8240 SP 31 D

KONINKL. NEDERL. AKADEMIE VAN WETENSCHAPPEN – AMSTERDAM
Reprinted from Proceedings, Series A, 57, No. 5 and Indag. Math., 16, No. 5, 1954

MATHEMATICS

AN UPPER BOUND FOR THE DEVIATION BETWEEN THE DISTRIBUTION OF WILCOXON'S TEST STATISTIC FOR THE TWO-SAMPLE PROBLEM AND ITS LIMITING NORMAL DISTRIBUTION FOR FINITE SAMPLES. I

BY

D. J. STOKER

(*Mathematical Centre, Amsterdam*) [1])

(Communicated by Prof. D. VAN DANTZIG at the meeting of September 25, 1954)

I. *Introduction and summary*

Given two independent random samples [2]) $\mathbf{x}_1, \ldots, \mathbf{x}_m$ and $\mathbf{y}_1, \ldots, \mathbf{y}_n$ which are drawn from two unknown populations with (cumulative) distribution functions F and G, respectively. In the two-sample problem the hypothesis to be tested is

$$H_0: \qquad F(t)=G(t) \text{ for all } t, \ -\infty<t<\infty.$$

Wilcoxon's two-sample test is based on the statistic $\mathbf{U}$ defined by the number of pairs (i, j) $(i=1, \ldots, m;\ j=1, \ldots, n)$ with $\mathbf{y}_j<\mathbf{x}_i$. We confine ourselves at first to the case of distribution functions F and G which are continuous, or at least have no point of discontinuity in common [3]). In these cases we have [4])

$$\mathbf{U} \overset{\text{def}}{=} \sum_{i=1}^{m} \sum_{j=1}^{n} \iota(\mathbf{x}_i - \mathbf{y}_j) \tag{1}$$

where

$$\iota(z) \overset{\text{def}}{=} \begin{cases} 1 \\ 0 \end{cases} \text{if } z \gtrless 0. \tag{2}$$

This test statistic was first suggested by WILCOXON (1945). In the following we give a short historical summary of the development of the above-mentioned test statistic for the two-sample problem and conclude this section with a summary of the contents of this paper.

In 1947 MANN and WHITNEY determined by recursion the distribution of $\mathbf{U}$ under the hypothesis H_0 and proved also that $\mathbf{U}$ is asymptotically normally distributed when m and n tend to infinity.

---

[1]) Report SP 31 of the Statistical Department.

[2]) Random variables will be distinguished from numbers (e.g. from the values they take in an experiment) by printing them in bold type.

[3]) See for instance VAN DANTZIG [1], p. 305.

[4]) In this and following sections the notation of VAN DANTZIG [2] will be used to a great extent.

LEHMANN proved in 1951 that the distribution of **U** is also asymptotical normal under any hypothesis

$$\text{H}: \qquad (F, G)$$

when m and n tend to infinity with m/n = constant, provided

$$0 < \theta \overset{\text{def}}{=} P[\mathbf{y} < \mathbf{x} | H] < 1. \tag{3}$$

(If $\theta = 0$ or $\theta = 1$ the limiting normal distribution is degenerate). His proof includes the case when F and G are not continuous, although he did not mention this fact explicitly.

For the case of continuous functions F and G, SUNDRUM (1953) illustrated the same result concerning the asymptotic normality of the distribution of **U** under any alternative hypothesis H (when m and n tend to infinity while m/n = constant) that was proved by LEHMANN (1951). SUNDRUM calculated the third and fourth moments of Wilcoxon's **U**-statistic under any continuous hypothesis (F, G).

As far as I know no direct attempt has been made to determine the maximum) difference between the probability distribution of Wilcoxon's **U**-statistic and its limiting normal distribution. The primary object of this paper is to find an upper bound for this difference when m and n are finite. More precisely, the following will be proved:

Theorem A.

Using the above notation and considering the case when F and G have no point of discontinuity in common, we have

$$D(\xi) \overset{\text{def}}{=} |P[\mathbf{U} - \mathscr{E}(\mathbf{U}| F, G) \leqslant \xi \cdot \sigma_{\mathbf{U}} | F, G] - \Phi(\xi)| \leqslant \Delta(m, n, \xi; F, G) \tag{4}$$

where

$$\sigma^2_{\mathbf{U}} \overset{\text{def}}{=} \text{var}(\mathbf{U} | F, G) \tag{5}$$

$$\Phi(\xi) \overset{\text{def}}{=} (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\xi} e^{-\frac{1}{2}x^2} dx \tag{6}$$

and $\Delta(m, n, \xi; F, G)$ is a function of m, n, ξ and a functional of F and G. The function Δ will be specified later.

Also an upper bound for $\Delta(m, n, \xi; F, G)$ for all values of ξ, $-\infty < \xi < \infty$, will be given.

Theorem B.

Using the notation of theorem A, we have

$$\sup_{-\infty<\xi<\infty} D(\xi) \leqslant \sup_{-\infty<\xi<\infty} \Delta(m, n, \xi; F, G) \leqslant \Delta'(m, n; F, G) \tag{7}$$

where also $\Delta'(m, n; F, G)$ will be specified later.

Theorems of the types A and B will be proved in sections III and IV. It will be proved that the functions Δ and Δ' vanish asymptotically when

m and n tend to infinity, no matter how, except in a trivial case. When either m or n does not tend to infinity, it is easy to prove that the distribution function of $\mathbf{U}$ does not tend to a normal distribution function.

Some more general remarks are made (section V) concerning the class of statistics introduced by HOEFFDING (1948) and also a few words are said on the case when F and G have at least one point of discontinuity in common.

Finally, an example, based on the class of non-parametric alternative hypotheses, introduced by LEHMANN (1953), is given.

II. *Notation and an auxiliary function*

We define

(8) $$x_{ij} \stackrel{\text{def}}{=} \iota(x_i - y_j) \qquad (i = 1, \ldots, m;\ j = 1, \ldots, n).$$

Therefore, from (3),

(9) $$\theta = \mathscr{E}\,\mathbf{x}_{ij} = \int_0^1 G\,dF$$

when F and G have no point of discontinuity in common. Defining

(10) $$\tilde{x}_{ij} \stackrel{\text{def}}{=} x_{ij} - \theta,$$

(11) $$\tilde{G}(x) \stackrel{\text{def}}{=} G(x) - \theta,$$

(12) $$\tilde{R}(y) \stackrel{\text{def}}{=} R(y) - \theta \stackrel{\text{def}}{=} 1 - F(y) - \theta,$$

(13) $$\beta_{hk} \stackrel{\text{def}}{=} \mathscr{E}\,\tilde{G}^h(\mathbf{x}_i)\,\tilde{R}^k(\mathbf{y}_j)\,\tilde{\mathbf{x}}_{ij}$$

so that

$$\beta_{r0} = \mathscr{E}\,\tilde{G}^{r+1}(\mathbf{x}) = \int (G(x) - \theta)^{r+1}\,dF(x)$$

and

$$\beta_{0r} = \mathscr{E}\,\tilde{R}^{r+1}(\mathbf{y}) = \int (1 - F(y) - \theta)^{r+1}\,dG(y).$$

