The reader may find it useful to keep this alternative interpretation in mind throughout the present chapter.

1.2. A special central limit theorem. Let A be the set of all non-vanishing real vectors $a = (a_1, \ldots, a_N)$ of all finite dimensions $N \geqq 1$. We shall say that the statistics T_a, $a \in A$, are asymptotically normal (μ_a, σ_a^2) for

$$\sum_{i=1}^{N} a_i^2 / \max_{1 \leqq i \leqq N} a_i^2 \to \infty\,, \tag{1}$$

if (1) entails

$$F(T_a \leqq \mu_a + x\sigma_a) \to (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{x} \exp\left(-\tfrac{1}{2}y^2\right) \mathrm{d}y\,, \quad -\infty < x < \infty\,. \tag{2}$$

Thus asymptotic normality (μ_a, σ_a^2) is equivalent to convergence in distribution of $(T_a - \mu_a)/\sigma_a$ to a standardized normal random variable. Obviously, asymptotic normality (μ_a, σ_a^2) is equivalent to the same property with (μ_a^*, σ_a^{*2}), if

$$\sigma_a^*/\sigma_a \to 1\,, \quad (\mu_a^* - \mu_a)/\sigma_a \to 0\,. \tag{3}$$

Theorem. *Let $Y_1, Y_2, \ldots$ be independent copies of a random variables with finite expectation μ and finite variance σ^2. Put*

$$T_a = \sum_{i=1}^{N} a_i Y_i\,, \quad a \in A\,. \tag{4}$$

Then, for (1), the statistics T_a are asymptotically normal (μ_a, σ_a^2) with

$$\mu_a = \mu \sum_{i=1}^{N} a_i \tag{5}$$

and

$$\sigma_a^2 = \sigma^2 \sum_{i=1}^{N} a_i^2\,. \tag{6}$$

Proof. The Lindeberg condition (LOÈVE (1955), p. 295) takes on the form

$$\sigma_a^{-2} \sum_{i=1}^{N} \int_{|x| > \varepsilon\sigma_a} x^2 \,\mathrm{d}P(a_i(Y_i - \mu) \leqq x) \to 0\,, \tag{7}$$

where σ_a^2 is given by (6). Upon substituting $a_i y$ for x, we obtain

$$\int_{|x| > \varepsilon\sigma_a} x^2 \,\mathrm{d}P(a_i(Y_i - \mu) \leqq x) = a_i^2 \int_{|ya_i| > \varepsilon\sigma_a} y^2 \,\mathrm{d}P(Y_i - \mu \leqq y) \leqq \tag{8}$$

$$\leqq a_i^2 \int_{|y| > \varepsilon\sigma v_a} y^2 \,\mathrm{d}P(Y_i - \mu \leqq y)$$

where

$$v_a^2 = \sum_{i=1}^{N} a_i^2 / \max_{1 \leq i \leq N} a_i^2 .$$

Consequently, the Y_i's having the same distribution,

(9) $$\sigma_a^{-2} \sum_{i=1}^{N} \int_{|x| > \varepsilon\sigma_a} x^2 \, dP(a_i(Y_i - \mu) \leqq x) \leqq \sigma^{-2} \int_{|y| > \varepsilon\sigma v_a} y^2 \, dP(Y_1 - \mu \leqq y) .$$

However, the variance of Y_1 is supposed finite and $v_a \to \infty$ in view of (1), so that

$$\sigma^{-2} \int_{|y| > \varepsilon v \sigma_a} y^2 \, dP(Y_1 - \mu \leqq y) \to 0 , \quad \varepsilon > 0 ,$$

which, in accordance with (9), entails (7). Q.E.D.

Remark. The above theorem could be reformulated as follows: For every $\varepsilon > 0$ there exists an n_0 such that

(10) $$\sum_{i=1}^{N} a_i^2 / \max_{1 \leq i \leq N} a_i^2 > n_0$$

entails

(11) $$\sup_x \left| P\left(\sum_{i=1}^{N} a_i Y_i \leqq \mu_a + x\sigma_a\right) - \Phi(x) \right| < \varepsilon ,$$

where Φ denotes the standardized normal distribution function.

