

# NAG Toolbox for MATLAB

## g08aa

### 1 Purpose

g08aa performs the Sign test on two related samples of size  $n$ .

### 2 Syntax

```
[isgn, n1, p, ifail] = g08aa(x, y, 'n', n)
```

### 3 Description

The Sign test investigates the median difference between pairs of scores from two matched samples of size  $n$ , denoted by  $\{x_i, y_i\}$ , for  $i = 1, 2, \dots, n$ . The hypothesis under test,  $H_0$ , often called the null hypothesis, is that the medians are the same, and this is to be tested against a one- or two-sided alternative  $H_1$  (see below).

g08aa computes:

- (a) the test statistic  $S$ , which is the number of pairs for which  $x_i < y_i$ ;
- (b) the number  $n_1$  of non-tied pairs ( $x_i \neq y_i$ );
- (c) the lower tail probability  $p$  corresponding to  $S$  (adjusted to allow the complement  $(1 - p)$  to be used in an upper one-tailed or a two-tailed test).  $p$  is the probability of observing a value  $\leq S$  if  $S < \frac{1}{2}n_1$ , or of observing a value  $< S$  if  $S > \frac{1}{2}n_1$ , given that  $H_0$  is true. If  $S = \frac{1}{2}n_1$ ,  $p$  is set to 0.5.

Suppose that a significance test of a chosen size  $\alpha$  is to be performed (i.e.,  $\alpha$  is the probability of rejecting  $H_0$  when  $H_0$  is true; typically  $\alpha$  is a small quantity such as 0.05 or 0.01). The returned value of  $p$  can be used to perform a significance test on the median difference, against various alternative hypotheses  $H_1$ , as follows

- (i)  $H_1$ : median of  $x \neq$  median of  $y$ .  $H_0$  is rejected if  $2 \times \min(p, 1 - p) < \alpha$ .
- (ii)  $H_1$ : median of  $x >$  median of  $y$ .  $H_0$  is rejected if  $p < \alpha$ .
- (iii)  $H_1$ : median of  $x <$  median of  $y$ .  $H_0$  is rejected if  $1 - p < \alpha$ .

### 4 References

Siegel S 1956 *Non-parametric Statistics for the Behavioral Sciences* McGraw-Hill

### 5 Parameters

#### 5.1 Compulsory Input Parameters

- 1: **x(n)** – double array
- 2: **y(n)** – double array

$x(i)$  and  $y(i)$  must be set to the  $i$ th pair of data values,  $\{x_i, y_i\}$ , for  $i = 1, 2, \dots, n$ .

#### 5.2 Optional Input Parameters

- 1: **n** – int32 scalar

*Default:* The dimension of the arrays **x**, **y**. (An error is raised if these dimensions are not equal.)  
 $n$ , the size of each sample.

*Constraint:*  $n \geq 1$ .

### 5.3 Input Parameters Omitted from the MATLAB Interface

None.

### 5.4 Output Parameters

- 1: **isgn – int32 scalar**  
The Sign test statistic,  $S$ .
- 2: **n1 – int32 scalar**  
The number of non-tied pairs,  $n_1$ .
- 3: **p – double scalar**  
The lower tail probability,  $p$ , corresponding to  $S$ .
- 4: **ifail – int32 scalar**  
0 unless the function detects an error (see Section 6).

## 6 Error Indicators and Warnings

Errors or warnings detected by the function:

**ifail** = 1

On entry, **n** < 1.

**ifail** = 2

**n1** = 0, i.e., the samples are identical.

## 7 Accuracy

The tail probability,  $p$ , is computed using the relationship between the binomial and beta distributions. For  $n_1 < 120$ ,  $p$  should be accurate to at least 4 significant figures, assuming that the machine has a precision of 7 or more digits. For  $n_1 \geq 120$ ,  $p$  should be computed with an absolute error of less than 0.005. For further details see g01ee.

## 8 Further Comments

The time taken by g08aa is small, and increases with  $n$ .

## 9 Example

```
x = [4;
      4;
      5;
      5;
      3;
      2;
      5;
      3;
      1;
      5;
      5;
      5;
      4;
      5;
```

```
5;  
5;  
5];  
y = [2;  
3;  
3;  
3;  
3;  
3;  
3;  
3;  
2;  
3;  
2;  
2;  
2;  
5;  
2;  
5;  
3;  
1];  
[isgn, n1, p, ifail] = g08aa(x, y)  
  
isgn =  
      3  
n1 =  
     14  
p =  
  0.0287  
ifail =  
      0
```

---