Further,

(14) $$r(x) \stackrel{\text{def}}{=} \mathscr{E}\,\{\tilde{R}(\mathbf{y})\,\iota(\mathbf{x} - \mathbf{y})\,|\,\mathbf{x} = x\}$$

and

(15) $$g(y) \stackrel{\text{def}}{=} \mathscr{E}\,\{\tilde{G}(\mathbf{x})\,\iota(\mathbf{x} - \mathbf{y})\,|\,\mathbf{y} = y\}.$$

Let

(16) $$\mathbf{V} \stackrel{\text{def}}{=} \{mn(m+n)\}^{-\frac{1}{2}}\{\mathbf{U} - \mathscr{E}(\mathbf{U}\,|\,F, G)\},$$

then we have, according to VAN DANTZIG [2],

(17) $$\mathscr{E}(\mathbf{U}\,|\,F, G) = mn\theta$$

(18) $$\mathscr{E}(\mathbf{V}\,|\,F, G) = 0$$

and

(19) $$\sigma_{\mathbf{V}}^2 \stackrel{\text{def}}{=} \operatorname{var}(\mathbf{V}\,|\,F, G) = (m+n)^{-1}\,[m\beta_{01} + n\beta_{10} + \theta(1-\theta) - (\beta_{10} + \beta_{01})].$$

We now introduce a new random variable $\mathbf{Y}$ defined by

(20) $$\mathbf{Y} \overset{\text{def}}{=} \{mn/(m+n)\}^{\frac{1}{2}} \, [m^{-1} \sum_{i=1}^{m} \tilde{G}(\mathbf{x}_i) + n^{-1} \sum_{j=1}^{n} \tilde{R}(\mathbf{y}_j)].$$

Note that $\mathbf{Y}$ is the sum of $m+n$ independent random variables. We have

(21) $$\mathscr{E}(\mathbf{Y} \mid F, G) = 0$$

and

(22) $$\sigma_{\mathbf{Y}}^2 \overset{\text{def}}{=} \operatorname{var}(\mathbf{Y} \mid F, G) = (m\beta_{01} + n\beta_{10})\,(m+n)^{-1}.$$

Thus

$$\sigma_{\mathbf{V}}^2 - \sigma_{\mathbf{Y}}^2 = [\theta(1-\theta) - (\beta_{10} + \beta_{01})]\,(m+n)^{-1}.$$

III. *A theorem on the values of $\Delta(m, n, \xi; F, G)$ and $\Delta'(m, n; F, G)$*

Theorem 1:

In the above notation we have, under the hypothesis (F, G), for any $\delta > 0$,

(23) $$|P\,[\mathbf{V} \leqslant \xi \sigma_{\mathbf{V}}] - \Phi(\xi)| \leqslant H(m, n) + \varepsilon(\xi, \delta; \sigma_{\mathbf{V}}, \sigma_{\mathbf{Y}}) + \mu(m, n, \delta)$$

where, with C a numerical constant,

(24) $$H(m,n) \overset{\text{def}}{=} \frac{C \max\{n(\beta_{30}/\beta_{10})^{\frac{1}{2}};\, m(\beta_{03}/\beta_{01})^{\frac{1}{2}}\}}{\{mn(m\beta_{01} + n\beta_{10})\}^{\frac{1}{2}}} = O(\max(m^{-\frac{1}{2}}, n^{-\frac{1}{2}}))$$

(25) $$\begin{cases} \varepsilon(\xi, \delta; \sigma_{\mathbf{V}}, \sigma_{\mathbf{Y}}) \overset{\text{def}}{=} [2 - \sigma_{\mathbf{Y}}/\sigma_{\mathbf{V}}]\,\delta\,(\sqrt{2\pi}\,\sigma_{\mathbf{Y}})^{-1} \quad \text{if } |\xi| \leqslant \delta/\sigma_{\mathbf{V}} \\ \text{and} \qquad = \dfrac{1}{\sqrt{2\pi}\,\sigma_{\mathbf{Y}}} \cdot \left[\delta + |v| \dfrac{\sigma_{\mathbf{V}} - \sigma_{\mathbf{Y}}}{\sigma_{\mathbf{V}}}\right] \exp\left[-\tfrac{1}{2}\min\left(\dfrac{v^2}{\sigma_{\mathbf{V}}^2}; \dfrac{(|v|-\delta)^2}{\sigma_{\mathbf{Y}}^2}\right)\right] \\ \qquad\qquad \text{if } |v| = |\xi|\,\sigma_{\mathbf{V}} > \delta; \end{cases}$$

and

(26) $$\mu(m, n, \delta) \overset{\text{def}}{=} \begin{cases} \dfrac{M_4 - M_2^2}{(\delta^2 - M_2)^2 + (M_4 - M_2^2)} & \text{if } \delta^2 \geqslant M_4/M_2 \\ M_2/\delta^2 & \text{if } M_2 \leqq \delta^2 \leqq M_4/M_2. \end{cases}$$

with the M_r defined in the lemma mentioned below.

Proof of theorem 1.

We prove this theorem with the aid of the following lemma:

Lemma.

Let

(27) $$M_r \overset{\text{def}}{=} \mathscr{E}(\mathbf{V} - \mathbf{Y})^r,$$

then

(28) $$M_1 = 0$$

(29) $$M_2 = [\theta(1-\theta) - (\beta_{10} + \beta_{01})]\,(m+n)^{-1} = \sigma_{\mathbf{V}}^2 - \sigma_{\mathbf{Y}}^2 \quad (\geqslant 0),$$

so that $\mathbf{V}-\mathbf{Y}$ and $\mathbf{Y}$ are uncorrelated,

(30) $M_3=[(1-2\theta)\{\theta(1-\theta)-3(\beta_{10}+\beta_{01})\}+2(\beta_{20}+\beta_{02})+6\beta_{11}](mn(m+n)^3)^{-\frac{1}{2}}$

and finally

$$M_4=[a+b/m+c/n+d/(mn)]\,(m+n)^{-2} \tag{31}$$

with for $i\neq i'$, $j\neq j'$

$$S\overset{\text{def}}{=}\mathscr{E}\tilde{\mathbf{x}}_{ij}\,\tilde{\mathbf{x}}_{ij'}\,\tilde{\mathbf{x}}_{i'j}\,\tilde{\mathbf{x}}_{i'j'}, \tag{32}$$

$$\left\{\begin{aligned}a=3[\theta(1-\theta)-(\beta_{10}+\beta_{01})]^2+6S-6(\beta_{10}^2+\beta_{01}^2)-12\operatorname{var}r(\mathbf{x})-\\-12\operatorname{var}g(\mathbf{y}),\end{aligned}\right. \tag{33}$$

$$\left\{\begin{aligned}b=3[\beta_{30}-2(1-2\theta)\beta_{20}+(1-2\theta)^2\beta_{10}+\beta_{10}^2+\\+2\beta_{01}^2-4\beta_{10}\beta_{01}-4(1-2\theta)\beta_{11}+4\beta_{21}+4\operatorname{var}g(\mathbf{y})+8\operatorname{var}r(\mathbf{x})-2S]\end{aligned}\right. \tag{34}$$

