,\infty)\},$$

$$\Omega_{2N} = \{|U_N^*(F)| \leq Mq_1(F) \text{ on } \bar{\Delta}_{N1}; |V_N^*(G)| \leq Mq_2(G) \text{ on } \bar{\Delta}_{N2}\},$$

$$\Omega_{3N} = \{|U_N^*(F)-U_N(F)| \leq \zeta_N q_1(F) \text{ on } \bar{\Delta}_{N1};$$
$$|V_N^*(G)-V_N(G)| \leq \zeta_N q_2(G) \text{ on } \bar{\Delta}_{N2}\},$$

$$\Omega_{4N} = \{r_1(F_N^*) \leq Mr_1(F), \tilde{r}_1(F_N^*) \leq M\tilde{r}_1(F) \text{ on } \bar{\Delta}_{N1};$$
$$r_2(G_N^*) \leq Mr_2(G), \tilde{r}_2(G_N^*) \leq M\tilde{r}_2(G) \text{ on } \bar{\Delta}_{N2}\}.$$

Here $r_1, \tilde{r}_1, r_2, \tilde{r}_2, q_1, q_2$ are the same as in Assumptions 2.1.1 and 2.1.2. Once for all let us notice the following property of the set Ω_{4N}. For each ω let $\tilde{\Phi}_N = \tilde{\Phi}_N(\omega)$ and $\tilde{\Psi}_N = \tilde{\Psi}_N(\omega)$ be functions defined on $\bar{\Delta}_{N1}$ and $\bar{\Delta}_{N2}$ respectively, satisfying

$$\min\{F, F_N^*\} \leq \tilde{\Phi}_N \leq \max\{F, F_N^*\} \text{ on } \bar{\Delta}_{N1},$$

$$\min\{G, G_N^*\} \leq \tilde{\Psi}_N \leq \max\{G, G_N^*\} \text{ on } \bar{\Delta}_{N2}.$$

Then, independently of H in H, we have

$$r_1(\tilde{\Phi}_N) \leq Mr_1(F), \tilde{r}_1(\tilde{\Phi}_N) \leq M\tilde{r}_1(F) \text{ on } \bar{\Delta}_{N1},$$

$$r_2(\tilde{\Psi}_N) \leq Mr_2(G), \tilde{r}_2(\tilde{\Psi}_N) \leq M\tilde{r}_2(G) \text{ on } \bar{\Delta}_{N2},$$

for each ω in Ω_{4N}. By way of an example let us prove the first of the above inequalities. Obviously on $\bar{\Delta}_{N1}$ both $r_1(\min\{F, F_N^*\})$ and $r_1(\max\{F, F_N^*\})$ are bounded by $\max\{r_1(F), r_1(F_N^*)\}$. For ω restricted to Ω_{4N} we have in turn $\max\{r_1(F), r_1(F_N^*)\} \leq Mr_1(F)$ on $\bar{\Delta}_{N1}$, provided $M \geq 1$. The definition of $\tilde{\Phi}_N$ and the properties of r_1 imply for each ω that $r_1(\tilde{\Phi}_N) \leq \max\{r_1(F), r_1(F_N^*)\}$, and hence for each ω restricted to Ω_{4N} that $r_1(\tilde{\Phi}_N) \leq Mr_1(F)$ on $\bar{\Delta}_{N1}$. A typical example of such functions $\tilde{\Phi}_N, \tilde{\Psi}_N$ are the functions Φ_N, Ψ_N occurring in (2.2.4). We may think of $\Phi_N(\Psi_N)$ as defined on $\Omega \times \bar{\Delta}_{N1}(\Omega \times \bar{\Delta}_{N2})$ by setting e.g. $\Phi_N = F$ ($\Psi_N = G$) where not defined. We shall occasionally also simply take $\tilde{\Phi}_N = F_N^*$, $\tilde{\Psi}_N = G_N^*$. The choice will always be clear from the context.

According to Lemmas 1.3.1, 2.3.1 and 2.3.2 (ii) for an arbitrary $\varepsilon > 0$ the constant $1 \leq M < \infty$ and the sequence ζ_N, decreasing to zero as $N \to \infty$, may be chosen in dependence of ε such that the set

$$\text{(2.4.1)} \qquad \Omega_N = \cap_{i=1}^{4} \Omega_{iN}$$

has probability $P_H(\Omega_N) = P(\Omega_N) \geq 1 - \varepsilon$ for all N and for all H in H. This set will play a major role in each of the corollaries without explicit reference. Note that B_{2N} is symmetric to B_{1N}, so that B_{2N} need not be treated in the sequel.

Let us first give a further decomposition of the rv B_{1N} which may be written as $B_{1N} = \sum_{i=1}^{5} C_{iN}$, where

$$C_{1N} = \chi(\Omega^*_{\gamma N})\iint_{S_\gamma} U^*_N(F)[J^{(1,0)}(\Phi_N,\Psi_N)-J^{(1,0)}(F,G)]dH_N,$$

$$C_{2N} = \chi(\Omega^*_{\gamma N})\iint_{S_\gamma} [U^*_N(F)-U_N(F)]J^{(1,0)}(F,G)dH_N,$$

$$C_{3N} = \chi(\Omega^*_{\gamma N})\iint_{S_\gamma} U_N(F)J^{(1,0)}(F,G)d(H_N-H),$$

$$C_{4N} = -\chi(\Omega^*_{\gamma N})\iint_{S^c_\gamma} U_N(F)J^{(1,0)}(F,G)dH,$$

$$C_{5N} = -\chi(\Omega^{*c}_{\gamma N})\iint U_N(F)J^{(1,0)}(F,G)dH.$$

COROLLARY 2.4.1. As $\gamma \downarrow 0$ and $N \to \infty$, $C_{4N} \to_P 0$ and $B_{3N} \to_P 0$, uniformly on H'.

PROOF. Let us first consider C_{4N}. From Assumption 2.1.1 it follows immediately that

(2.4.2) $$\chi(\Omega_N)|C_{4N}| \leq M\iint_{S^c_\gamma} q_1(F)\tilde{r}_1(F)r_2(G)dH.$$

Next we consider B_{3N} and start by pointing out how the mean value theorem, as applied in (2.2.4), may be modified to meet the needs of this somewhat different situation. For any ω in Ω, by $\tau = \tau(\omega,x,y,N)$ denote the number of points that the open line segment leading from $(F(x),G(y))$ to $(F^*_N(x),G^*_N(y))$ has in common with the lines $s = s_1$ and $t = t_1$, where (x,y) is in $\bar{\Delta}_N$ (see Figure 2.4.1). Evidently for

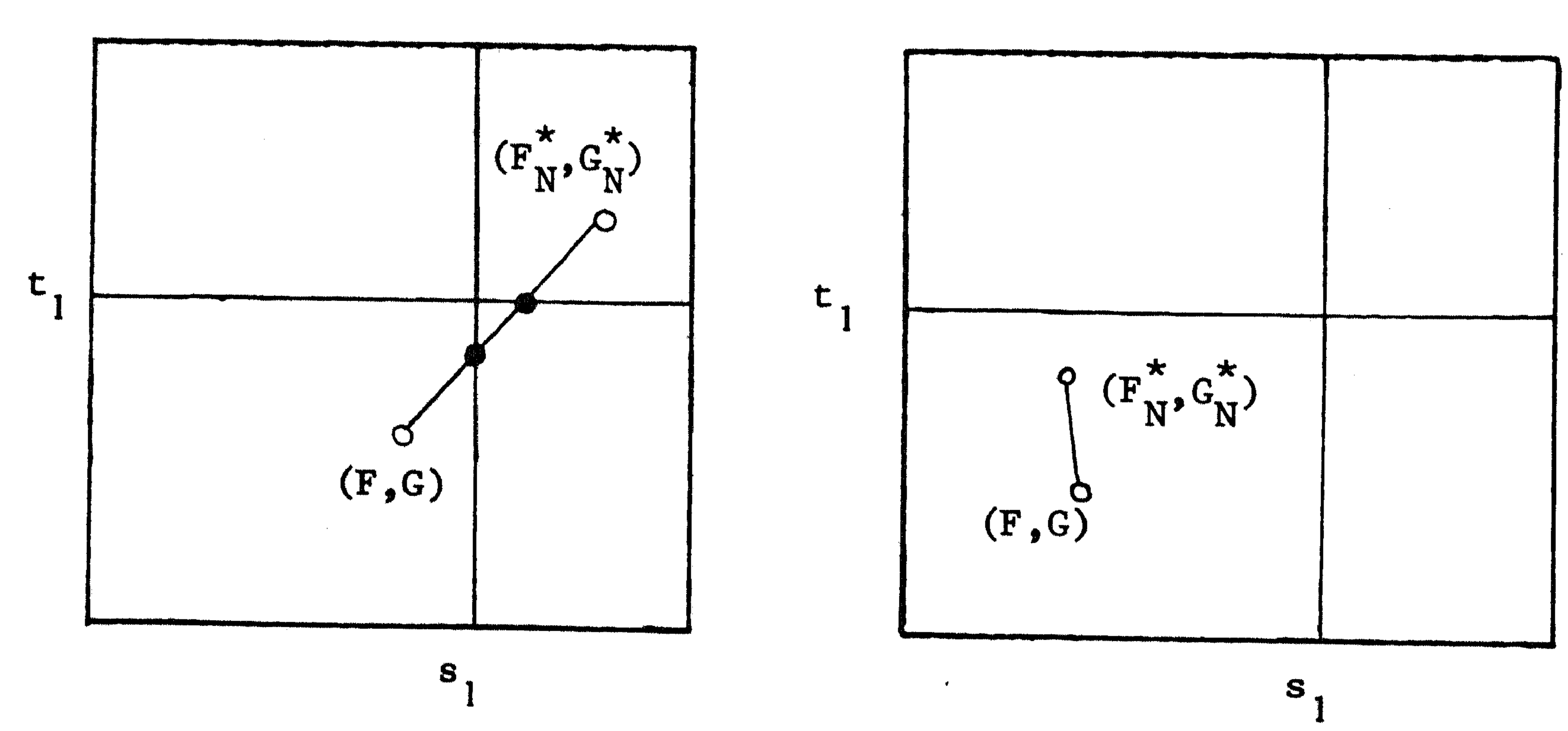

Figure 2.4.1

each ω we have $0 \le \tau \le 2$. In the special case where ω is in $\Omega^*_{\gamma N}$ and (x,y) is restricted to $\bar{\Delta}_N \cap S_\gamma$, the number τ is equal to zero and the mean value theorem applies at once. More generally, following the closed line segment starting at the point $(F,G) = (F_{0N},G_{0N})$ up to the random point $(F^*_N,G^*_N) = (F_{\tau+1N},G_{\tau+1N})$ we encounter successively the points of intersection $(F_{1N},G_{1N}),\ldots,(F_{\tau N},G_{\tau N})$ with the two lines if $\tau \neq 0$. Since J is continuous on the closed segment and continuously differentiable on each of the open subsegments from (F_{iN},G_{iN}) to (F_{i+1N},G_{i+1N}) for $i = 0,\ldots,\tau$, the mean value theorem may be applied step-wise. This leads to

$$N^{1/2}J(F^*_N,G^*_N) = N^{1/2}J(F,G) + \sum_{i=0}^{\tau} N^{1/2}(F_{i+1N}-F_{iN})J^{(1,0)}(\Phi_{iN},\Psi_{iN})$$

$$+ \sum_{i=0}^{\tau} N^{1/2}(G_{i+1N}-G_{iN})J^{(0,1)}(\Phi_{iN},\Psi_{iN}),$$

for each ω and (x,y) in $\bar{\Delta}_N$. Here (Φ_{iN},Ψ_{iN}) is a random point on the open line segment between (F_{iN},G_{iN}) and (F_{i+1N},G_{i+1N}) and hence (Φ_{iN},Ψ_{iN}) is a random point on the open line segment between (F,G) and (F^*_N,G^*_N) for $i = 0,\ldots,\tau$, where defined. For this reason the properties of the set Ω_{4N} may be applied with $\tilde{\Phi}_N = \Phi_{iN}$ and $\tilde{\Psi}_N = \Psi_{iN}$. Furthermore, where defined, $|F_{i+1N}-F_{iN}| \le |F^*_N-F|$ and $|G_{i+1N}-G_{iN}| \le |G^*_N-G|$ for $i = 0,\ldots,\tau$. Combining the above results with Assumption 2.1.1 one obtains the bound $\chi(\Omega_N)|B_{3N}| \le B_{31N} + B_{32N}$, where

$$B_{31N} = 3M^3\iint_{S^c_\gamma} q_1(F)\tilde{r}_1(F)r_2(G)dH_N,$$

$$B_{32N} = 3M^3\iint_{S^c_\gamma} q_2(G)r_1(F)\tilde{r}_2(G)dH_N.$$

By symmetry between the last two expressions it suffices to prove that $B_{31N} \to_P 0$ as $\gamma \downarrow 0$ and $N \to \infty$, uniformly on H'. This can be achieved by showing that $E(B_{31N}) \to 0$ as $\gamma \downarrow 0$ and $N \to \infty$, uniformly on H'. Since

$$(2.4.3) \qquad E(B_{31N}) = 3M^3\iint_{S^c_\gamma} q_1(F)\tilde{r}_1(F)r_2(G)dH,$$

and in view of (2.4.2), the corollary is proved if it can be shown that $\iint_{S^c_\gamma} q_1(F)\tilde{r}_1(F)r_2(G)dH \to 0$ as $\gamma \downarrow 0$, uniformly on H'.

In the case where $H' = \{H\}$ Assumption 2.1.2 holds with $\delta = 0$, and it follows at once by the dominated convergence theorem that the right-hand side of (2.4.3) converges to zero as $\gamma \downarrow 0$, for this fixed df H. If H' is an arbitrary subclass Assumption 2.1.2 is given to hold for some $\delta > 0$, so that $\sup_{H\in H'} \iint [q_1(F)\tilde{r}_1(F)r_2(G)]^{1+\delta} dH = \tilde{M}_{H'} < \infty$. Applying Hölder's inequality we find that

$$\iint_{S_\gamma^c} q_1(F)\tilde{r}_1(F)r_2(G)dH \leq \{\iint [q_1(F)\tilde{r}_1(F)r_2(G)]^{1+\delta} dH\}^{1/(1+\delta)} \times$$

$$[\iint_{S_\gamma^c} dH]^{\delta/(1+\delta)} \leq [\tilde{M}_{H'}]^{1/(1+\delta)} [8\gamma]^{\delta/(1+\delta)} \to 0$$

as $\gamma \downarrow 0$, independently of H in H'. □

COROLLARY 2.4.2. For fixed γ, $C_{5N} \to_P 0$, $B_{4N} \to_P 0$ and $B_{5N} \to_P 0$ as $N \to \infty$, uniformly on H.

PROOF. For fixed γ, $P(\Omega_{\gamma N}^{*c}) \to 0$ as $N \to \infty$, uniformly on H by the Glivenko-Cantelli theorem and because the probability distribution of $\sup |F_N - F|$, $\sup |G_N - G|$ does not depend on H in H. □

COROLLARY 2.4.3. For fixed γ, $C_{1N} \to_P 0$ as $N \to \infty$, uniformly on H.

PROOF. Let us note that

$$\chi(\Omega_N)|C_{1N}| \leq \tilde{M} \sup_{\bar{\Delta}_N \cap S_\gamma} |J^{(1,0)}(\Phi_N,\Psi_N) - J^{(1,0)}(F,G)|,$$

where $\tilde{M} = M \max_{0\leq s\leq 1} q_1(s)$. For fixed γ the function $J^{(1,0)}$ is uniformly continuous on the closed set $\{[\gamma/2, s_1-\gamma/2] \cup [s_1+\gamma/2, 1-\gamma/2]\} \times \{[\gamma/2, t_1-\gamma/2] \cup [t_1+\gamma/2, 1-\gamma/2]\}$. Since $|\Phi_N - F| \leq |F_N^* - F|$ and $|\Psi_N - G| \leq |G_N^* - G|$ where Φ_N and Ψ_N are defined, the Glivenko-Cantelli theorem yields the convergence to zero in probability of the right-hand side of the above inequality, uniformly on H. □

COROLLARY 2.4.4. For fixed γ, $C_{2N} \to_P 0$ as $N \to \infty$, uniformly on H.

PROOF. As $J^{(1,0)}$ is continuous on the closed set $\{[\gamma, s_1-\gamma] \cup [s_1+\gamma, 1-\gamma]\} \times \{[\gamma, t_1-\gamma] \cup [t_1+\gamma, 1-\gamma]\}$ there exists a finite positive constant $\tilde{M}_\gamma$ such that $|J^{(1,0)}(F,G)| \leq \tilde{M}_\gamma$ on S_γ for any H in H. Hence, with $\tilde{\zeta}_N = \zeta_N \max_{0\leq s\leq 1} q_1(s)$,

$$\chi(\Omega_N)|C_{2N}| \leq \tilde{\zeta}_N \tilde{M}_\gamma \to 0,$$

as $N \to \infty$, uniformly on H. □

COROLLARY 2.4.5. For fixed γ, $C_{3N} \to_P 0$ as $N \to \infty$, uniformly on H.

PROOF. For any positive integer m the rv $|C_{3N}|$ is bounded by $\sum_{i=1}^{3} C_{3iN}$, where

$$C_{31N} = \iint_{S_\gamma} |U_N(F)J^{(1,0)}(F,G) - U_N(I_m(F))J^{(1,0)}(I_m(F),I_m(G))|\,dH_N,$$

$$C_{32N} = |\iint_{S_\gamma} U_N(I_m(F))J^{(1,0)}(I_m(F),I_m(G))\,d(H_N - H)|,$$

$$C_{33N} = \iint_{S_\gamma} |U_N(I_m(F))J^{(1,0)}(I_m(F),I_m(G)) - U_N(F)J^{(1,0)}(F,G)|\,dH,$$

(see (2.3.3)). It suffices to show that each of the above rvs can be made arbitrarily small with arbitrarily high probability for some common positive integer m, provided N is large enough.

Consider C_{31N} and C_{33N}, which are both bounded by the supremum of the integrand over the set S_γ. The function $J^{(1,0)}$ is uniformly continuous on the closed set $\{[\gamma,s_1-\gamma] \cup [s_1+\gamma,1-\gamma]\} \times \{[\gamma,t_1-\gamma] \cup [t_1+\gamma,1-\gamma]\}$ so that independently of H in H we have $|J^{(1,0)}(F,G)| \le \tilde{M}_\gamma$ and $|J^{(1,0)}(F,G) - J^{(1,0)}(I_m(F),I_m(G))| \le \tilde{\zeta}_{m\gamma}$ on S_γ. Here $\tilde{M}_\gamma$ is a finite constant and $\tilde{\zeta}_{m\gamma} \to 0$ as $m \to \infty$. Application of Lemma 2.3.3 gives the existence of constants $\tilde{\zeta}_{mN} \to 0$ as $m,N \to \infty$, such that $\tilde{\Omega}_{mN} = \{\sup |U_N(F) - U_N(I_m(F))| \le \tilde{\zeta}_{mN}\}$ has probability larger than $1 - \varepsilon$ for all m,N and all H in H. Hence for i = 1,3

$$\chi(\Omega_N \cap \tilde{\Omega}_{mN})C_{3iN} \le \tilde{\zeta}_{mN}\tilde{M}_\gamma + \tilde{M}\tilde{\zeta}_{m\gamma} \to 0$$

for fixed γ as $m,N \to \infty$. Besides $P(\Omega_N \cap \tilde{\Omega}_{mN}) \ge 1 - 2\varepsilon$ for all N and all H in H by (2.4.1) and the above property of $\tilde{\Omega}_{mN}$.

Let us next consider C_{32N} for a fixed value m. For each ω the integrand in the expression for this rv is a simple step function assuming the value $Z_{ijmN}(\omega)$ on the rectangle

$$S_{\gamma ij} = (\{x : F(x) \in [(i-1)/m, i/m)\} \times \{y : G(y) \in [(j-1)/m, j/m)\}) \cap S_\gamma,$$

for i = 1,...,m and j = 1,...,m. Because $|Z_{ijmN}| \le \tilde{M}(\tilde{M}_\gamma + \tilde{\zeta}_{m\gamma})$ on Ω_N, we have

$$\chi(\Omega_N)C_{32N} = \chi(\Omega_N)\left|\sum_{i=1}^{m} \sum_{j=1}^{m} Z_{ijmN}\iint_{S_{\gamma ij}} d(H_N-H)\right|$$

$$\le 4m^2\tilde{M}(\tilde{M}_\gamma+\tilde{\zeta}_{m\gamma})\sup |H_N-H| \to_P 0$$

for fixed γ and m as $N \to \infty$, uniformly on H. Here Lemma 1.3.3 has been applied only in a weak form.

The conclusion of the corollary follows by straightforward combination of these partial results. □

In order to show how the results of these corollaries may be combined to complete the proof of Theorem 2.1.1, first use Corollary 2.4.1 to choose a fixed γ_0 and an index N_0 to ensure that $P(\{|B_{3N}|, |C_{4N}| \le \varepsilon\}) \ge 1 - \varepsilon$ for all $N \ge N_0$ and all dfs in H'. Next application of Corollaries 2.4.2–2.4.5 with the above fixed γ_0 gives the existence of an index $N_1 = N_1(\gamma_0) > N_0$ such that $P(\{|B_{4N}|, |B_{5N}|, |C_{1N}|, |C_{2N}|, |C_{3N}|, |C_{5N}| \le \varepsilon\}) \ge 1 - \varepsilon$ for all $N \ge N_1$ and all H in H. This implies that $P(\{|B_{1N}+B_{3N}+B_{4N}+B_{5N}| \le 8\varepsilon\}) \ge 1 - 2\varepsilon$ for all $N \ge N_1$ and all dfs in H'.

Chapter 3

ASYMPTOTIC NORMALITY WHEN THE LIMITING SCORE FUNCTION IS NO LONGER CONTINUOUS

3.1. STATEMENT OF THE MAIN RESULTS

In this chapter the results of Chapter 2 will be complemented by considering the case where the limiting score function J may exhibit discontinuities on the lines $s = s_1,\ldots,s = s_k$ and $t = t_1,\ldots,t = t_l$ in the open unit square. This weakening of the smoothness conditions entails, as can be expected (see e.g. Dupač and Hájek [8]), a local differentiability condition on the underlying continuous df H. Again we consider standardized versions $N^{1/2}(T_N-\mu)$ of the statistics $T_N = \iint J_N(F_N,G_N)dH_N$, where the asymptotic mean $\mu = \mu(H)$ is given by (2.1.1). An expression for the asymptotic variance cannot be given until the assumptions and some of their basic implications are formulated.

REMARK. As in the previous chapter, throughout Assumptions 3.1.1-3.1.4 the points $0 = s_0 < s_1 < \ldots < s_k < s_{k+1} = 1$ and $0 = t_0 < t_1 < \ldots < t_l < t_{l+1} = 1$ are the same fixed elements of the unit interval, the functions $r_1,\tilde{r}_1,r_2,\tilde{r}_2$ the same fixed members of $\mathcal{R}$, and $\mathcal{H}'$ the same subclass of $\mathcal{H}$.

The first three assumptions are directly comparable with Assumptions 2.1.1-2.1.3.

ASSUMPTION 3.1.1. The limiting score function J(s,t) is equal to J_{ij} on $\{[s_{i-1},s_i) \cap (0,1)\} \times \{[t_{j-1},t_j) \cap (0,1)\}$ for $i = 1,\ldots,k+1$ and $j = 1,\ldots,l+1$. Here J_{ij} is defined and continuous on $\{[s_{i-1},s_i] \cap (0,1)\} \times \{[t_{j-1},t_j] \cap (0,1)\}$ and possesses a continuous first partial derivative $J_{ij}^{(1,0)}(s,t) = \partial J_{ij}(s,t)/\partial s$ for $(s,t) \in (s_{i-1},s_i) \times \{[t_{j-1},t_j] \cap (0,1)\}$ and a continuous first partial derivative $J_{ij}^{(0,1)}(s,t) = \partial J_{ij}(s,t)/\partial t$ for $(s,t) \in \{[s_{i-1},s_i] \cap (0,1)\} \times (t_{j-1},t_j)$, $i = 1,\ldots,k+1$ and $j = 1,\ldots,l+1$.

The function J, defined above, satisfies

$$|J(s,t)| \leq r_1(s)r_2(t),\quad |J^{(1,0)}(s,t)| \leq \tilde{r}_1(s)r_2(t),$$

$$|J^{(0,1)}(s,t)| \leq r_1(s)\tilde{r}_2(t),$$

at those $(s,t) \in (0,1) \times (0,1)$ where the functions involved are defined.

ASSUMPTION 3.1.2. See Assumption 2.1.2.

ASSUMPTION 3.1.3. See Assumption 2.1.3.

The last assumption concerns the local differentiability property to be imposed on the df H (for $\bar{h}$ see also (1.2.12)).

ASSUMPTION 3.1.4. The following holds for the subclass $\mathcal{H}' \subset \mathcal{H}$. There is an open set O_1 containing the points $s_1,\ldots,s_k$ and an open set O_2 containing the points $t_1,\ldots,t_l$ such that for $H \in \mathcal{H}'$ the density $\bar{h}(s,t) = \partial^2 H(F^{-1}(s), G^{-1}(t))/\partial s \partial t$ exists and is continuous on $O = O_1 \times (0,1) \cup (0,1) \times O_2$. Moreover the subclass $\mathcal{H}'$ satisfies the equicontinuity conditions

$$\sup_{H\in\mathcal{H}'} |\bar{h}(s,t)-\bar{h}(s_i,t)| \to 0$$

$$\text{as } s \to s_i \text{ for all } t \in (0,1),\ i = 1,\ldots,k,$$

$$\sup_{H\in\mathcal{H}'} |\bar{h}(s,t)-\bar{h}(s,t_j)| \to 0$$

$$\text{as } t \to t_j \text{ for all } s \in (0,1),\ j = 1,\ldots,l.$$

Furthermore there exist functions $\bar{f}$ and $\bar{g}$ on $(0,1)$ such that

$$\sup_{H\in\mathcal{H}'} \bar{h}(s,t) \le \bar{f}(s) \text{ for all } (s,t) \in (0,1) \times O_2,$$

$$\text{with } \int_0^1 r_1(s)\bar{f}(s)ds < \infty,$$

$$\sup_{H\in\mathcal{H}'} \bar{h}(s,t) \le \bar{g}(t) \text{ for all } (s,t) \in O_1 \times (0,1),$$

$$\text{with } \int_0^1 r_2(t)\bar{g}(t)dt < \infty.$$

We shall next present two lemmas which are of general importance for the sequel and which first of all enable us to give an expression for the asymptotic variance.

LEMMA 3.1.1. Suppose the function J satisfies Assumption 3.1.1. Then, for some positive integer p, this function can be written in the form

(3.1.1) $$J(s,t) = J_c(s,t) + \sum_{m=1}^{p} K_m(s) \times L_m(t),$$

for all $(s,t) \in (0,1) \times (0,1)$. For some finite positive constant M' these functions have the following properties.

(i) The function J_c is defined and continuous on $(0,1) \times (0,1)$. Its first partial derivatives $J_c^{(1,0)}$ and $J_c^{(0,1)}$ exist and are continuous on $\cup_{i=1}^{k+1} \cup_{j=1}^{l+1} (s_{i-1},s_i) \times (t_{j-1},t_j)$.

With $r_1,\tilde{r}_1,r_2,\tilde{r}_2 \in \mathcal{R}$ as in Assumption 3.1.1 these functions satisfy

$$|J_c(s,t)| \leq M'r_1(s)r_2(t), \quad |J_c^{(1,0)}(s,t)| \leq M'\tilde{r}_1(s)r_2(t),$$

$$|J_c^{(0,1)}(s,t)| \leq M'r_1(s)\tilde{r}_2(t),$$

at those $(s,t) \in (0,1) \times (0,1)$ where they are defined.

(ii) The functions K_m and L_m are defined on $(0,1)$ and can be decomposed into $K_m = K_{mc} + K_{md}$ and $L_m = L_{mc} + L_{md}$ on $(0,1)$. Here $K_{md}(s) = \sum_{i=1}^{k} a_{mi}c(s-s_i)$ for $s \in (0,1)$ and constants $a_{m1},\ldots,a_{mk}$, and $L_{md}(t) = \sum_{j=1}^{l} b_{mj}c(t-t_j)$ for $t \in (0,1)$ and constants $b_{m1},\ldots,b_{ml}$, $m = 1,\ldots,p$. For $c(z)$ see (1.1.1). Further K_{mc} and L_{mc} are continuous on $(0,1)$ and have continuous derivatives $K'_{mc} = K'_m$ and $L'_{mc} = L'_m$ on $\cup_{i=1}^{k+1} (s_{i-1},s_i)$ and $\cup_{j=1}^{l+1} (t_{j-1},t_j)$ respectively.

With $r_1,\tilde{r}_1,r_2,\tilde{r}_2 \in \mathcal{R}$ as in Assumption 3.1.1 these functions satisfy

$$|K_m(s)| \leq M'r_1(s), \quad |K'_m(s)| \leq M'\tilde{r}_1(s),$$

$$|L_m(t)| \leq M'r_2(t), \quad |L'_m(t)| \leq M'\tilde{r}_2(t),$$

at those $s \in (0,1)$ and $t \in (0,1)$ respectively where they are defined, $m = 1,\ldots,p$.

