Abstract
The presence of aberrant observations (i.e. outliers) in cell lineage data is quite common. Such as, it is desirable to have an outlier-resistance estimation procedure as an alternative to least squares estimation (maximum likelihood estimation under normality). In this work, we consider rank-based estimates of the parameters of a first order bifurcating autoregressive (BAR(1)) model. The BAR(1) model was proposed by Cowan and Staudte (1986) for cell lineage data. It in each line of descendents follows a first order autoregressive (AR(1)) model and allows sister cells from the same mother to be correlated. Real examples and a simulation study are performed in order to examine the behavior of these rank-based estimation procedures. More specifically, we compute finite sample relative efficiencies with respect to least squares estimate. The results indicate that the rank-based estimation procedures are more efficient when outlying observations are present.

First-Order Bifurcating Autoregressive Model
Let \( X_t \) be an observed cell in a culture of some quantitative characteristic at time \( t \), starting with the initial value \( x_0 \), the zero mean BAR(1) model which proposed by Cowan and Staudte (1986), is defined as
\[
X_t = \phi X_{t-1} + \epsilon_t, \quad \text{for all } t \geq 1
\]
where \( \phi \) denotes the largest integer less than or equal to \( \phi \). It is assumed that \( \{x_0, \epsilon_1, \epsilon_2, \ldots \} \) is a sequence of independent and identically bivariate random vectors with common mean zero, and common variance-covariance structure
\[
\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}
\]
where \( \rho \) is the correlation between \( (x_t, \epsilon_t) \) they follow a distribution \( F[\mu, \sigma] \) denotes the mother-daughter correlation coefficient which need to be estimated. The sister-sister correlation is defined as \( \rho = \phi^2 (1 - \phi) \).

Comments on The Previous Works
Most of previous works used estimation methods such as Maximum Likelihood estimation. Most Maximum Likelihood estimation to robust against outliers, and Least squares estimation. The presence of aberrant observations in cell lineage data is quite common. The existence of outliers in data makes the data deviate from normality. Classical methods often have very poor performance in presence of outliners. Then outliers have a highly affect on estimated BAR parameters by using these estimations.

The Goal of This Work
This work aims to propose a robust estimation to the BAR model depending on Rank-Based procedure, or sometimes is called Weighted Wilcoxon (WW) procedure. Beyond the theoretical proves and their complications, we will consider an empirical study depends on real examples and a simulation study in order to examine the behavior of these rank-based estimation procedures. More specifically, we compute finite sample relative efficiencies with respect to least squares estimate.

In fact this work is an extension to Terpstra and et al. (1997, 2000, 2001a, 2001b, 2001c) works. Terpstra and others proposed robust estimations of Autoregressive model based on Weighted Wilcoxon (WW) procedure.

Example(1): EMT6 cells (tree 41 of Staudte, Guiget and Collin d’Hooge(1984)), lifetimes in Tenths of Hours

Example(2): EMT6 cells (tree 29 of Staudte, Guiget and Collin d’Hooge(1984)), lifetimes in Tenths of Hours


Let the observations numbers 11 and 21 are outliers comparing to the rest observations and practically the initial cell often is missing.

Simulation Study
The behavior of the Rank-Based estimates is studied via Monte Carlo. Our primary interest is the behavior of the mother-daughter parameter, only the zero mean BAR(1) is considered. The distribution of errors is determined according to a contaminated normal distribution. 


correlation is defined as \( \rho = \phi^2 (1 - \phi) \). Theoretically proves and their complications, we will consider an empirical study depends on real examples and a simulation study in order to examine the behavior of these rank-based estimation procedures. More specifically, we compute finite sample relative efficiencies with respect to least squares estimate.

In fact this work is an extension to Terpstra and et al. (1997, 2000, 2001a, 2001b, 2001c) works. Terpstra and others proposed robust estimations of Autoregressive model based on Weighted Wilcoxon (WW) procedure.

References: