STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

S 97 (V 3)

Asymptotische Eigenschappen van Wilcoxon's Toets

Voordracht van Prof.dr. D. van Dantzig

Verslag door J. Hemelrijk

Asymptotische eigenschappen van Wilcoxon's toets

Voordracht van Prof. Dr D. van Dantzig
Verslag door J. Hemelrijk
Mathematisch Centrum, Amsterdam

Prof. Van Dantzig's betoog was ongeveer het volgende.
Gegeven:

1e: m waarnemingen x_i $(i = 1, \ldots, n)$ van een stochastische grootheid $\underline{x}$ [1]) en n waarnemingen y_j $(j = 1, \ldots, m)$ van $\underline{y}$; deze waarnemingen zijn onderling volledig onafhankelijk.

2e: de verdelingsfuncties F van $\underline{x}$ en G van $\underline{y}$ zijn continu.

De getoetste hypothese is:

$$H_0 : \quad F(z) = G(z) \text{ voor iedere } z. \tag{1}$$

De vraag, die hier behandeld wordt is: voor *welke alternatieve hypothesen, d.w.z. paren (F,G) van verdelingsfuncties, die niet aan H_0 voldoen, kan men F en G met behulp van de toets van elkaar onderscheiden?*

Mann en Whitney [4] hebben voor de links-éénzijdige toetsing (dus met een kritieke zône van de vorm $\underline{U} \leqq U_0$ [2])) bewezen, dat o.a. alle paren van de klasse H' met:

$$H' \begin{cases} F(z) \geqq G(z) \text{ voor iedere } z \\ \text{en } F(z) > G(z) \text{ voor minstens één } z, \end{cases} \tag{2}$$

hiertoe behoren. Scherper geformuleerd: indien aan (2) voldaan is, nadert de kans op verwerping van H_0, voor onbeperkt toenemende m en n (terwijl m/n en n/m begrensd blijven) tot 1. Voor dergelijke paren heet de toets *bruikbaar* (of *asymptotisch onderscheidend*). H' bevat echter niet alle alternatieve hypothesen, waarvoor de toets asymptotisch onderscheidend is.

Bij vergelijking met de toets van Student voor het analoge probleem bij normaal verdeelde grootheden, zien wij, dat de klasse der bij die toets toegelaten alternatieve hypothesen een deelklasse van H' is. Immers bij Students toets wordt gewoonlijk ondersteld, dat de gemiddelden van $\underline{x}$ en $\underline{y}$ eventueel verschillend zijn, maar dat de spreidingen in ieder geval gelijk zijn. Interessant is nu echter de vraag, wat er gebeurt, indien de spreidingen van $\underline{x}$ en $\underline{y}$ ongelijk zijn. Zijn $\underline{x}$ en $\underline{y}$ normaal verdeeld, dan bezitten F en G precies één snijpunt, zodat de stelling van Mann en Whitney niets meer over dit geval zegt. Prof. Van Dantzig bewees echter in [2] het volgende:

[1]) Stochastische grootheden zijn onderstreept.

[2]) Uit de stelling voor linker-eenzijdige toetsing volgen direct analoge stellingen voor rechter-eenzijdige en voor tweezijdige toetsing.

De grootheid

$$p = \mathrm{P}[\underline{y} < \underline{x}] \tag{3}$$

bepaalt, of de toets bij voldoende grote n en m tot verwerping van H_0 zal leiden of niet. Voor een paar

$$H = (F, G) \tag{4}$$

is de links-éénzijdige toets asymptotisch onderscheidend indien de bij H behorende waarschijnlijkheid p van (3) kleiner dan $\frac{1}{2}$ is.

Een analoge stelling geldt uiteraard voor de rechtséénzijdige toetsing ($p < \frac{1}{2}$) en voor de tweezijdige ($p \neq \frac{1}{2}$).

Deze klasse van alternatieve hypothesen is veel ruimer dan de door Mann en Whitney aangegeven klasse H'.

