BYLAAG A.

Indien

(2.40') $\lim_{N\to\infty} N^{-1}\sum_h (a_h - \bar{\Psi}_N)^2 \geq \varepsilon > 0$ en

(2.42') $N^{-1}\sum_h (\Psi_{Nh} - \bar{\Psi}_N)^4 = O(N)$

dan voldoen die variante $\hat{t}_i$ aan die voorwaarde

(2.50) $\lim_{N\to\infty} E|\hat{t}_i|^{2+\delta} < \infty$ vir $i=1,2,\ldots,k$.

Bewys: Stel $M_N = \sum_h (a_h - \bar{\Psi}_N)^2$ en

$$P_N = \sum_h (a_h - \bar{\Psi}_N)^4 \leq \sum_h (\Psi_{Nh} - \bar{\Psi}_N)^4 .$$

Op analoë wyse as in lemma 2.2 volg dat

$$E(\underline{t}_i - E\underline{t}_i)^4 = \frac{n_i(N-n_i)(N^2+N-6Nn_i+6n_i^2)}{N(N-1)(N-2)(N-3)} \cdot P_N + \frac{3n_i(n_i-1)(N-n_i)(N-n_i-1)}{N(N-1)(N-2)(N-3)} \cdot M_N^2 .$$

Uit (2.4) volg

$$\hat{t}_i = (N-n_i)^{\frac{1}{2}}[\underline{t}_i - E\underline{t}_i][n_i(N-n_i)N^{-1}(N-1)^{-1}M_N]^{-\frac{1}{2}}$$

$$= [\underline{t}_i - E\underline{t}_i][n_i(N-1)^{-1}M_N]^{-\frac{1}{2}}.$$

$$E|\hat{\underline{t}}_i|^4 = E[\underline{t}_i - E\underline{t}_i]^4[n_i(N-1)^{-1}M_N]^{-2}$$

$$= \frac{(N-1)(N-n_i)(N^2+N-6Nn_i+6n_i^2)}{n_iN(N-2)(N-3)M_N^2} \cdot P_N + \frac{3(N-1)(N-n_i)(N-n_i-1)(n_i-1)}{n_iN(N-2)(N-3)}$$

$$= O(1)$$

want $P_N M_N^{-2} \leq N^{-1}\sum_h (\Psi_{Nh} - \bar{\Psi}_N)^4 \, N^{-1}[N^{-1}\sum_h (a_h - \bar{\Psi}_N)^2]^{-2}$

$= O(1)$ uit (2.40') en (2.42').

Gevolglik geld $E|\hat{t}_i|^{2+\delta} = O(1)$ met $0 < \delta \leq 2$.

Opmerking.

Onder voorwaarde (2.40) geld voorwaarde (2.40') met waarskynlikheid een.

BYLAAG B.

Onder die voorwaardes (3.10) - (3.13), (3.28) en

(3.14') $$|J(H)| \leq K[H(1-H)]^{-\frac{1}{4}+\delta}$$
$$|J^{(m)}(H)| \leq K[H(1-H)]^{-m-\frac{1}{2}+\delta} \quad , \ m=1,2 \quad \text{en } \delta > 0$$

geld die voorwaarde

(3.76) $\lim\limits_{N\to\infty} E|\hat{t}_i|^{2+\delta} < \infty$ vir $i=1,2,\ldots,k$ en $0 < \delta \leq 2$.

Bewys: $$E|\hat{t}_i|^{2+\delta} = \left[\frac{N-n_i}{N\mathrm{var}(t_i)}\right]^{1+\frac{1}{2}\delta} E|t_i - Et_i|^{2+\delta} \text{ uit (3.56).}$$

Aangesien $0 < \lim\limits_{N\to\infty} N\mathrm{var}(t_i) < \infty$ uit (3.68), is dit voldoende om te bewys dat $N^{1+\frac{1}{2}\delta}E|t_i - Et_i|^{2+\delta} = O(1)$.

Uit §3.4 opmerking 3, (3.60) en (3.62) volg:
$E(t_i) = A^{(i)} + o(N^{-\frac{1}{2}})$ met $A^{(i)}$ gedefinieer deur (3.61). Nou is

$$N^{1+\frac{1}{2}\delta}E|t_i - Et_i|^{2+\delta} = N^{1+\frac{1}{2}\delta}E|B_{1N}^{(i)} + B_{2N}^{(i)} + o_p(N^{-\frac{1}{2}})|^{2+\delta}$$

$$= N^{1+\frac{1}{2}\delta}E|\sum_g \nu_{Ng}[n_i^{-1}\sum_{j=1}^{n_i}\tilde{B}_g(x_{ij}) - n_g^{-1}\sum_{j=1}^{n_g}\tilde{B}_i(x_{gj})] + o_p(N^{-\frac{1}{2}})|^{2+\delta}.$$