$$\left\{\begin{aligned}c=3[\beta_{03}-2(1-2\theta)\beta_{02}+(1-2\theta)^2\beta_{01}+2\beta_{10}^2+\beta_{01}^2-4\beta_{10}\beta_{01}-4(1-2\theta)\beta_{11}\\+4\beta_{12}+8\operatorname{var}g(\mathbf{y})+4\operatorname{var}r(\mathbf{x})-2S]\end{aligned}\right. \tag{35}$$

and

$$\left\{\begin{aligned}d=(1-2\theta)^2[\theta(1-\theta)-7(\beta_{10}+\beta_{01})]-2\theta(1-\theta)[\theta(1-\theta)-4(\beta_{10}+\beta_{01})]-\\-6(\beta_{30}+\beta_{03})+12(1-2\theta)(\beta_{20}+\beta_{02})-6(\beta_{10}^2+\beta_{01}^2)+24\beta_{10}\beta_{01}+\\+36(1-2\theta)\beta_{11}-24\beta_{21}-24\beta_{12}-24\operatorname{var}g(\mathbf{y})-24\operatorname{var}r(\mathbf{x})+6S.\end{aligned}\right. \tag{36}$$

Proof of lemma.

We define

$$\tilde{X}_{ij}\overset{\text{def}}{=}\tilde{x}_{ij}-\tilde{G}(x_i)-\tilde{R}(y_j). \tag{37}$$

Then from equations (16), (20), (10), (11), (12) and (37)

$$\mathbf{V}-\mathbf{Y}=\{mn\,(m+n)\}^{-\frac{1}{2}}\sum_{i=1}^{m}\sum_{j=1}^{n}\tilde{\mathbf{X}}_{ij}. \tag{38}$$

It is easy to derive the following equalities under the general hypothesis (F,G)

(39) $\mathscr{E}\tilde{\mathbf{X}}_{ij}=0$

(40) $\mathscr{E}\mathbf{X}_{ij}^2=\theta(1-\theta)-(\beta_{10}+\beta_{01})$

(41) $\mathscr{E}(\tilde{\mathbf{X}}_{ij}\,|\,x_i)=0$

(42) $\mathscr{E}(\tilde{\mathbf{X}}_{ij}\,|\,y_j)=0$

(43) $\mathscr{E}\tilde{\mathbf{X}}_{ij}^3=(1-2\theta)[\theta(1-\theta)-3(\beta_{10}+\beta_{01})]+2(\beta_{20}+\beta_{02})+6\beta_{11}$

$$\left\{\begin{aligned}&\mathscr{E}\tilde{\mathbf{X}}_{ij}^4=(1-3\theta+3\theta^2)[\theta(1-\theta)-4(\beta_{10}+\beta_{01})]-3(\beta_{30}+\beta_{03})+\\&\quad+6(1-2\theta)(\beta_{20}+\beta_{02})+6\theta(1-\theta)(\beta_{10}+\beta_{01})+6\beta_{10}\beta_{01}+12(1-2\theta)\beta_{11}-\\&-12(\beta_{21}+\beta_{12}),\end{aligned}\right. \tag{44}$$

and for $i \neq i'$, $j \neq j'$

$$(45)\quad \begin{cases} \mathscr{E}\,\tilde{\mathbf{X}}_{ij}^2\,\tilde{\mathbf{X}}_{ij'}^2 = \beta_{30} - 2(1-2\theta)\beta_{20} + (1-2\theta)^2\beta_{10} + \beta_{01}^2 - 2\beta_{01}\beta_{10} + \\ +\theta(1-\theta)[\theta(1-\theta) - 2(\beta_{10}+\beta_{01})] - 4(1-2\theta)\beta_{11} + 4\beta_{21} + 4 \operatorname{var} r(\mathbf{x}). \end{cases}$$

$$(46)\quad \begin{cases} \mathscr{E}\,\tilde{\mathbf{X}}_{ij}^2\,\tilde{\mathbf{X}}_{i'j}^2 = \beta_{03} - 2(1-2\theta)\beta_{02} + (1-2\theta)^2\beta_{01} + \beta_{10}^2 - 2\beta_{10}\beta_{01} + \\ +\theta(1-\theta)[\theta(1-\theta) - 2(\beta_{10}+\beta_{01})] - 4(1-2\theta)\beta_{11} + 4\beta_{12} + 4 \operatorname{var} g(\mathbf{y}), \end{cases}$$

and finally,

$$(47)\qquad \mathscr{E}\,\tilde{\mathbf{X}}_{ij}\,\tilde{\mathbf{X}}_{ij'}\,\tilde{\mathbf{X}}_{i'j}\,\tilde{\mathbf{X}}_{i'j'} = S - \beta_{10}^2 - \beta_{01}^2 - 2 \operatorname{var} \mathrm{r}(\mathbf{x}) - 2 \operatorname{var} g(\mathbf{y}).$$

We are now in a position to calculate M_r for $r = 1, 2, 3, 4$ and find thus from equations (27), (37), (38) and the identities (39), (41) and (42):

$$(48)\qquad M_1 = 0$$

$$(49)\qquad M_2 = (m+n)^{-1}\,\mathscr{E}\,\tilde{\mathbf{X}}_{ij}^2$$

$$(50)\qquad M_3 = \{mn\,(m+n)^3\}^{-\frac{1}{2}}\,\mathscr{E}\,\tilde{\mathbf{X}}_{ij}^3$$

$$(51)\quad \begin{cases} M_4 = (mn)^{-1}(m+n)^{-2}\,[\mathscr{E}\,\tilde{\mathbf{X}}_{ij}^4 + 3\,(m-1)\,(n-1)\,\mathscr{E}\,\tilde{\mathbf{X}}_{ij}^2\,\tilde{\mathbf{X}}_{i'j'}^2 + \\ \qquad + 3\,(n-1)\,\mathscr{E}\,\tilde{\mathbf{X}}_{ij}^2\,\tilde{\mathbf{X}}_{ij'}^2 + 3\,(m-1)\,\mathscr{E}\,\tilde{\mathbf{X}}_{ij}^2\,\tilde{\mathbf{X}}_{i'j}^2 + \\ \qquad + 6\,(m-1)\,(n-1)\,\mathscr{E}\,\tilde{\mathbf{X}}_{ij}\,\tilde{\mathbf{X}}_{ij'}\,\tilde{\mathbf{X}}_{i'j}\,\tilde{\mathbf{X}}_{i'j'}]. \end{cases}$$

Substitution of equations (40), (43)–(47) into equations (48)–(51) gives the required results [5]).