Wij hebben n.l.:

$$p = \int_{-\infty}^{+\infty} G(z)\,\mathrm{d}F(z). \tag{5}$$

Zijn $\underline{x}$ en $\underline{y}$ normaal verdeeld met gemiddelden μ_1 en μ_2 en spreidingen σ_1 en σ_2, dan is dus:

$$p = \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}} \mathrm{d}x \int_{-\infty}^{x} e^{-\frac{1}{2}\frac{(y-\mu_2)^2}{\sigma_2^2}} \mathrm{d}y.$$

Substitueren wij hierin $v = y - \mu_2$ en $u = x - \mu_1$, dan komt er:

$$p = \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}\frac{u^2}{\sigma_1^2}} \mathrm{d}u \int_{-\infty}^{u+(\mu_1-\mu_2)} e^{-\frac{1}{2}\frac{v^2}{\sigma_2^2}} \mathrm{d}v, \tag{6}$$

zodat p, behalve van σ_1 en σ_2, slechts van $(\mu_1 - \mu_2)$ afhankelijk is. Verder ziet men uit (6) direct, dat

$$\left\{\begin{array}{l} p > \frac{1}{2} \text{ voor } \mu_1 - \mu_2 > 0. \\ p = \frac{1}{2} \text{ voor } \mu_1 - \mu_2 = 0. \\ p < \frac{1}{2} \text{ voor } \mu_1 - \mu_2 < 0. \end{array}\right. \tag{7}$$

Derhalve is de toets voor dit geval asymptotisch onderscheidend zodra $\mu_1 \neq \mu_2$ is (bij tweezijdige toetsing), dat wil zeggen zodra de gemiddelden verschillend zijn, onafhankelijk van de spreidingen.

Het resultaat kan, voor het algemene geval ook aldus worden uitgedrukt:

De toets is bruikbaar, indien de mediaan van de stochastische grootheid $\underline{x} - \underline{y}$ van 0 verschilt.

De getoetste hypothese H_0 houdt echter meer in dan alleen, dat deze mediaan gelijk aan o is. Op grond daarvan alleen kan nl. de waarschijnlijkheidsverdeling van de toetsingsgrootheid $\underline{U}$ niet berekend worden.

Ten slotte vermeldde Prof. Van Dantzig, dat de klasse van alternatieve hypothesen, waarvoor de toets asymptotisch onderscheidend is, niet kan worden uitgebreid. Voor $p = \frac{1}{2}$ is de toets, voor niet te grote α, niet meer asymptotisch onderscheidend, ook al zijn F en G verschillend. Hoe groot men dan n en m ook neemt, er is steeds een positieve waarschijnlijkheid, die niet tot o nadert, dat H_0 niet zal worden verworpen. De door de spreker aangegeven klasse van alternatieve hypothesen bevat dus precies alle hypothesen, waarvoor de toets bruikbaar is.

LITTERATUUR.

[1] H. R. van der Vaart, *Some remarks on the power function of Wilcoxon's test for the problem of two samples*, I and II, Proc. Kon. Ned. Ak. v.Wet. **53** (1950) pp. 494—506 and 507—520; also in Indagationes Mathematicae **12** (1950) pp. 146—158 and 159—172.

[2] D. van Dantzig, *On the consistency and the power of Wilcoxons two sample test*, Proc. Kon. Ned. Ak. v. Wet., Series A, vol. **54** (1951) pp. 3—10; also in Indagationes Mathematicae **13** (1951) pp. 1—8.

[3] F. Wilcoxon, *Individual comparisons by ranking methods*, Biometrics Bull. **1** (1945) pp. 80—83.

[4] H. B. Mann and D. R. Whitney, *On a test of whether one of two random variables is stochastically larger than the other*, Ann. of Math. Stat. **18** (1947) pp. 50—60.

[5] H. R. van der Vaart, *Gebruiksaanwijzing voor de toets van Wilcoxon*, Rapport S 32 (M 4), Mathematisch Centrum, Amsterdam (1950).

[6] E.L. Lehmann, *Consistency and unbiasedness of certain nonparametric tests*, Ann. Math. Stat. **22** (1951) pp. 165—179.

[7] J. Hemelrijk en H. R. van der Vaart, *Het gebruik van één- en tweezijdige overschrijdingskansen voor het toetsen van hypothesen*, Statistica **4** (1950) pp. 54—66.