PROOF. It suffices to prove the above representation for each of the components J_{ij} of J separately. Hence let us consider $\chi((0,s_1) \times (0,t_1))J_{11}$ and define the function J_c as follows

$$J_c(s,t) = J_{11}(s,t) \qquad \text{for } (s,t) \in (0,s_1) \times (0,t_1),$$

$$J_c(s,t) = J_{11}(s_1,t) \qquad \text{for } (s,t) \in [s_1,1) \times (0,t_1),$$

$$J_c(s,t) = J_{11}(s_1,t_1) \qquad \text{for } (s,t) \in [s_1,1) \times [t_1,1),$$

$$J_c(s,t) = J_{11}(s,t_1) \qquad \text{for } (s,t) \in (0,s_1) \times [t_1,1).$$

Note that J_c equals J_{11} on the set I, does not depend on s on the set II, is constant on the set III and does not depend on t on the set IV (see Figure 3.1.1). Let us further define

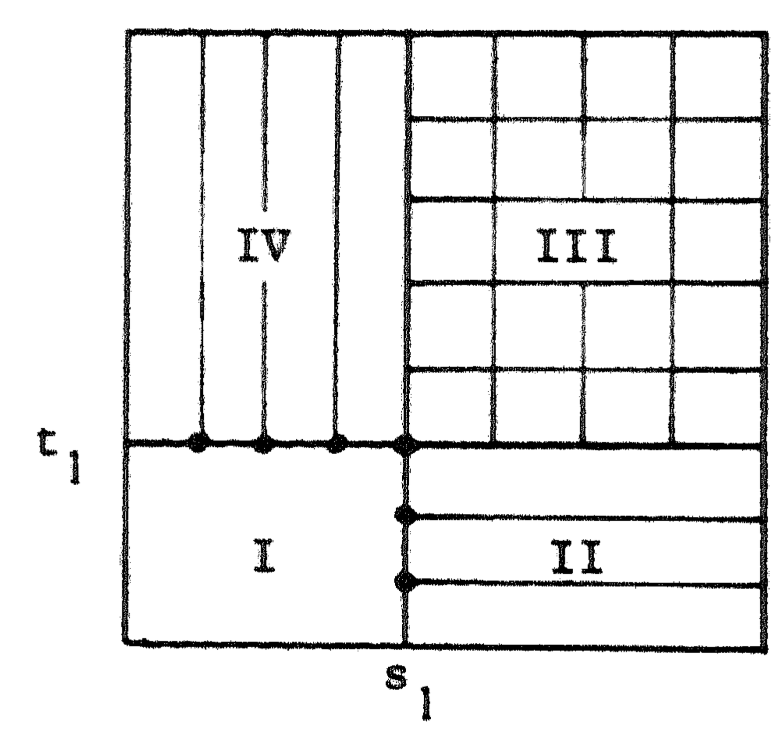

Figure 3.1.1

$$K_1(s) = -c(s-s_1) \qquad \text{for } s \in (0,1),$$

$$L_1(t) = J_{11}(s_1,t) \text{ for } t \in (0,t_1),\ L_1(t) = 0 \quad \text{for } t \in [t_1,1),$$

$$K_2(s) = c(s-s_1)J_{11}(s_1,t_1) \qquad \text{for } s \in (0,1),$$

$$L_2(t) = -c(t-t_1) \qquad \text{for } t \in (0,1),$$

$$K_3(s) = J_{11}(s,t_1) \text{ for } s \in (0,s_1),\ K_3(s) = 0 \quad \text{for } s \in [s_1,1),$$

$$L_3(t) = -c(t-t_1) \qquad \text{for } t \in (0,1).$$

It follows that $\chi((0,s_1) \times (0,t_1))J_{11} = J_c + \sum_{m=1}^{3} K_m \times L_m$ on $(0,1) \times (0,1)$. By straightforward verification one finds that J_c and K_m, L_m, $m = 1,2,3$, enjoy the properties listed under (i) and (ii) respectively. □

<u>LEMMA 3.1.2</u>. Let for $H \in \mathcal{H}$ Assumption 3.1.4 be satisfied with $\mathcal{H}' = \{H\}$. Let ϕ and ψ be functions on the unit interval such that $\int_0^1 |\phi(s)|ds$, $\int_0^1 |\phi(s)\bar{f}(s)|ds$, $\int_0^1 |\psi(t)|dt$, $\int_0^1 |\psi(t)\bar{g}(t)|dt < \infty$. Here $\bar{f}$ and $\bar{g}$ are the same as in Assumption 3.1.4. Then, if (X,Y) has df H,

(i) $E(\psi(G(Y))|F(X) = s)$ has a version continuous on O_1, to be denoted by $E_c(\psi(G(Y))|F(X) = s)$;
(ii) $E(\phi(F(X))|G(Y) = t)$ has a version continuous on O_2, to be denoted by $E_c(\phi(F(X))|G(Y) = t)$.

PROOF. It suffices to prove (i). Since (X,Y) has df H, $(F(X),G(Y))$ has df $H(F^{-1},G^{-1})$; the latter df has $Un(0,1)$ marginals. Consequently the function $\int_0^1 \psi(t)\bar{h}(s,t)dt$ is a version of the conditional expectation considered in (i), restricted to O_1. Moreover this version is continuous on O_1, for let $s,s+\zeta \in O_1$ and consider $\int_0^1 \psi(t)[\bar{h}(s+\zeta,t) - \bar{h}(s,t)]dt$. By continuity of the function $\bar{h}$ we have $\bar{h}(s+\zeta,t) - \bar{h}(s,t) \to 0$ as $\zeta \to 0$ for each $t \in (0,1)$. By the assumptions of the lemma we have furthermore that
$|\psi(t)|\ |\bar{h}(s+\zeta,t) - \bar{h}(s,t)| \leq 2|\psi(t)\bar{g}(t)|$ for each $t \in (0,1)$, and $\int_0^1|\psi(t)\bar{g}(t)|dt < \infty$. Finally, by the dominated convergence theorem, we obtain $\int_0^1 \psi(t)[\bar{h}(s+\zeta,t) - \bar{h}(s,t)]dt \to 0$ as $\zeta \to 0$. □

We are now in a position to give the expression for the limiting variance. The representation and notation given in Lemma 3.1.1 will be used. Because the functions K_m and L_m are easily seen to satisfy the conditions of Lemma 3.1.2, if (X,Y) has df H, we may write

$$(3.1.2)\qquad \sigma^2 = \sigma^2(H) = \mathrm{Var}\{J(F(X),G(Y)) + \iint[c(F(x)-F(X))-F(x)]\ J^{(1,0)}(F(x),G(y))dH(x,y) + \iint[c(G(y)-G(Y))-G(y)]\ J^{(0,1)}(F(x),G(y))dH(x,y) + \sum_{m=1}^{p} \sum_{i=1}^{k} a_{mi}[c(s_i-F(X))-s_i]E_c(L_m(G(Y))|F(X) = s_i) + \sum_{m=1}^{p} \sum_{j=1}^{l} b_{mj}[c(t_j-G(Y))-t_j]E_c(K_m(F(X))|G(Y) = t_j)\}.$$

THEOREM 3.1.1. If for the score functions J_N and the limiting score function J Assumptions 3.1.1-3.1.4 are satisfied with $H' = \{H\}$ for some fixed df $H \in H$ and with $\delta = 0$, then $N^{1/2}(T_N-\mu) \to_d N(0,\sigma^2)$ as $N \to \infty$. Here $\mu = \mu(H)$ and $\sigma^2 = \sigma^2(H)$ are finite and given by (2.1.1) and (3.1.2) respectively.

Suppose that for the score functions J_N and the limiting score function J Assumptions 3.1.1-3.1.4 are satisfied for some fixed subclass of dfs

$H' \subset H$ and with $\delta > 0$. Provided $\sigma^2 = \sigma^2(H)$ is bounded away from zero on H', the above mentioned convergence in distribution is uniform on H'.

A discussion of the assumptions can be found in Section 5.1. Under the null hypothesis the asymptotic mean and variance are independent of the particular underlying df $H = F \times G$. These numbers are denoted by μ_0 and σ_0^2 respectively and are equal to

$$(3.1.3) \qquad \mu_0 = \int_0^1\int_0^1 J(u,v)\,du\,dv,$$

$$(3.1.4) \qquad \sigma_0^2 = \int_0^1\int_0^1 [\mu_0 + J(u,v) - \int_0^1 J(u,t)\,dt - \int_0^1 J(s,v)\,ds]^2\,du\,dv,$$

where J is the limiting score function, defined in Assumption 3.1.1. We arrive at these expressions, which are the same as for continuous limiting score functions (see (2.1.5) and (2.1.6)), by calculations similar to those of Section 2.1, applied to the various parts into which J may be decomposed according to Lemma 3.1.1. The expression for μ_0 is immediate again and the variance may be found by observing that (cf. Section 2.1) $\int\int[c(F(x)-F(X))-F(x)]J_c^{(1,0)}(F(x),G(y))dF(x)dG(y) = \int_0^1\int_0^1 J_c(s,t)dsdt - \int_0^1 J_c(F(X),t)dt$ and $\int\int[c(F(x)-F(X))-F(x)]K'_{mc}(F(x))L_m(G(y))dF(x)dG(y) + \sum_{i=1}^k a_{mi}[c(s_i-F(X))-s_i]E_c(L_m(G(Y))|F(X) = s_i) = \int_0^1\int_0^1 K_m(s)L_m(t)dsdt - K_{mc}(F(X))\int_0^1 L_m(t)dt - \sum_{i=1}^k a_{mi}[1-s_i-c(s_i-F(X))+s_i]\int_0^1 L_m(t)dt$, for $m = 1,\ldots,p$. Because moreover $-\sum_{i=1}^k a_{mi}[1-c(s_i-F(X))]$ may be replaced by $-K_{md}(F(X))$ in the expression for the variance and since similar expressions may be obtained for the remaining terms that occur in the variance, we arrive at the expression for σ_0^2 in (3.1.4). The expression for the variance is studied more generally in Section 4.2.

For simplicity of presentation we did not attempt to formulate Theorem 3.1.1 in such a way that it would contain Theorem 2.1.1 as a special case. When J happens to be continuous the smoothness condition imposed on this function is slightly stronger than in Assumption 2.1.1, and hence stronger than necessary. Moreover Assumption 3.1.4, which is superfluous in this case, is automatically included.

The proof of the theorem covers Sections 3.2-3.4.

3.2. PROOF OF THEOREM 3.1.1: ASYMPTOTIC NORMALITY OF THE LEADING TERMS

For notational conventions we refer to Sections 2.1 and 2.2. Using the representation of Lemma 3.1.1 we arrive at the following basic decomposition

$$(3.2.1) \qquad N^{1/2}(T_N-\mu) = \sum_{i=0}^{2} A_{ciN} + \sum_{i=0}^{2} A_{iN} + \sum_{i=1}^{2} A_{diN} + B_{0N}^{*} + B_{cN} + \sum_{i=1}^{3}(B_{iN}+B_{diN}),$$

holding true for each ω. Here B_{0N}^{*} is defined in (2.1.4) and, when the a_{mi} and b_{mj} are the numbers occurring in Lemma 3.1.1,

$$A_{c0N} = N^{1/2}\int\int J_c(F,G)d(H_N-H),$$

$$A_{c1N} = \int\int U_N(F)J_c^{(1,0)}(F,G)dH, \quad A_{c2N} = \int\int V_N(G)J_c^{(0,1)}(F,G)dH,$$

$$A_{0N} = \sum_{m=1}^{p} N^{1/2}\int\int K_m(F)L_m(G)d(H_N-H),$$

$$A_{1N} = \sum_{m=1}^{p}\int\int U_N(F)K_m'(F)L_m(G)dH, \quad A_{2N} = \sum_{m=1}^{p}\int\int V_N(G)K_m(F)L_m'(G)dH,$$

$$A_{d1N} = \sum_{m=1}^{p} \sum_{i=1}^{k} a_{mi}\, E_c(L_m(G(Y))|F(X)=s_i)U_N(s_i),$$

$$A_{d2N} = \sum_{m=1}^{p} \sum_{j=1}^{l} b_{mj}\, E_c(K_m(F(X))|G(Y)=t_j)V_N(t_j),$$

$$B_{cN} = N^{1/2}[\int\int J_c(F_N^{*},G_N^{*})-J_c(F,G)]dH_N - \sum_{i=1}^{2} A_{ciN},$$

$$B_{1N} = \sum_{m=1}^{p} N^{1/2}\int\int[K_{mc}(F_N^{*})-K_{mc}(F)]L_m(G)dH_N - A_{1N},$$

$$B_{2N} = \sum_{m=1}^{p} N^{1/2}\int\int K_m(F)[L_{mc}(G_N^{*})-L_{mc}(G)]dH_N - A_{2N},$$

$$B_{3N} = \sum_{m=1}^{p} N^{1/2}\int\int[K_{mc}(F_N^{*})-K_{mc}(F)][L_m(G_N^{*})-L_m(G)]dH_N,$$

$$B_{d1N} = \sum_{m=1}^{p} N^{1/2}\int\int[K_{md}(F_N^{*})-K_{md}(F)]L_m(G)dH_N - A_{d1N},$$

$$B_{d2N} = \sum_{m=1}^{p} N^{1/2}\int\int K_m(F)[L_{md}(G_N^{*})-L_{md}(G)]dH_N - A_{d2N},$$

$$B_{d3N} = \sum_{m=1}^{p} N^{1/2}\int\int[K_{md}(F_N^{*})-K_{md}(F)][L_m(G_N^{*})-L_m(G)]dH_N.$$

In this section attention will be restricted to the asymptotic normality of the A-terms. Since J_c satisfies the assumptions of Section 2.1 the results of that section may be applied to the A_c-terms, determined by the continuous function J_c. Thus, according to (2.2.9) we may write

$$(3.2.2) \qquad \sum_{i=0}^{2} A_{ciN} = N^{-1/2} \sum_{n=1}^{N} \sum_{i=0}^{2} A_{cinN},$$

where $\sum_{i=0}^{2} A_{cinN}$ is obtained from $\sum_{i=0}^{2} A_{inN}$ in Section 2.2 by consistently replacing J, $J^{(1,0)}$ and $J^{(0,1)}$ by J_c, $J_c^{(1,0)}$ and $J_c^{(0,1)}$ respectively. It should be noted that $\sum_{i=0}^{2} A_{cinN}$ depends on $(X_1,Y_1),\ldots,(X_N,Y_N)$ through (X_n,Y_n) only, $n = 1,\ldots,N$. They are iid with mean zero. For the fixed df H (the fixed subclass of dfs H') the rv $\sum_{i=0}^{2} A_{cinN}$ has a finite moment of order 2 (a finite absolute moment of order larger than 2, bounded on H') by Lemma 3.1.1 and Assumption 3.1.2.

The method of Section 2.2 may be repeated for the A-terms, connected with the function $\sum_{m=1}^{p} K_m \times L_m$, to see that

$$(3.2.3) \qquad \sum_{i=0}^{2} A_{iN} = N^{-1/2} \sum_{n=1}^{N} \sum_{i=0}^{2} A_{inN}.$$

Here $A_{0nN} = \sum_{m=1}^{p} [K_m(F(X_n))L_m(G(Y_n)) - \iint K_m(F)L_m(G)dH]$, $A_{1nN} = \sum_{m=1}^{p} \iint [c(F-F(X_n))-F]K_m'(F)L_m(G)dH$ and $A_{2nN} = \sum_{m=1}^{p} \iint [c(G-G(Y_n))-G] \times K_m(F)L_m'(G)dH$ for $n = 1,\ldots,N$. The $\sum_{i=0}^{2} A_{inN}$ are iid with mean zero. For the fixed df H (the fixed subclass of dfs H') the rv $\sum_{i=0}^{2} A_{inN}$ has a finite moment of order 2 (a finite absolute moment of order larger than 2, bounded on H') by Lemma 3.1.1 and Assumption 3.1.2.

As to the A_d-terms, connected with the function $\sum_{m=1}^{p} (K_{md} \times L_m + K_m \times L_{md})$, note that for each ω in Ω we have $c(F^{-1}(s_i)-X_n) = c(s_i-F(X_n))$ so that $U_N(s_i) = N^{-1/2} \sum_{n=1}^{N} [c(s_i-F(X_n))-s_i]$. Using a similar expression for $V_N(t_j)$ we obtain

$$(3.2.4) \qquad \sum_{i=1}^{2} A_{diN} = N^{-1/2} \sum_{n=1}^{N} \sum_{i=1}^{2} A_{dinN},$$

where $A_{d1nN} = \sum_{m=1}^{p} \sum_{i=1}^{k} a_{mi}[c(s_i-F(X_n))-s_i]E_c(L_m(G(Y))|F(X)=s_i)$ and $A_{d2nN} = \sum_{m=1}^{p} \sum_{j=1}^{1} b_{mj}[c(t_j-G(Y_n))-t_j]E_c(K_m(F(X))|G(Y)=t_j)$ for $n = 1,\ldots,N$. The $\sum_{i=1}^{2} A_{dinN}$ are iid with mean zero. The absolute moments of any order exist for the fixed df H (are bounded on the fixed subclass of dfs H').

Combining (3.2.2)-(3.2.4) we see that $\sum_{i=0}^{2} A_{cinN} + \sum_{i=0}^{2} A_{inN} + \sum_{i=1}^{2} A_{dinN}$ still depends on $(X_1,Y_1),\ldots,(X_N,Y_N)$ through (X_n,Y_n) only, $n = 1,\ldots,N$. Moreover they are iid with mean zero and finite variance $\sigma^2 = \sigma^2(H) = \mathrm{Var}(\sum_{i=0}^{2} A_{cinN} + \sum_{i=0}^{2} A_{inN} + \sum_{i=1}^{2} A_{dinN})$, for any $n = 1,\ldots,N$. Consequently the central limit theorem applies and yields

$\sum_{i=0}^{2} A_{ciN} + \sum_{i=0}^{2} A_{iN} + \sum_{i=1}^{2} A_{diN} = N^{-1/2} \sum_{n=1}^{N} (\sum_{i=0}^{2} A_{cinN} + \sum_{i=0}^{2} A_{inN} + \sum_{i=1}^{2} A_{dinN}) \to_d N(0,\sigma^2(H))$ as $N \to \infty$, for the fixed df H. Given the fixed subclass of dfs H', the absolute moment of an order larger than 2 is bounded whereas the variance is bounded away from zero on H'. Hence by Esseen's theorem the above asymptotic normality is uniform on H'.

3.3. PROOF OF THEOREM 3.1.1: SOME LEMMAS

The lemmas of Section 2.3 will also suffice for the B_c- and B-terms. In this section some lemmas specific for the B_d-terms will be given. For $0 < s < 1$, $0 < u < 1$ and fixed constants a and b let us define the functions $c_N(a,b;s,u)$ by

$$c_N(a,b;s,u) = \sum_{m=0}^{N} \binom{N}{m} s^m(1-s)^{N-m}|c((m+a)/(N+b)-u)-c(s-u)|, \tag{3.3.1}$$

$N = 1,2,\ldots$ Here c(z) is defined in (1.1.1).

<u>LEMMA 3.3.1</u>. Let a and b be fixed constants and let $c_N(a,b;s,u)$ be defined as in (3.3.1). Then, as $N \to \infty$,
(i) $c_N(a,b;s,u) = O(\exp(-2N(s-u)^2))$, uniformly for $s,u \in (0,1)$;
(ii) $\int_0^1 c_N(a,b;s,u)ds = O(N^{-1/2})$, uniformly for $u \in (0,1)$.

<u>PROOF</u>. (i) The function $c_N(a,b;s,u)$ is unequal to zero only if $s < u$ and $m \geq (N+b)u - a$, or $s \geq u$ and $m < (N+b)u - a$. Suppose $s < u$. Then $c_N(a,b;s,u) = \Pr(Z \geq (N+b)u-a)$, where Z is a $\mathcal{Bi}(N,s)$ distributed rv. Since $(N+b)u - a = N(s+[u-s+(bu-a)/N])$, we have by (1.3.9) as a and b are fixed

$$\Pr(Z \geq (N+b)u-a) \leq M_0 \exp(-2N[u-s+(bu-a)/N]^2)$$
$$\leq M_1 \exp(-2N(s-u)^2),$$

provided $u - s + (bu-a)/N \geq 0$. Now consider the set $D = \{(s,u) : u + (bu-a)/N < s < u\}$. If D is empty there is nothing left to prove, hence suppose D is not empty. Then $\sup_{(s,u)\in D,N=1,2,\ldots} \exp(2N(s-u)^2) \leq \max_{N=1,2,\ldots} \exp(2(bu-a)^2/N) = \exp(2(bu-a)^2) = M_2$, say. Since c_N is a probability it is always bounded by 1 and hence by $M_2 \exp(-2N(s-u)^2)$ for all $(s,u) \in D$ and all $N = 1,2,\ldots$ We thus have, letting $M = \max\{M_1,M_2\}$, that c_N is bounded by $M \exp(-2N(s-u)^2)$

for all $s < u$ and $N = 1,2,\ldots$ This inequality can similarly be shown to hold for $s \geq u$.
(ii) This follows at once from part (i) by $\int_0^1 c_N(a,b;s,u)ds \leq M\int_{-\infty}^{\infty} \exp(-2N(s-u)^2)ds = O(N^{-1/2})$ as $N \to \infty$, independently of $0 < u < 1$. □

The next lemma illustrates the use of the functions $c_N(a,b;s,u)$.

LEMMA 3.3.2. Let ψ be a function on the unit interval with $\int_0^1|\psi(t)|dt < \infty$. Then for any $H \in \mathcal{H}$ and $u \in (0,1)$ the following holds
(i) $E\iint|[c(F_N^*-u)-c(F-u)]\psi(G)|dH \leq \int_0^1 c_N(0,1;s,u)E(|\psi(G(Y))||F(X) = s)ds$;
(ii) $E\iint|[c(F_N^*-u)-c(F-u)]\psi(G)|dH_N \leq \int_0^1 c_{N-1}(1,2;s,u)E(|\psi(G(Y))||F(X) = s)ds$.

PROOF. (i) Because $P(\{(N+1)F_N^*(x) = m\}) = \binom{N}{m}[F(x)]^m[1-F(x)]^{N-m}$ for $m = 0,1,\ldots,N$ we obtain

$$E\iint|[c(F_N^*-u)-c(F-u)]\psi(G)|dH$$

$$\leq \iint\sum_{m=0}^{N} \binom{N}{m}F^m(1-F)^{N-m}|[c(m/(N+1)-u)-c(F-u)]\psi(G)|dH$$

$$= \int_0^1 c_N(0,1;s,u)E(|\psi(G(Y))||F(X) = s)ds.$$

(ii) Similarly, since $P(\{(N+1)F_N^*(X_n) = m\}|F(X_n)) = \binom{N-1}{m-1} \times [F(X_n)]^{m-1}[1-F(X_n)]^{N-m}$ for $m = 1,\ldots,N$, we have

$$E\iint|[c(F_N^*-u)-c(F-u)]\psi(G)|dH_N$$

$$= N^{-1}\sum_{n=1}^{N} E\{E(|[c(F_N^*(X_n)-u)-c(F(X_n)-u)]\psi(G(Y_n))||F(X_n),G(Y_n))\}$$

$$= N^{-1}\sum_{n=1}^{N} E\{|\psi(G(Y_n))| \times E(|c(F_N^*(X_n)-u)-c(F(X_n)-u)||F(X_n))\}$$

$$= \iint\sum_{m=1}^{N} \binom{N-1}{m-1}F^{m-1}(1-F)^{N-m}|[c(m/(N+1)-u)-c(F-u)]\psi(G)|dH$$

$$= \int_0^1 c_{N-1}(1,2;s,u)E(|\psi(G(Y))||F(X) = s)ds. \quad \square$$

For any positive integer N and real number u in $(0,1)$ the positive integer N_u is uniquely determined by

$$(3.3.2) \qquad (N+1)u \leq N_u < (N+1)u + 1.$$

LEMMA 3.3.3. As $N \to \infty$ we have, uniformly for $H \in \mathcal{H}$ and $u \in (0,1)$,
(i) $|F_N^*(F^{-1}(u))-u| = O_P(N^{-1/2})$;
(ii) $|F(X_{N_u:N})-u| = O_P(N^{-1/2})$.

PROOF. Part (i) follows immediately from the properties of the empirical process formulated in Lemma 1.3.1, and the fact that $|F_N^*(x)-F_N(x)| \leq 1/(N+1)$ for all x in $(-\infty,\infty)$.

Similarly part (ii) is a direct consequence of the properties of order statistics given in Lemma 1.3.5, and the fact that $|(N_u/N)-u| \leq 2/N$. □

Furthermore for any positive integer N and real number u in $(0,1)$ we define the random interval

$$\Gamma_{N1} = \{x : \min\{X_{N_u:N}, F^{-1}(u)\} \leq x < \max\{X_{N_u:N}, F^{-1}(u)\}\}. \tag{3.3.3}$$

Let sgn (z) be the function defined in (1.2.6).

LEMMA 3.3.4. For all $N = 1,2,\ldots$, $\omega \in \Omega$, $x \in (-\infty,\infty)$, $u \in (0,1)$ and $H \in \mathcal{H}$, the equality $c(F_N^*(x)-u)-c(F(x)-u) = \text{sgn}(F^{-1}(u)-X_{N_u:N})\chi(\Gamma_{N1};x)$ holds true, with $c(z)$ defined in (1.1.1).

PROOF. Because of the properties of F^{-1} listed in Section 1.1 it follows that the left-hand side equals -1 on the set $\{F_N^*(x) < u, x \geq F^{-1}(u)\}$, 0 on the set $\{F_N^*(x) \geq u, x \geq F^{-1}(u)\} \cup \{F_N^*(x) < u, x < F^{-1}(u)\}$ and 1 on the set $\{F_N^*(x) \geq u, x < F^{-1}(u)\}$.

First let us suppose that ω is such that $X_{N_u:N} < F^{-1}(u)$ and hence sgn $(F^{-1}(u)-X_{N_u:N}) = 1$ and $\Gamma_{N1} = [X_{N_u:N}, F^{-1}(u))$. Consequently for such ω the right-hand side of the expression in the lemma equals 0 if $x < X_{N_u:N}$. For these x-values moreover $F_N^*(x) \leq (N_u-1)/(N+1) < u$ by (3.3.2), and $x < F^{-1}(u)$, so that according to the remarks in the first paragraph of the proof also the left-hand side equals 0. If $X_{N_u:N} \leq x < F^{-1}(u)$ the right-hand side equals 1. Furthermore $F_N^*(x) \geq N_u/(N+1) \geq u$ by (3.3.2) for such x-values, so that also the left-hand side equals 1. If finally $x \geq F^{-1}(u)$ the right-hand side equals 0 again and, since for these x-values $F_N^*(x) \geq u$, also the left-hand side equals 0.

Analogously the equality of the left- and right-hand side may be seen in the cases where ω is such that $X_{N_u:N} = F^{-1}(u)$ and $X_{N_u:N} > F^{-1}(u)$. □

LEMMA 3.3.5. As $N \to \infty$, $\sup_{I^*_{N2}} |H_N\{\Gamma_{N1} \times I^*_{N2}\} - H\{\Gamma_{N1} \times I^*_{N2}\}| = O_P(N^{-3/4})$, uniformly for $H \in \mathcal{H}$. Here the supremum is taken over all intervals $I^*_{N2} \subset (-\infty,\infty)$.

PROOF. Let an arbitrary $\varepsilon > 0$ be given. Lemma 3.3.3 (ii) implies the existence of a finite positive constant $M_1 = M_{1\varepsilon}$ such that the set

$$\Omega_{1N} = \{F^{-1}(u-M_1N^{-1/2}) \leq X_{N_u:N} \leq F^{-1}(u+M_1N^{-1/2})\}$$

has probability $P(\Omega_{1N}) \geq 1 - \varepsilon/2$, for all N, all u in (0,1) and all $H \in \mathcal{H}$. For brevity denote $I_{N1} = [F^{-1}(u-M_1N^{-1/2}),\ F^{-1}(u+M_1N^{-1/2})]$. Applying Lemma 1.3.4 with $I_N = I_{N1} \times (-\infty,\infty)$, and hence with $H\{I_N\} = 2M_1N^{-1/2}$, we find that there exists a finite positive $M_2 = M_{2\varepsilon}$ such that the set

$$\Omega_{2N} = \{\sup_{I^*_{N2}} |H_N\{I_{N1} \times I^*_{N2}\} - H\{I_{N1} \times I^*_{N2}\}| \leq (2M_1)^{1/2}M_2N^{-3/4}\}$$

has probability $P(\Omega_{2N}) \geq 1 - \varepsilon/2$, for all N, all u in (0,1) and all H in $\mathcal{H}$. Here the supremum is taken over all intervals $I^*_{N2} \subset (-\infty,\infty)$. Consequently for each ω in $\Omega_{1N} \cap \Omega_{2N}$ the supremum formulated in the lemma is bounded by $(2M_1)^{1/2}M_2N^{-3/4}$, and $P(\Omega_{1N} \cap \Omega_{2N}) \geq 1-\varepsilon$ for all N, u in (0,1) and H in $\mathcal{H}$. □

3.4. PROOF OF THEOREM 3.1.1: ASYMPTOTIC NEGLIGIBILITY OF THE REMAINDER TERMS

A substantial reduction of both number and complexity of the B-terms to be considered is possible. First it should be noted that B_{cN} can be dealt with by the methods of Section 2.4 because of the properties of the function J_c. As B_{1N} and B_{d1N} are symmetric to B_{2N} and B_{d2N} respectively, the latter rvs need not be considered in the sequel. Finally without loss of generality we may take $k = l = 1$ (see Assumption 3.1.1), $p = 1$ and $M' = a_{11} = b_{11} = 1$ (see Lemma 3.1.1) henceforth. We shall write K, K_c, K_d, L, L_c, L_d instead of $K_1, K_{1c}, K_{1d}, L_1, L_{1c}, L_{1d}$. Note that by these conventions we have in particular $K_d(u) = c(u-s_1)$ and $L_d(u) = c(u-t_1)$ for $0 < u < 1$. The results of this section depend heavily on the smoothness and boundedness properties of the functions J_c, K and L, derived in Lemma 3.1.1.

The method of the previous chapter applies essentially also to the rvs B_{1N} and B_{3N}. Instead of the bivariate mean value theorem we just use the univariate mean value theorem in order to expand the first factor (depending on K_c) in the integrand. In this manner we obtain for any ω in

$\Omega^*_{\gamma N}$ (see (2.2.3))

(3.4.1) $$N^{1/2}\, K_c(F^*_N) = N^{1/2}\, K_c(F) + U^*_N(F)K'_c(\Phi_N),$$

for all x in $\bar{\Delta}_{N1} \cap S_{\gamma 1}$ (see (2.2.1) and (2.2.2)). In (3.4.1) the random number Φ_N lies in the open interval with end points F and F^*_N. Thus we are lead to the further decomposition $B_{1N} = \sum_{i=1}^{7} C_{iN}$ of the rv B_{1N}, where

$$C_{1N} = \chi(\Omega^*_{\gamma N})\iint_{S_\gamma} U^*_N(F)[K'_c(\Phi_N)-K'_c(F)]L(G)dH_N,$$

$$C_{2N} = \chi(\Omega^*_{\gamma N})\iint_{S_\gamma} [U^*_N(F)-U_N(F)]K'_c(F)L(G)dH_N,$$

$$C_{3N} = \chi(\Omega^*_{\gamma N})\iint_{S_\gamma} U_N(F)K'_c(F)L(G)d(H_N-H),$$

$$C_{4N} = -\chi(\Omega^*_{\gamma N})\iint_{S^c_\gamma} U_N(F)K'_c(F)L(G)dH,$$

$$C_{5N} = -\chi(\Omega^{*c}_{\gamma N})\iint U_N(F)K'_c(F)L(G)dH,$$

$$C_{6N} = \chi(\Omega^*_{\gamma N})N^{1/2}\iint_{S^c_\gamma}[K_c(F^*_N)-K_c(F)]L(G)dH_N,$$

$$C_{7N} = \chi(\Omega^{*c}_{\gamma N})N^{1/2}\iint[K_c(F^*_N)-K_c(F)]L(G)dH_N.$$

As to the rv B_{3N} we obtain the decomposition $B_{3N} = \sum_{i=8}^{10} C_{iN}$, where

$$C_{8N} = \chi(\Omega^*_{\gamma N})\iint_{S_\gamma} U^*_N(F)K'_c(\Phi_N)[L(G^*_N)-L(G)]dH_N,$$

$$C_{9N} = \chi(\Omega^*_{\gamma N})N^{1/2}\iint_{S^c_\gamma}[K_c(F^*_N)-K_c(F)][L(G^*_N)-L(G)]dH_N,$$

$$C_{10N} = \chi(\Omega^{*c}_{\gamma N})N^{1/2}\iint[K_c(F^*_N)-K_c(F)][L(G^*_N)-L(G)]dH_N.$$

The proof of the asymptotic negligibility of each of these C-terms can be obtained from a proof presented in Section 2.4 by a straightforward modification. We shall content ourselves with a reference to that proof and, if necessary, briefly indicate the desired modification.