Following Cramér (p. 254–255), we have, with $\boldsymbol{\eta} = \mathbf{V} - \mathbf{Y}$ and $\delta > 0$ [6])

$$(52)\qquad P[\mathbf{V} \leqslant v] = P[\mathbf{V} \leqslant v \wedge |\boldsymbol{\eta}| \leqslant \delta] + P[\mathbf{V} \leqslant v \wedge |\boldsymbol{\eta}| > \delta]$$

$$(53)\qquad P[\mathbf{Y} \leqslant v] = P[\mathbf{Y} \leqslant v \wedge |\boldsymbol{\eta}| \leqslant \delta] + P[\mathbf{Y} \leqslant v \wedge |\boldsymbol{\eta}| > \delta]$$

and

$$(54)\qquad P[\mathbf{Y} \leqslant v - \delta \wedge |\boldsymbol{\eta}| \leqslant \delta] \leqslant P[\mathbf{V} \leqslant v \wedge |\boldsymbol{\eta}| \leqslant \delta] \leqslant P[\mathbf{Y} \leqslant v + \delta \wedge |\boldsymbol{\eta}| \leqslant \delta].$$

From these three equations it is easy to find for any $\delta > 0$

$$(55)\quad \begin{cases} P[\mathbf{V} \leqslant v] \leqslant P[\mathbf{Y} \leqslant v + \delta] + P[\mathbf{V} \leqslant v \wedge |\boldsymbol{\eta}| > \delta] \leqslant \\ \qquad \leqslant P[\mathbf{Y} \leqslant v + \delta] + P[|\boldsymbol{\eta}| > \delta] \end{cases}$$

and

$$(56)\quad \begin{cases} P[\mathbf{V} \leqslant v] \geqslant P[\mathbf{Y} \leqslant v - \delta] - P[\mathbf{Y} \leqslant v - \delta \wedge |\boldsymbol{\eta}| > \delta] \geqslant \\ \qquad \geqslant P[\mathbf{Y} \leqslant v - \delta] - P[|\boldsymbol{\eta}| > \delta]. \end{cases}$$

[5]) Recently Dr R. M. Sundrum [2] kindly sent me a copy of a paper in which he calculated the third and fourth moments of Wilcoxon's U-statistic for the two-sample problem in the general case.

[6]) The symbol " $\wedge$ " is the usual abbreviation for "and".

Consequently

$$(57)\quad \begin{cases} \left|P\,[\mathbf{V} \leqslant v] - \Phi\left(\frac{v}{\sigma_\mathbf{V}}\right)\right| \leqslant \max\limits_{\pm}\left|P\,[\mathbf{Y} \leqslant v \pm \delta] - \Phi\left(\frac{v \pm \delta}{\sigma_\mathbf{Y}}\right)\right| + \\ \qquad + \max\limits_{\pm}\left|\Phi\left(\frac{v\pm\delta}{\sigma_\mathbf{Y}}\right) - \Phi\left(\frac{v}{\sigma_\mathbf{V}}\right)\right| + P\,[|\mathbf{V}-\mathbf{Y}| > \delta]. \end{cases}$$

The right-hand side of (57) is the sum of three separate terms. We are now going to find upper bounds for them:

(i) As mentioned before, $\mathbf{Y}$ is equal to the sum of $m+n$ independent random variables, more specifically, to the sum of m isomorous [7]) random variables and n isomorous random variables. By means of papers by ESSEEN (1945) and BERRY (1941), it is possible to find an upper bound for the deviation between the distribution function of $\mathbf{Y}$ and its limiting normal distribution function.

Following BERRY we have, where C is a numerical constant,

$$(58)\quad \begin{cases} \max\limits_{\infty \leqslant v' \leqslant \infty}\left|P\,[\mathbf{Y} \leqslant v'] - \Phi\left(\frac{v'}{\sigma_\mathbf{Y}}\right)\right| \leqslant \frac{C \max\{n\mathscr{E}\,|\tilde{G}(\mathbf{x})|^3/\beta_{10};\, m\mathscr{E}\,|\tilde{R}(\mathbf{y})|^3/\beta_{01}\}}{\{mn\,(m+n)\}^{\frac{1}{2}}\,\sigma_\mathbf{Y}} \\ \qquad \leqslant H\,(m,n) = O\,(\max\,(m^{-\frac{1}{2}}, n^{-\frac{1}{2}})) \end{cases}$$ [8])

because of equations (13), (22), (24) and

$$\mathscr{E}\,|\tilde{G}(\mathbf{x})|^3 \leqslant \sqrt{\beta_{30}\,\beta_{10}}$$
$$\mathscr{E}\,|\tilde{R}(\mathbf{y})|^3 \leqslant \sqrt{\beta_{03}\,\beta_{01}}\,.$$

Hence the first term in (57) is $\leqslant H(m, n)$.

(ii) From (6) it follows, with

$$I \overset{\text{def}}{=} \max_{\pm}\left|\Phi\left(\frac{v\pm\delta}{\sigma_\mathbf{Y}}\right) - \Phi\left(\frac{v}{\sigma_\mathbf{V}}\right)\right| = (2\pi)^{-\frac{1}{2}} \max_{\pm}\{|\int\limits_{v/\sigma_\mathbf{V}}^{(v\pm\delta)/\sigma_\mathbf{Y}} e^{-\frac{1}{2}x^2}\,dx|\}.$$

$$(59)\quad \begin{cases} \sqrt{2\pi}\,I \leqslant \max\limits_{\pm}\left|\frac{v\pm\delta}{\sigma_\mathbf{Y}} - \frac{v}{\sigma_\mathbf{V}}\right| \leqslant \left[\frac{\delta}{\sigma_\mathbf{Y}} + |v|\left(\frac{1}{\sigma_\mathbf{Y}} - \frac{1}{\sigma_\mathbf{V}}\right)\right] \\ \qquad \leqslant \frac{\delta}{\sigma_\mathbf{Y}}\left[1 + \frac{\sigma_\mathbf{V} - \sigma_\mathbf{Y}}{\sigma_\mathbf{V}}\right] \quad \text{if} \quad |v| \leqslant \delta, \end{cases}$$

where $\sigma_\mathbf{V} - \sigma_\mathbf{Y} \leqslant \{8(m+n)\,\sigma_\mathbf{Y}\}^{-1} = O((m+n)^{-1})$ for $\sigma_\mathbf{Y} > 0$; and

$$(60)\quad \begin{cases} \sqrt{2\pi}\,I \leqslant \max\limits_{\pm}\left|\frac{v\pm\delta}{\sigma_\mathbf{Y}} - \frac{v}{\sigma_\mathbf{V}}\right| \exp\left[-\tfrac{1}{2}\min\{v^2/\sigma_\mathbf{V}^2;\, (|v|-\delta)^2/\sigma_\mathbf{Y}^2\}\right] \\ \qquad \leqslant \left[\frac{\delta}{\sigma_\mathbf{Y}} + |v|\left(\frac{1}{\sigma_\mathbf{Y}} - \frac{1}{\sigma_\mathbf{V}}\right)\right] \exp\left[-\tfrac{1}{2}\min\{v^2/\sigma_\mathbf{V}^2;\, (|v|-\delta)^2/\sigma_\mathbf{Y}^2\}\right] \end{cases}$$

if $|v| > \delta$.

---

7) "Isomorous": having the same distribution function.

8) It is interesting to note that, according to ESSEEN (p. 44), the order of the term $H\,(m, n)$ cannot be improved even if the moments of all orders of the $m + n$

Consequently, with $v=\xi\sigma_V$

$$\max_{\pm}\left|\Phi\left(\frac{v\pm\delta}{\sigma_Y}\right)-\Phi\left(\frac{v}{\sigma_V}\right)\right|\leqslant\varepsilon(\xi,\delta;\sigma_V,\sigma_Y)=O(\delta) \tag{61}$$

where $\varepsilon(\xi, \delta; \sigma_V, \sigma_Y)$ is defined by equation (25).

