The Wilcoxon rank-sum test

Natalija Jovanovic 226/15

Neka su X i Y dva nezavisna obelezja. Pretpostavimo da su raspodele X i Y istog oblika, sa istom dispezijom (i simetricne su).

$H0 : m_X = m_Y$, gde je $m_X = E(X)$ i $m_Y = E(Y)$.

Napravimo objedinjen uzorak $X_1, X_2, \ldots, X_n, Y_{n+1}, \ldots, Y_{n+m}$. Neka je r_i rang elementa u objedinjenom uzorku, a z_i indikator da je element sa rangom r_i iz prvog uzorka. Tada je test statistika:

$$T = \sum_{i=1}^{n+m} \underbrace{r_i z_i}_{u_i}$$

$$u_i = \begin{pmatrix} i & 0 \\ \frac{n}{n+m} & \frac{m}{n+m} \end{pmatrix}$$

$E(u_i) = i\frac{n}{n+m}$

$E(T) = \sum_{i=1}^{n+m} E(u_i) = \frac{n}{n+m}\sum_{i=1}^{n+m} i = \frac{n}{n+m}\frac{(n+m)(n+m+1)}{2} = \frac{n(n+m+1)}{2}$

$$u_i^2 = \begin{pmatrix} i^2 & 0 \\ \frac{n}{n+m} & \frac{m}{n+m} \end{pmatrix} w_{ij} = \begin{pmatrix} ij & 0 \\ \frac{n(n-1)}{(n+m)(n+m-1)} & 1 - \frac{n(n-1)}{(n+m)(n+m-1)} \end{pmatrix}$$

$E(u_i^2) = i^2\frac{n}{n+m}$

$E(w_{ij}) = ij\frac{n(n-1)}{(n+m)(n+m-1)}$

$E(T^2) = E((\sum_{i=1}^{n+m} r_i z_i)^2) = E(\sum_{i=1}^{n+m} (r_i z_i)^2 + 2\sum_{i=1}^{n+m}\sum_{j=i+1}^{n+m} r_i z_i r_j z_j) =$
$\sum_{i=1}^{n+m} E((r_i z_i)^2) + 2\sum_{i=1}^{n+m}\sum_{j=i+1}^{n+m} E(\underbrace{r_i z_i r_j z_j}_{w_{ij}}) = \frac{n}{n+m}\sum_{i=1}^{n+m} i^2 + \frac{2n(n-1)}{(n+m)(n+m-1)}\sum_{i=1}^{n+m}\sum_{j=i+1}^{n+m} ij =$
$\frac{n}{n+m}\frac{(n+m)(n+m+1)(2(n+m)+1)}{6} + \frac{2n(n-1)}{(n+m)(n+m-1)}\sum_{i=1}^{n+m} i \sum_{j=i+1}^{n+m} j =$
$\frac{n(n+m+1)(2(n+m)+1)}{6} + \frac{2n(n-1)}{(n+m)(n+m-1)}\sum_{i=1}^{n+m} i(\frac{(n+m)(n+m+1)}{2} - \frac{i(i+1)}{2}) =$
$\frac{n(n+m+1)(2(n+m)+1)}{6} + \frac{2n(n-1)}{(n+m)(n+m-1)}(\frac{(n+m)(n+m+1)}{2}\sum_{i=1}^{n+m} i - \sum_{i=1}^{n+m}\frac{i^2(i+1)}{2}) =$
$\frac{n(n+m+1)(2(n+m)+1)}{6} + \frac{2n(n-1)}{(n+m)(n+m-1)}(\frac{(n+m)^2(n+m+1)^2}{4} - \frac{(n+m)^2(n+m+1)^2}{8} - \frac{(n+m)(n+m+1)(2(m+n)+1)}{12})$

$D(T) = E(T^2) - E^2(T) = \frac{n(n+m+1)(2(n+m)+1)}{6} + \frac{2n(n-1)}{(n+m)(n+m-1)}(\frac{(n+m)^2(n+m+1)^2}{4} - \frac{(n+m)^2(n+m+1)^2}{8} - \frac{(n+m)(n+m+1)(2(m+n)+1)}{12}) - \frac{n^2(n+m+1)^2}{4} = \text{WolframAlpha} = \frac{mn(n+m+1)}{12}$