COROLLARY 3.4.1. As $\gamma \downarrow 0$ and $N \to \infty$, $C_{4N} \to_P 0$, $C_{6N} \to_P 0$ and $C_{9N} \to_P 0$, uniformly on H'.

PROOF. See the proof of Corollary 2.4.1. As far as C_{6N} and C_{9N} are concerned, the univariate mean value theorem should be applied step-wise to the first factor (depending on K_c) in the integrand, instead of step-wise application of the bivariate mean value theorem. □

COROLLARY 3.4.2. For fixed γ, $C_{5N} \to_P 0$, $C_{7N} \to_P 0$ and $C_{10N} \to_P 0$ as $N \to \infty$, uniformly on H.

PROOF. See the proof of Corollary 2.4.2. □

COROLLARY 3.4.3. For fixed γ, $C_{1N} \to_P 0$ and $C_{8N} \to_P 0$ as $N \to \infty$, uniformly on H.

PROOF. See the proof of Corollary 2.4.3. With respect to C_{1N} note that K_c' is uniformly continuous on $[\gamma/2, s_1-\gamma/2] \cup [s_1+\gamma/2, 1-\gamma/2]$. With respect to C_{8N} note that L is uniformly continuous on $[\gamma/2, t_1-\gamma/2] \cup [t_1+\gamma/2, 1-\gamma/2]$. □

COROLLARY 3.4.4. For fixed γ, $C_{2N} \to_P 0$ as $N \to \infty$, uniformly on H.

PROOF. See the proof of Corollary 2.4.4. Here we have to use the continuity of $K_c' \times L$ on the closed set $\{[\gamma, s_1-\gamma] \cup [s_1+\gamma, 1-\gamma]\} \times \{[\gamma, t_1-\gamma] \cup [t_1+\gamma, 1-\gamma]\}$. □

COROLLARY 3.4.5. For fixed γ, $C_{3N} \to_P 0$ as $N \to \infty$, uniformly on H.

PROOF. See the proof of Corollary 2.4.5, consistently replacing $J^{(1,0)}$ by $K_c' \times L$. □

Let us now turn to the B_d-terms. Using $S_{\gamma 2}$ defined in (2.2.2) we arrive at the decomposition $B_{d1N} = \sum_{i=1}^{4} D_{iN}$, where

$$D_{1N} = N^{1/2} \iint [c(F_N^* - s_1) - c(F - s_1)] L(G) dH - E_c(L(G(Y)) | F(X) = s_1) U_N(s_1),$$

$$D_{2N} = N^{1/2} \iint_{(-\infty,\infty) \times S_{\gamma 2}} [c(F_N^* - s_1) - c(F - s_1)] L(G) d(H_N - H),$$

$$D_{3N} = -N^{1/2} \iint_{(-\infty,\infty) \times S_{\gamma 2}^c} [c(F_N^* - s_1) - c(F - s_1)] L(G) dH,$$

$$D_{4N} = N^{1/2} \iint_{(-\infty,\infty) \times S_{\gamma 2}^c} [c(F_N^* - s_1) - c(F - s_1)] L(G) dH_N.$$

Furthermore let us write $B_{d3N} = \sum_{i=5}^{6} D_{iN}$, where

$$D_{5N} = N^{1/2}\iint_{(-\infty,\infty)\times S_{\gamma 2}} [c(F_N^*-s_1)-c(F-s_1)][L(G_N^*)-L(G)]dH_N,$$

$$D_{6N} = N^{1/2}\iint_{(-\infty,\infty)\times S_{\gamma 2}^c} [c(F_N^*-s_1)-c(F-s_1)][L(G_N^*)-L(G)]dH_N.$$

Some notation, introduced in the previous section, will be adapted to the present situation. More explicitly the arbitrary number u in (0,1) will be replaced by the fixed number s_1, the point where the function K has its unique simple discontinuity. We shall write N_1 for the index N_{s_1} defined in (3.3.2), and the random set Γ_{N1} will be the set defined in (3.3.3) with N_u replaced by $N_{s_1} = N_1$. Throughout this section the sets $\Omega_{1N},\ldots,\Omega_{4N}$ and the numbers M and ζ_N, figuring in their definition, have the same meaning as in Section 2.4. It is convenient to add the subsets

$$\Omega_{5N} = \{F^{-1}(F_N^*(F^{-1}(s_1))),X_{N_1:N} \in [F^{-1}(s_1-MN^{-1/2}),F^{-1}(s_1+MN^{-1/2})]\},$$

$$\Omega_{6N} = \{\sup_{I_{N2}^*} |H_N\{\Gamma_{N1} \times I_{N2}^*\}-H\{\Gamma_{N1} \times I_{N2}^*\}| \leq MN^{-3/4}\},$$

where the supremum is taken over all intervals $I_{N2}^* \subset (-\infty,\infty)$. From Lemmas 1.3.1, 2.3.1, 2.3.2 (ii) and Lemmas 3.3.3, 3.3.5 it is clear that for an arbitrary $\varepsilon > 0$ the finite positive constant $M \geq 1$ and the sequence ζ_N (decreasing to zero as $N \to \infty$) may be chosen in dependence of ε such that the set

$$(3.4.2) \qquad \Omega_N = \cap_{i=1}^{6} \Omega_{iN}$$

has probability $P_H(\Omega_N) = P(\Omega_N) \geq 1-\varepsilon$ for all N and all H in $\mathcal{H}$.

<u>COROLLARY 3.4.6</u>. As $\gamma \downarrow 0$ and $N \to \infty$, $D_{3N} \to_P 0$, $D_{4N} \to_P 0$ and $D_{6N} \to_P 0$, uniformly on $\mathcal{H}'$.

<u>PROOF</u>. Let $\alpha > 0$ be a fixed number small enough to ensure that $[s_1-\alpha,s_1+\alpha] \subset O_1$ (see Assumption 3.1.4). For small positive γ let us introduce the function

$$r_{2\gamma}(t) = r_2(t) \text{ for } t \in (0,\gamma) \cup (t_1-\gamma,t_1+\gamma) \cup (1-\gamma,1), r_{2\gamma}(t) = 0 \text{ elsewhere.}$$

Because by Assumption 3.1.4 the functions $r_{2\gamma}$ satisfy the condition of Lemma 3.1.2 for such γ, the conditional expectations $E(r_{2\gamma}(G(Y))|F(X) = s)$ have versions continuous on the open set 0_1, by convention denoted by $E_c(r_{2\gamma}(G(Y))|F(X) = s)$. Since $r_{2\gamma} \downarrow 0$ on $(0,1)$ as $\gamma \downarrow 0$, by Assumption 3.1.4 and the dominated convergence theorem we have as $\gamma \downarrow 0$

(3.4.3) $$\sup_{s\in[s_1-\alpha,s_1+\alpha],H\in H'} E_c(r_{2\gamma}(G(Y))|F(X) = s)$$
$$= \tilde{\zeta}_\gamma \le \int_0^1 r_{2\gamma}(t)\bar{g}(t)dt \to 0.$$

As to D_{3N}, application of Lemma 3.3.2 (i) yields (see also (3.3.1))

(3.4.4) $$E(|D_{3N}|) \le N^{1/2} \int_0^1 c_N(0,1;s,s_1)E_c(r_{2\gamma}(G(Y))|F(X) = s)ds.$$

As to D_{4N}, application of Lemma 3.3.2 (ii) yields

(3.4.5) $$E(|D_{4N}|) \le N^{1/2} \int_0^1 c_{N-1}(1,2;s,s_1)E_c(r_{2\gamma}(G(Y))|F(X) = s)ds.$$

Finally consider D_{6N} and note that on Ω_{4N} we have $|L(G_N^*)-L(G)| \le 2Mr_2(G)$ for x in $\bar{\Delta}_{N2}$ (see (2.2.1)). Using Lemma 3.3.2 (ii) we obtain

(3.4.6) $$E(\chi(\Omega_N)|D_{6N}|) \le 2MN^{1/2} \int_0^1 c_{N-1}(1,2;s,s_1)E_c(r_{2\gamma}(G(Y))|F(X)=s)ds.$$

Because of the similarity between the right-hand sides of (3.4.4)-(3.4.6) and because $P(\Omega_N) \ge 1-\varepsilon$ for all N and H in H, it suffices to investigate the behavior of the right-hand side of (3.4.4) as $\gamma \downarrow 0$ and $N \to \infty$. This expression is bounded by

$$N^{1/2}[\sup_{s\in[0,s_1-\alpha]\cup[s_1+\alpha,1]} c_N(0,1;s,s_1)][\int_0^1 r_2(t)dt] +$$
$$N^{1/2}[\int_0^1 c_N(0,1;s,s_1)ds]\tilde{\zeta}_\gamma \to 0$$

as $\gamma \downarrow 0$ and $N \to \infty$, uniformly for H in H', by Lemma 3.3.1. □

<u>COROLLARY 3.4.7</u>. As $N \to \infty$, $D_{1N} \to_P 0$, uniformly on H'.

<u>PROOF</u>. As in the proof of the preceding corollary, let $\alpha > 0$ be a fixed number small enough to ensure that $[s_1-\alpha,s_1+\alpha] \subset 0_1$ (see Assumption 3.1.4).

Let us consider only values N large enough to ensure that $[F^{-1}(s_1-MN^{-1/2}), F^{-1}(s_1+MN^{-1/2})] \subset [F^{-1}(s_1-\alpha), F^{-1}(s_1+\alpha)]$. Using Lemma 3.3.4 (with $u = s_1$) we may write $D_{1N} = N^{1/2}\iint_{\Gamma_{N1}\times(-\infty,\infty)} \operatorname{sgn}(F^{-1}(s_1)-X_{N_1:N})L(G)dH - E_c(L(G(Y))|F(X) = s_1)U_N(s_1) = \sum_{i=1}^{3} D_{1iN}$, where

$$D_{11N} = \chi(\Omega_N^c)D_{1N}$$

$$D_{12N} = \chi(\Omega_N)[N^{1/2}\int_{2s_1-F_N^*(F^{-1}(s_1))}^{s_1} E_c(L(G(Y))|F(X) = s)ds - E_c(L(G(Y))|F(X) = s_1)U_N(s_1)],$$

$$D_{13N} = \chi(\Omega_N)N^{1/2}\int_{F(X_{N_1:N})}^{2s_1-F_N^*(F^{-1}(s_1))} E_c(L(G(Y))|F(X) = s)ds.$$

Because of the properties of the set Ω_N it follows that $P(\{D_{11N} \neq 0\}) \leq \varepsilon$ for all N and all H in $\mathcal{H}$.

From Assumption 3.1.4 it follows that

$$(3.4.7) \qquad \sup_{s\in[s_1-\alpha,s_1+\alpha],H\in\mathcal{H}'} |E_c(L(G(Y))|F(X) = s)| = \tilde{M} < \infty,$$

and since $E_c(L(G(Y))|F(X) = s)$ is continuous on $[s_1-\alpha,s_1+\alpha]$ the mean value theorem for integrals applies. We thus obtain, writing $\Phi_N(s_1)$ for a random point between s_1 and $2s_1 - F_N^*(F^{-1}(s_1))$ and using (3.4.7),

$$|D_{12N}| \leq \chi(\Omega_N)N^{1/2}|F_N^*(F^{-1}(s_1))-s_1| \times$$

$$|E_c(L(G(Y))|F(X) = \Phi_N(s_1)) - E_c(L(G(Y))|F(X) = s_1)| +$$

$$\chi(\Omega_N)\tilde{M}N^{1/2}|F_N(F^{-1}(s_1))-F_N^*(F^{-1}(s_1))|.$$

As $\Omega_N \subset \Omega_{5N}$, for each ω in Ω_N the random point $\Phi_N(s_1)$ satisfies $|\Phi_N(s_1)-s_1| \leq MN^{-1/2}$, so that the equicontinuity condition concerning the densities $\bar{h}$ corresponding to the H in the subclass $\mathcal{H}'$ (see Assumption 3.1.4) yields that the first term in the bound for $|D_{12N}|$ converges to zero as $N \to \infty$, uniformly for all H in $\mathcal{H}'$. The same is true for the second term in this bound, since $|F_N(F^{-1}(s_1))-F_N^*(F^{-1}(s_1))| = 1/(N+1)$.

The rv D_{13N} is bounded by

$$|D_{13N}| \leq \chi(\Omega_N)\tilde{M}N^{1/2}|F_N(F^{-1}(s_1))-F_N(X_{N_1:N}-)+F(X_{N_1:N})-s_1| +$$

$$\chi(\Omega_N)\tilde{M}N^{1/2}|F_N^*(F^{-1}(s_1))-F_N(F^{-1}(s_1))+F_N(X_{N_1:N}-)-s_1|$$

$$\leq \chi(\Omega_N)\tilde{M}N^{1/2}|H_N\{\Gamma_{N1}\times(-\infty,\infty)\}-H\{\Gamma_{N1}\times(-\infty,\infty)\}| +$$

$$\chi(\Omega_N)\tilde{M}N^{1/2}[1/(N+1)+|(N_1-1)/N-s_1|] \to 0$$

as $N \to \infty$, uniformly for H in H', by (3.3.2) with $u = s_1$, (3.4.7) and because $\Omega_N \subset \Omega_{6N}$. □

<u>COROLLARY 3.4.8</u>. For fixed γ, $D_{2N} \to_P 0$ as $N \to \infty$, uniformly on H.

<u>PROOF</u>. For each positive integer m, consider the function $L(I_m(t))$ for t in (0,1) (see (2.3.3)). For any such m, with the aid of Lemma 3.3.4, let us make the decomposition $D_{2N} = D_{21N} + \sum_{i=2}^{4} D_{2imN}$, where

$$D_{21N} = \chi(\Omega_N^c)D_{2N},$$

$$D_{22mN} = \chi(\Omega_N)N^{1/2}\ \text{sgn}\ (F^{-1}(s_1)-X_{N_1:N})\iint_{\Gamma_{N1}\times S_{\gamma 2}} L(I_m(G))d(H_N-H),$$

$$D_{23mN} = \chi(\Omega_N)N^{1/2}\iint_{(-\infty,\infty)\times S_{\gamma 2}}[c(F_N^*-s_1)-c(F-s_1)] \times [L(G)-L(I_m(G))]dH_N,$$

$$D_{24mN} = \chi(\Omega_N)N^{1/2}\iint_{(-\infty,\infty)\times S_{\gamma 2}}[c(F_N^*-s_1)-c(F-s_1)] \times [L(I_m(G))-L(G)]dH.$$

Because of the properties of the set Ω_N it follows that $P(\{D_{21N} \neq 0\}) \leq \varepsilon$ for all N and all H in H.

The function $L(I_m(G))$ assumes the value $L((j-1)/m)$ on the set $S_{\gamma j2}$, where

$$S_{\gamma j2} = \{y : G(y) \in [(j-1)/m, j/m)\} \cap S_{\gamma 2},$$

for $j = 1,\ldots,m$. Let $\tilde{M}_\gamma = \max_{S_{\gamma 2}}|L(G)|$, then because $\Omega_N \subset \Omega_{6N}$ for every ω we have

$$|D_{22mN}| = \chi(\Omega_N)N^{1/2}|\sum_{j=1}^{m} L((j-1)/m)\iint_{\Gamma_{N1}\times S_{\gamma j2}} d(H_N-H)|$$
$$\leq \chi(\Omega_N)N^{1/2}\tilde{M}_\gamma \sum_{j=1}^{m}|H_N\{\Gamma_{N1}\times S_{\gamma j2}\}-H\{\Gamma_{N1}\times S_{\gamma j2}\}|$$
$$\leq m\tilde{M}_\gamma MN^{-1/4} \to 0$$

for fixed m as $N \to \infty$, uniformly on H.

Since L(G) is bounded and continuous on $S_{\gamma 2}$ we have $\sup_{S_{\gamma 2}}|L(G)-L(I_m(G))| = \tilde{\zeta}_{\gamma m} \to 0$ for fixed γ as $m \to \infty$, uniformly on H. Application of Lemma 3.3.2 (ii) and (i) with $\psi(G) = \tilde{\zeta}_{\gamma m}$ shows that the expectations of $|D_{23mN}|$ and $|D_{24mN}|$ are bounded by

$$E(|D_{23mN}|) \leq N^{1/2}\tilde{\zeta}_{\gamma m}\int_0^1 c_{N-1}(1,2;s,s_1)ds,$$
$$E(|D_{24mN}|) \leq N^{1/2}\tilde{\zeta}_{\gamma m}\int_0^1 c_N(0,1;s,s_1)ds$$

respectively. Application of Lemma 3.3.1 produces the convergence to zero of both expectations for fixed γ as $m, N \to \infty$, uniformly on H. Combination of these partial results leads to the conclusion of the corollary. □

<u>COROLLARY 3.4.9</u>. For fixed γ, $D_{5N} \to_P 0$ as $N \to \infty$, uniformly on H.

<u>PROOF</u>. Application of Lemma 3.3.2 (ii) with $\psi(G) = 1$ gives (see (2.2.1))

$$|D_{5N}| \leq \sup_{\bar{\Delta}_{N2}\cap S_{\gamma 2}}|L(G_N^*)-L(G)|N^{1/2}\int_0^1 c_{N-1}(1,2;s,s_1)ds.$$

The function L is uniformly continuous on $[\gamma/2, t_1-\gamma/2] \cup [t_1+\gamma/2, 1-\gamma/2]$ and $|G_N^*-G| \leq 1/(N+1) + |G_N-G|$. Hence by the Glivenko-Cantelli theorem we have $\sup_{\bar{\Delta}_{N2}\cap S_{\gamma 2}}|L(G_N^*)-L(G)| \to_P 0$ as $N \to \infty$, uniformly on H. The proof may be concluded by applying Lemma 3.3.1 (ii).

To show the asymptotic negligibility of the B-terms, these corollaries should be combined in a similar fashion as at the end of Section 2.4.

Chapter 4

SOME ASYMPTOTIC PROPERTIES OF THE RANK TESTS

4.1. CONSISTENCY AND VARIANCE STABILIZING TRANSFORMATIONS

In this section we give some immediate statistical consequences of the results obtained in the preceding chapters. Let us consider one-sided tests based on rank statistics of the type $T_N = \int\int J_N(F_N, G_N)dH_N$ with critical regions of the form $\{T_N \geq C_{\alpha,N}\}$, where the numbers $C_{\alpha,N}$ are defined in (1.1.8). For α we choose a natural significance level for T_N in (0,1). It has been explained in Section 1.1 that there exists a natural number $N(\alpha)$ such that for any rank test based on $N \geq N(\alpha)$ observations the level α can be attained without randomization. It will be tacitly understood that the sample sizes N satisfy $N \geq N(\alpha)$ if necessary. The limiting score function J is supposed to satisfy the conditions of Theorem 3.1.1. The case where J satisfies the conditions of Theorem 2.1.1 can be dealt with in the same way. An asymptotic expression for the numbers $C_{\alpha,N}$ is obtained with the aid of Theorem 3.1.1, assuming that the underlying df $H_{(0)}$ of the sequence $(X_1,Y_1),(X_2,Y_2),\ldots$ belongs to the null hypothesis $\mathcal{H}_0$. The asymptotic behavior of the power $P_H(\{T_N \geq C_{\alpha,N}\})$ for the fixed member H of the alternative $\mathcal{H}_1$ is studied with the aid of Theorem 3.1.1 under the assumption that the underlying df of the above sequence is H.

<u>THEOREM 4.1.1</u>. Suppose that for the score functions J_N and the limiting score function J Assumptions 3.1.1–3.1.4 are satisfied with $\mathcal{H}' = \{H_{(0)}, H\}$ and with $\delta = 0$. Here $H_{(0)} \in \mathcal{H}_0$ and $H \in \mathcal{H}_1$ are fixed. Then the asymptotic means $\mu_0 = \mu(H_{(0)})$, $\mu(H)$ (see (2.1.5) and (2.1.1) respectively) and variances $\sigma_0^2 = \sigma^2(H_{(0)})$, $\sigma^2(H)$ (see (3.1.4) and (3.1.2) respectively) are finite. Moreover for tests with critical regions $\{T_N \geq C_{\alpha,N}\}$ we have
(i) $N^{1/2}(C_{\alpha,N}-\mu_0) \to \sigma_0\Phi^{-1}(1-\alpha)$ as $N \to \infty$. Here Φ is the standard normal df.
(ii) $P_H(\{T_N \geq C_{\alpha,N}\}) \to 1$ as $N \to \infty$, provided $\mu_0 < \mu(H)$.

<u>PROOF</u>. By Theorem 3.1.1 the finiteness of the quantities μ_0, $\mu = \mu(H)$, σ_0^2 and $\sigma^2 = \sigma^2(H)$ follows immediately.
(i) If $\sigma_0 = 0$, then under the null hypothesis $N^{1/2}(T_N-\mu) \to_P 0$ by Theorem 3.1.1, which implies (i) in this case. Note that α is a natural significance level in (0,1) which implies that $\mathrm{Var}(T_N) \neq 0$ for finite N and that $\sigma_0\Phi^{-1}(1-\alpha)$ is well defined. Assume now that $\sigma_0 > 0$. If $H_{(0)}$ is the under-

lying df of the random sequence, application of Theorem 3.1.1 yields that $P_0(\{N^{1/2}(T_N-\mu_0)/\sigma_0 \geq \Phi^{-1}(1-\alpha)\}) \to \alpha$ as $N \to \infty$. On the other hand in view of (1.1.8) the numbers $C_{\alpha,N}$ satisfy $P_0(\{N^{1/2}(T_N-\mu_0)/\sigma_0 \geq N^{1/2}(C_{\alpha,N}-\mu_0)/\sigma_0\}) = \alpha$ for all N. From this the assertion in the first part of the theorem follows.
(ii) If H is the underlying df of the random sequence, application of Theorem 3.1.1 yields that $P_H(\{T_N \geq C_{\alpha,N}\}) = P_H(\{N^{1/2}(T_N-\mu) \geq N^{1/2}(C_{\alpha,N}-\mu)\}) = P_H(\{N^{1/2}(T_N-\mu) \geq N^{1/2}(C_{\alpha,N}-\mu_0) + N^{1/2}(\mu_0-\mu)\}) \to 1$ as $N \to \infty$, since by part (i) $N^{1/2}(C_{\alpha,N}-\mu_0) \to \sigma_0\Phi^{-1}(1-\alpha)$ and by the assumption of the theorem $N^{1/2}(\mu_0-\mu) \to -\infty$ as $N \to \infty$. □

The property (ii) of the power is called *consistency* of the sequence of one-sided tests based on $T_N(N=1,2,\ldots)$, at level α, against all fixed alternatives $H \in \mathcal{H}_1$ with $\mu(H) > \mu_0$.

For a further application let us consider a parametric class of alternatives $\{H_\theta \in \mathcal{H}, 0 \leq \theta < \theta_0\}$, depending on a single real parameter, and let

$$(4.1.1) \qquad \{H_\theta^* \in \mathcal{H} : H_\theta^* = \bar{H}_\theta(F^*,G^*), F^* \times G^* \in \mathcal{H}_0, 0 \leq \theta < \theta_0\}$$

be the more general corresponding class of non-parametric alternatives (see (1.2.11) and the end of Section 1.2). In principle the formulas for the asymptotic mean and variance enable us to express $\mu_\theta = \mu(H_\theta^*)$ and $\sigma_\theta^2 = \sigma^2(H_\theta^*)$ as functions of θ. It may turn out that there exists a function g, continuous and positive on the domain of interest, such that the variance can be written as

$$\sigma_\theta^2 = g(\mu_\theta), \qquad 0 \leq \theta < \theta_0.$$

Suppose that the function f is a solution of the equation

$$(4.1.2) \qquad [f'(\mu)]^2 g(\mu) = 1.$$

Then (see [26], [32], Section 6g, or [42])

$$(4.1.3) \qquad N^{1/2}[f(T_N)-f(\mu_\theta)] \to_d \mathcal{N}(0,1), \qquad 0 \leq \theta < \theta_0.$$

Hence the statistic is transformed in such a way that the asymptotic variance equals 1, whatever the underlying df in the class (4.1.1) may be. Such a transformation is therefore called a *variance stabilizing transformation.* In Section 5.3 an example, in which a variance stabilizing transformation can be calculated explicitly, will be indicated.

4.2. ASYMPTOTIC POWER AND EFFICIENCY

So far all asymptotic results were based on a given sequence of mutually independent and identically distributed two-dimensional random vectors. In this section we shall be concerned with the more general situation that a triangular array of two-dimensional random vectors is given. For each index $\nu = 1,2,\ldots$ let be given a set

$$(4.2.1) \qquad (X_1^{(\nu)},Y_1^{(\nu)}),\ldots,(X_{N_\nu}^{(\nu)},Y_{N_\nu}^{(\nu)})$$

of N_ν iid two-dimensional random vectors, having common df $H_{(\nu)} \in \mathcal{H}$, the marginal dfs of which will be denoted by $F_{(\nu)}$ and $G_{(\nu)}$. The empirical df of this sample will be denoted by $H_{N_\nu}^{(\nu)}$ and its marginal empirical dfs by $F_{N_\nu}^{(\nu)}$ and $G_{N_\nu}^{(\nu)}$. It will be assumed throughout that the sample sizes satisfy $N_\nu \to \infty$ as $\nu \to \infty$. Without loss of generality for each ν the set of rvs in (4.2.1) may be thought of as defined on a single probability space $(\Omega_\nu, A_\nu, P_\nu)$, where $P_\nu = P_{H_{(\nu)}}$ is such that

$$P_\nu(\{X_n^{(\nu)} \le x, Y_n^{(\nu)} \le y\}) = H_{(\nu)}(x,y)$$

for all $-\infty < x < \infty$, $-\infty < y < \infty$. The basic difference with the previous set-up is that for different ν the samples may have different underlying dfs.

For each index $\nu = 1,2,\ldots$ we consider a score function J_ν defined on $(0,1) \times (0,1)$ on which a statistic

$$(4.2.2) \qquad T_\nu = \iint J_\nu\,(F_{N_\nu}^{(\nu)}, G_{N_\nu}^{(\nu)})\,dH_{N_\nu}^{(\nu)}$$

will be based. The asymptotic behavior of the above rank statistic will be investigated, a suitable standardization of which will be $N_\nu^{1/2}(T_\nu - \mu(H_{(\nu)}))$. Here for some limiting score function J, defined on $(0,1) \times (0,1)$, the parameter $\mu(H_{(\nu)})$ equals

$$(4.2.3) \qquad \mu(H_{(\nu)}) = E[J(F_{(\nu)}(X), G_{(\nu)}(Y))].$$

Here (X,Y) has bivariate df $H_{(\nu)}$. Asymptotic normality can be obtained under the assumption that the underlying df $H_{(\nu)}$ of the ν-th sample converges weakly to an arbitrary df $H \in \mathcal{H}$ as $\nu \to \infty$. In most applications this limiting df H will be a member of $\mathcal{H}_0$.

The asymptotic normality will be established for limiting score functions J that are not necessarily continuous, and Theorem 3.1.1 will be used. A similar result can be given in the case where J is continuous using Theorem 2.1.1. In the latter case Assumption 4.2.4 below may be replaced by a simpler assumption asserting the weak convergence of $H_{(\nu)}$ to H only.

REMARK. Throughout Assumptions 4.2.1-4.2.4 the points $0 = s_0 < s_1 < \ldots < s_k < s_{k+1} = 1$ and $0 = t_0 < t_1 < \ldots < t_l < t_{l+1} = 1$ are the same fixed elements of the unit interval, and the functions $r_1, \tilde{r}_1, r_2, \tilde{r}_2$ are the same fixed members of $\mathcal{R}$ (see Definition 1.3.2).

ASSUMPTION 4.2.1. See Assumption 3.1.1.

ASSUMPTION 4.2.2. See Assumption 2.1.2.

ASSUMPTION 4.2.3. Let a subclass $\mathcal{H}' \subset \mathcal{H}$ be given. As $\nu \to \infty$, $B^*_{0\nu} = N_\nu^{1/2} \iint [J_\nu(F_{N_\nu}^{(\nu)}, G_{N_\nu}^{(\nu)}) - J(F_{N_\nu}^{(\nu)*}, G_{N_\nu}^{(\nu)*})] dH_{N_\nu}^{(\nu)} \to_{P_\nu} 0$, uniformly on $\mathcal{H}'$.

ASSUMPTION 4.2.4. As $\nu \to \infty$, $H_{(\nu)}(x,y) \to H(x,y)$ for all $(x,y) \in (-\infty,\infty) \times (-\infty,\infty)$, where $H, H_{(1)}, H_{(2)}, \ldots \in \mathcal{H}$. Furthermore Assumption 3.1.4 is satisfied with $\mathcal{H}' = \{H, H_{(1)}, H_{(2)}, \ldots\}$. If $\bar{h}, \bar{h}_1, \bar{h}_2, \ldots$ are the densities of the transformed dfs corresponding to $H, H_{(1)}, H_{(2)}, \ldots$ we have $\bar{h}_\nu(s_i,t) \to \bar{h}(s_i,t)$ for all $t \in (0,1)$, $i = 1,\ldots,k$ and $\bar{h}_\nu(s,t_j) \to \bar{h}(s,t_j)$ for all $s \in (0,1)$, $j = 1,\ldots,l$, as $\nu \to \infty$.

