

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

S 47 (M9)

De toets van Student voor 2 steekproeven.



De toets van Student voor 2 steekproeven¹⁾.

Deze toets van Student wordt gebruikt voor het toetsen van de hypothese $\mathcal{H}_0$, dat 2 steekproeven $x_1, x_2, \dots, x_n$ en $y_1, y_2, \dots, y_m$ beschouwd kunnen worden als steekproeven uit eenzelfde normale verdeling.

De grootheid $\underline{t}$, welke hierbij als toetsingsgrootheid gebruikt wordt en waarvan Student de verdelingsfunctie berekend heeft, is de volgende:

$$\underline{t} = \frac{\underline{m}_x - \underline{m}_y}{S} \sqrt{\frac{nm}{n+m}}$$

Hierin zijn $\underline{m}_x$ en $\underline{m}_y$ de gemiddelden van de 2 steekproeven terwijl S een schatting is van de spreiding σ van de normale verdeling, waaruit volgens $\mathcal{H}_0$ de steekproeven getrokken zijn. De bij de gevonden steekproeven behorende waarde van $\underline{m}_x$, $\underline{m}_y$ en S zijn op volgende wijze uit de waarnemingen te berekenen:

$$\underline{m}_x = \frac{1}{n} (x_1 + x_2 + \dots + x_n)$$

$$\underline{m}_y = \frac{1}{m} (y_1 + y_2 + \dots + y_m)$$

$$S^2 = \frac{\sum_{i=1}^n (x_i - \underline{m}_x)^2 + \sum_{j=1}^m (y_j - \underline{m}_y)^2}{n+m-2}$$

Bovengenoemde grootheid $\underline{t}$ bezit als verdeling de verdeling van "Student" met $\nu = n+m-2$ vrijheidsgraden.

Deze verdeling is getabelleerd voor $\nu = 1$ t/m 20 [1] voor $\nu = 20$ t/m 100 zijn waarden van $\underline{t}$ berekend, welke behoren bij de onbetrouwbaarheidsdrempels $\alpha = 0,01$ en $\alpha = 0,05$. [2]

Is $\mathcal{H}_0$ juist, dan zullen dicht bij nul gelegen waarden vaker voorkomen, dan ver van nul gelegen waarden. De kritieke zône kiezen we dus zodanig, dat ze gevormd worden door grote en kleine waarden van $\underline{t}$. De kritieke zône bestaat dus, bij gegeven waarde van α , uit twee stukken: $\underline{t} < -t_\alpha$ en $\underline{t} > t_\alpha$, waarin t_α de bij α en ν behorende kritieke waarde is, die in [1] en [2] getabelleerd is. Voor éénzijdige toetsing is de kritieke zône van de vorm $\underline{t} \leq -t_\alpha$ of $\underline{t} \geq t_\alpha$.

1) Dit memorandum is slechts bedoeld ter orientatie en streeft niet naar volledigheid of volledige exactheid.

Literatuur:

- [1]. M.G.Kendall, The Advanced Theory of Statistics,
London 1946, Vol. II p.109, tabellen in
deel I p. 440-41;
Opmerking: het bij de tabellen vermelde
aantal vrijheidsgraden (aangegeven door
 ν) is gelijk aan $n + m - 2$.
- [2]. Elizabeth M.Baldin, Table of Percentage Points of the
t-distribution, Biometrika 33 (1943),
p. 362. (Deze tabel worden tweezijdige
overschrijdingskansen gegeven).