1.3. A convergence theorem

Theorem. *Let $(\Omega, \mathscr{A}, \mu)$ be a measure space with a σ-finite measure μ. Consider a sequence $\{h_\nu\}$ of square integrable functions converging almost everywhere to a square integrable function h. Assume that*

(1) $$\limsup_{\nu \to \infty} \int h_\nu^2 \, d\mu \leqq \int h^2 \, d\mu .$$

Then

(2) $$\lim_{\nu \to \infty} \int (h_\nu - h)^2 \, d\mu = 0 .$$

Proof. Fatou's lemma together with (1) implies

(3) $$\lim_{\nu \to \infty} \int h_\nu^2 \, d\mu = \int h^2 \, d\mu .$$

Furthermore, the Schwartz inequality yields

$$\int |h_\nu h| \, d\mu \leqq \left[\int h_\nu^2 \, d\mu \int h^2 \, d\mu \right]^{\frac{1}{2}} \tag{4}$$

so that

$$\limsup_{\nu \to \infty} \int |h_\nu h| \, d\mu \leqq \int h^2 \, d\mu .$$

Consequently, according to Theorem II.4.2

$$\lim_{\nu \to \infty} \int h_\nu h \, d\mu = \int h^2 \, d\mu . \tag{5}$$

Now (3) and (5) imply (2).

1.4. Further preliminaries. Consider a probability space $(\Omega, \mathscr{A}, P)$ and a sequence of sub σ-fields $\mathscr{F}_1 \subset \mathscr{F}_2 \subset \ldots \subset \mathscr{A}$. Denote by $\mathscr{F}_\infty$ the smallest σ-field containing the field $\bigcup_1^\infty \mathscr{F}_N$. For every event $A \in \mathscr{F}_\infty$ and every $\varepsilon > 0$ there exists an N and $A_N \in \mathscr{F}_N$ such that

$$P(A \div A_N) < \varepsilon , \tag{1}$$

where $\div$ denotes the symmetric difference. Actually, the assertion is trivially true for $A \in \bigcup_1^\infty \mathscr{F}_N$, and the events having this property obviously consist a σ-field. Denoting by I_A and I_{A_N} the respective indicators, (1) may be rewritten as follows:

$$\mathsf{E}(I_A - I_{A_N})^2 < \varepsilon . \tag{2}$$

If Y is a $\mathscr{F}_\infty$-measurable function such that $\mathsf{E}Y^2 < \infty$, then there exists for every N a $\mathscr{F}_N$-measurable random variable Y_N such that

$$\mathsf{E}(Y_N - Y)^2 \leqq \mathsf{E}(Y_N^* - Y)^2 \tag{3}$$

for any other $\mathscr{F}_N$-measurable random variable Y_N^*. It is well-known that this property is possessed by the conditional expectation with respect to $\mathscr{F}_N$,

$$Y_N = \mathsf{E}(Y \mid \mathscr{F}_N) . \tag{4}$$

If $\mathscr{F}_N$ is generated by a statistic T_N, then $\mathsf{E}(Y \mid \mathscr{F}_N) = \psi(T_N)$, where $\psi(t_N) = \mathsf{E}(Y \mid T_N = t_N)$. Now, if $I_A^N = \mathsf{E}(I_A \mid \mathscr{F}_N)$, then (2), (3) and (4) imply that

$$\mathsf{E}(I_A - I_A^N)^2 < \varepsilon \tag{5}$$

for N sufficiently large, and, consequently,

$$\lim_{N\to\infty} \mathsf{E}(I_A - I_A^N)^2 = 0\,. \tag{6}$$

Before generalizing the relation (6) to all random variables with finite variance, let us recall that

$$\mathsf{E}Y_N^2 \leqq \mathsf{E}Y^2 \tag{7}$$

for Y_N given by (4).

Lemma a. *Let Y be a $\mathscr{F}_\infty$-measurable random variable such that $\mathsf{E}Y^2 < \infty$, and let Y_N be given by (4). Then*

$$\lim_{N\to\infty} \mathsf{E}(Y_N - Y)^2 = 0 \tag{8}$$

and

$$\lim_{N\to\infty} \mathsf{E}Y_N^2 = \mathsf{E}Y^2\,. \tag{9}$$

Proof. Fix an $\varepsilon > 0$ and find a $\mathscr{F}_\infty$-measurable simple function $\sum_{i=1}^{n} c_i I_{A_i}$ such that

$$\mathsf{E}(Y - \sum_{i=1}^{n} c_i I_{A_i})^2 < \tfrac{1}{6}\varepsilon\,.$$

Denoting $I_{A_i}^N = \mathsf{E}(I_{A_i} \mid \mathscr{F}_N)$ and noting (7), we have

$$\begin{aligned}\mathsf{E}(Y_N - Y)^2 &\leqq 3\mathsf{E}(Y_N - \sum_{i=1}^{n} c_i I_{A_i}^N)^2 + \\ &+ 3\mathsf{E}(Y - \sum_{i=1}^{n} c_i I_{A_i})^2 + 3\mathsf{E}[\sum_{i=1}^{n} c_i (I_{A_i} - I_{A_i}^N)]^2 \leqq \\ &\leqq 6\mathsf{E}(Y - \sum_{i=1}^{n} c_i I_{A_i})^2 + 3\sum_{i=1}^{n} c_i^2 \sum_{i=1}^{n} \mathsf{E}(I_{A_i} - I_{A_i}^N)^2 < \\ &< \varepsilon + 3\sum_{i=1}^{n} c_i^2 \sum_{i=1}^{n} \mathsf{E}(I_{A_i} - I_{A_i}^N)^2\,.\end{aligned}$$

Since, in view of (6), the last sum converges to 0 as $N \to \infty$, we conclude that

$$\mathsf{E}(Y_N - Y)^2 < \varepsilon$$

for N sufficiently large. This proves (8).(9) follows from the well-known relation

$$\mathsf{E}(Y_N - Y)^2 = \mathsf{E}Y^2 - \mathsf{E}Y_N^2\,, \tag{10}$$

holding for any conditional expectation. Q.E.D.

Now let $U_1, U_2, \ldots$ be independent random variables, each uniformly distributed over $(0, 1)$. Let R_{Ni} denote the rank of U_i, $1 \leqq i \leqq N$, in the partial sequence $U_1, \ldots, U_N$. Let $\varphi(u)$, $0 < u < 1$, be some square integrable function

(11) $$\int_0^1 \varphi^2(u)\, \mathrm{d}u < \infty ,$$

and put

(12) $$a_N^{\varphi}(i) = \mathsf{E}[\varphi(U_1) \mid R_{N1} = i] , \quad 1 \leqq i \leqq N < \infty .$$

Theorem a. *Under assumption* (11),

(13) $$\lim_{N \to \infty} \mathsf{E}[a_N^{\varphi}(R_{N1}) - \varphi(U_1)]^2 = 0 ,$$

holds, where $a_N^{\varphi}(i)$ *is defined by* (12).

Proof. Let $\mathscr{F}_N$ be the sub σ-field generated by $(R_{N1}, \ldots, R_{NN})$. Note that $\mathscr{F}_N \subset \subset \mathscr{F}_{N+1} \subset \ldots$ and recall that $\mathscr{F}_\infty$ denotes the smallest σ-field containing $\bigcup_1^\infty \mathscr{F}_N$. We first show that $\varphi(U_1)$ is equivalent to a $\mathscr{F}_\infty$-measurable random variable. In view of (II.1.2.12), we have

$$\mathsf{E}\left(U_1 - \frac{R_{N1}}{N+1}\right)^2 = \frac{1}{N} \sum_{j=1}^{N} \mathsf{E}\left[\left(U_1 - \frac{j}{N+1}\right)^2 \,\middle|\, R_{N1} = j\right] =$$

$$= \frac{1}{N} \sum_{j=1}^{N} \operatorname{var} U_N^{(j)} = \frac{1}{N} \sum_{j=1}^{N} \frac{j(N-j+1)}{(N+1)^2 (N+2)} < \frac{1}{N} ,$$

so that

$$\lim_{\nu \to \infty} \frac{R_{N_\nu 1}}{N_\nu + 1} = U_1$$

with probability 1 for some properly chosen subsequence $\{N_\nu\}$. Consequently U_1, and hence also $\varphi(U_1)$, is equivalent to a $\mathscr{F}_\infty$-measurable random variable. Now it remains to apply the above lemma with $\varphi(U_1) \equiv Y$ and $a_N^{\varphi}(R_{N1}) \equiv Y_N$. The proof is thus concluded.

Lemma b. (D.K. Faddeev.) *Let the functions* $f_N(t, u)$, $N \geqq 1$, $0 < t$, $u < 1$, *be densities in* t *for each fixed* u, *such that for every* $\varepsilon > 0$

(14) $$\lim_{N \to \infty} \int_{u-\varepsilon}^{u+\varepsilon} f_N(t, u)\, \mathrm{d}t = 1 , \quad 0 < u < 1 .$$

Moreover, assume that

$$f_N(t, u) \leqq g_N(t, u), \quad N \geqq 1, 0 < t, u < 1, \tag{15}$$

where the functions $g_N(t, u)$ *are increasing in* $t \in (0, u)$ *and decreasing in* $t \in (u, 1)$ *for every fixed* $N \geqq 1$ *and* $0 < u < 1$, *and*

$$\sup_N \int_0^1 g_N(t, u)\, \mathrm{d}t < \infty, \quad 0 < u < 1. \tag{16}$$

Then for every integrable function $\varphi(u)$

$$\lim_{N\to\infty} \int_0^1 \varphi(t) f_N(t, u)\, \mathrm{d}t = \varphi(u) \tag{17}$$

in almost all points $u \in (0, 1)$.

Proof. See I. P. NATANSON (1957), Theorem 3, § 2, Chapter X, and Theorem 5, § 4, Chapter IX.

Theorem b. *Let* $\varphi(u)$, $0 < u < 1$, *be square integrable and let* $a_N^\varphi(i)$ *be given by* (12). *Then*

$$\lim_{N\to\infty} \int_0^1 [a_N^\varphi(1 + [uN]) - \varphi(u)]^2\, \mathrm{d}u = 0, \tag{18}$$

with $[uN]$ *denoting the largest integer not exceeding* uN.

Proof. Since, in accordance with (7), where $Y = \varphi(U_1)$,

$$\int_0^1 [a_N^\varphi(1 + [uN])]^2\, \mathrm{d}u \leqq \int_0^1 \varphi^2(u)\, \mathrm{d}u, \tag{19}$$

it suffices to prove that

$$\lim_{N\to\infty} a_N^\varphi(1 + [uN]) = \varphi(u) \tag{20}$$

almost everywhere and then apply Theorem 1.3.

Now (20) follows from Lemma b, if we put

$$f_N(t, u) = N \binom{N-1}{i-1} t^{i-1}(1-t)^{N-i}, \quad \frac{i-1}{N} \leqq u < \frac{i}{N}, \quad 0 < t < 1 \tag{21}$$

and

$$g_N(t, u) = N\binom{N-1}{i-1}\left(\frac{i-1}{N-1}\right)^{i-1}\left(\frac{N-i}{N-1}\right)^{N-i}, \quad \frac{i-1}{N} \leqq t, u < \frac{i}{N},$$

$$= f_N(t, u), \quad \text{otherwise}.$$

Actually, then (see (II.1.2.10))

$$a_N^{\varphi}(1 + [uN]) = \mathsf{E}\,\varphi(U_N^{(i)}) = \int_0^1 \varphi(t)\, N\binom{N-1}{i-1} t^{i-1}(1-t)^{N-i}\,\mathrm{d}t =$$