THEOREM 4.2.1. Suppose that the sample with index ν has underlying df $H_{(\nu)} \in \mathcal{H}, \nu = 1,2,\ldots$ and let $H \in \mathcal{H}$. Let for the score functions J_ν and the limiting score function J Assumptions 4.2.1-4.2.4 be satisfied with $\mathcal{H}' = \{H, H_{(1)}, H_{(2)}, \ldots\}$ and $\delta > 0$. Then $N_\nu^{1/2}(T_\nu - \mu(H_{(\nu)})) \to_d \mathcal{N}(0, \sigma^2(H))$ as $\nu \to \infty$, provided $\sigma^2(H) > 0$. Here $\mu(H_{(\nu)})$ and $\sigma^2(H)$, given by (2.1.1) and (3.1.2) respectively, are finite.

PROOF. By Theorem 3.1.1 the finiteness of $\mu(H_{(\nu)})$, $\mu(H)$, $\sigma_\nu^2 = \sigma^2(H_{(\nu)})$ and $\sigma^2 = \sigma^2(H)$ follows at once. It suffices to prove that $\sigma_\nu^2 \to \sigma^2$ as $\nu \to \infty$. Then $\sigma_\nu^2 \geq \sigma^2/2 > 0$ for $\nu \geq \nu_0$ say, and all conditions, necessary for the application of the part of Theorem 3.1.1 concerning the uniformity with

$H' = \{H_{(\nu_0)}, H_{(\nu_0+1)}, \ldots\}$, are covered by the conditions of the present theorem. We may therefore conclude that the convergence $N_\nu^{1/2}(T_\nu - \mu(\tilde{H})) \to_d N(0,\sigma^2(\tilde{H}))$ is uniform for $\tilde{H}$ in $\{H_{(\nu_0)}, H_{(\nu_0+1)}, \ldots\}$ as $\nu \to \infty$. But if $\sigma_\nu^2 \to \sigma^2$, $N(0,\sigma_\nu^2)$ converges weakly to $N(0,\sigma^2)$, and thus we finally obtain $N_\nu^{1/2}(T_\nu - \mu(H_\nu)) \to_d N(0,\sigma^2(H))$ as $\nu \to \infty$.

As in Section 3.4 let us assume that $k = l = 1$ (see Assumption 4.2.1), $p = 1, M' = 1$ and $a_{11} = b_{11} = 1$ (see Lemma 3.1.1). We shall write K, L instead of K_1, L_1. For a function $\phi(F_{(\nu)}, G_{(\nu)})$, integrable with respect to $H_{(\nu)}$, we have $\iint \phi(F_{(\nu)}(x), G_{(\nu)}(y)) dH_{(\nu)}(x,y) = \iint \phi(s,t) d\bar{H}_{(\nu)}(s,t)$, where $\bar{H}_{(\nu)}(s,t) = H_{(\nu)}(F_{(\nu)}^{-1}(s), G_{(\nu)}^{-1}(t))$ for (s,t) in $(0,1) \times (0,1)$. Note that $\bar{H}_{(\nu)}$ has $Un(0,1)$ marginal dfs. In order to derive a more explicit expression for the variance σ_ν^2 let us first observe that the expectation under $H_{(\nu)}$ of the rv within the brackets on the right-hand side of (3.1.2) equals $\iint J(u,v) d\bar{H}_{(\nu)}(u,v)$. Writing moreover the square of an integral as a repeated integral we arrive at the formula

$$(4.2.4) \qquad \sigma_\nu^2 = \int_0^1\int_0^1 \{J(s,t) - \iint J(u,v) d\bar{H}_{(\nu)}(u,v) +$$

$$\iint [c(u-s)-u] J^{(1,0)}(u,v) d\bar{H}_{(\nu)}(u,v) +$$

$$\iint [c(v-t)-v] J^{(0,1)}(u,v) d\bar{H}_{(\nu)}(u,v) +$$

$$[c(s_1-s)-s_1] E_c(L(G_{(\nu)}(Y)) | F_{(\nu)}(X) = s_1) +$$

$$[c(t_1-t)-t_1] E_c(K(F_{(\nu)}(X)) | G_{(\nu)}(Y) = t_1)\}^2 d\bar{H}_{(\nu)}(s,t)$$

$$= \sum_{i=1}^6 \sum_{j=1}^6 \iiiint\iint \phi_i(s,t,u,v) \phi_j(s,t,u',v')$$

$$d\bar{H}_{(\nu)}(u',v') d\bar{H}_{(\nu)}(u,v) d\bar{H}_{(\nu)}(s,t),$$

for $\nu = 1,2,\ldots$ Here s,t,u,v,u',v' are restricted to $(0,1)$, (X,Y) has bivariate df $H_{(\nu)}$, and

$$\phi_1(s,t,u,v) = J(s,t),$$

$$\phi_2(s,t,u,v) = -J(u,v),$$

$$\phi_3(s,t,u,v) = [c(u-s)-u]J^{(1,0)}(u,v),$$

$$\phi_4(s,t,u,v) = [c(v-t)-v]J^{(0,1)}(u,v),$$

$$\phi_5(s,t,u,v) = [c(s_1-s)-s_1]E_c(L(G_{(\nu)}(Y))|F_{(\nu)}(X) = s_1),$$

$$\phi_6(s,t,u,v) = [c(t_1-t)-t_1]E_c(K(F_{(\nu)}(X))|G_{(\nu)}(Y) = t_1).$$

According to Assumption 4.2.1, Lemma 2.2.1 (the finite positive constant M_i depends on q_i only, $i = 1,2$) and Assumption 4.2.4 these functions are bounded by

$$|\phi_1(s,t,u,v)| \leq r_1(s)r_2(t),$$

$$|\phi_2(s,t,u,v)| \leq r_1(u)r_2(v),$$

$$|\phi_3(s,t,u,v)| \leq M_1[q_1(s)]^{-1}q_1(u)\tilde{r}_1(u)r_2(v),$$

$$|\phi_4(s,t,u,v)| \leq M_2[q_2(t)]^{-1}q_2(v)r_1(u)\tilde{r}_2(v),$$

$$|\phi_5(s,t,u,v)| \leq \int_0^1 r_2(t)\bar{g}(t)dt,$$

$$|\phi_6(s,t,u,v)| \leq \int_0^1 r_1(s)\bar{f}(s)ds.$$

The convergence $H_{(\nu)}(x,y) \to H(x,y)$ for all x,y as $\nu \to \infty$ (see Assumption 4.2.4) entails the convergence $\bar{H}_{(\nu)}(u',v')\bar{H}_{(\nu)}(u,v)\bar{H}_{(\nu)}(s,t) \to \bar{H}(u',v')\bar{H}(u,v)\bar{H}(s,t)$ for all $u',v',u,v,s,t \in (0,1)$, as $\nu \to \infty$. A further application of Assumption 4.2.4 combined with the dominated convergence theorem yields

$$E_c(L(G_{(\nu)}(Y))|F_{(\nu)}(X) = s_1) \to \int_0^1 L(t)\bar{h}(s_1,t)dt,$$

$$E_c(K(F_{(\nu)}(X))|G_{(\nu)}(Y) = t_1) \to \int_0^1 K(s)\bar{h}(s,t_1)ds,$$

as $\nu \to \infty$. Convergence of each of the summands on the right in (4.2.4) suffices for the convergence of the σ_ν^2 to σ^2. The functions $\phi_1,\ldots,\phi_4$ are independent of ν, and the functions ϕ_5 and ϕ_6 depend on ν only through multiplicative constants that converge properly. It follows from Billingsley [5], Theorem 5.4, that a sufficient condition for the convergence of σ_ν^2 to σ^2 is that for some $\zeta > 0$

$$(4.2.5) \qquad \sup_{\nu=1,2,\ldots} \int\int\int\int\int\int |\phi_i(s,t,u,v)\phi_j(s,t,u',v')|^{1+\zeta} d\bar{H}_{(\nu)}(u',v')d\bar{H}_{(\nu)}(u,v)d\bar{H}_{(\nu)}(s,t) < \infty,$$

for $1 \leq i \leq j \leq 6$. By Schwarz's inequality, the nature of the bounds for the $|\phi_i|$ (in particular the boundedness of ϕ_5 and ϕ_6) and the similarity between ϕ_3 and ϕ_4 it follows that we have to verify (4.2.5) only for $i = j = 1,2,3$.

Henceforth let us choose $\zeta = \delta/2 > 0$ (see the conditions of the theorem) and let us first take $i = j = 1$. Since ϕ_1 is a function of s and t only, the supremum in (4.2.5) is bounded by

$$\sup_{\nu=1,2,\ldots} \int\int [r_1(s)r_2(t)]^{2+\delta} d\bar{H}_{(\nu)}(s,t) < \infty,$$

by Assumption 4.2.2. The function ϕ_2 does not depend on s,t so that for $i = j = 2$ the supremum in (4.2.5) is bounded by

$$\sup_{\nu=1,2,\ldots} \{\int\int [r_1(u)r_2(v)]^{1+\delta/2} d\bar{H}_{(\nu)}(u,v)\}^2 < \infty,$$

by Assumption 4.2.2. Finally for $i = j = 3$ we see that the supremum in (4.2.5) is bounded by

$$\sup_{\nu=1,2,\ldots} M_1^{2+\delta} \int\int\int\int\int\int [q_1(s)]^{-2-\delta}[q_1(u)\tilde{r}_1(u)r_2(v)]^{1+\delta/2} \times$$
$$[q_1(u')\tilde{r}_1(u')r_2(v')]^{1+\delta/2} d\bar{H}_{(\nu)}(u',v')d\bar{H}_{(\nu)}(u,v)d\bar{H}_{(\nu)}(s,t)$$
$$\leq \sup_{\nu=1,2,\ldots} M_1^{2+\delta} \int_0^1 [q_1(s)]^{-2-\delta} ds \times$$
$$\{\int\int [q_1(u)\tilde{r}_1(u)r_2(v)]^{1+\delta/2} d\bar{H}_{(\nu)}(u,v)\}^2 < \infty,$$

again by Assumption 4.2.2. □

Next Theorem 4.2.1 will be applied to calculate the asymptotic power of the one-sided test based on T_ν with critical region $\{T_\nu \geq C_{\alpha,\nu}\}$ (see

(1.1.8)). Here α is a natural significance level in $(0,1)$ and ν is supposed to be sufficiently large in order that this level can be attained without randomization. In the present application the underlying df $H_{(\nu)}$ of the ν-th sample belongs to the alternative H_1 for $\nu = 1,2,\ldots$, and now $H = H_{(0)}$ is a member of the null hypothesis H_0. As usual we write $\mu(H_{(0)}) = \mu_0$ and $\sigma^2(H_{(0)}) = \sigma_0^2$. An additional assumption concerning the limiting behavior of $N_\nu^{1/2}(\mu(H_{(\nu)})-\mu_0)$ will be useful.

<u>THEOREM 4.2.2</u>. Let for the score functions J_ν and the limiting score function J Assumptions 4.2.1-4.2.4 be satisfied with $H' = \{H_{(0)}, H_{(1)}, H_{(2)}, \ldots\}$ and $\delta > 0$. Here $H_{(0)} \in H_0$ (null hypothesis) and $H_{(\nu)} \in H_1$ (alternative) for $\nu = 1,2,\ldots$ Then the numbers $\mu(H_{(\nu)})$ and σ_0^2, given by (2.1.1) and (3.1.4) respectively, are finite for $\nu = 0,1,2,\ldots$ Suppose that $\sigma_0^2 > 0$ and that $N_\nu^{1/2}(\mu(H_{(\nu)})-\mu_0)/\sigma_0 \to e$ as $\nu \to \infty$, for some finite constant e. Then for the tests with critical regions $\{T_\nu \geq C_{\alpha,\nu}\}$ the power function satisfies $P_\nu(\{T_\nu \geq C_{\alpha,\nu}\}) \to 1-\Phi(\Phi^{-1}(1-\alpha)-e)$ as $\nu \to \infty$. Here Φ is the standard normal df.

<u>PROOF</u>. The finiteness of $\mu_\nu = \mu(H_{(\nu)})$ and $\sigma_\nu^2 = \sigma^2(H_{(\nu)})$ for $\nu = 0,1,2,\ldots$ follows at once from Theorem 3.1.1. Notice that $N_\nu^{1/2}(T_\nu-\mu_0) = N_\nu^{1/2}(T_\nu-\mu_\nu) + N_\nu^{1/2}(\mu_\nu-\mu_0)$, so that $P_\nu(\{T_\nu \geq C_{\alpha,\nu}\}) = P_\nu(\{N_\nu^{1/2}(T_\nu-\mu_\nu)/\sigma_0 \geq N_\nu^{1/2}(C_{\alpha,\nu}-\mu_0)/\sigma_0 - N_\nu^{1/2}(\mu_\nu-\mu_0)/\sigma_0\})$. Application of Theorem 4.1.1 (i) yields $N_\nu^{1/2}(C_{\alpha,\nu}-\mu_0)/\sigma_0 \to \Phi^{-1}(1-\alpha)$ as $\nu \to \infty$ and application of Theorem 4.2.1 with $H = H_{(0)}$ leads to the conclusion of the present theorem. □

The limit $1-\Phi(\Phi^{-1}(1-\alpha)-e)$ occurring in the above theorem is called the *asymptotic power* of the sequence of one-sided tests based on T_ν, at level α, **against** the sequence of alternatives $H_{(\nu)} \in H_1$ $(\nu = 1,2,\ldots)$, see Hájek and Šidák [17]. It should be noted that the number e not only depends on the statistics T_ν but also on the special choice of the sequence of alternatives $H_{(\nu)}$ and their relation to the sample sizes N_ν.

The number e has an interpretation in the context of asymptotic relative efficiencies. For each ν consider a pair of rank statistics $T_{i\nu}$ $(i = 1,2)$ based on samples of size N_ν from the df $H_{(\nu)}$. Here $H_{(\nu)} \in H_1$ for $\nu = 1,2,\ldots$ and $H_{(\nu)} \to H_{(0)} \in H_0$ on $(-\infty,\infty) \times (-\infty,\infty)$ as $\nu \to \infty$. Let us consider the one-sided rank tests with critical regions $\{T_{i\nu} \geq C_{i,\alpha,\nu}\}$, $i = 1,2$, based on these statistics. Let for both statistics the conditions of Theorem 4.2.2 be satisfied, but let us, in order to obtain asymptotic powers strictly between α and 1, more restrictively assume that $N_\nu^{1/2}(\mu_i(H_{(\nu)})-\mu_{i0})/\sigma_{i0} \to e_i > 0$ as $\nu \to \infty$, for $i = 1,2$. Then the *asymptotic relative efficiency* of the $T_{1\nu}$-tests with respect to the $T_{2\nu}$-tests for the sequence of alternatives $\{H_{(\nu)}, \nu = 1,2,\ldots\}$ at level α in $(0,1)$ equals

$$\text{(4.2.6)} \qquad \text{ARE}(T_{1\nu},T_{2\nu};\{H_{(\nu)}\},\alpha) = (e_1/e_2)^2.$$

Actually, if $\tilde{N}_{i\nu}$ is any pair of sample sizes $(i = 1,2)$ for which the asymptotic powers are equal and strictly between α and 1, it follows that for some finite constant $\rho > 0$ we have $\lim_{\nu\to\infty} \tilde{N}_{1\nu}^{1/2}(\mu_1(H_{(\nu)})-\mu_{10})/\sigma_{10} = \lim_{\nu\to\infty} \tilde{N}_{2\nu}^{1/2}(\mu_2(H_{(\nu)})-\mu_{20})/\sigma_{20} = \rho$. From the conditions of the theorem it follows that for $i = 1,2$ we have $(\tilde{N}_{i\nu}/N_\nu) \to (\rho/e_i)^2$ as $\nu \to \infty$, so that $(\tilde{N}_{2\nu}/\tilde{N}_{1\nu}) \to (e_1/e_2)^2$ as $\nu \to \infty$. This is the definition of asymptotic relative efficiency. In a similar fashion the asymptotic relative efficiency of a rank test with respect to an arbitrary (e.g. parametric) test can be calculated, provided the statistic on which the latter test is based has an asymptotic behavior like that of the rank test in Theorem 4.2.2.

4.3. ASYMPTOTIC OPTIMALITY

For the material of this section our basic references are Witting and Nölle [42] and Behnen [2,3]. Let us consider a parametric class of alternatives $\{H_\theta \in \mathcal{H}, 0 \le \theta < \theta_0; H_0 \in \mathcal{H}_0\}$, depending on a single real parameter, and let

$$\text{(4.3.1)} \qquad \{H_\theta^* \in \mathcal{H}: H_\theta^* = \bar{H}_\theta(F^*,G^*), F^* \times G^* \in \mathcal{H}_0, 0 \le \theta < \theta_0; H_0^* \in \mathcal{H}_0\}$$

be the corresponding more general class of non-parametric alternatives (see (1.2.11) and the end of Section 1.2). For each $\nu = 1,2,\ldots$ let $0 < \theta_\nu < \theta_0$ and suppose that $\theta_\nu \to 0$ as $\nu \to \infty$. To carry through the notation of the previous section, let us write $H_{\theta_\nu}^* = H_{(\nu)}^*$ and $H_0^* = H_{(0)}^*$.

From now on let us fix the choice $F^* \times G^* \in \mathcal{H}_0$. In Section 1.2 we have seen that $H_{(\nu)}^*$ has marginal dfs F^* and G^*, for every $\nu = 0,1,2,\ldots$ Furthermore it will be supposed that a random sample of size N_ν ($N_\nu \to \infty$ as $\nu \to \infty$) from the df $H_{(\nu)}^*$ is given. Let us assume that the locally most powerful rank test against the alternatives (4.3.1) has scores generating function J (see Section 1.2) and that approximate score functions will be used, so that the statistics may be written as

(4.3.2) $$T_\nu = \iint J([N_\nu/(N_\nu+1)]F_{N_\nu}^{(\nu)}, [N_\nu/(N_\nu+1)]G_{N_\nu}^{(\nu)})dH_{N_\nu}^{(\nu)}.$$

Among all possible tests there usually exists a locally most powerful test in the restricted sense of Schmetterer [35], page 237, 238. It is based on the statistic

(4.3.3) $$\tilde{T}_\nu = \iint J(F^*, G^*)dH_{N_\nu}^{(\nu)}.$$

It should be noted that in (4.3.2) and (4.3.3) the function J is the same. Without going into details we state that asymptotic normality of T_ν under fixed alternatives is easy to prove because we may rely directly on the central limit theorem. For the underlying df $H^*_{(\nu)}$ we introduce

(4.3.4)
$$\tilde{\mu}_\nu = \tilde{\mu}(H^*_{(\nu)}) = E(\tilde{T}_\nu) = \iint J(F^*, G^*)dH^*_{(\nu)},$$
$$\tilde{\sigma}^2_\nu = \tilde{\sigma}^2(H^*_{(\nu)}) = N\,\mathrm{Var}(\tilde{T}_\nu) = \iint [J(F^*, G^*) - \tilde{\mu}_\nu]^2 dH^*_{(\nu)},$$

and because the marginal dfs of each $H^*_{(\nu)}$ are F^* and G^* we have

(4.3.5) $$\tilde{\mu}_\nu = \mu_\nu,$$

for $\nu = 0,1,2,\ldots$ Here μ_ν (and σ^2_ν) are defined in the previous section.

Hence under the assumptions of Theorem 4.2.2 it follows from (4.3.5) that the asymptotic powers of the tests with critical regions $\{T_\nu \geq C_{\alpha,\nu}\}$ and $\{\tilde{T}_\nu \geq \tilde{C}_{\alpha,\nu}\}$ are equal to

(4.3.6) $$1 - \Phi(\Phi^{-1}(1-\alpha)-e) \text{ and } 1 - \Phi(\Phi^{-1}(1-\alpha)-e\sigma_0/\tilde{\sigma}_0)$$

respectively. Consequently equality of these asymptotic powers occurs if $\tilde{\sigma}^2_0 = \sigma^2_0$ and in view of formula (3.1.4) this will be the case if (cf. [3])

(4.3.7)
$$\int_0^1 J(u,t)dt = 0 \quad \text{for all } u \in (0,1),$$
$$\int_0^1 J(s,v)ds = 0 \quad \text{for all } v \in (0,1).$$

This condition is frequently satisfied. Because the tests $\{\tilde{T}_\nu \geq \tilde{C}_{\alpha,\nu}\}$ are easily seen to be asymptotically most powerful against $H^*_{(1)},H^*_{(2)},\ldots$, it follows that under (4.3.7) the rank tests $\{T_\nu \geq C_{\alpha,\nu}\}$ are also *asymptotically most powerful* against $H^*_{(1)},H^*_{(2)},\ldots$

Chapter 5

DISCUSSION OF THE ASSUMPTIONS, COMPARISON WITH EARLIER RESULTS AND SOME EXAMPLES

5.1. DISCUSSION OF THE ASSUMPTIONS

In Section 1.2 it has been pointed out that in many important situations the score functions J_N are in a natural way related to a limiting score function J which is of the product type $J = K \times L$, or which is a finite sum of functions of product type. For this reason we shall restrict ourselves in this chapter to limiting score functions J of product type, a restriction that simplifies the verification of the conditions on J.

Our first step is to formulate a new assumption which will take the place of Assumptions 2.1.1, 3.1.1 and which, after a simple extension, covers Assumptions 2.1.1, 3.1.1 and 2.1.2 (3.1.2). Let us recall the definition of the function R in (1.3.8) and remember that for arbitrary finite constants $D \geq 0$ and $\tau \geq 0$ the function DR^{τ} is an element of the class $\mathcal{R}$ (see Definition 1.3.2).

ASSUMPTION 5.1.1. For fixed $0 = s_0 < s_1 < \dots < s_k < s_{k+1} = 1$, $0 = t_0 < t_1 < \dots < t_l < t_{l+1} = 1$ and fixed finite constants $a_1,\dots,a_k,b_1,\dots,b_l$ the limiting score function $J = K \times L$ can be written as (see (1.1.1))

$$(5.1.1) \qquad J(s,t) = [K_c(s) + \textstyle\sum_{i=1}^{k} a_i c(s-s_i)][L_c(t) + \sum_{j=1}^{l} b_j c(t-t_j)],$$

$0 < s < 1$, $0 < t < 1$. Here K_c and L_c are continuous throughout (0,1) and have continuous first derivatives $K_c^{(1)}$ and $L_c^{(1)}$ on $\cup_{i=1}^{k+1}(s_{i-1},s_i)$ and $\cup_{j=1}^{l+1}(t_{j-1},t_j)$ respectively.

The above functions satisfy $|K| \leq r_1$, $|K'| \leq \tilde{r}_1$, $|L| \leq r_2$, $|L'| \leq \tilde{r}_2$, where they are defined on (0,1). Here

$$(5.1.2) \qquad r_1 = D_1R^{\alpha},\ \tilde{r}_1 = D_1R^{\alpha+1},\ r_2 = D_2R^{\beta},\ \tilde{r}_2 = D_2R^{\beta+1},$$

on (0,1), for fixed finite constants D_1, $D_2 > 0$ and $\alpha,\beta \in (0,1/2)$.

LEMMA 5.1.1. Suppose J satisfies Assumption 5.1.1. Then J satisfies Assumption 3.1.1, and in the special case where $a_1 = \dots = a_k = b_1 = \dots = b_l = 0$

Assumption 2.1.1 is also fulfilled.

PROOF. The proof is immediate. □

Let us next turn to the verification of Assumption 2.1.2 (3.1.2) in the case where the reproducing u-shaped functions are given by (5.1.2) and consider the two possibilities

$$\alpha + \beta < 1/2,$$

$$\alpha + \beta \geq 1/2,$$

for the exponents in (5.1.2).

If $\alpha + \beta < 1/2$, Assumption 2.1.2 may be verified for some $\delta > 0$ and $H' = H$. The proof relies on Hölder's inequality in the form

$$(5.1.3) \qquad \iint |\phi(F)\psi(G)|\,dH \leq [\textstyle\int_0^1 |\phi(s)|^{\xi} ds]^{1/\xi} [\int_0^1 |\psi(t)|^{\eta} dt]^{1/\eta},$$

where ϕ and ψ are functions on (0,1) such that the above integrals exist, and where $\xi > 1$, $\eta > 1$ satisfy $\xi^{-1} + \eta^{-1} = 1$. A counterexample shows that the condition $\alpha + \beta < 1/2$ is necessary for the assumption to hold with $\delta = 0$ and $H' = H$. For suppose that $F \times G \in H_0$ and consider the continuous bivariate df

$$(5.1.4) \qquad H(x,y) = \min\{F(x), G(y)\},$$

$-\infty < x < \infty$, $-\infty < y < \infty$. This df has marginal dfs F and G and concentrates mass 1 on the curve defined by $F(x) = G(y)$. The df H in (5.1.4) is called *Fréchet's maximal distribution* with given marginal dfs (see Feller [11], page 162, 163). It follows that for this df $\iint [R(F)]^{2\alpha}[R(G)]^{2\beta} dH = \int_0^1 [R(u)]^{2\alpha+2\beta} du < \infty$ if and only if $\alpha + \beta < 1/2$. A more natural example is provided by the bivariate Cauchy distribution, given by the density $(2\pi)^{-1}(1+x^2+y^2)^{-3/2}$ (see e.g. Feller [11], page 69, or Mardia [28]). Then $F(x)(1-F(x)) \sim \pi^{-1}|x|^{-1}$ as $|x| \to \infty$, and a similar formula holds true for $G(y)$. Using polar coordinates it follows that $\iint [R(F)]^{2\alpha}[R(G)]^{2\beta} dH < \infty$ if and only if $\iint |x|^{2\alpha}|y|^{2\beta}(1+x^2+y^2)^{-3/2} dx dy < \infty$ if and only if $\int_0^\infty r^{2\alpha+2\beta+1}(1+r^2)^{-3/2} dr < \infty$, which is the case if and only if $\alpha + \beta < 1/2$.

If $\alpha + \beta \geq 1/2$, the bound on the limiting score function is allowed to be of essentially larger order than in the previous case. The above counterexamples show that the assumption can no longer be satisfied for every df H in H, so that we shall have to restrict ourselves to an appropriate subclass of H. Several choices are possible, but we shall only consider the particular subclass

$$(5.1.5) \qquad H_{C,a,b} = \{H \in H : E([R(G(Y))]^{1-b} | F(X)) \le C[R(F(X))]^{a},$$
$$E([R(F(X))]^{1-a} | G(Y)) \le C[R(G(Y))]^{b}\}.$$

Here $1 \le C < \infty$, $0 < a < 1-2\alpha$, $0 < b < 1-2\beta$ are arbitrary but fixed constants. Evidently for the null hypothesis H_0 and Gumbel's class defined in (1.2.17) we have

$$H_0 \subset \{FG[1+\theta(1-F)(1-G)], F \times G \in H_0, -1 < \theta < 1\} \subset H_{C,a,b},$$

for any $a > 0$, $b > 0$ and some $C = C(a,b)$ satisfying $1 \le C < \infty$. Another important subclass of (5.1.5) is obtained if one considers the normal class defined in (1.2.15). Given fixed $0 < a < 1-2\alpha$, $0 < b < 1-2\beta$, the existence of a constant $1 \le C = C(a,b) < \infty$ may be shown such that

$$\{\bar{\Phi}_\theta(F,G), F \times G \in H_0,$$
$$\theta^2 < \max\{(2\alpha+2a-1)(2\beta+2b-1)[(1-a)(1-b)]^{-1}, 1\}\} \subset H_{C,a,b}.$$

Here we use the formula $\Phi(z)(1-\Phi(z)) \sim (2\pi)^{-1/2}|z|^{-1}\exp(-z^2/2)$ as $|z| \to \infty$ (see Feller [10], page 166). Note that $(2\alpha+2a-1)(2\beta+2b-1)[(1-a)(1-b)]^{-1} < 1$ for $\alpha + \beta \ge 1/2$. It follows that the set of admissible values of θ gradually decreases as $\alpha + \beta$ increases and, moreover, that the bound on θ^2 is close to 1 as long as $\alpha + \beta$ is close to 1/2. Choosing e.g. $\alpha = \beta = 1/4$ and $a = b = 1/2-\zeta$ for some small $\zeta > 0$, we find that θ^2 should be bounded by $[(1/2-2\zeta)/(1/2+\zeta)]^2$.

<u>LEMMA 5.1.2</u>. In Assumption 2.1.2 (3.1.2) choose $r_1, \tilde{r}_1, r_2, \tilde{r}_2$ as in (5.1.2), with $0 < \alpha < 1/2$, $0 < \beta < 1/2$.

(A) If $\alpha + \beta < 1/2$, then Assumption 2.1.2 (3.1.2) is satisfied for some $\delta > 0$ and $H' = H$.

(B) If $\alpha + \beta \ge 1/2$, then Assumption 2.1.2 (3.1.2) is satisfied for some $\delta > 0$ and $H' = H_{C,a,b}$ for arbitrary but fixed $1 \le C < \infty$, $0 < a < 1-2\alpha$,

$0 < b < 1-2\beta$.

PROOF. Let us for the moment think of δ as a small positive number, the actual value of which will be determined later on, and let us choose $q_1(u) = q_2(u) = [R(u)]^{-1/2+\delta}$, $0 < u < 1$, throughout this proof. These functions satisfy the last requirement of Assumption 2.1.2 (3.1.2), since

$$\int_0^1 [R(u)]^{(1/2-\delta)(2+\delta)} du < \infty,$$

for δ sufficiently small. Without loss of generality we may suppose that $D_1 = D_2 = 1$.

(A) If $\alpha + \beta < 1/2$ let us first apply (5.1.3) with $\xi = (\alpha+\beta)/\alpha$ and $\eta = (\alpha+\beta)/\beta$. Then we have independently of H in H, provided δ is taken sufficiently small,

$$\iint \{[R(F)]^{\alpha}[R(G)]^{\beta}\}^{2+\delta} dH$$

$$\leq \{\int_0^1 [R(u)]^{(\alpha+\beta)(2+\delta)} du\}^{\alpha/(\alpha+\beta)+\beta/(\alpha+\beta)} < \infty,$$

because $(\alpha+\beta)(2+\delta) < 1$ for δ sufficiently small.