(iii) Application of the general Bienaymé–Tchebychef's inequality gives

$$P[|\mathbf{V}-\mathbf{Y}|>\delta]\leqslant M_r/\delta^r \tag{62}$$

where M_r is defined by (27). This inequality is in general very rough. A better approximation (for $r\leqslant 4$) can be given by Cantelli's [9]) generalization of the above inequality, i.e. (with $N=m+n$)

$$P[|\mathbf{V}-\mathbf{Y}|>\delta]\leqslant\begin{cases}\dfrac{M_4-M_2^2}{(\delta^2-M_2)^2+(M_4-M_2^2)}=O(N^{-2}\delta^{-4}) & \text{if } \delta^2>M_4/M_2=O(N^{-1})\\ M_2/\delta^2=O(N^{-1}\delta^{-2}) & \text{if } M_2\leqslant\delta^2\leqslant M_4/M_2.\end{cases} \tag{63}$$

Consequently

$$P[|\mathbf{V}-\mathbf{Y}|>\delta]\leqslant\mu(m,n,\delta) \tag{64}$$

where $\mu(m, n, \delta)$ is defined by equation (26).

Substitution of the inequalities (58), (61) and (64) into (57) proves theorem 1.

---

independent random variables, which compose the statistic **Y**, exist and are finite. According to P. L. Hsu (Ann. of Math. Stat., 16, 3 (1945)), Berry has made a mistake in his calculation of the value of C, which he has found to be equal to 1.88.

[9]) Cf. Chapter IV of Fréchet's book.

AN UPPER BOUND FOR THE DEVIATION BETWEEN THE DISTRIBUTION OF WILCOXON'S TEST STATISTIC FOR THE TWO-SAMPLE PROBLEM AND ITS LIMITING NORMAL DISTRIBUTION FOR FINITE SAMPLES. II

BY

D. J. STOKER

(*Mathematical Centre, Amsterdam*)

(Communicated by Prof. D. VAN DANTZIG at the meeting of September 25, 1954)

IV. *Further theorems on the values of the functions Δ and Δ'*

For the function $\varepsilon(\xi, \delta; \sigma_V, \sigma_Y)$, defined by equation (25), it is also possible to find an upper bound independent of ξ.

Theorem 2.

$$\sup_{-\infty<\xi<\infty} \varepsilon(\xi, \delta; \sigma_V, \sigma_Y) \leqslant \varepsilon'(\delta; \sigma_V, \sigma_Y) \tag{65}$$

where $\varepsilon(\xi, \delta; \sigma_V, \sigma_Y)$ is defined by equation (25) and

$$\varepsilon'(\delta; \sigma_V, \sigma_Y) \stackrel{\text{def}}{=} \frac{\delta}{\sqrt{2\pi}\,\sigma_Y}\left[1 + \frac{\sigma_V - \sigma_Y}{\sigma_Y} + \frac{(\sigma_V - \sigma_Y)^2}{\delta^2}\right]. \tag{66}$$

Proof:

First we consider the case

$$|v| = |\xi|\,\sigma_V > \delta.$$

Then, we have

$$\min\left(\frac{v^2}{\sigma_V^2}; \frac{(|v|-\delta)^2}{\sigma_Y^2}\right) = (|v| - \delta)^2/\sigma_Y^2 \tag{67}$$

if

$$\frac{\delta\,\sigma_V}{\sigma_V + \sigma_Y} \leqslant |v| \leqslant \frac{\delta\,\sigma_V}{\sigma_V - \sigma_Y}.$$

The function

$$\sqrt{2\pi}\, I_1 \stackrel{\text{def}}{=} \left[\frac{\delta}{\sigma_Y} + |v|\left(\frac{1}{\sigma_Y} - \frac{1}{\sigma_V}\right)\right] \exp\left[-\tfrac{1}{2}(|v| - \delta)^2/\sigma_Y^2\right] \tag{68}$$

reaches its maximum value when

$$|v| = \frac{\delta\,\sigma_Y}{2(\sigma_V - \sigma_Y)}\left[-1 + \left\{1 + 4\left(\frac{\sigma_V(\sigma_V - \sigma_Y)}{\sigma_Y^2} + \frac{(\sigma_V - \sigma_Y)^2}{\delta^2}\right)\right\}^{\frac{1}{2}}\right]$$
$$< \delta\,\sigma_V/\sigma_Y + \sigma_Y(\sigma_V - \sigma_Y)/\delta$$

and thus

$$(69)\qquad \sqrt{2\pi}\, I_1 < \frac{\delta}{\sigma_Y}\left[1 + \frac{\sigma_V - \sigma_Y}{\sigma_Y}\right] + \frac{(\sigma_V - \sigma_Y)^2}{\delta\, \sigma_V} \text{ if } \delta < |v| \leqslant \frac{\delta\, \sigma_V}{\sigma_V - \sigma_Y}.$$

On the other hand, for the function

$$(70)\qquad \sqrt{2\pi}\, I_2 \stackrel{\text{def}}{=} \left[\frac{\delta}{\sigma_Y} + |v| \left(\frac{1}{\sigma_Y} - \frac{1}{\sigma_V}\right)\right] \exp\left[-\tfrac{1}{2} v^2/\sigma_V^2\right],$$

we have similarly

$$(71)\qquad \sqrt{2\pi}\, I_2 < \delta/\sigma_Y + (\sigma_V - \sigma_Y)^2/(\delta\, \sigma_Y) \text{ if } |v| > \frac{\delta\, \sigma_V}{\sigma_V - \sigma_Y}.$$

Thus we have, from (59), (60), (69) and (71), for all values of ξ,

$$\varepsilon(\xi, \delta; \sigma_V, \sigma_Y) < \frac{\delta}{\sqrt{2\pi}\, \sigma_Y}\left[1 + \frac{\sigma_V - \sigma_Y}{\sigma_Y} + \frac{(\sigma_V - \sigma_Y)^2}{\delta^2}\right].$$

The last theorem, which will be stated and proved in this section, gives an improvement of the term $H(m, n)$ (cf. equations (23), (24) and (58)) for values of ξ far from 0, under the hypothesis (F, G).