$$= \int_0^1 \varphi(t) f_N(t, u)\,\mathrm{d}t, \quad \frac{i-1}{N} \leqq u < \frac{i}{N},$$

while (15) is satisfied since $f_N(t, u)$ is unimodal with mode at $(i-1)/(N-1)$, which lies within the interval $((i-1)/N, i/N)$. Also (16) holds true, since

$$\int_0^1 g_N(t, u)\,\mathrm{d}t \leqq \int_0^1 f_N(t, u)\,\mathrm{d}t + \binom{N-i}{i-1}\left(\frac{i-1}{N-1}\right)^{i-1}\left(\frac{N-i}{N-1}\right)^{N-i} \leqq 2.$$

Thus (17) holds, which is equivalent to (20). Q.E.D.

1.5. Locally optimum rank-test statistics for H_0. Now we are prepared to prove easily all the theorems needed. Consider real vectors $c = (c_1, \ldots, c_N)$ such that

(1) $$\sum_{i=1}^{N}(c_i - \bar{c})^2 > 0$$

where

(2) $$\bar{c} = \frac{1}{N}\sum_{i=1}^{N} c_i.$$

Let C be the set of real vectors of all finite dimensions $N \geqq 1$, satisfying (1). We shall consider limiting distributions of statistics indexed by $c \in C$ for

(3) $$\frac{\sum_{i=1}^{N}(c_i - \bar{c})^2}{\max_{1 \leqq i \leqq N}(c_i - \bar{c})^2} \to \infty.$$

Take a square integrable function $\varphi(u)$, $0 < u < 1$, and denote by $a_N^{\varphi}(i)$ the scores associated with φ by (1.4.12). Put

(4) $$S_c = \sum_{i=1}^{N} c_i\, a_N^{\varphi}(R_{Ni}), \quad c \in C$$

where R_{Ni} is the rank of X_i in a set of N independent observations $X_1, \ldots, X_N$, each with density f. If $U_i = F(X_i)$, $F(x) = \int_{-\infty}^{x} f(y)\,dy$, then the random variables U_i will be uniformly distributed and R_{Ni} may be interpreted as the rank of U_i in the set $U_1, \ldots, U_N$ as well. As we know from § II. 4, the test statistics generating locally most powerful rank tests are just of the type (4).

Theorem a. *Let the scores $a_N^{\varphi}(i)$ be associated with a square integrable function $\varphi(u)$ by* (1.4.12). *Put $\bar{\varphi} = \int_0^1 \varphi(u)\,du$ and assume $\int_0^1 [\varphi(u) - \bar{\varphi}]^2\,du > 0$. Assume H_0.*
Then, for (3), *the statistics* (4) *are asymptotically normal (μ_c, σ_c^2) with*

$$\mu_c = \bar{c} \sum_{i=1}^{N} a_N(i) \tag{5}$$

and

$$\sigma_c^2 = [\sum_{i=1}^{N} (c_i - \bar{c})^2] \int_0^1 [\varphi(u) - \bar{\varphi}]^2\,du\,, \tag{6}$$

or $\sigma_c^2 = \operatorname{var} S_c$.

Proof. Rewrite S_c in the following form:

$$S_c = \sum_{i=1}^{N} (c_i - \bar{c})\, a_N(R_{Ni}) + \bar{c} \sum_{i=1}^{N} a_N(i)\,. \tag{7}$$