The boundedness on H of the two remaining integrals in the assumption for some $\delta > 0$ is obtained by applying (5.1.3) with $\xi = (\alpha+1/2+2\delta)^{-1}$, $\eta = (1/2-\alpha-2\delta)^{-1}$ and $\xi = (1/2-\beta-2\delta)^{-1}$, $\eta = (\beta+1/2+2\delta)^{-1}$ respectively. By symmetry we only need consider the first of the two remaining integrals, for which we obtain independently of H in H, provided δ is sufficiently small,

$$\iint \{[R(F)]^{\alpha+1/2+\delta}[R(G)]^{\beta}\}^{1+\delta} dH$$

$$\leq \{\int_0^1 [R(s)]^{(\alpha+1/2+\delta)(1+\delta)(\alpha+1/2+2\delta)^{-1}} ds\}^{\alpha+1/2+2\delta} \times$$

$$\{\int_0^1 [R(t)]^{\beta(1+\delta)(1/2-\alpha-2\delta)^{-1}} dt\}^{1/2-\alpha-2\delta} < \infty,$$

because $(\alpha+1/2+\delta)(\alpha+1/2+2\delta)^{-1}(1+\delta) < 1$ and $\beta(1+\delta)(1/2-\alpha-2\delta)^{-1} = (\beta+\beta\delta)(\beta+(1/2-\alpha-\beta)-2\delta)^{-1} < 1$ for δ sufficiently small.

(B) Suppose $\alpha + \beta \geq 1/2$. For any H in $H_{C,a,b}$ we have

(5.1.6) $$E([R(G(Y))]^{(1-b)\zeta} | F(X)) \leq C[R(F(X))]^{a\zeta},$$

for arbitrary $0 < \zeta < 1$. This follows from Jensen's inequality for conditional expectations.

Applying (5.1.6) with $\zeta = \beta(2+\delta)(1-b)^{-1}$ (< 1 for δ sufficiently small) and using (1.3.4) we have independently of H in $H_{C,a,b}$, provided δ is chosen sufficiently small,

$$\int\int\{[R(F)]^{\alpha}[R(G)]^{\beta}\}^{2+\delta}dH$$

$$\leq C \int_0^1[R(s)]^{\alpha(2+\delta)+a\beta(2+\delta)/(1-b)}ds < \infty,$$

because $\alpha(2+\delta) + a\beta(2+\delta)(1-b)^{-1} < \alpha(2+\delta) + a < 1$ for δ sufficiently small.

By symmetry we need only consider the first of the two remaining integrals. Independently of H in $H_{C,a,b}$ we find by application of (5.1.6) with $\zeta = \beta(1+\delta)(1-b)^{-1}$ and using (1.3.4), provided δ is small enough,

$$\int\int\{[R(F)]^{\alpha+1/2+\delta}[R(G)]^{\beta}\}^{1+\delta}dH$$

$$\leq C \int_0^1[R(s)]^{(\alpha+1/2+\delta)(1+\delta)+a\beta(1+\delta)/(1-b)}ds < \infty,$$

because $(\alpha+1/2+\delta)(1+\delta) + a\beta(1+\delta)(1-b)^{-1} < (\alpha+1/2+\delta)(1+\delta) + a(1/2+\delta/2) = (2\alpha+1+a)/2 + O(\delta) < 1 + O(\delta)$. □

The present setup of the asymptotic theory is particularly well suited to the use of approximate score functions (see (1.2.3)).

<u>LEMMA 5.1.3</u>. When J_N (N=1,2,...) are the approximate score functions derived from J, we have $B^*_{0N} = 0$ for all $H \in H$ and any limiting score function J defined and finite on $(0,1) \times (0,1)$ (N=1,2,...). Hence in this case Assumption 2.1.3 (3.1.3) is trivially satisfied with $H' = H$.

<u>PROOF</u>. The proof follows immediately from formulas (1.2.3) and (2.1.4). □

The assumption may also be verified when J_N (N = 1,2,...) are the exact score functions derived from J (see (1.2.2)), provided J is of product type $J = K \times L$, and K and L are continuous on (0,1). Moreover the functions K and L have to possess continuous second derivatives on (0,1) except at an at most finite number of points. This result will be given in Theorem 5.2.4 of the next section.

Let us finally consider Assumption 3.1.4. The functions $\bar{h}$, $\bar{f}$, $\bar{g}$ and the sets O, O_1, O_2 will have the same meaning as in Assumption 3.1.4. The sets O_1 and O_2 can - and will - be chosen such that $O_1, O_2 \subset (c, 1-c)$ for some $0 < c < 1/2$. When they exist on O, we write $\bar{h}^{(1,0)}$ and $\bar{h}^{(0,1)}$ for the first partial derivatives of $\bar{h}$. Again a subclass of H will be introduced for which the assumption is easy to verify. Let us define

$$(5.1.7) \qquad \bar{H}_{C,\alpha+a,\beta+b} = \{H \in H : \bar{h}^{(i,j)}(s,t) \leq C[R(s)]^{\alpha+a+i} \times [R(t)]^{\beta+b+j}, (s,t) \in O \text{ and } i+j = 0,1\}.$$

As in (5.1.5) the numbers $1 \leq C < \infty$, $0 < a < 1-2\alpha$, $0 < b < 1-2\beta$ are arbitrary but fixed constants. The numbers C, a and b are supposed to be chosen equal to those in (5.1.5). It should be observed that, as far as the boundedness conditions in (5.1.7) are concerned, the restrictions imposed are relatively weak because nothing is said about the behavior of $\bar{H}$ in neighborhoods of the four vertices of the unit square (see Figure 5.1.1). It is not hard to see that for the

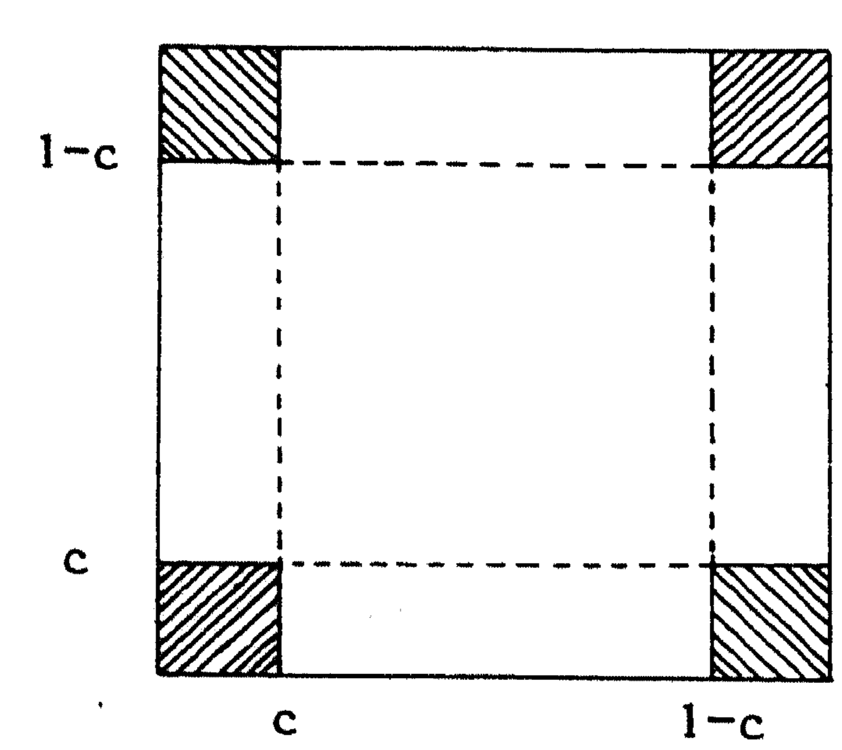

Figure 5.1.1

null hypothesis H_0 and Gumbel's class defined in (1.2.17) we have

$$H_0 \subset \{FG[1+\theta(1-F)(1-G)], F \times G \in H_0, -1 < \theta < 1\} \subset \bar{H}_{C,\alpha+a,\beta+b},$$

for C = 2 and arbitrary $\alpha, a, \beta, b > 0$. Moreover, for any $0 < \zeta < 1/2$, the normal class defined in (1.2.15) satisfies

$$\{\bar{\Phi}_\theta(F,G), F \times G \in H_0, -1+\zeta \leq \theta \leq 1-\zeta\} \subset \bar{H}_{C,\alpha+a,\beta+b}.$$

Here the constants $\alpha, a, \beta, b > 0$ may be chosen arbitrarily. The constant C not only depends on ζ and c, but also on the particular choice of the numbers α, a, β, b.

LEMMA 5.1.4. In Assumption 3.1.4 choose r_1 and r_2 as in (5.1.2), with $0 < \alpha < 1/2$, $0 < \beta < 1/2$. Then this assumption is satisfied with $H' = \bar{H}_{C,\alpha+a,\beta+b}$, as defined in (5.1.7). Here $1 \le C < \infty$, $0 < a < 1-2\alpha$, $0 < b < 1-2\beta$ are arbitrary but fixed constants.

PROOF. Without loss of generality we may take $D_1 = D_2 = 1$. Let us start with the equicontinuity condition. Provided s is sufficiently close to s_i, the mean value theorem applies and ensures the existence of a number $\tilde{s}_i$ between s and s_i such that

$$\sup_{H \in \bar{H}_{C,\alpha+a,\beta+b}} |\bar{h}(s,t)-\bar{h}(s_i,t)| \le$$

$$\le C|s-s_i|[R(\tilde{s}_i)]^{\alpha+a+1}[R(t)]^{\alpha+b} \to 0,$$

for all $0 < t < 1$, as $s \to s_i$. The second equicontinuity condition may be verified in the same way.

As to the boundedness conditions, note that

$$\sup_{H \in \bar{H}_{C,\alpha+a,\beta+b}} \bar{h}(s,t) \le C[R(s)]^{\alpha+a}[R(c)]^{\beta+b},$$

for all (s,t) in $(0,1) \times O_2$. Hence for $\bar{f}$ we may choose $C_1[R(s)]^{\alpha+a}$, where $C_1 = C[R(c)]^{\beta+b}$, and we have

$$\int_0^1 r_1(s)\bar{f}(s)ds = C_1\int_0^1[R(s)]^{2\alpha+a}ds < \infty,$$

because $2\alpha + a < 1$. Similarly a function $\bar{g}$ may be constructed for which the finiteness of $\int_0^1 r_2(t)\bar{g}(t)dt$ follows in the same way. □

The results of this section will be applied to formulate some special cases of Theorems 2.1.1 and 3.1.1 under relatively simple conditions. Before listing these special cases it is worth noticing that in particular the asymptotic variance can be given a nicer expression when $J = K \times L$ is of product type. If $J = K \times L$ satisfies Assumption 5.1.1 for arbitrary $a_1,\ldots,a_k$, $b_1,\ldots,b_l$, it follows from formulas (2.1.1) and (3.1.2) that

$$(5.1.8) \qquad \mu = \mu(H) = E[K(F(X))L(G(Y))],$$

$$(5.1.9) \qquad \sigma^2 = \sigma^2(H) = \mathrm{Var}\{K(F(X))L(G(Y)) + \int_0^1[c(s-F(X))-s]E(L(G(Y))|F(X) = s)dK(s) + \int_0^1[c(t-G(Y))-t]E(K(F(X))|G(Y) = t)dL(t)\},$$

provided these quantities exist. This simplification is obtained by using formula (1.3.4) and the usual notation for Lebesgue-Stieltjes integrals of functions of one real variable.

To begin with, some theorems for approximate score functions will be given.

THEOREM 5.1.1. Suppose that the limiting score function $J = K \times L$ satisfies Assumption 5.1.1 with $a_1 = \dots = a_k = b_1 = \dots = b_l = 0$, and let J_N $(N = 1,2,\dots)$ be the approximate score functions derived from J. Fix the numbers $0 < \alpha < 1/2$, $0 < \beta < 1/2$ and $1 \leq C < \infty$, $0 < a < 1-2\alpha$, $0 < b < 1-2\beta$.
(A) If $\alpha + \beta < 1/2$, then for any fixed underlying df $H \in \mathcal{H}$ we have $N^{1/2}(T_N-\mu) \to_d N(0,\sigma^2)$ as $N \to \infty$. Here $\mu = \mu(H)$ and $\sigma^2 = \sigma^2(H)$ are finite and given by (5.1.8) and (5.1.9) respectively.

This convergence in distribution is uniform on each subclass $\mathcal{H}' \subset \mathcal{H}$ on which $\sigma^2 = \sigma^2(H)$ is bounded away from zero.
(B) If $\alpha + \beta \geq 1/2$, then for any fixed underlying df $H \in \mathcal{H}_{C,a,b}$ we have $N^{1/2}(T_N-\mu) \to_d N(0,\sigma^2)$ as $N \to \infty$. Here $\mu = \mu(H)$ and $\sigma^2 = \sigma^2(H)$ are finite and given by (5.1.8) and (5.1.9) respectively.

This convergence in distribution is uniform on each subclass $\mathcal{H}' \subset \mathcal{H}_{C,a,b}$ on which $\sigma^2 = \sigma^2(H)$ is bounded away from zero.

PROOF. The proof follows immediately from Theorem 2.1.1 and Lemmas 5.1.1-5.1.3. □

THEOREM 5.1.2. Suppose that the limiting score function $J = K \times L$ satisfies Assumption 5.1.1 with $a_1,\dots,a_k$, $b_1,\dots,b_l$ arbitrary, and let J_N $(N = 1,2,\dots)$ be the approximate score functions derived from J. Fix the numbers $0 < \alpha < 1/2$, $0 < \beta < 1/2$ and $1 \leq C < \infty$, $0 < a < 1-2\alpha$, $0 < b < 1-2\beta$.
(A) If $\alpha + \beta < 1/2$, then for any fixed underlying df $H \in \overline{\mathcal{H}}_{C,\alpha+a,\beta+b}$ we have $N^{1/2}(T_N-\mu) \to_d N(0,\sigma^2)$ as $N \to \infty$. Here $\mu = \mu(H)$ and $\sigma^2 = \sigma^2(H)$ are finite

and given by (5.1.8) and (5.1.9) respectively.

This convergence in distribution is uniform on each subclass $H' \subset \bar{H}_{C,\alpha+a,\beta+b}$ on which $\sigma^2 = \sigma^2(H)$ is bounded away from zero.

(B) If $\alpha + \beta \geq 1/2$, then for any fixed underlying df $H \in H_{C,a,b} \cap \bar{H}_{C,\alpha+a,\beta+b}$ we have $N^{1/2}(T_N-\mu) \to_d N(0,\sigma^2)$ as $N \to \infty$. Here $\mu = \mu(H)$ and $\sigma^2 = \sigma^2(H)$ are finite and given by (5.1.8) and (5.1.9) respectively.

This convergence in distribution is uniform on each subclass $H' \subset H_{C,a,b} \cap \bar{H}_{C,\alpha+a,\beta+b}$ on which $\sigma^2 = \sigma^2(H)$ is bounded away from zero.

(C) If $0 < \alpha < 1/2$, $0 < \beta < 1/2$ without any further restriction, then $N^{1/2}(T_N-\mu_0) \to_d N(0,\sigma_0^2)$ as $N \to \infty$, uniformly on the class H_0 of all null hypothesis dfs. Here μ_0 and σ_0^2 are finite and given by $\mu_0 = [\int K][\int L]$ and $\sigma_0^2 = [\int K^2-(\int K)^2][\int L^2-(\int L)^2]$.

<u>PROOF</u>. Part (A) and (B) are an immediate consequence of Theorem 3.1.1 and Lemmas 5.1.1-5.1.4.

Part (C) follows from the fact that $H_0 \subset H_{C,a,b} \cap \bar{H}_{C,\alpha+a,\beta+b}$, provided the constant C is chosen sufficiently large, as has been observed in the examples accompanying the definitions of the above classes of dfs in (5.1.5) and (5.1.7). For μ_0 and σ_0^2 see also (3.1.3) and (3.1.4). □

Next let us give a theorem for arbitrary score functions, in the case where the limiting score function is continuous.

<u>THEOREM 5.1.3</u>. Suppose that the limiting score function $J = K \times L$ satisfies Assumption 5.1.1 with $a_1 = \ldots = a_k = b_1 = \ldots = b_l = 0$ and let J_N $(N = 1,2,\ldots)$ be arbitrary score functions. Fix the numbers $0 < \alpha < 1/2$, $0 < \beta < 1/2$ and $1 \leq C < \infty$, $0 < a < 1-2\alpha$, $0 < b < 1-2\beta$.

(A) If $\alpha + \beta < 1/2$, then $N^{1/2}(T_N-\mu) \to_d N(0,\sigma^2)$ as $N \to \infty$, for any fixed underlying df $H \in H$ for which $B^*_{0N} \to_P 0$ as $N \to \infty$. Here $\mu = \mu(H)$ and $\sigma^2 = \sigma^2(H)$ are finite and given by (5.1.8) and (5.1.9) respectively.

This convergence in distribution is uniform on each subclass $H' \subset H$ on which $B^*_{0N} \to_P 0$ uniformly, as $N \to \infty$ and on which $\sigma^2 = \sigma^2(H)$ is bounded away from zero.

(B) If $\alpha + \beta \geq 1/2$, then $N^{1/2}(T_N-\mu) \to_d N(0,\sigma^2)$ as $N \to \infty$, for any fixed underlying df $H \in H_{C,a,b}$ for which $B^*_{0N} \to_P 0$ as $N \to \infty$. Here $\mu = \mu(H)$ and $\sigma^2 = \sigma^2(H)$ are finite and given by (5.1.8) and (5.1.9) respectively.

This convergence in distribution is uniform on each subclass

$H' \subset H_{C,a,b}$ on which $B^*_{0N} \to_P 0$ uniformly, as $N \to \infty$ and on which $\sigma^2 = \sigma^2(H)$ is bounded away from zero.

PROOF. This is an immediate consequence of Theorem 2.1.1 and Lemmas 5.1.1-5.1.3. □

In the conclusion of the theorems the limiting variance is allowed to be zero, i.e. the limiting distribution may occasionally be degenerate. It is illustrative to consider more closely a specific situation where this occurs. It is intuitively clear that this will be the case when the underlying df is Fréchet's maximal distribution, defined in (5.1.4).

Let us suppose that $J = K \times L$ satisfies Assumption 5.1.1 with $a_1 = \ldots = a_k = b_1 = \ldots = b_l = 0$ and with $\alpha + \beta < 1/2$. Furthermore, let J_N $(N = 1,2,\ldots)$ be the approximate score functions derived from J, so that Theorem 5.1.1 (A) applies and yields that

$$N^{1/2}(T_N-\mu) \to_d N(0,\sigma^2),$$

as $N \to \infty$, when the underlying df is the one in (5.1.4).

On the other hand, for any sample from the latter df we have $P(\{R_{nN}=Q_{nN}\}) = 1$ for $n = 1,\ldots,N$, so that $T_N = [\sum_{n=1}^N K(n/(N+1))L(n/(N+1))]/N$. Since in this case the limiting mean equals $\mu = \int_0^1 K(u)L(u)du$, T_N may be considered as a Riemann-sum approximation of the integral representing μ. Because under the present conditions the error of this approximation is of order $o(N^{-1/2})$ (see Lemma (6.3.2)) it follows that

$$N^{1/2}(T_N-\mu) \to_{a.s.} 0.$$

rom this it might be concluded that σ^2 must be equal to zero, but we shall ive a direct verification.

Indeed, in this case the limiting variance equals (see (5.1.9))

$$\int_0^1\{K(u)L(u)-\mu + \int_0^1[c(s-u)-s]L(s)dK(s) + \int_0^1[c(s-u)-s]K(s)dL(s)\}^2du$$

$$= \int_0^1\{K(u)L(u)-\mu + \int_0^1[c(s-u)-s]d[K(s) \times L(s)]\}^2du$$

$$= \int_0^1 \{K(u)L(u)-\mu - uK(u)L(u) + \int_0^u K(s)L(s)ds -$$

$$(1-u)K(u)L(u) + \int_u^1 K(s)L(s)ds\}^2 du = 0.$$

In a similar fashion relatively simple special cases of Theorem 4.2.1 can be formulated. By way of an example let us give a theorem for continuous limiting score functions, when approximate score functions are used.

THEOREM 5.1.4. Suppose that the limiting score function $J = K \times L$ satisfies Assumption 5.1.1 with $a_1 = \ldots = a_k = b_1 = \ldots = b_l = 0$ and let J_ν ($\nu = 1,2,\ldots$) be the approximate score functions derived from J for the sample sizes N_ν ($N_\nu \to \infty$ as $\nu \to \infty$). Fix the numbers $0 < \alpha < 1/2$, $0 < \beta < 1/2$ and $1 \leq C < \infty$, $0 < a < 1-2\alpha$, $0 < b < 1-2\beta$.

(A) If $\alpha + \beta < 1/2$, then for any sequence $H, H_{(1)}, H_{(2)}, \ldots \in \mathcal{H}$ of underlying dfs such that $H_{(\nu)} \to H$ on $(-\infty,\infty) \times (-\infty,\infty)$ as $\nu \to \infty$ we have $N_\nu^{1/2}(T_\nu - \mu(H_{(\nu)})) \to_d \mathcal{N}(0,\sigma^2(H))$ as $\nu \to \infty$. Here $\mu(H_{(\nu)})$ and $\sigma^2(H)$, given by (5.1.8) and (5.1.9) respectively, are finite.

(B) If $\alpha + \beta \geq 1/2$, then for any sequence $H, H_{(1)}, H_{(2)}, \ldots \in \mathcal{H}_{C,a,b}$ of underlying dfs such that $H_{(\nu)} \to H$ on $(-\infty,\infty) \times (-\infty,\infty)$ as $\nu \to \infty$ we have $N_\nu^{1/2}(T_\nu - \mu(H_{(\nu)})) \to_d \mathcal{N}(0,\sigma^2(H))$ as $\nu \to \infty$. Here $\mu(H_{(\nu)})$ and $\sigma^2(H)$, given by (5.1.8) and (5.1.9) respectively, are finite.

PROOF. See Theorem 4.2.1 and Lemmas 5.1.1-5.1.3. □

5.2. COMPARISON WITH EARLIER RESULTS

In order to facilitate a comparison of the present results with earlier work on this subject, we shall first state the equivalence of Assumption 2.1.3 and a certain modification of this assumption. For the score functions J_N and the limiting score function J consider B_{0N}^* (see (2.1.4)) and define (for Δ_N see (2.2.1))

(5.2.1) $$B_{01N} = N^{1/2} \iint_{\Delta_N} [J_N(F_N,G_N) - J(F_N,G_N)]dH_N,$$

(5.2.2) $$B_{02N} = N^{1/2} \iint_{\Delta_N^c} [J_N(F_N,G_N) - J(F,G)]dH_N,$$

(5.2.3) $$B_{0N} = B_{01N} + B_{02N}.$$

REMARK. In any of the assumptions mentioned in Theorem 5.2.1, the points $0 < s_1 < \ldots < s_k < 1$, $0 < t_1 < \ldots < t_l < 1$ are the same fixed elements of $(0,1)$, the functions $r_1, \tilde{r}_1, r_2, \tilde{r}_2$ are the same fixed elements of $\mathcal{R}$ and $\mathcal{H}'$ is the same subclass of H.

THEOREM 5.2.1. Let Assumption 2.1.2 be satisfied for $H' = \{H\}$, where the df $H \in H$ is fixed, and $\delta = 0$ (for some fixed subclass $H' \subset H$ and $\delta > 0$).
(i) Let the limiting score function J satisfy Assumption 2.1.1. Then, as $N \to \infty$, $B^*_{ON} \to_P 0$ uniformly on H' if and only if $B_{ON} \to_P 0$ uniformly on H'.
(ii) Let the limiting score function J satisfy Assumption 3.1.1 and suppose that Assumption 3.1.4 is satisfied for H'. Then, as $N \to \infty$, $B^*_{ON} \to_P 0$ uniformly on H' if and only if $B_{ON} \to_P 0$ uniformly on H'.

PROOF. In both cases it suffices to prove that $B_{ON} - B^*_{ON} \to_P 0$ uniformly on H'. Without loss of generality take $k = 1 = 1$.
(i) The following decomposition of $B_{ON} - B^*_{ON}$ is based on a twofold application of the mean value theorem. First, for (x,y) restricted to $\Delta_N \cap S_\gamma$ and for each ω in $\Omega_{\gamma N}$ (see (2.2.2) and (2.2.3)), defined by

$$\Omega_{\gamma N} = \{\omega : \sup|F_N - F| < \gamma/2,\ \sup|G_N - G| < \gamma/2\} \cap \Omega^*_{\gamma N},$$

we have

$$(5.2.4) \qquad N^{1/2} J(F^*_N, G^*_N) = N^{1/2} J(F_N, G_N) + N^{1/2}(F^*_N - F_N) J^{(1,0)}(\Phi^*_N, \Psi^*_N) + N^{1/2}(G^*_N - G_N) J^{(0,1)}(\Phi^*_N, \Psi^*_N).$$

In (5.2.4) the random point (Φ^*_N, Ψ^*_N) lies in the open line segment joining the random points (F_N, G_N) and (F^*_N, G^*_N). Secondly, for (x,y) restricted to $\Delta^c_N \cap S_\gamma$ we use the expression (2.2.4), valid for all ω in $\Omega_{\gamma N}$. Thus we arrive at the decomposition $B_{ON} - B^*_{ON} = \sum_{i=1}^{8} C_{iN}$, where

$$C_{1N} = \chi(\Omega_{\gamma N}) N^{1/2} \iint_{\Delta_N \cap S_\gamma} (F^*_N - F_N) J^{(1,0)}(\Phi^*_N, \Psi^*_N)\, dH_N,$$

$$C_{2N} = \chi(\Omega_{\gamma N}) N^{1/2} \iint_{\Delta_N \cap S_\gamma} (G^*_N - G_N) J^{(0,1)}(\Phi^*_N, \Psi^*_N)\, dH_N,$$

$$C_{3N} = \chi(\Omega_{\gamma N}) N^{1/2} \iint_{S^c_\gamma} [J(F^*_N, G^*_N) - J(F,G)]\, dH_N,$$

$$C_{4N} = \chi(\Omega_{\gamma N}) N^{1/2} \iint_{\Delta_N \cap S^c_\gamma} [J(F,G) - J(F_N, G_N)]\, dH_N,$$

$$C_{5N} = \chi(\Omega^c_{\gamma N}) N^{1/2} \iint_{\Delta_N} [J(F^*_N, G^*_N) - J(F_N, G_N)]\, dH_N,$$

$$C_{6N} = \chi(\Omega_{\gamma N})\iint_{\Delta_N^c \cap S_\gamma} U_N^*(F) J^{(1,0)}(\Phi_N, \Psi_N) dH_N,$$

$$C_{7N} = \chi(\Omega_{\gamma N})\iint_{\Delta_N^c \cap S_\gamma} V_N^*(G) J^{(0,1)}(\Phi_N, \Psi_N) dH_N,$$

$$C_{8N} = \chi(\Omega_{\gamma N}^c) N^{1/2} \iint_{\Delta_N^c} [J(F_N^*, G_N^*) - J(F,G)] dH_N.$$

The proof of the asymptotic negligibility of these C-terms relies on the same methods that were used in Section 2.4. The proof will be divided into four parts, to be combined in the usual way. We shall only give an outline.

(1) As $\gamma \downarrow 0$ and $N \to \infty$, $C_{3N} \to_P 0$ and $C_{4N} \to_P 0$ uniformly on H'. For the proof we may refer to Corollary 2.4.1. When dealing with C_{4N} we only have to use Lemma 2.3.2(i) instead of Lemma 2.3.2(ii).

(2) As $N \to \infty$, $C_{6N} \to_P 0$ and $C_{7N} \to_P 0$ uniformly on H'. By symmetry we only have to consider C_{7N}. Let an arbitrary $\varepsilon > 0$ be given and put $\gamma_N = (\log N)^{-1}$. Denote $D_N = \{x : F(x) \in [\gamma_N, 1-\gamma_N]\} \times \{y : G(y) \in [\gamma_N, 1-\gamma_N]\}$ and let $\tilde{\Omega}_{1N} = \{\Delta_N^c \subset D_N^c\} = \{\Delta_N \supset D_N\}$. Then $P(\tilde{\Omega}_{1N}) \to 1$ as $N \to \infty$, uniformly on H, because

$$P(\tilde{\Omega}_{1N}^c) = P(\{(X_n, Y_n) \in D_N, n = 1,\ldots,N\}) \leq [1-2(\log N)^{-1}]^N \to 0,$$

as $N \to \infty$, uniformly on H. With Ω_N as in (2.4.1) for some index N_1 we have $P(\Omega_N \cap \tilde{\Omega}_{1N}) \geq 1-2\varepsilon$ for all $N \geq N_1$, uniformly on H, and

$$E(\chi(\Omega_N \cap \tilde{\Omega}_{1N}) |C_{7N}|) \leq M^3 \iint_{D_N^c} q_1(F) \tilde{r}_1(F) r_2(G) dH.$$

The proof that the latter integral converges to zero as $N \to \infty$, uniformly on H', can be given by an argument similar to that given at the end of the proof of Corollary 2.4.1 (see (2.4.3)).

(3) For fixed γ, $C_{5N} \to_P 0$ and $C_{8N} \to_P 0$ as $N \to \infty$, uniformly on H. See the proof of Corollary 2.4.2.

(4) For fixed γ, $C_{1N} \to_P 0$ and $C_{2N} \to_P 0$ as $N \to \infty$, uniformly on H. By symmetry we only need consider C_{1N}. Note that the function $J^{(1,0)}$ is continuous on the closed set $\{[\gamma/2, s_1-\gamma/2] \cup [s_1+\gamma/2, 1-\gamma/2]\} \times \{[\gamma/2, t_1-\gamma/2] \cup [t_1+\gamma/2, 1-\gamma/2]\}$ and hence assumes a maximum, say $\tilde{M}_\gamma$, on this set. The properties of the set $\Omega_{\gamma N}$ and the random point (Φ_N^*, Ψ_N^*) combined with the fact that $\sup|F_N^* - F_N| \leq (N+1)^{-1}$, ensure that $|C_{1N}| \leq N^{1/2}(N+1)^{-1}\tilde{M}_\gamma \to 0$ as $N \to \infty$, independently of H in H.