Theorem 3:

$$(72)\qquad |P[\mathbf{Y} \leqslant \xi\sigma_Y] - \Phi(\xi)| \leqslant \min\{H(m, n)\ ;\ H'(m, n, \xi)\}$$

where $H(m, n)$ is defined by equation (24) and

$$(73)\qquad H'(m, n, \xi) \stackrel{\text{def}}{=} [c\{\ln H(m, n)\}^2 \cdot H(m, n) + K(m, n)]\, (1 + \xi^4)^{-1}$$

with c a numerical constant

$$\leqslant \begin{cases} 30.9 & \text{if } H(m, n) \leqslant \frac{1}{2} \\ 18.9 & \text{if } H(m, n) \leqslant 1/10 \end{cases}$$

and

$$(74)\qquad K(m, n) \stackrel{\text{def}}{=} \frac{(\beta_{30} - 3\beta_{10}^2)\, n^2/m + (\beta_{03} - 3\beta_{01}^2)\, m^2/n}{(m\beta_{01} + n\beta_{10})^2} = O(\max(m^{-1}, n^{-1})).$$

Proof. [10])

From (58) we have

$$(75)\qquad |P(\xi) - \Phi(\xi)| \leqslant H(m, n)$$

where

$$P(\xi) \stackrel{\text{def}}{=} P[\mathbf{Y} \leqslant \xi\sigma_Y],$$

with

$$(76)\qquad \int \xi^2 d\Phi(\xi) = 1,\ \int \xi^4 d\Phi(\xi) = 3 \text{ and } \int \xi^2 dP(\xi) = 1.$$

It is easy to prove that (from (20) and (22))

$$(77)\qquad \mathscr{E}(\mathbf{Y}/\sigma_Y)^4 = 3 + K(m, n)$$

[10]) We are using a method given by Esseen (p. 68–70).

where $K(m, n)$ is defined by (74). Let a be a number $\geqslant 1$ to be determined later. Without loss of generality we may suppose that $P(\xi)$ is continuous at $\xi = \pm a$. For k even we have

$$\begin{aligned}\int_{-a}^{a} \xi^k dP(\xi) &= \int_{-a}^{a} \xi^k d(P(\xi)-\Phi(\xi)) + \int_{-a}^{a} \xi^k d\Phi(\xi) = \\ &= a^k(P(a)-\Phi(a)) - a^k(P(-a)-\Phi(-a)) - \\ &\qquad - k\int_{-a}^{a} \xi^{k-1}(P(\xi)-\Phi(\xi))d\xi + \int_{-a}^{a} \xi^k d\Phi(\xi)\end{aligned}$$

Thus

$$\int_{-a}^{a} \xi^k dP(\xi) \geqslant -4a^k H(m, n) + \int_{-a}^{a} \xi^k d\Phi(\xi). \tag{78}$$

From (76), (77) and (78) it follows (taking $k=4$)

$$\begin{cases} \int\limits_{|\xi|\geqslant a} \xi^4 dP(\xi) \leqslant 4a^4 H(m, n) + 3 + K(m, n) - \int_{-a}^{a} \xi^4 d\Phi(\xi) \\ \qquad = 4a^4 H(m, n) + K(m, n) + \int\limits_{|\xi|\geqslant a} \xi^4 d\Phi(\xi). \end{cases} \tag{79}$$

Now the following relations hold

$$\int\limits_{|\xi|\geqslant a} \xi^4 \, dP(\xi) \geqslant \begin{cases} \zeta^4(1 - P(\zeta)) & \text{for } \zeta \geqslant a \\ \zeta^4 P(\zeta) & \text{for } \zeta \leqslant -a \end{cases} \tag{80}$$

$$\int\limits_{|\xi|\geqslant a} \xi^4 \, d\Phi(\xi) \geqslant \begin{cases} \zeta^4(1 - \Phi(\zeta)) & \text{for } \zeta \geqslant a \\ \zeta^4 \Phi(\zeta) & \text{for } \zeta \leqslant -a \end{cases} \tag{81}$$

and

$$\begin{aligned}\int\limits_{|\xi|\geqslant a} \xi^4 \, d\Phi(\xi) &= \frac{1}{\sqrt{2\pi}} \int\limits_{|\xi|\geqslant a} \xi^4 e^{-\frac{1}{2}\xi^2} d\xi \\ &= \frac{2}{\sqrt{2\pi}} [-\xi^3 e^{-\frac{1}{2}\xi^2}]_a^\infty + \frac{3}{\sqrt{2\pi}} \int\limits_{|\xi|\geqslant a} \xi^2 e^{-\frac{1}{2}\xi^2} d\xi \\ &= \sqrt{\frac{2}{\pi}} e^{-\frac{1}{2}a^2} [a^3 + 3a + 3/a] - \frac{3}{\sqrt{2\pi}} \int\limits_{|\xi|\geqslant a} \xi^{-2} e^{-\frac{1}{2}\xi^2} d\xi.\end{aligned}$$

Thus

$$\int\limits_{|\xi|\geqslant a} \xi^4 \, d\Phi(\xi) \leqslant \sqrt{\frac{2}{\pi}} e^{-\frac{1}{2}a^2} \cdot [a^3 + 3a + 3/a]. \tag{82}$$

From (79), (80) and (81) we obtain, for $|\xi| \geqslant a$,

$$\begin{aligned}\xi^4|P(\xi)-\Phi(\xi)| &\leqslant \xi^4 \max\,[\min\{P(\xi)\;;\;1-P(\xi)\}, \min\{\Phi(\xi)\;;\;1-\Phi(\xi)\}] \\ &\leqslant 4a^4 H(m, n) + K(m, n) + \int\limits_{|\xi|\geqslant a} \xi^4 d\Phi(\xi)\end{aligned}$$

and hence from (75) and (82), for $|\xi| \geqslant a$,

$$\begin{cases} |P(\xi)-\Phi(\xi)| \leqslant (1+\xi^4)^{-1}[(4a^4+1)H(m, n) + K(m, n) + \\ \qquad\qquad + \sqrt{\dfrac{2}{\pi}} e^{-\frac{1}{2}a^2}(a^3 + 3a + 3/a)]. \end{cases} \tag{83}$$

This inequality obviously holds not only for $|\xi| \geqslant a$ but also for all other values of ξ. In (83) we choose a such that the first and the last term on the right-hand side have the same order of magnitude in m, n. Suppose m_0 and n_0 are so large that $H(m, n) \leqslant \frac{1}{2}$ for $m \geqslant m_0$ and $n \geqslant n_0$. Minimizing the two terms with respect to a, we obtain

$$a = O((2 \ln H^{-1})^{\frac{1}{2}})$$

where

$$H \stackrel{\text{def}}{=} H(m, n).$$

Putting

$$a = (2 \ln H^{-1})^{\frac{1}{2}}, \text{ thus } a = O(\{\ln \min (m, n)\}^{\frac{1}{2}})$$

we obtain

$$\begin{aligned}(1+\xi^4)|P(\xi)-\Phi(\xi)| &\leqslant (2 \ln H^{-1})^2 \cdot H[4+(2 \ln H^{-1})^{-2}+ \\ &\quad + \sqrt{2/\pi}(2 \ln H^{-1})^{-\frac{1}{2}} + 3\sqrt{2/\pi}(2 \ln H^{-1})^{-3/2} + \\ &\quad + 3\sqrt{2/\pi}(2 \ln H^{-1})^{-5/2}] + K(m, n). \\ &\leqslant \begin{cases} 7.72(2 \ln H^{-1})^2 \cdot H + K(m, n) & \text{for } H \leqslant \frac{1}{2} \\ 4.71(2 \ln H^{-1})^2 \cdot H + K(m, n) & \text{for } H \leqslant 1/10. \end{cases}\end{aligned}$$

The numerical constants are calculated by taking $H = \frac{1}{2}$ and $H = 1/10$, respectively. This proves the theorem.

