

一般化 Wilcoxon 検定

青木繁伸

1 目的

一般化 Wilcoxon 検定を行う。

2 使用法

```
from generalized_wilcoxon_test import generalized_wilcoxon_test
generalized_wilcoxon_test(group, event, time, verbose=True)
```

2.1 引数

group	群を識別するベクトル (2 種類の値を取るなら整数値に限らない)
event	死亡なら 1, 生存なら 0 の値をとるベクトル
time	生存期間ベクトル
verbose	必要最小限のプリント出力をする

2.2 戻り値の名前

"W"	検定統計量
"V(W)"	検定統計量の分散
"Z"	正規化得点
"p value"	p 値

3 使用例

```
group = [1, 1, 2, 2, 1, 1, 2, 2, 1, 2, 2, 1, 1, 1,
         2, 2, 2, 1, 1, 2, 1, 2, 1, 2, 1, 2, 2, 1, 1, 1]
event = [1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0,
         0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0]
time = [2, 20, 5, 1, 3, 17, 2, 3, 15, 14, 12, 13, 11,
        11, 10, 8, 8, 3, 7, 3, 6, 2, 5, 4, 2, 3, 1, 3, 2, 1]

import sys
sys.path.append("statlib")
from generalized_wilcoxon_test import generalized_wilcoxon_test

a = generalized_wilcoxon_test(group, event, time)

generalized Wilcoxon test
```

$W = 63$, $V(W) = 1320.3$, $Z = 1.7338$, $p \text{ value} = 0.083$

グループの指定を文字列で行った場合は、辞書順で早い方が第1群になる。

下の例では "female" が "male" より辞書順で早いので、"female" が第1群とされる。この場合、上の例と W と Z の符号が逆になる。しかし p 値は同じである（どちらを第1群にしても検定の結果は変わらない）。

```
group2 = ["male", "male", "female", "female", "male", "male",
          "female", "female", "male", "female", "female", "male",
          "male", "male", "female", "female", "female", "male",
          "male", "female", "male", "female", "male", "female",
          "male", "female", "female", "male", "male", "male"]
event = [1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 1, 1, 0, 0,
         0, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0, 0]
time = [2, 20, 5, 1, 3, 17, 2, 3, 15, 14, 12, 13, 11,
        11, 10, 8, 8, 3, 7, 3, 6, 2, 5, 4, 2, 3, 1, 3, 2, 1]
a = generalized_Wilcoxon_test(group2, event, time)
```

generalized Wilcoxon test

$W = -63$, $V(W) = 1320.3$, $Z = -1.7338$, $p \text{ value} = 0.083$