(ii) Using the representation of Lemma 3.1.1 we obtain the decomposition $B_{ON} - B^*_{ON} = \sum_{i=1}^{2} C_{ciN} + \sum_{i=1}^{4} C'_{iN} + \sum_{i=1}^{4} D_{iN}$, where

$$C_{c1N} = N^{1/2}\iint_{\Delta_N}[J_c(F^*_N,G^*_N)-J_c(F_N,G_N)]dH_N,$$

$$C_{c2N} = N^{1/2}\iint_{\Delta^c_N}[J_c(F^*_N,G^*_N)-J_c(F,G)]dH_N,$$

$$C'_{1N} = \sum_{m=1}^{p} N^{1/2}\iint_{\Delta_N}[K_{mc}(F^*_N)-K_{mc}(F_N)]L_m(G^*_N)dH_N,$$

$$C'_{2N} = \sum_{m=1}^{p} N^{1/2}\iint_{\Delta_N}K_m(F_N)[L_{mc}(G^*_N)-L_{mc}(G_N)]dH_N,$$

$$C'_{3N} = \sum_{m=1}^{p} N^{1/2}\iint_{\Delta^c_N}[K_{mc}(F^*_N)-K_{mc}(F)]L_m(G^*_N)dH_N,$$

$$C'_{4N} = \sum_{m=1}^{p} N^{1/2}\iint_{\Delta^c_N} K_m(F)[L_{mc}(G^*_N)-L_{mc}(G)]dH_N,$$

$$D_{1N} = \sum_{m=1}^{p} N^{1/2}\iint_{\Delta_N}[K_{md}(F^*_N)-K_{md}(F_N)]L_m(G^*_N)dH_N,$$

$$D_{2N} = \sum_{m=1}^{p} N^{1/2}\iint_{\Delta_N}K_m(F_N)[L_{md}(G^*_N)-L_{md}(G_N)]dH_N,$$

$$D_{3N} = \sum_{m=1}^{p} N^{1/2}\iint_{\Delta^c_N}[K_{md}(F^*_N)-K_{md}(F)]L_m(G^*_N)dH_N,$$

$$D_{4N} = \sum_{m=1}^{p} N^{1/2}\iint_{\Delta^c_N}K_m(F)[L_{md}(G^*_N)-L_{md}(G)]dH_N.$$

The proof of the asymptotic negligibility of these terms relies essentially on the same methods that were used in Section 3.4. Again the proof is divided into four parts. Only a sketch will be given. Without loss of generality let us moreover take $p = 1$, $a_{11} = b_{11} = 1$ and $M' = 1$ (see Lemma 3.1.1). We shall write K,K_c,K_d,L,L_c,L_d instead of $K_1,K_{1c},K_{1d},L_1,L_{1c},L_{1d}$ respectively. Hence $K_d(u) = c(u-s_1)$, $L_d(u) = c(u-t_1)$ for $0 < u < 1$.

(1) As $N \to \infty$, $C_{c1N} \to_p 0$ and $C_{c2N} \to_p 0$, uniformly on H'. These terms involve only the continuous part J_c of the score function, satisfying Assumption 2.1.1, and hence the proof follows from part (i) of the present theorem.

(2) As $N \to \infty$, $C'_{iN} \to_p 0$, uniformly on H', $i = 1,2,3,4$. For the proof the univariate mean value theorem should be used instead of the bivariate one. Then the results follow from obvious modifications of the proofs of the first five corollaries in Section 3.4.

(3) As $N \to \infty$, $D_{1N} \to_p 0$ and $D_{2N} \to_p 0$, uniformly on H. By symmetry we only have to consider D_{1N}. Note that $\sum_{n=1}^{N}|c(F_N(X_n)-s_1)-c(F^*_N(X_n)-s_1)| \leq 1$, because

at most one term in the summation is unequal to zero. Consequently $|D_{1N}| \leq N^{-1/2} r_2(1/(N+1)) \to 0$ as $N \to \infty$, independently of H in $\mathcal{H}$, since Assumption 2.1.2 implies square integrability of the function r_2.

(4) As $N \to \infty$, $D_{3N} \to_P 0$ and $D_{4N} \to_P 0$, uniformly on $\mathcal{H}$. For reasons of symmetry we only need consider D_{3N}. With Ω_N as in (2.4.1) and γ_N, $\tilde{\Omega}_{1N}$ as in part (i,2) of the proof of the present theorem, by Lemma 3.3.2(ii) we have

$$E(\chi(\Omega_N \cap \tilde{\Omega}_{1N})|D_{3N}|)$$

$$\leq N^{1/2} M[\int_0^{\gamma_N} c_{N-1}(1,2;s,s_1) E(r_2(G(Y))|F(X) = s)ds +$$

$$\int_{1-\gamma_N}^1 c_{N-1}(1,2;s,s_1) E(r_2(G(Y))|F(X) = s)ds]$$

$$= O(N^{1/2}[\exp(-2N(\gamma_N - s_1)^2) + \exp(-2N(1-\gamma_N-s_1)^2)]) \to 0$$

as $N \to \infty$, independently of H in $\mathcal{H}$. Here (X,Y) has df H and Lemma 3.3.1(i) is used. Assumption 2.1.2 implies square integrability, and hence integrability of the function r_2. In turn this implies integrability of the conditional expectation, which is used here. □

The assumption of the asymptotic negligibility of B^*_{0N} guarantees the simplest approach to asymptotic theory. Moreover, as has been pointed out in the previous section, this assumption has the advantage of being trivially satisfied in the practically important case of approximate scores. However, the earliest theorem on asymptotic normality of linear rank tests for independence under fixed alternatives, due to Bhuchongkul [4], is based on the assumption that B_{01N} is asymptotically negligible. Our next step is to show that the conditions of the theorem generally are such that B_{02N} is automatically negligible. To show this we need some lemmas. Let λ and ν be the random indices $1 \leq \lambda(\omega) \leq N$, $1 \leq \nu(\omega) \leq N$ such that

(5.2.5) $\qquad X_\lambda = X_{N:N},\ Y_\nu = Y_{N:N}.$

<u>LEMMA 5.2.1</u>. As $N \to \infty$, $P(\{\gamma_N \leq F(X_\lambda) \leq 1-\gamma_N\} \cap \{\gamma_N \leq G(Y_\lambda) \leq 1-\gamma_N\}) \to 1$, uniformly on $\mathcal{H}$, provided $\gamma_N = o(N^{-1})$.

<u>PROOF</u>. The probability of the complementary event is bounded above by $\gamma_N^N + 3[1-(1-\gamma_N)^N] \to 0$ as $N \to \infty$, independently of H in $\mathcal{H}$. □

LEMMA 5.2.2. Let $C \geq 1$, $0 < a < 1$, $0 < b < 1$ be fixed given numbers, satisfying $a + b < 1$. For the definition of the subclass $\mathcal{H}_{C,a,b}$ see (5.1.5).
(i) As $N \to \infty$, $P(\{Y_\lambda = Y_{N:N}\}) \to 0$, uniformly on $\mathcal{H}_{C,a,b}$.
(ii) As $N \to \infty$, $P(\{\delta_N \leq G(Y_\lambda) \leq 1-\delta_N\}) \to 1$, uniformly on $\mathcal{H}_{C,a,b}$, provided $\delta_N = o(N^{-a/(1-b)})$.

PROOF. Throughout this proof let Γ denote the gamma-function.
(i) $P(\{Y_\lambda = Y_{N:N}\}) = P(\cup_{n=1}^N \{(X_n,Y_n) = (X_{N:N},Y_{N:N})\}) = N\int\int H^{N-1}dH$. Note that for all x,y we have $H(x,y) \leq F(x)$ and $H(x,y) \leq G(y)$, so that $H^{N-1} \leq F^{(N-1)/2} \times G^{(N-1)/2}$. Furthermore the function $t^{(N-1)/2}[t(1-t)]^{1-b}$ attains for $0 \leq t \leq 1$ a maximum of order $O(N^{-(1-b)})$. By the properties of the class $\mathcal{H}_{C,a,b}$ this implies that $E([G(Y)]^{(N-1)/2}|F(X) = s) \leq c_N C R^a(s)$, where $c_N = O(1/N^{1-b})$ as $N \to \infty$. Thus we obtain

$$N\int\int[H(x,y)]^{N-1}dH(x,y) \leq c_N CN \int_0^1 s^{(N-1)/2}[R(s)]^a ds$$

$$= c_N CN\Gamma(N/2 + 1/2-a)\Gamma(1-a)/\Gamma(N/2 + 3/2-2a)$$

$$= O(N^{a+b-1}) \to 0$$

as $N \to \infty$, because by assumption $a + b < 1$. The convergence is obviously uniform on $\mathcal{H}_{C,a,b}$.
(ii) This probability equals $1-P(\{G(Y_\lambda) < \delta_N\})-P(\{G(Y_\lambda) > 1-\delta_N\})$, provided N is so large that $\delta_N < 1/2$. For such N we have moreover $\chi((0,\delta_N);t) \leq \delta_N^{1-b}[R(t)]^{1-b}$, for $0 < t < 1$. By the properties of the class $\mathcal{H}_{C,a,b}$ this implies that $E(\chi((0,\delta_N);G(Y))|F(X) = s) \leq \delta_N^{1-b}C[R(s)]^a$. Because of the independence of the sample elements we have

$$P(\{G(Y_\lambda) < \delta_N\}) = NP((\cap_{n=1}^{N-1}\{F(X_n) < F(X_N)\}) \cap \{G(Y_N) < \delta_N\})$$

$$= N\int\int[F(x_N)]^{N-1}\chi((0,\delta_N);G(y_N))dH(x_N,y_N)$$

$$\leq \delta_N^{1-b} CN \int_0^1 s^{N-1}[R(s)]^a ds$$

$$= \delta_N^{1-b} O(N\Gamma(N-a)\Gamma(1-a)/\Gamma(N+1-2a)) = \delta_N^{1-b} O(N^a) \to 0$$

as $N \to \infty$. Obviously the convergence is uniform on $\mathcal{H}_{C,a,b}$. One can deal with $P(\{G(Y_\lambda) > 1-\delta_N\})$ in the same way. □

THEOREM 5.2.2. For each $N = 1,2,\ldots$ let J_N be a score function defined on $(0,1] \times (0,1]$ as described in (1.1.10), and let J be a function defined on $(0,1) \times (0,1)$. Suppose that for some fixed given numbers $0 < D < \infty$, $0 < \alpha < 1/2$, $0 < \beta < 1/2$ each of these functions is in absolute value bounded by $D[R(s)]^{\alpha}[R(t)]^{\beta}$, for $(s,t) \in (0,1) \times (0,1)$. Let B_{02N} be based on the above J_N and J.

(A) If $\alpha + \beta < 1/2$, then $B_{02N} \to_P 0$ as $N \to \infty$, uniformly on $\mathcal{H}$.

(B) If $\alpha + \beta \geq 1/2$, let $C \geq 1$, $0 < a < 1-2\alpha$, $0 < b < 1-2\beta$ be fixed given numbers. Then $B_{02N} \to_P 0$ as $N \to \infty$, uniformly on $\mathcal{H}_{C,a,b}$.

PROOF. Without loss of generality we may take $D = 1$. We have $|B_{02N}| \leq N^{1/2}\iint_{\Delta_N^c}[|J_N(F_N,G_N)| + |J(F,G)|]dH_N \leq \sum_{i=1}^{5} C_{iN}$, where

$$C_{1N} = N^{1/2}\iint_{\{X_{N:N}\}\times\Delta_{N2}} |J_N(1,G_N)|dH_N,$$

$$C_{2N} = N^{1/2}\iint_{\{(X_{N:N},Y_{N:N})\}} |J_N(1,1)|dH_N,$$

$$C_{3N} = N^{1/2}\iint_{\Delta_{N1}\times\{Y_{N:N}\}} |J_N(F_N,1)|dH_N,$$

$$C_{4N} = N^{-1/2}[R(F(X_\lambda))]^{\alpha}[R(G(Y_\lambda))]^{\beta},$$

$$C_{5N} = N^{-1/2}[R(F(X_\nu))]^{\alpha}[R(G(Y_\nu))]^{\beta},$$

with λ and ν defined in (5.2.5). By the definition of J_N, in particular because J_N is a simple step function, it follows that $|J_N(1,t)| \leq [R(N^{-1})]^{\alpha}[R(t)]^{\beta}$ for $0 < t < 1$, $|J_N(s,1)| \leq [R(s)]^{\alpha}[R(N^{-1})]^{\beta}$ for $0 < s < 1$, and $\max_{(0,1)\times(0,1)}|J_N| \leq [R(N^{-1})]^{\alpha}[R(N^{-1})]^{\beta}$. By symmetry we only have to consider C_{1N}, C_{2N} and C_{4N}.

(A) From the above remark it follows immediately that $C_{iN} = \mathcal{O}(N^{\alpha+\beta-1/2}) = o(1)$ as $N \to \infty$, uniformly on $\mathcal{H}$, for $i = 1,2$. As to C_{4N} define $\tilde{\Omega}_{1N} = \{\gamma_N \leq F(X_\lambda) \leq 1-\gamma_N\} \cap \{\gamma_N \leq G(Y_\lambda) \leq 1-\gamma_N\}$, with $\gamma_N = (N \log N)^{-1}$. Thus

$$\chi(\tilde{\Omega}_{1N})C_{4N} \leq N^{-1/2}\ \gamma_N^{-\alpha-\beta}(1-\gamma_N)^{-\alpha-\beta}$$

$$= \mathcal{O}(N^{-1/2+\alpha+\beta}(\log N)^{\alpha+\beta}) \to 0$$

as $N \to \infty$, because $\alpha + \beta < 1/2$. Lemma 5.2.1 ensures that $P(\tilde{\Omega}_{1N}) \to 1$ as $N \to \infty$, uniformly on $\mathcal{H}$.

(B) With Ω_N as in (2.4.1) we find

$$\chi(\Omega_N)C_{1N} \leq MN^{-1/2}[R(N^{-1})]^{\alpha}[R(G(Y_\lambda))]^{\beta}.$$

Let $\delta_N = (N^{a/(1-b)}\log N)^{-1}$ and define $\tilde{\Omega}_{2N} = \{\delta_N \leq G(Y_\lambda) \leq 1-\delta_N\}$. Then

$$\chi(\Omega_N \cap \tilde{\Omega}_{2N})C_{1N} \leq MN^{-1/2}[R(N^{-1})]^{\alpha}\delta_N^{-\beta}(1-\delta_N)^{-\beta}$$

$$= O(N^{-1/2+\alpha+a\beta/(1-b)}(\log N)^{\beta}) \to 0$$

as $N \to \infty$, because $a\beta/(1-b) < a/2 < 1/2-\alpha$. Using the properties of Ω_N and Lemma 5.2.2(ii) it follows that $P(\Omega_N \cap \tilde{\Omega}_{2N}) \geq 1-2\varepsilon$ for N large enough, independently of H in $H_{C,a,b}$.

For C_{2N} we use Lemma 5.2.2(i) to see that the set on which this rv may assume a non-zero value has probability converging to zero as $N \to \infty$, uniformly on $H_{C,a,b}$.

Let us finally consider C_{4N} and introduce $\tilde{\Omega}_{3N} = \{\gamma_N \leq F(X_\lambda) \leq 1-\gamma_N\} \cap \{\delta_N \leq G(Y_\lambda) \leq 1-\delta_N\}$, with γ_N and δ_N as above. Consequently we have

$$\chi(\tilde{\Omega}_{3N})C_{4N} \leq N^{-1/2}\gamma_N^{-\alpha}\delta_N^{-\beta}(1-\gamma_N)^{-\alpha}(1-\delta_N)^{-\beta}$$

$$= O(N^{-1/2+\alpha+a\beta/(1-b)}(\log N)^{\alpha+\beta}) \to 0$$

as $N \to \infty$, for the same reason as above. Moreover, by Lemmas 5.2.1 and 5.2.2 (ii) we see that $P(\tilde{\Omega}_{3N}) \to 1$ as $N \to \infty$, uniformly on $H_{C,a,b}$. □

We mentioned already that the earliest theorem on asymptotic normality of rank statistics for testing the hypothesis of independence under fixed alternatives has been given by Bhuchongkul [4] in 1964, and is based on the assumption of asymptotic negligibility of B_{01N} rather than B^*_{0N}. In [4] only linear rank statistics, i.e. statistics having score functions J_N and limiting score function J which are products of functions of one variable $K_N \times L_N$ and $K \times L$ respectively, are considered. Moreover the paper is restricted to the case where both K and L are continuous throughout (0,1). In order to arrive at a comparison of Theorem 1 of Bhuchongkul [4] with our results, let us take Theorem 5.1.3 as a starting point.

<u>THEOREM 5.2.3</u>. Let the limiting score function $J = K \times L$ satisfy Assumption

5.1.1 with $a_1 = \dots = a_k = b_1 = \dots = b_l = 0$. Fix the numbers $0 < \alpha < 1/2$, $0 < \beta < 1/2$ and $1 \leq C < \infty$, $0 < a < 1-2\alpha$, $0 < b < 1-2\beta$ and suppose that for some $0 < D < \infty$ the score functions satisfy $|J_N(s,t)| \leq D[R(s)]^{\alpha}[R(t)]^{\beta}$ for $(s,t) \in (0,1) \times (0,1)$ and $N = 1,2,\dots$

(A) If $\alpha + \beta < 1/2$, then $N^{1/2}(T_N-\mu) \to_d N(0,\sigma^2)$ as $N \to \infty$, for any underlying df $H \in \mathcal{H}$ for which $B_{01N} \to_P 0$, or equivalently $B^*_{0N} \to_P 0$, as $N \to \infty$. Here $\mu = \mu(H)$ and $\sigma^2 = \sigma^2(H)$ are finite and given by (5.1.8) and (5.1.9) respectively.

This convergence in distribution is uniform on each subclass $\mathcal{H}' \subset \mathcal{H}$ on which $B_{01N_2} \to_P 0$ uniformly, or equivalently $B^*_{0N} \to_P 0$ uniformly, as $N \to \infty$ and on which $\sigma^2 = \sigma^2(H)$ is bounded away from zero.

(B) If $\alpha + \beta \geq 1/2$, then $N^{1/2}(T_N-\mu) \to_d N(0,\sigma^2)$ as $N \to \infty$, for any underlying df $H \in \mathcal{H}_{C,a,b}$ for which $B_{01N} \to_P 0$, or equivalently $B^*_{0N} \to_P 0$, as $N \to \infty$. Here $\mu = \mu(H)$ and $\sigma^2 = \sigma^2(H)$ are finite and given by (5.1.8) and (5.1.9) respectively.

This convergence in distribution is uniform on each subclass $\mathcal{H}' \subset \mathcal{H}_{C,a,b}$ on which $B_{01N_2} \to_P 0$ uniformly, or equivalently $B^*_{0N} \to_P 0$ uniformly, as $N \to \infty$ and on which $\sigma^2 = \sigma^2(H)$ is bounded away from zero.

PROOF. The result follows from straightforward combination of Theorems 5.1.3, 5.2.1 and 5.2.2. □

In Bhuchongkul's Theorem 1 the smoothness conditions imposed on K, L are stronger: these functions are supposed to be twice differentiable throughout (0,1). The growth conditions on K and L are also more restrictive. According to part (A) of the above theorem we have asymptotic normality for any df H in $\mathcal{H}$ (for which $B_{01N} \to_P 0$ as $N \to \infty$) e.g. if we choose $\alpha = \beta = 1/4-\zeta$ for some small $\zeta > 0$ as the exponents determining the growth of K,L, and $5/4-\zeta$ for those determining the growth of $K^{(1)}$, $L^{(1)}$. In [4] these exponents are $1/8-\zeta$ and 1 respectively; in fact the latter condition effectively reduces the orders of magnitude of K,L to $|\log(R)|$. In part (B) of the above theorem even exponents $\alpha = \beta = 1/2-\zeta$ for small $\zeta > 0$ are allowed, provided the underlying df H remains restricted to some subclass $\mathcal{H}_{C,a,b}$ of $\mathcal{H}$. In [4] this case is not considered. In Puri and Sen [30], Theorem 8.4.1, an immediate multivariate extension of Bhuchongkul's result is presented, based on the asymptotic negligibility of B^*_{0N} and under smoothness conditions on K and L that are considerably weaker than those of either Theorem 1 in [4] or our Theorem 5.2.3. Apart from this, the conditions and the content of their theorem are the same as those of [4], Theorem 1. Since they do not provide a complete proof, it is hard to judge whether their conditions do in fact suffice. In a joint paper [34] by Shorack, van Zwet and the author, asymptotic normality is established under the condition that either B_{01N} or B^*_{0N} is asymptotically negligible. This result is essentially the same as that of Theorem 5.2.3 above.

The proof of [4], Theorem 1, as well as the proof of our Theorem 2.1.1 is based on the method employed by Chernoff and Savage [7] in 1958, to prove asymptotic normality of linear rank statistics for the two-sample problem under fixed alternatives. However, both in [4] and [7] a Taylor-series expansion up to second order derivatives is used where we have needed only the mean value theorem. We have borrowed a number of ideas from Bhuchongkul, such as the representation (1.1.9) of T_N as an integral with respect to the bivariate empirical df and the outlines of the proof of Corollary 2.4.5.

In [4] the crucial condition concerning the asymptotic negligibility of B_{01N} is studied in Theorem 2. The next result is a generalization of this theorem.

<u>THEOREM 5.2.4</u>. **Let** the limiting score function be of the product type $J = K \times L$ with K and L defined on (0,1). Suppose that K and L are continuous on (0,1) and possess second derivatives $K^{(2)}$ and $L^{(2)}$ on $\cup_{i=1}^{k+1} (s_{i-1},s_i)$ and $\cup_{j=1}^{l+1} (t_{j-1},t_j)$ respectively. For given fixed $0 < D_1 < \infty$, $0 < D_2 < \infty$, $0 < \alpha < 1/2$, $0 < \beta < 1/2$ let $|K^{(i)}| \leq D_1 R^{\alpha+i}$, $|L^{(i)}| \leq D_2 R^{\beta+i}$ where defined on (0,1), i = 0,1,2. For each N = 1,2,... let J_N be either the exact (see (1.2.2)) or the approximate (see (1.2.3)) score function derived from $J = K \times L$. Then in both cases the following holds true.

On $(0,1] \times (0,1]$ the function J_N can be written as a product $J_N = K_N \times L_N$ of simple step functions K_N and L_N defined on (0,1]. There exist constants $0 < D_1' < \infty$, $0 < D_2' < \infty$ (not depending on N) such that $|K_N| \leq D_1' R^{\alpha}$, $|L_N| \leq D_2' R^{\beta}$ on (0,1). Moreover $|K_N(1)| = O(N^{\alpha})$ and $|L_N(1)| = O(N^{\beta})$.

(A) If $\alpha + \beta < 1/2$, then $B^*_{0N} \to_P 0$ and $B_{01N} \to_P 0$ as $N \to \infty$, uniformly on $\mathcal{H}$.

(B) If $\alpha + \beta \geq 1/2$, then $B^*_{0N} \to_P 0$ and $B_{01N} \to_P 0$ as $N \to \infty$, uniformly on $\mathcal{H}_{C,a,b}$, for given fixed $1 \leq C < \infty$, $0 < a < 1-2\alpha$, $0 < b < 1-2\beta$.

<u>PROOF</u>. By symmetry we only have to consider K_N. In the case of exact scores it is obvious to choose (see also (1.2.1))

(5.2.6) $$K_N(s) = \int_0^1 K(u) b_{m,N-m+1}(u)du,$$

for s in $((m-1)/N, m/N]$, $m = 1,\ldots,N$. In the case of approximate scores define

(5.2.7) $$K_N(s) = K(m/(N+1)),$$

for s in $((m-1)/N, m/N]$, $m = 1,\ldots,N$.

Because $|[R(m/N)]^\alpha - [R((m-1)/N)]^\alpha| \le D_{11}[R(m/N)]^\alpha$, for $m = 2,\ldots,N-1$ and some $0 < D_{11} < \infty$, and since in both cases K_N is a simple step function it suffices to show that $K_N(m/N) \le D_1^*[R(m/N)]^\alpha$ for $m = 1,\ldots,N-1$ and some $0 < D_1^* < \infty$, and that $K_N(1) = O(N^\alpha)$.

First suppose that K_N is given by (5.2.6). Then

$$|K_N(m/N)| \le D_1 \Gamma(N+1)[\Gamma(m)\Gamma(N-m+1)]^{-1} \int_0^1 s^{m-\alpha-1}(1-s)^{N-m-\alpha} ds$$

$$= D_1 \Gamma(m-\alpha)\Gamma(N-m-\alpha+1)[\Gamma(m)\Gamma(N-m+1)\Gamma(N-2\alpha+1)]^{-1}$$

$$\le \tilde{D}_1 m^{-\alpha}(N-m+1)^{-\alpha}(N+1)^{2\alpha} = \tilde{D}_1[R(m/(N+1))]^\alpha,$$

for all $m = 1,\ldots,N$, where $0 < \tilde{D}_1 < \infty$. The symbol Γ in the above formulas denotes the gamma-function. When K_N is given by (5.2.7) it follows at once that

$$|K_N(m/N)| \le D_1[R(m/(N+1))]^\alpha,$$

for all $m = 1,\ldots,N$. Clearly in both cases $|K_N(1)| = O(N^\alpha)$. Because for some $0 < D_{12} < \infty$ we have

$$[R(m/(N+1))]^\alpha \le D_{12}[R(m/N)]^\alpha,$$

we may conclude that in both cases $|K_N(m/N)| \le D_1^*[R(m/N)]^\alpha$, for $m = 1,\ldots,N-1$. Here $D_1^* = \max\{D_1 \times D_{12}, \tilde{D}_1 \times D_{12}\}$.

First let J_N be the exact score function. Because of the first part of the theorem we may write $B_{01N} = \sum_{i=1}^2 C_{iN}$, where

$$C_{1N} = N^{1/2} \iint_{\Delta_N} [K_N(F_N) - K(F_N)] L(G_N)\, dH_N,$$

$$C_{2N} = N^{1/2} \iint_{\Delta_N} K_N(F_N) [L_N(G_N) - L(G_N)]\, dH_N.$$

It follows from Chernoff and Savage [7], formulas (7.14) and (7.24), that under the conditions of the present theorem we have

$$\text{(5.2.8)} \qquad \sum_{m=1}^{N-1} |K_N(m/N) - K(m/N)| = O(N^{\alpha}), \quad \sum_{n=1}^{N-1} |L_N(n/N) - L(n/N)| = O(N^{\beta}).$$

The fact that K(L) fails to have a derivative at a finite number of fixed points in the open unit interval, which makes the condition on K(L) slightly weaker compared with Chernoff and Savage [7], does not affect the validity of the above conclusion. (If necessary the Taylor-series expansion is applied step-wise.)

If $\alpha + \beta < 1/2$, then independently of H in $\mathcal{H}$

$$|C_{1N}| \le D_2 N^{-1/2} \sum_{m=1}^{N-1} |K_N(m/N) - K(m/N)| [R(1/N)]^{\beta}$$

$$= O(N^{-1/2+\alpha+\beta}) \to 0$$

as $N \to \infty$. Here (5.2.8) is used. With the aid of the results of the first part of the present theorem we can deal with C_{2N} in a similar fashion.

Suppose $\alpha + \beta \ge 1/2$. Application of Lemma 2.3.2(i) with $r = D_2 R^{\beta}$ and $\tilde{\Psi}_N = G_N$ gives for an arbitrary $\varepsilon > 0$ the existence of a constant $M = M_{\varepsilon}$, such that $\tilde{\Omega}_N = \{D_2[R(G_N)]^{\beta} \le MD_2[R(G)]^{\beta}$ on $\Delta_{N2}\}$ has probability $P(\tilde{\Omega}_N) \ge 1-\varepsilon$ for all N and all H in $\mathcal{H}$. Let us for simplicity define $K_N^* = K$ on (0,1) and $K_N^*(1) = K_N(1)$ for N = 1,2,... Arguing as in the proof of Lemma 3.3.2(ii), we see that independently of H in $\mathcal{H}_{C,a,b}$

$$E(\chi(\tilde{\Omega}_N)|C_{1N}|)$$

$$\le MD_2 N^{-1/2} \sum_{n=1}^{N} E(|K_N(F_N(X_n)) - K_N^*(F_N(X_n))| [R(G(Y_n))]^{\beta})$$

$$= MD_2 N^{-1/2} \sum_{n=1}^{N} E\{[R(G(Y_n))]^{\beta} \times E(|K_N(F_N(X_n)) - K_N^*(F_N(X_n))| \,|\, F(X_n))\}$$

$$= MD_2 N^{1/2} \iint [R(G)]^{\beta} \sum_{m=1}^{N} |K_N(m/N) - K_N^*(m/N)| \binom{N-1}{m-1} F^{m-1} (1-F)^{N-m} dH$$

$$= O(N^{1/2} \sum_{m=1}^{N} |K_N(m/N) - K_N^*(m/N)| \quad \times$$

$$\binom{N-1}{m-1}\int_0^1 s^{m-1-a/2}(1-s)^{N-m-a/2}ds)$$

$$= \quad O(N^{1/2-a/2+a-1+\alpha}) \to 0$$

as $N \to \infty$, because $\alpha + a/2 < 1/2$. Again (5.2.8) is used. Since $\varepsilon > 0$ is arbitrary, this proves that $C_{1N} \to_P 0$ as $N \to \infty$, uniformly on $H_{C,a,b}$. By the first part of the present theorem a similar argument may be used for C_{2N}.

In case (A) and (B) respectively the conditions of Theorem 5.2.2 (A) and (B) respectively are satisfied, because of the results of the first part of the present theorem. Hence, in combination with Theorem 5.2.1 it follows that also $B^*_{0N} \to_P 0$ as $N \to \infty$, uniformly on H and on $H_{C,a,b}$ respectively.

Secondly let J_N be the approximate score function. Then by definition $B^*_{0N} = 0$ for all $N = 1,2,\ldots$ and all H in H. Again by the first part of the present theorem, in case (A) and (B) the conditions of Theorem 5.2.2 (A) and (B) respectively are satisfied. Hence, in combination with Theorem 5.2.1 it follows that also $B_{01N} \to_P 0$ as $N \to \infty$, uniformly on H and on $H_{C,a,b}$ respectively. □

We conjecture that the conclusions of the last theorem will remain valid in the case where K and L are allowed to have a finite number of discontinuities of the first kind. A restriction on the class of underlying dfs, in the spirit of Assumption 3.1.4, will be needed. A proof would appear to be very technical and laborious.

We shall not dwell upon a comparison of our results with existing theorems for the case of the null hypothesis, since our primary aim was the investigation under fixed alternatives. Of course our results apply under the null hypothesis but we have in no way exploited the independence. Let us nevertheless mention some results by Hájek and Šidák [17] and Jogdeo [19], and use Theorem 5.1.2 (C) or, more generally, Theorem 3.1.1 for comparison. In Hájek and Šidák [17], Theorem V.1.8, asymptotic normality of linear rank statistics in the case where $J_N = K_N \times L_N$, $J = K \times L$, and either exact or approximate scores derived from J are used, is shown under the null hypothesis. The only condition on K,L is measurability and square integrability in the case of exact scores; in the case of approximate scores both K and L should be expressible as a finite sum of monotone and square integrable functions. In Jodgeo [19], Theorem 4.2 and Section 5, asymptotic normality in the case where J is not necessarily a product of two univariate functions, and J_N is either the approximate score function or derived from J by piecewise integration, is shown under the null hypothesis. The statistics are of the more general type $N^{-2} \sum_{m=1}^{N} \sum_{n=1}^{N} c_{mn} \times J_N(F_N(X_m), G_N(Y_n))$, where the regression constants may be chosen $c_{mn} = \delta_{mn}$ (Kronecker's δ). The function J should be piecewise monotone and J^4 should be integrable.