Die resultaat volg m.b.v. lemma 3.11 saam met (3.1) en (3.3) indien bewys kan word dat

$$E|n_i^{-\frac{1}{2}}\sum_{j=1}^{n_i}\tilde{B}_g(x_{ij})|^{2+\delta} = O(1).$$

Onder voorwaarde (3.14') word hieraan voldoen vir $\delta = 2$, want:

$$E|n_i^{-\frac{1}{2}}\sum_j \tilde{B}_g(x_{ij})|^4$$

$$= E[n_i^{-2}\sum_j\sum_{j'}\sum_{j''}\sum_{j'''}\tilde{B}_g(x_{ij})\tilde{B}_g(x_{ij'})\tilde{B}_g(x_{ij''})\tilde{B}_g(x_{ij'''})] \quad \text{waar}$$

alle j 's die waardes $1,2,\ldots,n_i$ deurloop

$$= n_i^{-2}[\sum_j\sum_{j'} E\tilde{B}_g^2(x_{ij})E\tilde{B}_g^2(x_{ij'}) + \sum_j E\,\tilde{B}_g^4(x_{ij})] \quad \text{want } x_{ij} \text{ en}$$

$x_{ij'}$ is onderling onafhanklik vir $j \neq j'$ en $E\tilde{B}_g(x_{ij}) = 0$

$= O(1)$

aangesien:

$n_i^{-2}\sum_j\sum_{j'} E\,\tilde{B}_g^2(x_{ij})E\,\tilde{B}_g^2(x_{ij'}) = O(1)$ m.b.v. (3.27) en

onder op bladsy 72, en

$n_i^{-2}\sum_j E\,\tilde{B}_g^4(x_{ij}) = O(1)$ soos blyk uit die volgende:

$$|B_g(x_{ij})| = |\int_{x_0}^{x_{ij}} J'[H(x)]dF_g(x)|$$

$$\leq K\,|\int_{x_0}^{x_{ij}} J'[H(x)]dH(x)|$$

$$\leq K\,[\,|J[H(x_{ij})]|+|J[H(x_0)]|\,]$$

$$\leq K'[H(x_{ij})\{1-H(x_{ij})\}]^{\delta-\frac{1}{4}} + O(1) \text{ uit } (3.14').$$

$$E[B_g(x_{ij})]^4 \leq K''\int_0^1 [H(1-H)]^{4(\delta-\frac{1}{4})}dH + O(1)$$

$$= K''\int_0^1 [H(1-H)]^{-1+4\delta}dH + O(1)$$

$$= O(1), \text{ sodat}$$

$$E[\tilde{B}_g(x_{ij})]^4 = O(1).$$

Hiermee is die bewys voltooi.

Opmerkings.

1). Waarskynlik is die meer beperkende voorwaarde (3.14') in vergelyking met (3.14) onnodig. Let in hierdie verband op dat $\lim_{N\to\infty} E\tilde{B}_g^2(x_{ij}) < \infty$ onder (3.14) - sien C+S(1958). Verder is nodig vir bogenoemde afleiding dat $\lim_{N\to\infty} n_i^{-2}\sum_j E\tilde{B}_g^4(x_{ij}) < \infty$ waarby reeds

$\lim_{N\to\infty} n_i^{-1}\sum_j E\tilde{B}_g^4(x_{ij}) < \infty$ as (3.14') geld.

Daar kan ook op gewys word dat

$|B_g(x_{ij})| < K[J\{H(x_{ij})\}]$ en dat

(3.13) $J_N(1) = o(N^{\frac{1}{2}})$.

2). Alle spesiale gevalle van die algemene toetsingsgrootheid wat in hierdie hoofstuk beskou word, voldoen aan die meer beperkende voorwaarde (3.14').

BYLAAG C.

Aan die voorwaardes

i) $\lim_{N\to\infty} E|t'_{i.}|^{2+\delta} = O(1)$,

ii) $\lim_{N\to\infty} E|t'_{.j}|^{2+\delta} = O(1)$ en

iii) $\lim_{N\to\infty} E|t''_{ij}|^{2+\delta} = O(1)$ met $\delta > 0$

word voldoen onder voorwaardes (5.13), (5.14) en

(5.13') $N^{-1}\sum_p (a_p - a_{...})^4 = O(N)$.

Bewys: Die bewys dat aan voorwaardes i) en ii) voldoen word, geskied soos in bylaag A.

Deur te skrywe

$a_{ijh} - a_{i..} - a_{.j.} + a_{...} = (a_{ijh} - a_{...}) - (a_{i..} - a_{...}) - (a_{.j.} - a_{...})$,

kan geredelik aangetoon word dat (5.13') voldoende is vir die geldigheid van voorwaarde iii) deur op 'n soortgelyke wyse as in bylaag A te werk te gaan met gebruikmaking van die tegniek wat in die bewys van lemma 5.2 ontwikkel is.

