

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

S 73 (M 13a)

Trend-toets met behulp van rangcorrelatie



Statistische Afdeling.

S 73 (M 13a)

Trend-toets met behulp van rangcorrelatie¹⁾

Men kan de methode der rangcorrelatie toepassen om na te gaan of er in een reeks waarnemingen van een stochastische grootheid y een trend, d.w.z. een stijgend of dalend verloop aanwezig is. Men maakt daarbij gebruik van het feit, dat bij een systeem' van n waarnemingsparen (x_i, y_i) de grootte van S^2 en de verdeling van S onder de hypothese van onafhankelijkheid niet verandert als men de volgorde van de paren wijzigt. Men kan dus bij de berekening van S de volgorde der paren (x_i, y_i) zodanig kiezen, dat de rangnummers van bijvoorbeeld de x_i de rij 1, 2, 3, ..., n vormen. De hypothese van onafhankelijkheid komt dan overeen met de hypothese H_0 , dat alle mogelijke permutaties van de rangnummers der y_i even waarschijnlijk zijn.

Indien wij nu willen onderzoeken of er een trend aanwezig is in de rij waarnemingen $y_1, \dots, y_n$ van de stochastische grootheid y , dan voegen wij de rij 1, 2, ..., n toe en bepalen vervolgens S zoals aangegeven in memorandum S 47 (M 13) en toetsen bovengenoemde hypothese H_0 . Onder deze hypothese heeft S dezelfde verdeling als de overeenkomstige grootheid der rangcorrelatie-toets onder de hypothese van onafhankelijkheid; wij gebruiken hier ook dezelfde kritieke zônes. Indien S een positieve (resp. negatieve) waarde heeft behorende tot de kritieke zône spreken wij van een significant stijgende (resp. dalende) trend. Voor literatuur over deze toets zie men [1].

Deze toets kan worden uitgebreid tot meerdere onafhankelijke reeksen waarnemingen, die niet alle even groot behoeven te zijn. Bij voorbeeld:

$y_{1,1}, y_{1,2}, \dots, y_{1,n_1}$

$y_{2,1}, y_{2,2}, \dots, y_{2,n_2}$

$y_{3,1}, y_{3,2}, \dots, y_{3,n_3}$

$\vdots$

$y_{m,1}, y_{m,2}, \dots, y_{m,n_m}$

1) Dit memorandum is een aanvulling op memorandum S 47 (M 13). Het is slechts bedoeld ter orientatie en streeft niet naar volledigheid of volledige exactheid.

2) Zie voor de definitie van S memorandum S 47 (M 13).

Wij bepalen nu de S voor het stelsel paren $(1, y_{11})(2, y_{12}) \dots (n_1, y_{1n_1})$ en noemen deze S_1 . Evenzo behandelen wij de tweede rij en verkrijgen zo S_2 etc.

Wij bepalen dan de som

$$S_{\text{tot}} = S_1 + S_2 + \dots + S_m.$$

Eveneens berekenen wij bij de eerste rij de spreiding σ_1 , volgens formule (1) of formule (2) van memorandum S 47 (M 13). Evenzo $\sigma_2, \sigma_3, \dots, \sigma_m$.

Wij weten nu, dat onder de hypothese H_0 voor rij 1 $\underline{S}_1$ bij benadering normaal verdeeld is met gemiddelde 0 en spreiding σ_1 , indien H_0 geldt voor rij 2 is $\underline{S}_2$ bij benadering normaal verdeeld met gemiddelde 0 en spreiding σ_2 etc. Indien hypothese H_0 geldt voor alle rijen, dus indien in alle rijen alle permutaties der rangnummers van de waarnemingen even waarschijnlijk zijn, en indien deze rijen bovendien onafhankelijk zijn, dan geldt dat $\underline{S}_{\text{tot}}$ bij benadering normaal verdeeld is met gemiddelde 0 en spreiding:

$$\sigma_{\text{tot}} = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_m^2}.$$

Wij kiezen nu als kritieke zône weer $|S_{\text{tot}}| \geq S_0$ bij tweezijdige toetsing, $S_{\text{tot}} \geq S'_0$ bij rechtszijdige en $S_{\text{tot}} \leq S''_0$ bij linkszijdige toetsing. Voor literatuur over deze toets zie men [2].

Literatuur: [1] H.B.Mann: Non parametric tests against trend, *Econometrica* 13 (1945) blz. 245-259.

[2] G.Elfvig en J.H.Whitlock: A simple trend-test with application to erythrocyte data, *Biometrics* 6 (1950) Blz. 282-288.