Our methods are also not very well adapted to a treatment of contiguous alternatives. A result which points in this direction is contained in Theorem 5.1.4 or, more generally, in Theorem 4.2.1. In Hājek and Šidāk [17] a modification of Theorem V.1.8 for contiguous alternatives of type (1.2.4) is conjectured (see [17], page 222). Behnen [2] proves the contiguity of more general sequences of alternatives and establishes asymptotic normality under these contiguous alternatives in the case where $J_N = K_N \times L_N$, $J = K \times L$ and K_N, L_N are score functions derived from K,L respectively. The only condition on K,L is measurability and square integrability. In Behnen [3] the restriction that J is a product of two univariate functions is removed. A large amount of information on asymptotic theory of rank tests is also contained in Witting and Nölle [42].

Under fixed alternatives, no previous general result on asymptotic normality for limiting score functions J that are no longer continuous is known to the author. However, it is worth while to call attention to the papers by Hājek [14] and Dupač and Hājek [8], where asymptotic normality of simple linear rank statistics under fixed alternatives is shown. Possibly a useful relationship between the regression and the independence problem can also be established under fixed alternatives, as this has already been done in Hājek and Šidāk [17] under the null hypothesis. It is illuminating that the presence of a single discontinuity of the first kind in the limiting score function, considered in [8], necessitates a local differentiability condition on the underlying dfs and causes considerable technical difficulties, as it does in our case. Their proof, however, is based on a quite different approach. To handle the discontinuities in the limiting score function we mainly need Lemma 1.3.4, which is similar to a bivariate form of Lemma 1 by Bahadur [1], Theorem 2.1 by Sen [36], or Theorem 1 by Ghosh [12]. The results of Bahadur and Sen are for univariate dfs only but stronger in the sense that they provide "almost sure" statements while our result gives a statement "in probability". On the other hand Lemma 1.3.4 does not require any condition on the underlying bivariate df, which need not even be continuous, and the conclusion of the lemma is uniform in all sequences of intervals in the plane. It should be noted that Sen [37] applies his result [36] to multivariate rank order statistics for the two-sample problem in the case where the limiting score functions are simple step functions, where we use Lemma 1.3.4. Let us finally mention the paper [31] by Pyke and Shorack, where still another method is used to overcome the difficulties of discontinuous limiting score functions in the univariate two-sample problem. Their starting point is the representation of the statistic as an integral of some random process, called the two-sample empirical process, with respect to a non-random signed measure. For a similar approach in our case the representation (1.1.12) of the statistic could perhaps be a starting point.

5.3. THE EXAMPLES OF SECTION 1.2 REVISITED

In this section we shall briefly review the examples of Section 2.1 and show how our results apply. In all examples, except for Example 1.2.4, the boundedness conditions of Assumption 5.1.1 are satisfied with $\alpha + \beta < 1/2$. To see this note that

$$|[\Phi^{-1}]^{(i)}| = \mathcal{O}([(\log R)^{1/2}]^{(i)}) = \mathcal{O}(R^{\zeta+i}), \qquad i = 0,1,2, \tag{5.3.1}$$

for any $\zeta > 0$. In Example 1.2.4 for $\nu > 4$ this boundedness condition is still satisfied with $\alpha + \beta < 1/2$, but for $\nu > 2$ we have $\alpha + \beta \geq 1/2$.

According to Theorems 5.1.1 (A) and 5.2.4 (A), the statistics of Examples 1.2.2, 1.2.3, 1.2.6, 1.2.7 have a normal distribution in the limit for any underlying df $H \in \mathcal{H}$, both when exact and when approximate score functions are used.

The statistic of Example 1.2.4 has a limiting normal distribution for any underlying df $H \in \mathcal{H}$ if $\nu > 4$ and for any df $H \in \mathcal{H}_{C,a,b}$ with C sufficiently large and $0 < a = b < (\nu-2)/\nu$ if $\nu > 2$, both when exact and when approximate score functions are used. This follows from Theorems 5.1.1 (A) and 5.2.4 (A) (Theorems 5.1.1 (B) and 5.2.4 (B)).

When approximate score functions are used, the statistic derived from (1.2.10) is asymptotically normally distributed for any underlying df $H \in \mathcal{H}$. This follows from Theorem 2.1.1 and formula (5.3.1).

Finally the statistics of Examples 1.2.1, 1.2.5 have a normal law in the limit for any underlying df $H \in \bar{\mathcal{H}}_{C,1-\zeta,1-\zeta}$ (C sufficiently large and any $\zeta > 0$), provided approximate score functions are used. This follows from Theorem 5.1.2 (A), formula (5.1.3) and by noticing that equivalent statistics are obtained when the function sgn (z) is replaced by the function $-1 + 2c(z)$ in the expressions for J.

In conclusion of this section let us return to the statistic defined by the limiting score function given in Example 1.2.2 or in (1.2.18) (equivalent to Spearman's rank correlation), to be denoted by T_{1N}. At the same time we consider the statistic defined by the limiting score function in Example 1.2.1 (giving rise to the quadrant statistic), to be denoted by T_{2N}. Both statistics will be studied for underlying dfs belonging to Gumbel's class (see also (1.2.17))

$$\text{(5.3.2)} \qquad \{H_\theta \in \mathcal{H} : H_\theta = FG[1+\theta(1-F)(1-G)],\ F \times G \in \mathcal{H}_0,\ 0 \leq \theta < 1\},$$

against which tests based on T_{1N} are locally most powerful, as we have seen at the end of Section 1.2. For both statistics the limiting mean and variance are easy to calculate. The density of the transformed df $\bar{H}$ equals $\bar{h}_\theta(s,t) = 1+\theta(2s-1)(2t-1)$, for $0 < s < 1$, $0 < t < 1$. Thus we find

$$\mu_{1\theta} = \int_0^1\int_0^1 (2u-1)(2v-1)\bar{h}_\theta(u,v)du\,dv = \theta/9,$$

$$\mu_{2\theta} = \int_0^1\int_0^1 \operatorname{sgn}(u-1/2)\operatorname{sgn}(v-1/2)\bar{h}_\theta(u,v)du\,dv = \theta/4,$$

$0 \leq \theta < 1$. In particular $\mu_{10} = \mu_{20} = 0$. With the aid of formula (3.1.4) we obtain

$$\sigma_{10}^2 = \int_0^1\int_0^1 [(2u-1)(2v-1)]^2 du\,dv = 1/9,$$

$$\sigma_{20}^2 = \int_0^1\int_0^1 [\operatorname{sgn}(u-1/2)\operatorname{sgn}(v-1/2)]^2 du\,dv = 1.$$

According to Theorem 4.1.1 both the test based on T_{1N} and the one based on T_{2N} are consistent against the alternatives in (5.3.2) for $0 < \theta < 1$.

As to their asymptotic relative efficiency let us choose a sequence of alternatives in (5.3.2) determined by $\theta_N = \tilde{\theta}N^{-1/2}$ for some $0 < \tilde{\theta} < 1$. Then we have (writing $\mu_{i\theta_N} = \mu_{iN}$, $i = 1,2$, and $H_{\theta_N} = H_{(N)}$)

$$N^{1/2}(\mu_{1N}-\mu_{10})/\sigma_{10} = \tilde{\theta}/3,\quad N^{1/2}(\mu_{2N}-\mu_{20})/\sigma_{20} = \tilde{\theta}/4,$$

so that

$$\mathrm{ARE}(T_{1N},T_{2N};\{H_{(N)}\},\alpha) = 4/3,$$

independently of the choice for $0 < \tilde{\theta} < 1$.

We found that $\mu_{1\theta} = \theta/9$ and direct calculation shows that $\sigma^2_{1\theta} = (1/9) - (11/5)(\theta^2/81) = (1/9) - (11/5)\mu^2_{1\theta}$. According to (4.1.2) a variance stabilizing transformation f is any solution of the equation $f'(\mu) = 3[1-(99/5)\mu^2]^{-1/2}$. A solution of this equation is

$$f(\mu) = (6/c)\arctan[(1+c\mu)/(1-c\mu)]^{1/2}, \quad c = (99/5)^{1/2}. \tag{5.3.3}$$

Hence for the statistics

$$T_{1N} = [\textstyle\sum_{n=1}^{N} \{2NR_{nN}/(N+1)-1\}\{2NQ_{nN}/(N+1)-1\}]/N, \tag{5.3.4}$$

which are locally most powerful and consistent against Gumbel's alternatives given in (5.3.2), we have that

$$N^{1/2}[f(T_{1N})-f(\mu_\theta)] \to_d N(0,1), \tag{5.3.5}$$

as $N \to \infty$, irrespective of the particular underlying df H_θ in the class (5.3.2). Here f is given by (5.3.3).

As to the asymptotic optimality of the tests based on T_{1N} let us from now on fix the choice of $F \times G \in H_0$ in the class (5.3.2) and choose θ_N as before, with corresponding dfs $H_{(N)}$ in (5.3.2). The asymptotically most powerful test is based on the statistic

$$\tilde{T}_{1N} = \iint (2F-1)(2G-1)\,dH_N^{(N)},$$

see Witting and Nölle [42. page 152, 153]. Here the function $J(u,v) = (2u-1)(2v-1)$ clearly statisfies (4.3.7). As in [42] we arrive at the conclusion that the tests based on the T_{1N} are asymptotically most powerful against $H_{(1)}, H_{(2)}, \ldots$

Chapter 6

DISCRETE UNDERLYING DISTRIBUTION FUNCTIONS

6.1. INTRODUCTION

In this chapter we are exclusively concerned with the case where the sequence (X_1,Y_1), (X_2,Y_2),... of iid random vectors has an underlying bivariate df H that is entirely concentrated on a finite lattice of points in the plane. This lattice is given by $\{(x_i,y_j),\ i=1,\dots,q \text{ and } j=1,\dots,r\}$, where q and r are supposed to be fixed positive integers. (Hence throughout this chapter q(r) is not an element of $\mathcal{Q}(\mathcal{R})$.) Let us denote $P(\{X=x_i,\ Y=y_j\}) = \pi_{ij}$ and let us assume that

$$(6.1.1)\qquad \sum_{j=1}^{r} \pi_{ij} = \pi_{i.} > 0,\ \sum_{i=1}^{q} \pi_{ij} = \pi_{.j} > 0,$$

$i=1,\dots,q$ and $j=1,\dots,r$. The only purpose of this assumption is to guarantee that any x_i is a possible realization of X and that any y_j is a possible realization of Y. In other words, in this section we shall deal with $q \times r$ *contingency tables*, where the categories of the two attributes involved admit a natural ordering.

Some adaptations of rank statistics as constructed for underlying continuous bivariate dfs to the above mentioned discrete situation, and the asymptotic normality of the resulting statistics are our only concern. We do not consider asymptotic nor finite optimality properties. For the first the reader is referred to Vorlíčková [41] and for the latter to Krauth [24]. Both in [24] and [41] a countable infinite lattice of points is considered for underlying dfs belonging to, or contiguous to, the null hypothesis. In our case the underlying df is arbitrary and we conjecture that our results can be proved for countable lattices in essentially the same way.

In this chapter we shall only consider limiting score functions that are of the product type $J = K \times L$, and we shall only use approximate score functions (see (1.2.3)); i.e. we shall consider modifications of statistics of the relatively simple form

$$(6.1.2)\qquad T_N = N^{-1} \sum_{n=1}^{N} K(R_{nN}/(N+1))L(Q_{nN}/(N+1)),$$

for some functions K and L on (0,1). Here the definition of rank, given in

Section 1.1 for an arbitrary underlying bivariate df (continuous or not), is used which amounts to saying that

$$(6.1.3) \qquad R_{nN} = \sum_{m=1}^{N} c(X_n - X_m), \; Q_{nN} = \sum_{m=1}^{N} c(Y_n - Y_m),$$

$n = 1,\ldots,N$, with $c(z)$ defined in (1.1.1).

In a sample from such a discrete bivariate df, there is a positive probability that there will be groups of sample elements having both equal first and second coordinates. Any such group is called a *tie*. Each group of equal first (second) coordinates will be called a *marginal tie*. Consequently the ranks are no longer all different with probability 1. As a matter of fact, all observations (X,Y) in one and the same tie give rise to the same pair of ranks (R,Q).

More precisely, given a random sample of size N, let

$$(6.1.4) \qquad \begin{aligned} &I_{i.} = \{n : X_n = x_i,\ n=1,\ldots,N\},\ I_{.j} = \{n : Y_n = y_j,\ n=1,\ldots,N\} \\ &I_{ij} = I_{i.} \cap I_{.j}, \end{aligned}$$

$i = 1,\ldots,q$ and $j = 1,\ldots,r$. The random number of indices contained in $I_{i.}$, $I_{.j}$ and I_{ij} will be denoted by $\nu_{i.}$, $\nu_{.j}$ and ν_{ij} respectively. It will be convenient to define $\pi_{0.} = \pi_{.0} = \nu_{0.} = \nu_{.0} = 0$. Let us introduce the notation

$$(6.1.5) \qquad M(N,\pi_1,\ldots,\pi_k)$$

for a multinomial distribution based on N trials with success probabilities $0 \leq \pi_1 \leq 1,\ldots,\ 0 \leq \pi_k \leq 1$, satisfying $\pi_1 + \ldots + \pi_k = 1$. Then the random vector $(\nu_{11},\ldots,\nu_{1r},\ldots,\nu_{q1},\ldots,\nu_{qr})$ has a $M(N,\pi_{11},\ldots,\pi_{1r},\ldots,\pi_{q1},\ldots,\pi_{qr})$ distribution. For the ranks we may write

$$(6.1.6) \qquad \begin{aligned} &R_{nN} = \nu_{1.} + \ldots + \nu_{i.} \qquad \text{for all } n \in I_{i.}, \\ &Q_{nN} = \nu_{.1} + \ldots + \nu_{.j} \qquad \text{for all } n \in I_{.j}, \end{aligned}$$

for $i = 1,\dots,q$ and $j = 1,\dots,r$.

Later on we shall make use of the weak convergence

(6.1.7) $$N^{1/2}(N^{-1}\nu_{11}-\pi_{11},\dots,N^{-1}\nu_{1r}-\pi_{1r},\dots,N^{-1}\nu_{q1}-\pi_{q1},\dots,N^{-1}\nu_{qr}-\pi_{qr})\to_d \bar{N}(0,S),$$

as $N \to \infty$, where $\bar{N}(0,S)$ is a degenerate $q \times r$-dimensional normal distribution with mean vector 0 and covariance matrix S, having at most rank $q \times r-1$ because of the boundary condition $\sum_{i=1}^{q} \sum_{j=1}^{r} \nu_{ij} = N$, given by

(6.1.8) $$S_{ij,i'j'} = \pi_{ij}(\delta_{ii'}\delta_{jj'}-\pi_{i'j'}),$$

for $i,i' = 1,\dots,q$ and $j,j' = 1,\dots,r$. Here $\delta_{ii'}$ and $\delta_{jj'}$ are Kronecker-deltas.

In Hájek [15,16] and Vorlíčková [41] two possible ways of removing the ties are described. The first of these is based on randomization of the ranks, leaving the score functions unaltered, and the second on an averaging procedure applied to the score functions, leaving the ranks unaltered.

To start with the first method, suppose that (ξ,η) is a random vector on (Ω,A,P) having a bivariate uniform distribution on $(-d,d) \times (-d,d)$, where d is a fixed constant satisfying
$0 < d < \min\{x_{i+1}-x_i, y_{j+1}-y_j;\ i=1,\dots,q-1 \text{ and } j=1,\dots,r-1\}/2$ (see Figure 6.1.1). Let $(\xi_1,\eta_1),(\xi_2,\eta_2),\dots$ be a sequence of iid random vectors all defined on (Ω,A,P), all possessing the same bivariate uniform distribution

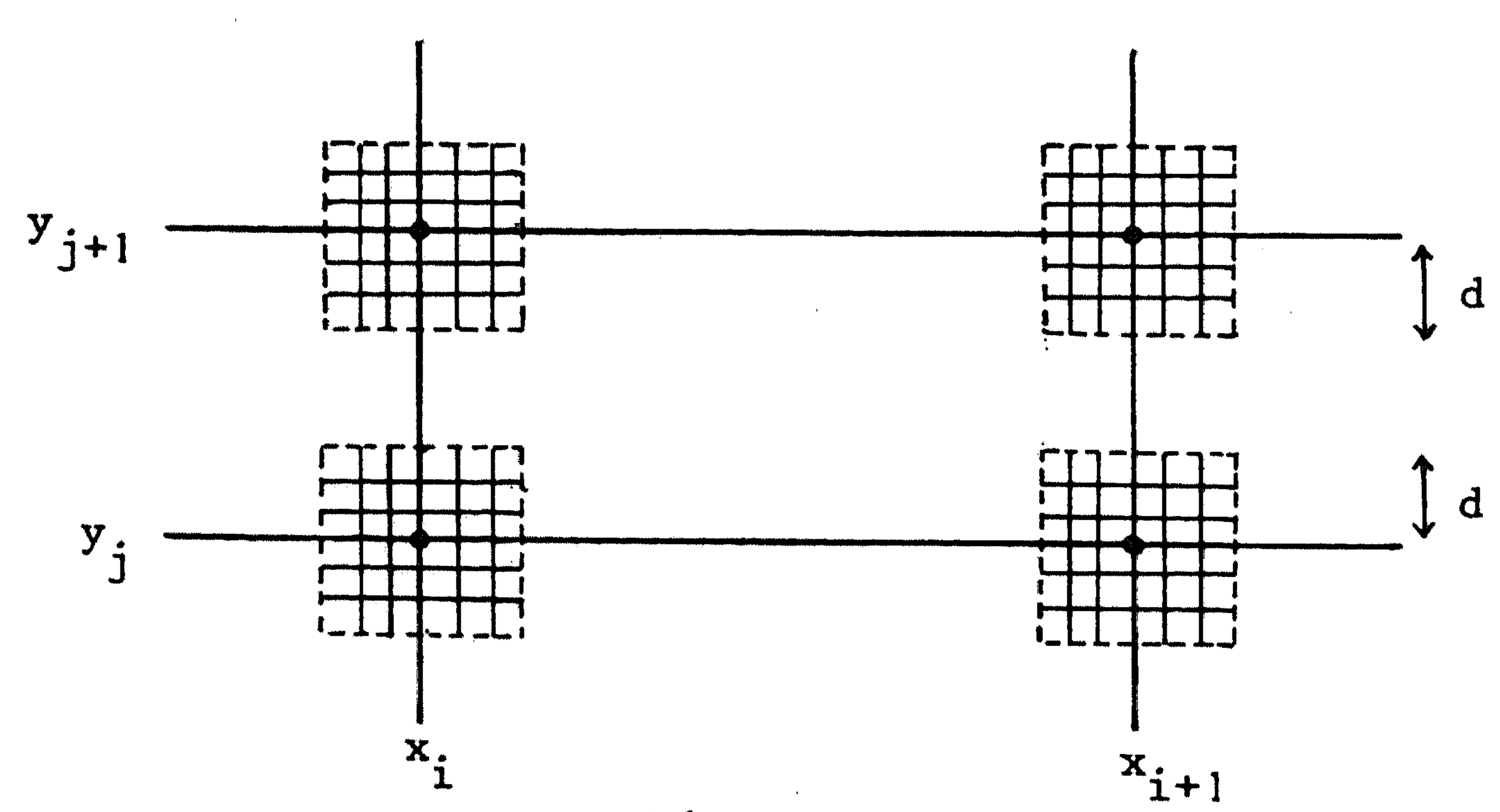

Figure 6.1.1

as (ξ,η), and all independent of the $(X_1,Y_1),(X_2,Y_2),\ldots$ The original random vectors (X_n,Y_n) will be replaced by

$$(X_n^*,Y_n^*) = (X_n+\xi_n,Y_n+\eta_n),$$

for $n = 1,2,\ldots$ The sample $(X_1^*,Y_1^*),\ldots,(X_N^*,Y_N^*)$ may be thought of as a random sample from the bivariate df H^* of the random vector $(X+\xi,Y+\eta)$. Since this df is easily seen to be continuous the corresponding ranks

$$(6.1.9) \qquad R_{nN}^* = \sum_{m=1}^N c(X_n^*-X_m^*),\ Q_{nN}^* = \sum_{m=1}^N c(Y_n^*-Y_m^*),$$

$n = 1,\ldots,N$, are with probability 1 permutations of the numbers $1,\ldots,N$. The ranks defined in (6.1.9) are called *randomized ranks*. In view of (6.1.6) we have

$$R_{nN}^* = \nu_{0.} + \ldots + \nu_{i-1.}+1,\ldots,\nu_{0.} + \ldots + \nu_{i.} \text{ as n varies in } I_{i.},$$

$$Q_{nN}^* = \nu_{.0} + \ldots + \nu_{.j-1}+1,\ldots,\nu_{.0} + \ldots + \nu_{.j} \text{ as n varies in } I_{.j}.$$

Let T_N be replaced by the statistic

$$(6.1.10) \qquad T_N^* = N^{-1} \sum_{n=1}^N K(R_{nN}^*/(N+1))L(Q_{nN}^*/(N+1)),$$

based on the original score functions and the randomized ranks. The distribution of T_N^* is the same for all H^* in H_0 and we shall make clear below that the asymptotic theory for T_N^* relies directly on the results of Chapter 3. A drawback of this method is of course that the statistician's conclusion depends on the randomization involved.

Before asserting the asymptotic normality of T_N^* let us consider more closely the df H^*. It has a density

$$\partial^2 H^*(x,y)/\partial x\partial y = (2d)^{-2} \sum_{i=1}^q \sum_{j=1}^r \pi_{ij}\ \chi((a_i-d,a_i+d);x) \times$$

$$\chi((b_j-d,b_j+d);y).$$

The transformed df $\bar{H}^*$ (see (1.2.11)) has a density $\bar{h}^*$ throughout $(0,1) \times (0,1)$ assuming the values

$$(6.1.11) \qquad \bar{h}^*(s,t) = \sum_{i=1}^q \sum_{j=1}^r (\pi_{i.}\pi_{.j})^{-1}\pi_{ij}\chi_{i.}(s)\chi_{.j}(t),$$

for (s,t) in (0,1) × (0,1). Here $\chi_{i.}$ and $\chi_{.j}$ are the indicator functions of the intervals $(\pi_{0.} + \dots + \pi_{i-1.}, \pi_{0.} + \dots + \pi_{i.})$ and $(\pi_{.0} + \dots + \pi_{.j-1}, \pi_{.0} + \dots + \pi_{.j})$ respectively. Hence $\bar{h}^*$ is bounded on (0,1) × (0,1), but discontinuous on the lines $s = \pi_{1.} + \dots + \pi_{i.}$ $(i=1,\dots,q-1)$ and $t = \pi_{.1} + \dots + \pi_{.j}$ $(j=1,\dots,r-1)$. Of course it follows that we may take $\bar{h}^* = 1$ on (0,1) × (0,1) if H (and hence H^*) is a null hypothesis df.

The limiting mean $\mu^* = \mu(H^*)$ and variance $\sigma^{*2} = \sigma^2(H^*)$ can be expressed in the probabilities π_{ij} by means of formulas (5.1.8) and (5.1.9). Let us write for brevity

(6.1.12) $$(\pi_{i.})^{-1} \int_0^1 \chi_{i.}(s)K(s)ds = \kappa_i, \quad (\pi_{.j})^{-1} \int_0^1 \chi_{.j}(t)L(t)dt = \lambda_j,$$

for $i = 1,\dots,q$ and $j = 1,\dots,r$. We find

(6.1.13) $$\mu^* = \mu(H^*) = \int_0^1\int_0^1 K(u)L(v)d\bar{H}^*(u,v) = \sum_{i=1}^q \sum_{j=1}^r \kappa_i\lambda_j\pi_{ij}.$$

As to the variance first note that

$$E(L(G^*(Y^*))|F^*(X^*) = s) = \int_0^1 L(t)\bar{h}^*(s,t)dt,$$

$$E(K(F^*(X^*))|G^*(Y^*) = t) = \int_0^1 K(s)\bar{h}^*(s,t)ds,$$

for $0 < s < 1$ and $0 < t < 1$. Hence we obtain

(6.1.14) $$\sigma^{*2} = \sigma^2(H^*) = \sum_{i=1}^q \sum_{j=1}^r (\pi_{i.}\pi_{.j})^{-1}\pi_{ij} \times \int_0^1\int_0^1\{K(u)L(v)-\mu^* + \int_0^1[c(s-u)-s][\int_0^1 L(t)\bar{h}^*(s,t)dt]dK(s) + \int_0^1[c(t-v)-t][\int_0^1 K(s)\bar{h}^*(s,t)ds]dL(t)\}^2 \chi_{i.}(u)\chi_{.j}(v)dudv.$$

<u>THEOREM 6.1.1</u>. Suppose that the limiting score function $J = K \times L$ satisfies Assumption 5.1.1 with arbitrary but fixed $a_1,\dots,a_k$, $b_1,\dots,b_l$ and $0 < \alpha < 1/2$, $0 < \beta < 1/2$. Let J_N $(N=1,2,\dots)$ be the approximate score function derived from J and let H be any fixed discrete underlying df,

satisfying (6.1.1) and $\{s_1,\dots,s_k\} \cap \{\pi_{1.},\dots,\sum_{i=1}^{q-1} \pi_{i.}\} = \{t_1,\dots,t_l\} \cap \{\pi_{.1},\dots,\sum_{j=1}^{r-1}\pi_{.j}\} = \emptyset$ and with corresponding continuous df $H^* \in \mathcal{H}$. Then $N^{1/2}(T_N^*-\mu^*) \to_d \mathcal{N}(0,\sigma^{*2})$. Here T_N^* is based on the randomized ranks (see (6.1.10)) and $\mu^* = \mu(H^*)$ and $\sigma^{*2} = \sigma^2(H^*)$, given by (6.1.13) and (6.1.14) respectively, are finite.

PROOF. For arbitrary discrete df H the corresponding transformed continuous df $\bar{H}^*$ has a density $\bar{h}^*$ which is bounded on $(0,1) \times (0,1)$, so that $\bar{H}^* \in \mathcal{H}_{C,a,b} \cap \bar{\mathcal{H}}_{C,\alpha+a,\beta+b}$ (see (5.1.5) and (5.1.7)) for sufficiently large C. Hence the conclusion follows immediately from Theorem 5.1.2 because approximate scores are used. □

Let us note that the probability distribution of T_N^* and the values of μ^* and σ^{*2} do not depend on the particular underlying discrete df H for which $\pi_{ij} = \pi_{i.}\pi_{.j}$ (i.e. for which X and Y are independent), because the corresponding continuous df H^* is an element of $\mathcal{H}_0$.

To describe the second method of removing the ties let us introduce the new score function $\bar{J}_{(N)} = \bar{K}_{(N)} \times \bar{L}_{(N)}$. Here

$$(6.1.15) \qquad \bar{K}_{(N)}(s) = (\nu_{i.})^{-1} \sum_{m=\nu_{0.}+\dots+\nu_{i-1.}+1}^{\nu_{0.}+\dots+\nu_{i.}} K(m/(N+1)) = \kappa_{Ni},$$

for $(\nu_{0.} + \dots + \nu_{i-1.})/(N+1) < s \leq (\nu_{0.} + \dots + \nu_{i.})/(N+1)$, $i = 1,\dots,q$, and

$$(6.1.16) \qquad \bar{L}_{(N)}(t) = (\nu_{.j})^{-1} \sum_{n=\nu_{.0}+\dots+\nu_{.j-1}+1}^{\nu_{.0}+\dots+\nu_{.j}} L(n/(N+1)) = \lambda_{Nj},$$

for $(\nu_{.0} + \dots + \nu_{.j-1})/(N+1) < t \leq (\nu_{.0} + \dots + \nu_{.j})/(N+1)$, $j = 1,\dots,r$.
The score function $\bar{J}_{(N)} = \bar{K}_{(N)} \times \bar{L}_{(N)}$ will be called *averaged score function*. This function is of a stochastic nature: it depends on the sizes of the marginal ties.

We now replace T_N by

$$(6.1.17) \qquad \bar{T}_N = N^{-1} \sum_{n=1}^{N} \bar{K}_{(N)}(R_{nN}/(N+1))\bar{L}_{(N)}(Q_{nN}/(N+1)),$$

based on the original ranks and the averaged score function. An advantage over the first method of dealing with ties is that no additional random experiment is involved.

Conditionally, given the sizes of the marginal ties, the (limiting) distribution of $\bar{T}_N$ does not depend on the particular underlying discrete null hypothesis df. The unconditional (limiting) distribution, however, does depend on the discrete null hypothesis df. Tests based on $\bar{T}_N$ are therefore carried out as conditional tests. The asymptotic theory for $\bar{T}_N$ relies mainly on the (conditional) multivariate normality of the multinomial distribution. In Section 6.2 we study the conditional limiting distribution of $\bar{T}_N$, and Section 6.3 is devoted to the unconditional asymptotic distribution. The latter, of course, is of minor statistical importance for the reasons we have mentioned above.

The last two sections will be presented rather loosely and we content ourselves to show the existence of (conditional) normal limiting distributions, without explicit calculation of the parameters.

6.2. THE CONDITIONAL LIMITING DISTRIBUTION

The weak convergence in (6.1.7) may be obtained as a corollary to the local limit theorem for multinomial densities (see e.g. Gnedenko [45] or Morgenstern [46]). Let us take this local limit theorem as a starting point and add the boundary conditions $\sum_{j=1}^{r} \nu_{ij} = \nu_{i.}$, $\sum_{i=1}^{q} \nu_{ij} = \nu_{.j}$ ($i=1,\dots,q$ and $j=1,\dots,r$) to the condition $\sum_{i=1}^{q} \sum_{j=1}^{r} \nu_{ij} = N$. Among these conditions $q + r - 1$ are independent. Conditional on the sizes of the marginal ties $\nu_{1.},\dots,\nu_{q.}$, $\nu_{.1},\dots,\nu_{.r}$ it may be seen from Stirling's formula that the asymptotic distribution of

$$(6.2.1) \qquad N^{1/2}(N^{-1}\nu_{11}-m_{N,11},\dots,N^{-1}\nu_{1r}-m_{N,1r},\dots,\\ N^{-1}\nu_{q1}-m_{N,q1},\dots,N^{-1}\nu_{qr}-m_{N,qr})$$

is

$$(6.2.2) \qquad \bar{N}(0,S_N),$$

as $N \to \infty$, provided

$$(6.2.3) \qquad N^{1/2}|N^{-1}\nu_{i.}-\pi_{i.}| \le C,\ N^{1/2}|N^{-1}\nu_{.j}-\pi_{.j}| \le C,$$

for $i = 1,\dots,q$ and $j = 1,\dots,r$. Here $m_N = \{m_{N,ij}\}$ is a $q \times r$-dimensional vector and $\bar{N}(0,S_N)$ is a degenerate $q \times r$-dimensional normal distribution with mean vector 0 and covariance matrix $S_N = \{S_{N,ij,i'j'}\}$, having at most rank $q \times r - q - r + 1$ because of the boundary conditions mentioned above.