V. *General remarks*

1) It follows from (23), and (65) that

$$(84) \qquad \begin{cases} \sup\limits_{-\infty<\xi<\infty} |P[\mathbf{V} \leqslant \xi \cdot \sigma_{\mathbf{V}}] - \Phi(\xi)| < H(m, n) + \\ \qquad + \min\limits_{\delta} \{\mu(m, n, \delta) + \varepsilon'(\delta; \sigma_{\mathbf{V}}, \sigma_{\mathbf{Y}})\}. \end{cases}$$

Hence, from (26) and (66), with

$$\sigma_{\mathbf{V}} - \sigma_{\mathbf{Y}} = O(N^{-1}), \; M_4/M_2 = O(N^{-1}) \text{ and } N = m + n,$$

it follows that

$$\delta = O(N^{-2/5}) \qquad (N \text{ large}).$$

Thus

$$\delta^2 > M_4/M_2 \text{ and } (\sigma_{\mathbf{V}} - \sigma_{\mathbf{Y}})/\delta = O(N^{-3/5})$$

so that

$$\mu(m, n, \delta) = (M_4 - M_2^2)/\delta^4 \cdot \{1 + O(N^{-1/5})\}$$

and

$$\varepsilon'(\delta; \sigma_{\mathbf{V}}, \sigma_{\mathbf{Y}}) = \frac{\delta}{\sqrt{2\pi}\,\sigma_{\mathbf{Y}}} \{1 + O(N^{-1})\}.$$

Consequently, for N sufficiently large,

$$(85) \qquad \begin{cases} \sup\limits_{-\infty<\xi<\infty} |P[\mathbf{V} \leqslant \xi \sigma_{\mathbf{V}}] - \Phi(\xi)| < H(m, n) + \\ \qquad + (1 + \lambda)\, {}^5/_4\, (\pi \sigma_{\mathbf{Y}}^2)^{-2/5} (M_4 - M_2^2)^{1/5}. \end{cases}$$

where λ is some positive number which tends to 0 when N tends to infinity.

Whereas the order of the term $H(m, n)$ cannot be improved (cf. note 8)), the order of the second and third terms (minimized with respect to δ) on the right-hand side of inequality (84) can be improved by calculation of higher moments M_r $(r>4)$, and, by taking r large enough, it seems that we can approximate the order $O((m+n)^{-\frac{1}{2}})$ arbitrarily near. On the other hand, the amount of work to obtain M_r increases very rapidly with increasing r. Hitherto we have not succeeded in finding an appropriate upper bound for a M_r $(r>4)$.

2) *Degeneration:* From (22), (24), (29), (31) and (85) it follows that

$$\sup_{-\infty<\xi<\infty} |P\,[\mathbf{V} \leqslant \xi\,\sigma_{\mathbf{V}}] - \Phi(\xi)|$$

tends to 0 when m and n tend to infinity, no matter how, unless

$$\sigma^2_{\mathbf{Y}} = (m\beta_{01} + n\beta_{10})(m+n)^{-1}$$

is equal to or tends to 0. In this case we have degeneration of the limiting normal distribution of Wilcoxon's statistic $\mathbf{V}$. Now (from (11), (12) and (13))

$$\sigma^2_{\mathbf{Y}} = \mathscr{E}\,\{\sqrt{n}\,\tilde{G}(\mathbf{x}) + \sqrt{m}\,\tilde{R}(\mathbf{y})\}^2 \cdot (m+n)^{-1}$$

(F and G have no point of discontinuity in common). Consequently,

$$\sigma^2_{\mathbf{Y}} = 0$$

is equivalent to

$$\{\sqrt{n}\,G(\mathbf{x}) + \sqrt{m}\,(1 - F(\mathbf{y}))\}\,(m+n)^{-\frac{1}{2}} = \text{constant} \qquad \text{spr } 0, \text{ [11]}$$

which can only be true for

$$\theta = P[\mathbf{y} < \mathbf{x}] = 0 \text{ or } 1 \text{ if } m/n \text{ and } n/m \text{ are bounded [12]}.$$

Theorem 4.

When F and G have no point of discontinuity in common, the condition

$$0 < \theta = P[\mathbf{y} < \mathbf{x}] < 1$$

is spr 0 equivalent to the condition

$$(86) \qquad \sigma^2_{\mathbf{Y}} = (n\beta_{10} + m\beta_{01})(m+n)^{-1} \neq 0 \text{ and } \not\to 0$$

when m and n tend to infinity with m/n and n/m bounded, where β_{hk} is defined by equation (13).

3) *Discontinuous case.* If the distribution functions F, G have at least one point of discontinuity in common, all the above-mentioned theorems remain true. This can easily be shown as follows:

---

[11]) "spr 0": except for a probability 0.

[12]) See also LEHMANN [1].

For such a point of discontinuity we have a positive probability that an $\mathbf{x}$ shall be equal to an $\mathbf{y}$. We replace $\iota(z)$, defined by equation (2), by

$$\tfrac{1}{2}(1+\operatorname{sgn} z)$$

where

$$\operatorname{sgn} z \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if} \quad z > 0 \\ 0 & \text{if} \quad z = 0 \\ -1 & \text{if} \quad z < 0. \end{cases} \tag{87}$$

We have thus (cf. (1) and HEMELRIJK (1952))

$$\mathbf{U} = \tfrac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \operatorname{sgn}(\mathbf{x}_i - \mathbf{y}_j) + \tfrac{1}{2} mn.$$

This method amounts to the method of averaging the ranks of the observations in each tie (cf. KENDALL or VAN DANTZIG and HEMELRIJK).

We now have the following type of expressions:

$$\begin{cases} x_{ij} = \tfrac{1}{2}(1 + \operatorname{sgn}(x_i - y_j)) \\ \theta = \mathscr{E} \mathbf{x}_{ij} = \tfrac{1}{2} \int_{-\infty}^{\infty} \{G(x+0) + G(x-0)\}\, dF(x) \end{cases} \tag{88}$$

where the integral is a Lebesque–Stieltjes-integral.

In the same way, with

$$\tilde{G}(x) \stackrel{\text{def}}{=} \tfrac{1}{2}\{G(x+0) + G(x-0)\} - \theta$$

and

$$\tilde{R}(y) \stackrel{\text{def}}{=} 1 - \tfrac{1}{2}\{F(y+0) + F(y-0)\} - \theta,$$

we have

$$\beta_{hk} \stackrel{\text{def}}{=} \mathscr{E}\, \tilde{G}^h(\mathbf{x}_i)\, \tilde{R}^k(\mathbf{y}_j)\, (\mathbf{x}_{ij} - \theta),$$

which is the same expression as given by (13).

With this type of modifications in the discontinuous case the proofs of the above theorems remain the same.

4) The results of this paper are of much wider application than only to Wilcoxon's statistic. In fact, Wilcoxon's statistic is a special case of a class of non-parametric statistics introduced by HOEFFDING [1] and LEHMANN [1]. In the proof of the asymptotic normality of this class of statistics also use is made of an auxiliary function that is a sum of r, say, independent random variables, where r tends to infinity. In the same way we can apply the method of this paper to obtain an upper bound for the (maximum) deviation between the distribution of any statistic of this class of statistics and its limiting normal distribution.

