

Lösung Aufgabe 27

Zur Theorie:

- **Definitionen:**

Sei (x_i, y_i) das Beobachtungstupel des i -ten Merkmalsträgers. Wir betrachten nun ein Paar von Beobachtungstupeln (x_i, y_i) und (x_j, y_j) . Sei o.B.d.A. $x_i < x_j$. Das Paar heißt **konkordant**, wenn auch $y_i < y_j$. Das Paar heißt **diskordant**, wenn $y_i > y_j$. Ist ein Paar weder konkordant noch diskordant, ist es ein **Tie** bzw. eine **Bindung**, d.h. entweder gilt $x_i = x_j$ oder $y_i = y_j$ oder beides.

- **Bezeichnungen:**

Insgesamt gibt es $\binom{n}{2} = \frac{n(n-1)}{2}$ Paare, die man überprüfen muss. Dabei gelten die folgenden Bezeichnungen:

- N_c = Anzahl der konkordanten Paare
- N_d = Anzahl der diskordanten Paare
- T_x = Anzahl der Paare mit $x_i = x_j$ aber $y_i \neq y_j$ (x-Ties)
- T_y = Anzahl der Paare mit $y_i = y_j$ aber $x_i \neq x_j$ (y-Ties)
- T_{xy} = Anzahl der Paare mit $x_i = x_j$ und zugleich $y_i = y_j$
- Da für jedes zu überprüfende Paar nur einer der oberen fünf Fälle in Frage kommt, gilt also: $\binom{n}{2} = \frac{n(n-1)}{2} = N_c + N_d + T_x + T_y + T_{xy}$

- Kendall's Tau beim Vorliegen von Ties (T_{xy} spielen keine Rolle beim Überprüfen der Korrelation!):

$$\tau = \frac{N_c - N_d}{\sqrt{(N_c + N_d + T_x)(N_c + N_d + T_y)}}$$

- **Interpretation:**

Es gilt: $\tau \in [-1; 1]$. Dabei zeigt eine positive Korrelation an, dass die Ränge beider Variablen gemeinsam steigen, während eine negative Korrelation anzeigt, dass die Ränge der einen Variable steigen, wenn die Ränge der anderen abnehmen.