



Tutorials and worked examples for simulation,
 curve fitting, statistical analysis, and plotting.
<http://www.simfit.org.uk>

Sometimes a robust estimate is required for the difference in location (with corresponding confidence limits) for two samples, not necessarily of the same size, but without assuming normality or any other distribution.

From the main SIMFIT menu choose [Statistics], navigate to [Data exploration] and open the option for [Robust analysis of two samples]. The two default test files are `ttest.tf4` and `ttest.tf5` with these values

ttest.tf4	ttest.tf5
134	70
146	118
104	101
119	85
124	107
161	132
107	94
83	
113	
129	
97	
123	

while analysis produces the following results.

Robust analysis of two samples	
X-sample size	12
Y-sample size	7
Difference in location	-18.501
Lower confidence limit	-40.009
Upper confidence limit	2.997
Percentage confidence limit	95.30%
Lower Mann-whitney U-value	19.000
Upper Mann-Whitney U-value	66.000

The procedure is based on the assumption that X of size n_x is distributed as $F(x)$ and Y of size n_y as $F(x - \theta)$, so an estimate $\hat{\theta}$ for the difference in location is calculated as

$$\hat{\theta} = \text{median}(y_j - x_i, i = 1, 2, \dots, n_x, j = 1, 2, \dots, n_y).$$

$100\alpha\%$ confidence limits U_L and U_H are then estimated by inverting the Mann-Whitney U statistic so that

$$P(U \leq U_L) \leq \alpha/2$$

$$P(U \leq U_L + 1) > \alpha/2$$

$$P(U \geq U_H) \leq \alpha/2$$

$$P(U \geq U_H - 1) > \alpha/2.$$