Furthermore $0 < C < \infty$ is an arbitrary fixed constant. Let us note that, given an arbitrary $\varepsilon > 0$, the constant $C = C_\varepsilon$ may be chosen sufficiently large to ensure that

$$P(\{N^{1/2}|N^{-1}\nu_{i.}-\pi_{i.}| \leq C,\ N^{1/2}|N^{-1}\nu_{.j}-\pi_{.j}| \leq C,$$

$$i=1,\ldots,q \text{ and } j=1,\ldots,r\}) \geq 1-\varepsilon.$$

In general m_N and S_N will not only depend on $\nu_{1.},\ldots,\nu_{q.},\nu_{.1},\ldots,\nu_{.r}$ but also on the cell probabilities $\pi_{11},\ldots,\pi_{1r},\ldots,\pi_{q1},\ldots,\pi_{qr}$. It has been shown in [46] that *under the null hypothesis*, i.e. when $\pi_{ij} = \pi_{i.}\pi_{.j}$ for all i, j, we have mean vector m_{N0} and covariance matrix S_{N0} depending on $\nu_{1.},\ldots,\nu_{q.},\nu_{.1},\ldots,\nu_{.r}$ but no longer on the π_{ij}.

A suitable standardization for $\bar{T}_N$ will be $N^{1/2}(\bar{T}_N-\bar{\mu}_N)$. Formal expressions for the limiting mean and variance are

$$(6.2.4) \qquad \bar{\mu}_N = \bar{\mu}_N(H) = \textstyle\sum_{i=1}^q \sum_{j=1}^r \kappa_{Ni}\lambda_{Nj}m_{N,ij},$$

$$(6.2.5) \qquad \bar{\sigma}_N^2 = \bar{\sigma}_N^2(H) = \textstyle\sum_{i=1}^q \sum_{j=1}^r \sum_{i'=1}^q \sum_{j'=1}^r \kappa_{Ni}\lambda_{Nj}\kappa_{Ni'}\lambda_{Nj'}S_{N,ij,i'j'}.$$

It is not hard to see that, if condition (6.2.3) is fulfilled, it follows that $N^{-1}\nu_{i.} \to \pi_{i.}$, $N^{-1}\nu_{.j} \to \pi_{.j}$ and that $\kappa_{Ni} \to \kappa_i$, $\lambda_{Nj} \to \lambda_j$ as $N \to \infty$.

<u>THEOREM 6.2.1</u>. Suppose that the limiting score function $J = K \times L$ satisfies Assumption 5.1.1 with arbitrary but fixed $a_1,\ldots,a_k$, $b_1,\ldots,b_l$ and $0 < \alpha < 1/2$, $0 < \beta < 1/2$. Let $\bar{J}_{(N)} = \bar{K}_{(N)} \times \bar{L}_{(N)}$ be the averaged score function, defined in (6.1.15) and (6.1.16). Then, conditional on the sizes of the marginal ties $\nu_{1.},\ldots,\nu_{q.},\ \nu_{.1},\ldots,\nu_{.r}$, the asymptotic distribution of $N^{1/2}(\bar{T}_N-\bar{\mu}_N)$ is $N(0,\bar{\sigma}_N^2)$ as $N \to \infty$, provided (6.2.3) is satisfied. For $\bar{T}_N$ see (6.1.17). Here the $\bar{\mu}_N = \bar{\mu}_N(H)$ and $\bar{\sigma}_N^2 = \bar{\sigma}_N^2(H)$, given by (6.2.4) and (6.2.5) respectively, are all finite. They do not depend on the π_{ij} provided $\pi_{ij} = \pi_{i.}\pi_{.j}$ for all i, j, i.e. in case of independence.

<u>PROOF</u>. Recalling the properties of R_{nN}, Q_{nN} in (6.1.6) and the definition of the set of indices I_{ij} in (6.1.4) we find that

(6.2.6) $$\bar{T}_N = N^{-1} \sum_{i=1}^{q} \sum_{j=1}^{r} \sum_{n\in I_{ij}} \bar{K}_{(N)}(R_{nN}/(N+1))\bar{L}_{(N)}(Q_{nN}/(N+1))$$

$$= N^{-1} \sum_{i=1}^{q} \sum_{j=1}^{r} \nu_{ij}\kappa_{Ni}\lambda_{Nj}.$$

Hence $N^{1/2}(\bar{T}_N-\bar{\mu}_N) = N^{1/2}(\nu_{ij}-m_{N,ij})\kappa_{Ni}\lambda_{Nj}$, from which the conditional weak convergence follows at once in view of (6.2.1), (6.2.2). Because m_{N0} and S_{N0} do not depend on the particular underlying discrete null hypothesis df the same may be said of the conditional asymptotic normality established in the theorem. □

There is an other way to compute directly sequences of numbers $\bar{\mu}_{N0}$ and $\bar{\sigma}^2_{N0}$ that may be used as asymptotic mean and variance under the null hypothesis. It is not hard to see that Theorem 3.1.1 continues to hold in some cases where the so called limiting score function depends on N, provided the underlying df is sufficiently smooth. In particular Theorem 5.1.2 (C) may be used in such cases, because the underlying null hypothesis df satisfies all kinds of relevant smoothness conditions. Let us next consider the sizes of the marginal ties as given numbers for any N. Then $\bar{J}_{(N)} = \bar{K}_{(N)} \times \bar{L}_{(N)}$ is a non-random limiting score function depending on N, where $\bar{K}_{(N)}(\bar{L}_{(N)})$ is a simple step function assuming at most q(r) different values. The points where the jumps take place converge properly to points in (0,1) and the heights of the jumps remain bounded under condition (6.2.3).

Using the randomized ranks, introduced in (6.1.9), we may write $\bar{T}_N$ equivalently as

$$\bar{T}_N = \sum_{n=1}^{N} \bar{K}_{(N)}(R^*_{nN}/(N+1))\bar{L}_{(N)}(Q^*_{nN}/(N+1)).$$

Moreover, conditionally given the $\nu_{i.}$ and the $\nu_{.j}$, the vector of ranks $(R^*_{1N},\ldots,R^*_{NN})$ remains independent of the vector of ranks $(Q^*_{1N},\ldots,Q^*_{NN})$, and still each of these vectors assumes any permutation of the numbers $1,\ldots,N$ with equal probability $1/N!$. Hence the modification of Theorem 5.1.2 (C) applies and yields asymptotic normality, to be interpreted as conditional asymptotic normality given the sizes of the marginal ties.

In this modified form the asymptotic null hypothesis mean and variance are

(6.2.7) $$\bar{\mu}_{N0} = \int_0^1\int_0^1 \bar{K}_{(N)}(u)\bar{L}_{(N)}(v)\,du\,dv,$$

$$(6.2.8) \qquad \bar{\sigma}^2_{N0} = \int_0^1\int_0^1[\bar{\mu}_{N0}+\bar{K}_{(N)}(u)\bar{L}_{(N)}(v)-\bar{K}_{(N)}(u)\int_0^1 \bar{L}_{(N)}(t)dt - \bar{L}_{(N)}(v)\int_0^1 \bar{K}_{(N)}(s)ds]^2 dudv.$$

It should be noted that, by substitution of $(N+1)^{-2}\nu_{i.}\nu_{.j}$ for $m_{N,ij}$ and of $(N+1)^{-2}\nu_{i.}\nu_{.j}[\delta_{ii'}\delta_{jj'}-(N+1)^{-2}\nu_{i'.}\nu_{.j'}]$ for $S_{N,ij,i'j'}$, formula (6.2.4) yields $\bar{\mu}_{N0}$ as given in (6.2.7) and formula (6.2.5) yields $\bar{\sigma}^2_{N0}$ as given in (6.2.8).

6.3. THE UNCONDITIONAL LIMITING DISTRIBUTION

For the unconditional limiting distribution we refer directly to (6.1.7). A suitable standardization for the statistic in (6.1.17) will be $N^{1/2}(\bar{T}_N-\bar{\mu})$, where the limiting mean depends on the underlying discrete df H and equals

$$(6.3.1) \qquad \bar{\mu} = \bar{\mu}(H) = \textstyle\sum_{i=1}^q \sum_{j=1}^r \kappa_i\lambda_j\pi_{ij}.$$

Let us note that $\bar{\mu} = \mu^*$ (see (6.1.13)), where H^* is the continuous df corresponding to the discrete df H in the way described in Section 6.1. The asymptotic variance, which also depends on H, is equal to

$$(6.3.2) \qquad \bar{\sigma}^2 = \bar{\sigma}^2(H) = \textstyle\sum_{i=1}^q \sum_{j=1}^r \sum_{i'=1}^q \sum_{j'=1}^r c_{ij}c_{i'j'}\pi_{ij}(\delta_{ii'}\delta_{jj'}-\pi_{i'j'}),$$

where the numbers c_{ij} depend on $J = K \times L$ and H in a way which will be made clear in the proof of Theorem 6.3.1.

<u>THEOREM 6.3.1</u>. Suppose that the limiting score function $J = K \times L$ satisfies Assumption 5.1.1 with arbitrary fixed $a_1,\ldots,a_k$, $b_1,\ldots,b_l$ and $0 < \alpha < 1/2$, $0 < \beta < 1/2$. Let H be any fixed discrete underlying df, satisfying (6.1.1) and $\{s_1,\ldots,s_k\} \cap \{\pi_{1.},\ldots,\sum_{i=1}^{q-1}\pi_{i.}\} = \{t_1,\ldots,t_l\} \cap \{\pi_{.1},\ldots,\sum_{j=1}^{r-1}\pi_{.j}\} = \emptyset$. Then $N^{1/2}(\bar{T}_N-\bar{\mu}) \to_d N(0,\bar{\sigma}^2)$. Here $\bar{T}_N$ is based on the averaged score function (see (6.1.17)) and $\bar{\mu} = \bar{\mu}(H)$ and $\bar{\sigma}^2 = \bar{\sigma}^2(H)$, given by (6.3.1) and (6.3.2) respectively, are finite.

The remaining part of this section is devoted to the proof of this theorem. Although $\bar{\mu}$ does not depend on the particular discrete underlying df H for which $\pi_{ij} = \pi_{i.}\pi_{.j}$ (i.e. for which X and Y are independent), the limiting variance $\bar{\sigma}^2$ depends on the particular values of the $\pi_{i.}$ and the $\pi_{.j}$.

In order to write $N^{1/2}(\bar{T}_N-\bar{\mu})$ in a more convenient form, let us introduce for brevity the random variables

$$(6.3.3)\qquad \begin{aligned} \tilde{\kappa}_{Ni} &= N(\nu_{i.})^{-1} \int_{(\nu_{0.}+\dots+\nu_{i-1.})/N}^{(\nu_{0.}+\dots+\nu_{i.})/N} K(s)ds, \\ \tilde{\lambda}_{Nj} &= N(\nu_{.j})^{-1} \int_{(\nu_{.0}+\dots+\nu_{.j-1})/N}^{(\nu_{.0}+\dots+\nu_{.j})/N} L(t)dt. \end{aligned}$$

Using (6.2.6) it follows that

$$(6.3.4)\qquad \begin{aligned} N^{1/2}(\bar{T}_N-\bar{\mu}) &= N^{1/2} \textstyle\sum_{i=1}^q \sum_{j=1}^r (N^{-1}\nu_{ij}\kappa_{Ni}\lambda_{Nj}-\pi_{ij}\kappa_i\lambda_j) \\ &= \textstyle\sum_{i=0}^2 A_{iN} + \sum_{i=1}^6 B_{iN}, \end{aligned}$$

where the κ_i, λ_j are defined in (6.1.12), the κ_{Ni}, λ_{Nj} in (6.1.15), (6.1.16) and where

$$A_{0N} = N^{1/2} \textstyle\sum_{i=1}^q \sum_{j=1}^r (N^{-1}\nu_{ij}-\pi_{ij})\kappa_i\lambda_j,$$

$$A_{1N} = N^{1/2} \textstyle\sum_{i=1}^q \sum_{j=1}^r \pi_{ij}(\tilde{\kappa}_{Ni}-\kappa_i)\lambda_j,$$

$$A_{2N} = N^{1/2} \textstyle\sum_{i=1}^q \sum_{j=1}^r \pi_{ij}(\tilde{\lambda}_{Nj}-\lambda_j)\kappa_i,$$

$$B_{1N} = N^{1/2} \textstyle\sum_{i=1}^q \sum_{j=1}^r (N^{-1}\nu_{ij}-\pi_{ij})(\tilde{\kappa}_{Ni}-\kappa_i)\lambda_j,$$

$$B_{2N} = N^{1/2} \textstyle\sum_{i=1}^q \sum_{j=1}^r (N^{-1}\nu_{ij}-\pi_{ij})(\tilde{\lambda}_{Nj}-\lambda_j)\kappa_i,$$

$$B_{3N} = N^{1/2} \textstyle\sum_{i=1}^q \sum_{j=1}^r N^{-1}\nu_{ij}(\tilde{\kappa}_{Ni}-\kappa_i)(\tilde{\lambda}_{Nj}-\lambda_j),$$

$$B_{4N} = N^{1/2} \textstyle\sum_{i=1}^q \sum_{j=1}^r N^{-1}\nu_{ij}(\kappa_{Ni}-\tilde{\kappa}_{Ni})\tilde{\lambda}_{Nj},$$

$$B_{5N} = N^{1/2} \textstyle\sum_{i=1}^q \sum_{j=1}^r N^{-1}\nu_{ij}\tilde{\kappa}_{Ni}(\lambda_{Nj}-\tilde{\lambda}_{Nj}),$$

$$B_{6N} = N^{1/2} \textstyle\sum_{i=1}^q \sum_{j=1}^r N^{-1}\nu_{ij}(\kappa_{Ni}-\tilde{\kappa}_{Ni})(\lambda_{Nj}-\tilde{\lambda}_{Nj}).$$

Our first step concerns the asymptotic behavior of the A-terms. The asymptotic normality of the A_{0N} is immediate from (6.1.7). The A_{1N} and A_{2N} will also be written as linear combinations of the $(\nu_{ij}/N)-\pi_{ij}$ and lower order terms. By symmetry we need only consider A_{1N} in detail. Let us observe that, because

(6.3.5) $$\tilde{\kappa}_{Ni} \to_p \kappa_i,$$

$$N^{1/2}(N^{-1}\nu_{i.}-\pi_{i.}) \to_d N(0,\pi_{i.}(1-\pi_{i.})),$$

as $N \to \infty$, we may write

(6.3.6) $$N^{1/2}(\tilde{\kappa}_{Ni}-\kappa_i)$$

$$= N^{1/2}\left[\int_{(\nu_{0.}+\dots+\nu_{i-1.})/N}^{(\nu_{0.}+\dots+\nu_{i.})/N} K(s)ds - \int_{\pi_{0.}+\dots+\pi_{i-1.}}^{\pi_{0.}+\dots+\pi_{i.}} K(s)ds\right]\pi_{i.}^{-1} -$$

$$N^{1/2}[N^{-1}\nu_{i.}-\pi_{i.}]\kappa_i\pi_{i.}^{-1} + o_p(1).$$

The next lemma will be helpful.

LEMMA 6.3.1. Let Z be a $\mathcal{Bi}(N,\pi)$ distributed rv, $0 < \pi < 1$, and let the function ϕ, defined on (0,1), be such that $\int_0^1|\phi(u)|du < \infty$. Suppose that ϕ is continuously differentiable on an interval $(\pi-\eta,\pi+\eta) \subset (0,1)$ for some small $\eta > 0$. Then we have, as $N \to \infty$,

$$\int_\pi^{Z/N}\phi(u)du = \phi(\pi)(N^{-1}Z-\pi) + O_p(N^{-1}).$$

PROOF. For any $\varepsilon > 0$, there exists an index $N_1 = N_{1\varepsilon}$ such that

(6.3.7) $$\Pr(|N^{-1}Z-\pi| \le \eta/3) \ge 1-\varepsilon/2$$

for all $N \ge N_1$. By the mean value theorem we have, for $|(Z/N)-\pi| < \eta$,

$$\int_\pi^{Z/N}\phi(u)du = \phi(\pi)(N^{-1}Z-\pi) + \int_\pi^{Z/N}(u-\pi)\phi'(\tilde{u})du,$$

where $\tilde{u}$ lies in the open random interval between the points π and Z/N.

Because ϕ' is continuous on $(\pi-\eta,\pi+\eta)$ there exists a constant $0 < \tilde{M} < \infty$ such that $\max_{[\pi-\eta/2,\pi+\eta/2]}|\phi'| \leq \tilde{M}$. Hence, for $|(Z/N)-\pi| \leq \eta/2$,

$$(6.3.8) \qquad |N^{-1}Z-\pi|\int_{\pi}^{Z/N}|\phi'(\tilde{u})|\,du \leq \tilde{M}(N^{-1}Z-\pi)^2.$$

Since $N^{1/2}[(Z/N)-\pi]$ has a normal distribution in the limit it follows that the expression in (6.3.8) is $O_P(N^{-1})$. Together with (6.3.7) this implies the statement of the lemma. □

Because $\{s_1,\ldots,s_k\} \cap \{\pi_{1.},\ldots,\sum_{i=1}^{q-1}\pi_{i.}\} = \emptyset$ and

$$\int_{(\nu_{0.}+\ldots+\nu_{i-1.})/N}^{(\nu_{0.}+\ldots+\nu_{i.})/N} K(s)ds - \int_{\pi_{0.}+\ldots+\pi_{i-1.}}^{\pi_{0.}+\ldots+\pi_{i.}} K(s)ds$$

$$= \int_{\pi_{0.}+\ldots+\pi_{i.}}^{(\nu_{0.}+\ldots+\nu_{i.})/N} K(s)ds - \int_{\pi_{0.}+\ldots+\pi_{i-1.}}^{(\nu_{0.}+\ldots+\nu_{i-1.})/N} K(s)ds,$$

Lemma 6.3.1 applies. From application of this lemma it follows easily that $N^{1/2}(\tilde{\kappa}_{Ni}-\kappa_i)\pi_{ij}\lambda_j$ is a sum of a linear combination of $N^{1/2}(N^{-1}\nu_{1.}-\pi_{1.}),\ldots,$ $N^{1/2}(N^{-1}\nu_{i.}-\pi_{i.})$ and a term of order $o_P(1)$. Consequently A_{1N} is a sum of a linear combination of the $N^{1/2}(N^{-1}\nu_{i.}-\pi_{i.})$, and hence of the $N^{1/2}(N^{-1}\nu_{ij}-\pi_{ij})$, and a term of order $o_P(1)$. The same may be said of A_{2N}, so that

$$(6.3.9) \qquad \sum_{i=0}^{2} A_{iN} = \sum_{i=1}^{q} \sum_{j=1}^{r} c_{ij}N^{1/2}(N^{-1}\nu_{ij}-\pi_{ij}).$$

Here

$$-\infty < c_{ij} = c_{ij}(J,H) < \infty,$$

$i = 1,\ldots,q$ and $j = 1,\ldots,r$, are finite constants depending on the limiting score function $J = K \times L$ and the underlying discrete df H only. The actual values of the c_{ij} follow by straightforward but tedious computation. From (6.1.7), (6.1.8) and (6.3.9) it follows that

$\sum_{i=0}^{2} A_{iN} \to_d N(0,\bar{\sigma}^2(H))$ as $N \to \infty$, for the fixed discrete underlying df H, with $\bar{\sigma}^2 = \bar{\sigma}^2(H)$ given in (6.3.2). See e.g. [32], page 108 and 371, or [42], page 54.

Our next step is to prove the asymptotic negligibility of the B-terms. By symmetry we need only consider B_{1N}, B_{3N}, B_{4N}, B_{6N}.

LEMMA 6.3.2. Let $0 = u_0 < u_1 < \ldots < u_k < u_{k+1} = 1$ and let ϕ be defined on $(0,1)$ and continuously differentiable on $\cup_{j=1}^{k+1} (u_{j-1},u_j)$, such that $|\phi^{(i)}| \leq DR^{\alpha+i}$ where defined on $(0,1)$, for $i = 0,1$. Here $0 < D < \infty$ and $0 < \alpha < 1/2$ are fixed constants and R is the function defined in (1.3.8). Let, for $j = 1,2$, Z_j be a $\mathcal{B}i(N,\pi_j)$ distributed rv with $0 \leq \pi_1 < 1$, $0 < \pi_2 \leq 1$, defined on the same probability space such that $Pr(0 \leq Z_1+Z_2 \leq N) = 1$. Defining $0/0 = 0$ we have, as $N \to \infty$,

$$(Z_2)^{-1} \sum_{n=Z_1+1}^{Z_1+Z_2} \phi(n/(N+1))-N(Z_2)^{-1} \int_{Z_1/N}^{(Z_1+Z_2)/N} \phi(u)du = O_P(N^{\alpha-1}).$$

PROOF. Let us first suppose that ϕ is continuously differentiable throughout $(0,1)$. By the mean value theorem we have

$$\phi(u) = \phi(n/(N+1)) + (u-n/(N+1))\phi'(u_{n,N}),$$

for $(n-1)/N < u < n/N$, where $u_{n,N}$ is a point between u and $n/(N+1)$. Hence

$$N(Z_2)^{-1} \int_{Z_1/N}^{(Z_1+Z_2)/N} \phi(u)du = (Z_2)^{-1} \sum_{n=Z_1+1}^{Z_1+Z_2} \phi(n/(N+1)) +$$

$$N(Z_2)^{-1} \sum_{n=Z_1+1}^{Z_1+Z_2} \int_{(n-1)/N}^{n/N} (u-n/(N+1))\phi'(u_{n,N})du.$$

Consequently the difference at the left in the expression of the lemma is in absolute value bounded by

$$D(Z_2)^{-1} \sum_{n=Z_1+1}^{Z_1+Z_2} \int_{(n-1)/N}^{n/N} [R(u_{n,N})]^{\alpha+1}du$$

$$= O_P(N^{-1}) \times O(N^{-1} \sum_{n=1}^{N-1} [R(n/N)]^{\alpha+1})$$

$$= O_P(N^{\alpha-1}).$$

The case where ϕ is of the more general form as described in the lemma can be dealt with in a similar fashion. We only have to add the fixed points $u_1,\dots,u_k$ to each partition with mesh $(1/N)$, and apply the mean value theorem for all open sub-intervals of this modified partition. □

Note that application of the lemma with $\pi_1 = 0$ and $\pi_2 = 1$ yields the error of the non-random Riemann-sum approximation on page 76.

The asymptotic negligibility of B_{1N}, as $N \to \infty$, follows at once from the fact that

$$N^{1/2}(N^{-1}\nu_{ij}-\pi_{ij}) \to_d N(0,\pi_{ij}(1-\pi_{ij})),$$

as $N \to \infty$, and from (6.3.5).

As to B_{3N}, note that $N^{-1}\nu_{ij} = O_P(1)$ and that $N^{1/2}(\tilde{\kappa}_{Ni}-\kappa_i)$ has a normal distribution in the limit, which has been shown when dealing with A_{1N}. Because $\tilde{\lambda}_{Nj} \to_P \lambda_j$, as $N \to \infty$, it follows that $B_{3N} = o_P(1)$, as $N \to \infty$.

Because $\kappa_{Ni}-\kappa_{Ni} = O_P(N^{\alpha-1})$ by Lemma 6.3.2 and since $N^{-1}\nu_{ij}\lambda_{Nj} = O_P(1)$, it follows that $B_{4N} = O_P(N^{\alpha-1/2})$ as $N \to \infty$.

Finally it is not hard to see that by Lemma 6.3.2 we have $B_{6N} = O_P(N^{\alpha+\beta-3/2})$, as $N \to \infty$. This concludes the proof of the theorem.

REFERENCES

[1] BAHADUR, R.R. (1966). A note on quantiles in large samples. *Ann. Math. Statist.* *37*, 577-580.

[2] BEHNEN, K. (1971). Asymptotic optimality and ARE of certain rank-order tests under contiguity. *Ann. Math. Statist.* *42*, 325-329.

[3] BEHNEN, K. (1972). A characterization of certain rank-order tests with bounds for the asymptotic relative efficiency. *Ann. Math. Statist.* *43*, 1839-1851.

[4] BHUCHONGKUL, S. (1964). A class of nonparametric tests for independence in bivariate populations. *Ann. Math. Statist.* *35*, 138-149.

[5] BILLINGSLEY, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[6] BLUM, J.R., J. KIEFER, and M. ROSENBLATT (1961). Distribution free tests of independence based on the sample distribution function. *Ann. Math. Statist.* *32*, 485-498.

[7] CHERNOFF, H. and I.R. SAVAGE (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Statist.* *29*, 972-994.

[8] DUPAČ, V. and J. HÁJEK (1969). Asymptotic normality of simple linear rank statistics under alternatives II. *Ann. Math. Statist.* *40*, 1992-2017.

[9] DURBIN, J. (1970). Asymptotic distributions of some statistics based on the bivariate sample distribution function. *Nonparametric Techniques in Statistical Inference* (M.L. PURI editor), 435-449. Cambridge Un. Press.

[10] FELLER, W. (1957). *An Introduction to Probability Theory and its Applications 1*, 2nd edition. Wiley, New York.

[11] FELLER, W. (1966). *An Introduction to Probability Theory and its Applications 2*. Wiley, New York.

[12] GHOSH, J.K. (1971). A new proof of the Bahadur representation of quantiles and an application. *Ann. Math. Statist.* *42*, 1957-1961.

[13] GUMBEL, E.J. (1958). Multivariate distributions with given margins and analytical examples. *Bull. de l'Inst. Int. de Statist.* *37-3*, 363-373.

[14] HÁJEK, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Ann. Math. Statist.* *39*, 325-346.

[15] HÁJEK, J. (1969). *A Course in Nonparametric Statistics*. Holden-Day, San Francisco.

[16] HÁJEK, J. (1970). Miscellaneous problems of rank test theory. *Nonparametric Techniques in Statistical Inference* (M.L. PURI editor), 3-17. Cambridge Un. Press.

[17] HÁJEK, J. and Z. ŠIDÁK (1967). *Theory of Rank Tests*. Academic Press, New York.

[18] HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Ass. 58*, 13-30.

[19] JOGDEO, K. (1968). Asymptotic normality in nonparametric methods. *Ann. Math. Statist. 39*, 905-922.

[20] KENDALL, M.G. (1948). *Rank Correlation Mehtods*. Griffin, London.

[21] KIEFER, J. (1961). On large deviations of the empiric d.f. of vector chance variables and a law of the iterated logarithm. *Pacific J. Math. 11*, 649-660.

[22] KONIJN, H.S. (1956). On the power of certain tests for independence in bivariate populations. *Ann. Math. Statist. 27*, 300-323.

[23] KONIJN, H.S. (1959). Positive and negative dependence of two random variables. *Sankhyā A-21*, 269-280.

[24] KRAUTH, J. (1971). A locally most powerful tied rank test in a Wilcoxon situation. *Ann. Math. Statist. 42*, 1949-1956.

[25] LEHMANN, E.L. (1953). The power of rank tests. *Ann. Math. Statist. 24*, 23-43.

[26] LEHMANN, E.L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

[27] LEHMANN, E.L. (1966). Some concepts of dependence. *Ann. Math. Statist. 37*, 1137-1153.

[28] MARDIA, K.V. (1970). *Families of Bivariate Distributions*. Griffin, London.

[29] NEUHAUS, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist. 42*, 1285-1295.

[30] PURI, M.L. and P.K. SEN (1971). *Nonparametric Methods in Multivariate Analysis*. Wiley, New York.

[31] PYKE, R. and G.R. SHORACK (1968). Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage theorems. *Ann. Math. Statist. 39*, 755-771.

[32] RAO, C.R. (1965). *Linear Statistical Inference and its Applications*. Wiley, New York.

[33] RUYMGAART, F.H. (1972). Asymptotic normality of nonparametric tests for independence. To appear in *Ann. Statist.*.

[34] RUYMGAART, F.H., G.R. SHORACK and W.R. VAN ZWET (1972). Asymptotic normality of nonparametric tests for independence. *Ann. Math. Statist. 43*, 1122-1135.

[35] SCHMETTERER, L. (1966). *Einführung in die Mathematische Statistik*, 2nd edition. Springer-Verlag, Wien.

[36] SEN, P.K. (1968). Asymptotic normality of sample quantiles for m-dependent processes. *Ann. Math. Statist. 39*, 1724-1730.

[37] SEN, P.K. (1970). Asymptotic distribution of a class of multivariate rank order statistics. *Calcutta Statist. Ass. 19*, 23-31.

[38] SHORACK, G.R. (1970). A uniformly convergent empirical process. Technical Report No. 20 Math. Dept., University of Washington.

[39] SHORACK, G.R. (1972). Functions of order statistics. *Ann. Math. Statist. 43*, 412-427.

[40] SKOROKHOD, A.V. (1956). Limit theorems for stochastic processes. *Theory Prob. and Appl.* (translated by SIAM) *1*, 261-290.

[41] VORLIČKOVÁ, D. (1970). Asymptotic properties of rank tests under discrete distributions. *Z. Wahrscheinlichkeitsth. und Verw. Geb. 14*, 275-289.

[42] WITTING, H. and G. NÖLLE (1970). *Angewandte Mathematische Statistik*. Teubner, Stuttgart.

[43] YANAGIMOTO, T. (1970). On measures of association and a related problem. *Ann. Inst. Statist. Math. 22*, 57-63.

[44] VAN ZWET, W.R.. Personal communication.

[45] GNEDENKO, B.V. (1962). *The Theory of Probability*. Chelsea, New York.

[46] MORGENSTERN, D. (1964). *Einführung in die Wahrscheinlichkeitsrechnung und Mathematische Statistik*. Springer-Verlag, Berlin.

INDEX