VI. *Example*

For an example we take

$$G = F^2$$

which is a special case of the class of non-parametric alternatives introduced by LEHMANN [2].

Then we have

$$\theta=1/3,\ \theta(1-\theta)=2/9,\ \beta_{10}=4/45,\ \beta_{01}=1/18,$$
$$\beta_{20}=16/945,\ \beta_{02}=1/135,\ \beta_{30}=16/945,\ \beta_{03}=1/135,$$
$$\beta_{11}=1/270,\ \beta_{21}=37/11340,\ \beta_{12}=2/567,$$
$$\operatorname{var} r(\mathrm{x})+\beta_{01}^2=4/945,\ \operatorname{var} g(\mathrm{y})+\beta_{10}^2=1/108,\ S=7/405.$$

Substitution of the above values in (33)–(36) gives

$$a=163/6300,\ b=-19/3150,\ c=-41/6300 \text{ and } d=-1/1575$$

and we have thus (from (31))

(89) $$\begin{cases} M_4=[163-38/m-41/n-4/(mn)]\,\{6300\,(m+n)^2\}^{-1}< \\ \quad <163/\{6300\,(m+n)^2\}. \end{cases}$$

From (29)

(90) $$M_2=7/\{90\,(m+n)\}.$$

Further, from the above values and eq. (22)

(91) $$\sigma_{\mathrm{Y}}^2=(4/5\,n+\tfrac{1}{2}m)/\{9\,(m+n)\}$$

and from eq. (24)

(92) $$\begin{cases} H(m,n)=C\max\{n\sqrt{4/21}\,;\ m\sqrt{2/15}\}\,3\,\{mn\,(4/5\,n+\tfrac{1}{2}m)\}^{-\frac{1}{2}} \\ \quad =1.15\,C\,n^{-\frac{1}{2}} \text{ for } m=n. \end{cases}$$

Also

(93) $$\begin{cases} \left(\dfrac{\sqrt{M_4-M_2^2}}{\pi\,\sigma_{\mathrm{Y}}^2}\right)^{2/5}\leqslant\left(\dfrac{1.27}{\pi\,(4/5\,n+\frac{1}{2}m)}\right)^{2/5} \\ \quad =0.63\,n^{-2/5} \text{ for } m=n. \end{cases}$$

Substitution of these values into the right-hand side of eq. (85), for instance, gives an expression that only depends on m and n and which gives us an upper bound for finite m and n.

Acknowledgement: The author wishes to express his sincere thanks to Prof. Dr D. van Dantzig for the helpful suggestions made during the study and for the help to bring the paper in its final form.

REFERENCES

Berry, A. C., The accuracy of the Gaussian approximation to the sum of independent variates, Trans. Amer. Math. Soc. **49**, 122–136 (1941).

Cramér, H., Mathematical methods of statistics (Princeton University Press, 1946).

Dantzig, D. van, 1) Kadercursus Mathematische Statistiek 1947–1950, Hoofdstuk VI, par. 3 (Mathematisch Centrum, Amsterdam).

——— 2) On the consistency and the power of Wilcoxon's two-sample test, Proc. Kon. Ned. Ak. van Wetensch. **A 54**, 3–10 (1951); Indagationes Mathematicae 13, 1–8 (1951).

——— and Hemelrijk, J., Statistical methods based on few assumptions, to appear in the International Statistical Institute (1953) (28th session, 6–12 September in Rome).

ESSEEN, C. G., Fourier analysis of distribution functions, A mathematical study of the Laplace-Gaussian law, Acta Math. **77**, 1–125 (1945).

FRÉCHET, M., Généralités sur les probabilités, Eléments aléatoires, 2nd ed.; Traité du calcul des probabilités et de ses applications, tome I, fasc. 3, premier livre, (Gauthier–Villars, Paris, 1950).

HEMELRIJK, J., Note on Wilcoxon's two-sample test when ties are present, Ann. Math. Stat. **23**, 133–135 (1952).

HOEFFDING, W., 1) A class of statistics with asymptotically normal distributions, Ann. Math. Stat. **19**, 293–325 (1948).

——— 2) A combinatorial central limit theorem, Ann. Math. Stat. **22**, 558–566 (1951).

KEMPERMANN, J. H., De verdelingsfunctie van het aantal inversies in de test van Mann en Whitney, Rapport T.W. no. 7 van de Afdeling Toegepaste Wiskunde, Mathematisch Centrum, (Amsterdam, 1950).

KENDALL, M. G., Rank correlation methods (London, 1948).

KRUSKAL, W. H., A non-parametric test for the several sample problem, Ann. Math. Stat. **23**, 525–540 (1952).

LEHMANN, E. L., 1) Consistency and unbiasedness of certain non-parametric tests, Ann. Math. Stat. **22**, 165–179 (1951).

——— 2) The power of rank tests, Ann. Math. Stat. **24**, 23–43 (1953).

——— and STEIN, C., On the theory of some non-parametric hypotheses, Ann. Math. Stat. **20**, 28–45 (1949).

MANN, H. B. and WHITNEY, D. R., On a test of whether one of two random variables is stochastically larger than the other, Ann. Math. Stat. **18**, 50–60 (1947).

NOETHER, G. E., On a theorem by Wald and Wolfowitz, Ann. Math. Stat. **20**, 455–458 (1949).

SUNDRUM, R. M., 1) The power or Wilcoxon's two-sample test, Journ. Roy. Stat. Soc. **15**–2, 246–252 (1953).

——— 2) A further approximation to the distribution of Wilcoxon's statistic in the general case, (1953), unpublished.

VAART, H. R. VAN DER, 1) Some remarks on the power function of Wilcoxon's test for the problem of two samples, I and II, Proc. Kon. Ned. Ak. van Wetensch. **53**, 494–506, 507–520 (1950); Indagationes Mathematicae **12** 146–158, 159–172 (1950).

——— 2) Cursus „Parametervrije Methoden", III. De toets van Wilcoxon voor het probleem van twee steekproeven, Mathematisch Centrum (Amsterdam, Mei 1951).

——— 3) An investigation on the power function of Wilcoxon's two-sample test if the underlying distributions are not normal, Proc. Kon. Ned. Ak. van Wetensch. **A 56**, 438–448 (1953); also Indagationes Mathematicae **15** (1953).

WAERDEN, B. L. VAN DER, 1) Order tests for the two-sample problem and their power, Proc. Kon. Ned. Ak. van Wetensch. **55**, 453–458 (1952); also Indagationes Mathematicae **14** (1952).

——— 2) Order tests for the two-sample problem, II and III, Proc. Kon. Ned. Ak. van Wetensch. **A 56**, 303–310, 310–316 (1953); also Indagationes Mathematicae **15** (1953).

WALD, A. and J. WOLFOWITZ, Statistical tests based on permutations of the observations, Ann. Math. Stat. **15**, 358–372 (1944).

WILCOXON, F., Individual comparisons by ranking methods, Biometrics Bull. **1**, 80–83 (1945).