Die geldigheid van voorwaarde (5.13') word verder implisiet aanvaar.

SUMMARY.

DISTRIBUTION-FREE TESTS FOR THE PROBLEM OF TWO OR MORE SAMPLES.

Distribution-free tests are constructed for the k-sample problem ($k \geq 2$). Some of these tests are extensions of two-sample tests proposed by WILCOXON(1945), TERRY(1952), MOOD(1954), VAN DER WAERDEN(1957) and LEMMER and STOKER(1960).

The test statistics have asymptotic χ^2 distributions with (k-1) degrees of freedom under the null hypothesis H_o (chapter II), and also under the general hypothesis H subject to certain general conditions (chapter III - see CHERNOFF and SAVAGE(1958)). A number of these k-sample distribution-free tests for difference in location are compared with the parametric F test and in the case of tests for dispersion with the Bartlett test for homogeneity of variance by means of their asymptotic relative efficiencies. Several of the k-sample tests compare favourably with the parametric tests and are highly recommended for practical application. Necessary and sufficient conditions are given for the consistency of these tests.

Different tests are suggested for the problem of m rankings (chapter IV) and their asymptotic properties derived.

The analysis of variance in the two-factor case is studied in chapter V. Different tests for row-, column- and interaction effects are constructed, which, under H_o, possess asymptotic χ^2 distributions and are distributed independent of each other.

LITERATUURLYS.

1. ANDREWS, F. C.(1954) „Asymptotic behavior of some rank order tests for analysis of variance", Ann. Math. Stat. 25, pp. 724-736.

2. ANSARI, A. R. en BRADLEY, R. A.(1960) „Rank-sum tests for dispersions", Ann. Math. Stat. 31, pp. 1174-1189.

3. BENARD, A en VAN ELTEREN, P.(1953) „A generalization of the method of m rankings", Indag. Math. 15, pp. 358-369.

4. CAPON, J.(1961) „Asymptotic efficiency of certain locally most powerful rank tests", Ann. Math. Stat. 32, pp. 88-100.

5. CHACKO, V. J.(1963) „Testing homogeneity against ordered alternatives", Ann. Math. Stat. 34, pp. 945-956.

6. CHERNOFF, H. en SAVAGE, I. R. (1958) „Asymptotic normality and efficiency of certain nonparametric test statistics", Ann. Math. Stat. 29, pp. 972-994.

7. CRAMéR, H.(1946) „Mathematical Methods of Statistics", Princeton University Press, Princeton.

8. CRAMéR, H.(1962) „Random Variables and Probability Distributions", 2de uitgawe, Cambridge University Press.

9. CROUSE, C. F.(1960) „Combinatorial tests, for differences in location and dispersion, for the case of m samples", Proefskrif, Universiteit van London.

10. DWASS, M.(1956) „The large-sample power of rank order tests in the two-sample problem", Ann. Math. Stat. 27, pp. 352-374.

11. FRASER, D. A. S.(1957) „Nonparametric Methods in Statistics", New York.

12. FREUND, J. E. en ANSARI, A. R.(1957) „Two-way rank-sum tests for variances", Technical Report No. 34 to Office of Ordnance Research and National Science Foundation, Virginia Polytechnic Institute.

13. FRIEDMAN, M.(1937) „The use of ranks to avoid the assumption of normality implicit in the analysis of variance", Jour. Am. Stat. Ass. 32, pp. 675-699.

14. GIBSON, G. A.(1954) „Advanced Calculus", MacMillan and Co. Ltd., London.

15. HODGES, J. L. en LEHMANN, E. L.(1956) „The efficiency of some nonparametric competitors of the t-test", Ann. Math. Stat. 27, pp. 324-335.

16. HODGES, J. L. en LEHMANN, E. L.(1961) „Comparison of the normal scores and Wilcoxon tests", Proc. 4 th. Berkeley Symp. on Math. Stat. and Prob. 1, pp. 307-318.

17. HOEFFDING, W.(1948) „A class of statistics with asymptotically normal distribution", Ann. Math. Stat. 19, pp. 293-325.

18. HOEFFDING, W.(1951) „A combinatorial central limit theorem", Ann. Math. Stat. 22, pp. 558-566.

19. HOEFFDING, W.(1952) „The large sample power of tests based on permutations of observations", Ann. Math. Stat. 23, pp. 169-192.

20. KENDALL, M. G.(1948) „Rank Correlation Methods", London.

21. KENDALL, M. G. en STUART, A.(1958) „The advanced Theory of Statistics", vol.I, Griffin and Co. Ltd., London.

22. KENDALL, M. G. en STUART, A.(1961) „The advanced Theory of Statistics", vol. II, Griffin and Co., Ltd., London.

23. KENNEY, J. F. en KEEPING, E. S.(1951) „Mathematics

of Statistics", vol. II, 2^de uitgawe, D Van Nostrand, New York.