Introduce

$$T_c = \sum_{i=1}^{N} (c_i - \bar{c})\, \varphi(U_i) + \bar{c} \sum_{i=1}^{N} a_N(i)\,, \tag{8}$$

where $U_i = F(X_i)$, $1 \leqq i \leqq N$. Now drop N in R_{Ni}, and recall that the distribution of $(R_1, \ldots, R_N)$ is independent of $U^{(\cdot)}$. Consequently, by (II.3.1.23), we obtain

$$\begin{aligned}
\mathsf{E}\{(T_c - S_c)^2 \mid U^{(\cdot)} = u^{(\cdot)}\} &= \\
&= \mathsf{E}\{\sum_{i=1}^{N} (c_i - \bar{c})\,(a_N(R_i) - \varphi(u^{(R_i)}))\}^2 = \\
&= \frac{1}{N-1} \sum_{i=1}^{N} (c_i - \bar{c})^2 \sum_{j=1}^{N} [a_N(j) - \varphi(u^{(j)}) - \bar{a}_N + \bar{\varphi}]^2 \leqq \\
&\leqq \frac{1}{N-1} \sum_{i=1}^{N} (c_i - \bar{c})^2 \sum_{j=1}^{N} [a_N(j) - \varphi(u^{(j)})]^2 \\
&= \frac{N}{N-1} \sum_{i=1}^{N} (c_i - \bar{c})^2\, \mathsf{E}\{[a_N(R_1) - \varphi(U_1)]^2 \mid U^{(\cdot)} = u^{(\cdot)}\}\,.
\end{aligned} \tag{9}$$

Consequently

$$\mathsf{E}(T_c - S_c)^2 \leqq \frac{N}{N-1} \sum_{i=1}^{N} (c_i - \bar{c})^2\, \mathsf{E}[a_N(R_1) - \varphi(U_1)]^2 \tag{10}$$

and

$$\mathsf{E}\left(\frac{T_c - S_c}{\sigma_c}\right)^2 \leqq \frac{N}{N-1}\left(\int_0^1 [\varphi(u) - \bar{\varphi}]^2 \,\mathrm{d}u\right)^{-1} \mathsf{E}[a_N(R_1) - \varphi(U_1)]^2 . \tag{11}$$

On the other hand

$$\frac{\sum_{i=1}^{N} (c_i - \bar{c})^2}{\max_{1 \leqq i \leqq N} (c_i - \bar{c})^2} \leqq N \tag{12}$$

so that (3) entails $N \to \infty$. This fact, together with Theorem 1.4.a and (11), implies

$$\lim_c \mathsf{E}\left(\frac{T_c - S_c}{\sigma_c}\right)^2 = 0 , \tag{13}$$

and a fortiori

$$\lim_c P\left(\left|\frac{T_c - S_c}{\sigma_c}\right| > \varepsilon\right) = 0 , \quad \varepsilon > 0 . \tag{14}$$

Now we know from Theorem 1.2 that the random variables T_c are asymptotically normal with parameters given by (5) and (6). Furthermore,

$$\frac{S_c - \mu_c}{\sigma_c} = \frac{T_c - \mu_c}{\sigma_c} + \frac{S_c - T_c}{\sigma_c} , \tag{15}$$

where the last term converges to 0 in probability according to (14). Thus asymptotic normality (0, 1) of $(T_c - \mu_c)/\sigma_c$ implies the same for $(S_c - \mu_c)/\sigma_c$, in view of a well-known lemma (see CRAMÉR (1945), Section 20.6).

Now σ_c^2 given by (6) equals $\operatorname{var} T_c$, and (13) implies $\operatorname{var} S_c/\operatorname{var} T_c \to 1$, since $|(\operatorname{var} S_c)^{\frac{1}{2}} - (\operatorname{var} T_c)^{\frac{1}{2}}| \leqq [\mathsf{E}(T_c - S_c)^2]^{\frac{1}{2}}$. Consequently, we may put $\sigma_c^2 = \operatorname{var} S_c$ as well. Q.E.D.

In the two-sample problem we consider statistics

$$S_{mn} = \sum_{i=1}^{m} a_{m+n}(R_{m+n,i}) \tag{16}$$

and we are concerned with their limiting properties for

$$\min(m, n) \to \infty . \tag{17}$$

Theorem b. *Let the scores* $a_N^{\varphi}(i)$ *be associated with a square integrable function* $\varphi(u)$ *by* (1.4.12) *and assume* $\int_0^1 [\varphi(u) - \bar{\varphi}]^2 \,\mathrm{d}u > 0$.