STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

S 73 (M 17a)

Het combineren van onafhankelijke toetsen

Dr. J. Hemelrijk Jr.

1952

Stichting
MATHEMATISCH CENTRUM
2e Boerhaavestraat 49
A m s t e r d a m - O.

April 1952.
(herzien april 1962)

Statistische Afdeling
S 73 (M 17 a)

Het combineren van onafhankelijke toetsen[1].

A. Tweezijdige toetsing van éénzelfde of een aantal soortgelijke hypothesen.

1. Eenvoudige beschrijving van de methode.

Wij beschouwen eerst een zeer eenvoudig geval, waarbij één hypothese H_o een aantal (en wel h) malen getoetst wordt door toepassing van één bepaalde toets op de resultaten van h onafhankelijke experimenten. Zij $\underline{t}$ de toetsingsgrootheid[2], waarop de toetsen berusten en laten de bij de h toetsen door $\underline{t}$ aangenomen waarden $t_1, t_2, \ldots, t_h$ zijn.

De linker-éénzijdige overschrijdingskansen, die bij deze uitkomsten behoren, zijn

$$k_1' = P[\underline{t} \leq t_1 \mid H_o], \ldots, k_h' = P[\underline{t} \leq t_h \mid H_o],$$

waarin het symbool H_o, achter de verticale streep, aangeeft, dat deze kansen berekend zijn in de onderstelling, dat H_o juist is. De rechter-éénzijdige overschrijdingskansen zijn

$$k_1'' = P[t \geq t_1 \mid H_o], \ldots, k_h'' = P[\underline{t} \geq t_h \mid H_o].$$

Wij berekenen nu de twee producten

$$k_1' . k_2' . \ldots . k_h' \quad \text{en} \quad k_1'' . k_2'' . \ldots . k_h'',$$

en nemen van deze twee het kleinste. Dit noemen wij K. Vervolgens berekenen wij

$$Q = - 2 \ln K$$

en zoeken in een tabel of nomogram van de χ^2-verdeling de overschrijdingskans

$$k_{\chi^2} = P[\underline{\chi}^2 \geq Q]$$

op, waarbij als aantal vrijheidsgraden 2h genomen wordt. De tweezijdige overschrijdingskans, behorende bij de gecombineerde toetsingen van H_o, is dan gelijk aan $2k_{\chi^2}$.

----

1) Dit memorandum dient slechts ter oriëntatie en streeft niet naar volledigheid of volledige exactheid.

2) De onderstreping geeft aan, dat de toetsingsgrootheid stochastisch is, d.w.z. een waarschijnlijkheidsverdeling bezit.

2. Opmerkingen.

In plaats van de kleinste van de twee producten van éénzijdige overschrijdingskansen te gebruiken, kan men ook alleen het product van de tweezijdige overschrijdingskansen der oorspronkelijke toetsingen berekenen en dit als K nemen. De overschrijdingskans k_{χ^2} behoeft dan niet meer verdubbeld te worden. Dit is het procédé, dat in de literatuur ([1], [2][3]) wordt aangegeven. De bovenbeschreven methode bezit echter vermoedelijk in vele gevallen een aanzienlijk groter onderscheidingsvermogen en is daarom te prefereren. Een publicatie hierover is in voorbereiding.

Indien de toetsingsgrootheid $\underline{t}$ een discontinue verdeling bezit, is bovenstaande methode in zoverre nog van toepassing, dat de gevonden waarde van $2k_{\chi^2}$ te groot is. Dat wil zeggen, indien men H_0 slechts dan verwerpt, als

$$2k_{\chi^2} \leq \alpha$$

is, is de kans op ten onrechte verwerpen van H_0 kleiner dan α (dus de onbetrouwbaarheid van de methode wordt op deze wijze overschat). Dat het verschil aanzienlijk kan zijn blijkt uit [3].

De opheffing van deze moeilijkheid kan o.a. bereikt worden door de verdeling van $\underline{t}$ continu te maken; een beschrijving daarvan vindt men in [4].

3. Het algemene geval.

In het algemene geval, waarin bovenstaande methode kan worden toegepast, zijn er h hypothesen $H_1, H_2, \ldots, H_h$, die getoetst worden met h toetsingsmethoden, $T_1, T_2, \ldots T_h$. De hypothesen en de toetsen kunnen nu verschillend zijn; de toetsen berusten echter steeds op h onderling onafhankelijke stelsels van waarnemingen. Van ieder der toetsen wordt ondersteld, dat de kritieke zone één- of tweezijdig kan zijn, dus dat een tweezijdige kritieke zone (bij gegeven onbetrouwbaarheidsdrempel α) in twee delen uiteenvalt, die wij de linker- en rechter-éénzijdige kritieke zone noemen.

Indien nu de verzameling van alternatieve hypothesen, waartegen het stelsel $H_1, H_2, \ldots, H_h$ getoetst wordt, van dien aard is, dat bij éénzijdige toetsing alle hypothesen links-éénzijdig of alle rechts-éénzijdig zouden worden getoetst, is de boven beschreven methode van toepassing.

----

3) Zie literatuurlijst aan het einde van dit memorandum.

4. Tabellen.

Tabellen van de χ^2-verdeling vindt men in vrijwel ieder boek over statistiek. B.v. in M.G. Kendall, The advanced theory of statistics, Griffin and Co., London 1947,
W.J. Dixon and F.J. Massey, Introduction to statistical analysis, Mc Graw-Hill, N.Y.- Toronto-London 1951.

Een speciale tabel voor het combineren van toetsen, die enkele (overigens eenvoudige) berekeningen bespaart, is te vinden in
F.N. David, On the P_λ test for randomness, Biometrika 26 (1934) pp. 1-11.

Literatuur.

[1] R.A. Fisher, Statistical methods for research workers, Oliver and Boyd, London, 4th ed., 1932.

[2] E.S. Pearson, The probability integral transformation for testing goodness of fit and combining independent tests of significance, Biometrika 30 (1938) pp. 134-148.

[3] W.A. Wallis, Compounding probabilities from independent significance tests, Econometrica 10 (1942) pp. 229-248.

[4] E.S. Pearson On questions raised by the combination of tests based on discontinuous distributions, Biometrika 37 (1950) pp. 383-398.