24. KLOTZ, J.(1962) „Nonparametric tests for scale", Ann. Math. Stat. 33, pp. 498-512.
25. KRUSKAL, W. H.(1952) „A non-parametric test for the several sample problem", Ann. Math. Stat. 23, pp. 525-540.
26. LEMMER, H. H. en STOKER, D. J.(1960) „'n Ondersoek na die doeltreffendheid van verdelingsvrye toetse vir die probleem van twee klein steekproewe om vir verskil in spreiding te toets", Tydskrif vir Wetenskap en Kuns, Okt. 1960, pp. 177-191.
27. LEMMER, H. H. en STOKER, D. J.(1961) „'n Klas van verdelingsvrye toetsingsgroothede vir die probleem van k onafhanklike steekproewe", Tydskrif vir Natuurwetenskappe, S. A. Akademie vir Wetenskap en Kuns, I, No. 4, Des. 1961, pp. 231-246.
28. LOèVE, M.(1955) „Probability Theory", 2^de uitgawe, D. Van Nostrand, New Jersey.
29. MOOD, A. M.(1950) „Introduction to the Theory of Statistics", McGraw-Hill Book Co., New York.
30. MOOD, A. M.(1954) „On the asymptotic efficiency of certain nonparametric two-sample tests", Ann. Math. Stat. 25, pp. 514-522.
31. NOETHER, G. E.(1949) „On a theorem by Wald and Wolfowitz", Ann. Math. Stat. 20, pp. 455-458.
32. PITMAN, E. J. G.(1937) „Significance tests which may be applied to samples from any populations", I, Jour. Roy. Stat. Soc. Suppl. 4, pp. 119-130.
33. RIJKOORT, P. J.(1952) „A generalisation of Wilcoxon's test", Indag. Math. 14, p. 394-404.
34. SCHEFFé, H.(1959) „The Analysis of Variance", New York.

35. SCHUTTE, H. J. en VAN ROOY, D. J.(1961) „Lineêre Algebra en Meetkunde", Pro Rege-Pers Bpk., Potchefstroom.

36. STOKER, D. J.(1955) „Oor 'n klas van toetsingsgroothede vir die probleem van twee steekproewe", Proefskrif, Universiteit van Amsterdam.

37. SUKHATME, B. V.(1957) „On certain two-sample nonparametric tests for variances", Ann. Math. Stat. 28, pp. 188-194.

38. SUKHATME, B. V.(1958) „Testing the hypothesis that two populations differ only in location", Ann. Math. Stat. 29, pp. 60-78.

39. TANG, P. C.(1938) „The power function of the analysis of variance tests with tables and illustrations of their use", Statistical Research Memoirs, vol. II, University College, London, pp. 126-149.

40. TATE, M. W. en CLELLAND, R. C.(1957) „Nonparametric and shortcut statistics in the social, biological and medical sciences", Interstate Printers and Publishers, Danville, Illinois.

41. TERPSTRA, T. J.(1955) „A generalization of Kendall's rank correlation statistic", I, Indag. Math. 17, pp. 690-696.

42. TERPSTRA, T. J.(1956) „A generalization of Kendall's rank correlation statistic", II, Indag. Math. 18, pp. 59-66.

43. TERPSTRA, T. J.(1962) „Interpretations and generalizations of Kendall's rank correlation test", Proefskrif, Rijksuniversiteit, Groningen.

44. TERRY, M. E.(1952) „Some rank order tests, which are most powerful against spesific parametric alternatives", Ann. Math. Stat. 23, pp. 346-366.

45. TURNBULL, W. H. (1947) „Theory of Equations", Oliver and Boyd, London.

46. USPENSKY, J. V. (1937) „Introduction to Mathematical Probability", McGraw-Hill, New York.

47. VAN DANTZIG, D. (1947-1950) „Kadercursus Mathematische Statistiek", Amsterdam, Mathematisch Centrum.

48. VAN DER WAERDEN, B. L. (1957) „Mathematische Statistik", Springer-Verlag, Berlyn.

49. VAN ELTEREN, P. (1957) „The asymptotic distribution for large m of Terpstra's statistic for the problem of m rankings", Indag. Math. 19, pp. 522-534.

50. VAN ELTEREN, P. en NOETHER, G. E. (1959) „The asymptotic efficiency of the χ_r^2-test for a balanced incomplete block design", Biometrika 46, p. 475.

51. WILCOXON, F. (1945) „Individual comparisons by ranking methods", Biometrics 1, pp. 80-83.

52. WILKS, S. S. (1962) „Mathematical Statistics", Wiley and Sons, New York.