On Robust Parameter Estimation for Phase I Linear Profile Monitoring

E. E. Nwezza[1]; C. K. Acha[2]

[1]Department of Mathematics/Computer Science/Statistics/Informatics
Federal University Ndufu-Alike, Ikwo.

[2]Department of Statistics,
[2]Micheal Okpara University of Agriculture, Umudike.

***Abstract*- We consider the use of t-distribution which is a heavy tailed distribution in the estimation of parameters of linear profile when the data are not normally distributed. The estimates of parameters of the linear profile obtained from this approach are compared with estimates obtained from two other approaches; Huber function (which has heavy tail than normal distribution) which is a robust approach and least square method. The results obtained indicate that the new approach produced better estimates at three (3) degree of freedom and the estimates with the least square approach at ten thousand (10000) degree of freedom.**

Keywords- *Profile monitoring, Statistical process control, simple linear profile, Phase I, robust estimates, $t-$ distribution.*

I. Introduction

In Statistical Process Control (SPC), some quality and process characteristics may be better explained as a function of some independent or explanatory variables. Such a situation in SPC is usually referred to as "profile". Monitoring of profile entails monitoring of the parameters of the profile over time to know if there is change in the profile as a result of change in the parameters of the profile. This involves two phases; phase I which is a retrospective phase in which the parameters of the profile are estimated from historical data sets which are used to construct a control chart to determine whether the profiles are in- statistical control, and phase II which involves the future profile monitoring based on the in-statistical control chart established in phase I.

Consider $Y_k = f(x_{jk}, \beta_{jk}) + \varepsilon_k$ where Y_k, is the quality or process characteristics (response variable) of the k^{th} profile $k = 1, 2, \ldots, m$, f defines the functional form between Y_k and x_{jk}. f may be linear or nonlinear. x_{jk} is the j^{th} explanatory variable of the k^{th} profile $j = 1, 2, \ldots, l$, β_{jk} is the effect of the j^{th} explanatory variable on the k^{th} profile, ε_k is the error term which is usually assumed to be independently identically distributed $(i.i.d)$ normal random variable with mean zero and variance σ^2. The error term accounts for other factors which cannot be explained by the explanatory variable but affect the response variable.

A number of researchers [1, 2, 3, 4, 5] have considered linear profile monitoring assuming that the functional form f of the response variable Y_k with respect to explanatory variable is linear and the error term is independent identically distributed normal. They used least square method of estimation "classical approach" to estimate the parameters of the profile for Phase I. The least square method is known for its computational ease, its estimated parameters are optimal and it is the maximum likelihood estimators for the parameters of the linear function when the error terms are normally distributed [6]. The assumptions of independence and normality of the error terms however, do not always hold [7, 8, 9]. [9] considered the effect of non-normality on phase I and concluded that the false alarm rate of phase I increases in the presence of non-normality of the error terms. [10] studied the effect of non-normality and autocorrelation on linear profile and noted that the violation of normality and independent identical distribution assumption affect the performance of the control chart and may lead to misjudgement of the process status.

Robust methods have been developed to reduce the effect of outliers on estimators and to produce estimates which are optimal around the neighbourhood of the assumed model [11] and [12].

In profile monitoring, [13] considered nonparametric $L-1$ regression methods, [14] considered the use of weighted functions, Huber and bisquare. The Huber $M-$estimate is known to have normal distribution between the interval $[-k,k]$ and exponential distribution outside the interval which provides the least favourable distribution [12]; where k is the tuning parameter. According to [11], the Huber least favourable distribution appears to have longer tail than the normal distribution and observation that is farther away from other observations that the Huber's least favourable distribution cannot accommodated may be discarded. And this may lead to having estimates which do not reflect all the possible information contained in the observations. In this paper, we will consider a distribution which has longer tail than the Huber's least favourable distribution with a view to accommodating far outlying observation(s) and at the same time make robust the effect of the outliers on the estimated parameters of the profile. Section 2 considers the formulation of the model. Section 3 deals with the estimation of parameters, sections 4, 5, and 6 deal with the construction of Phase I control chart, Implementation and conclusion respectively.

II. MATERIALS AND METHODS

Let $y_{ki} = f(x_{kij}, \beta_{kj}) + e_{ki}$ defines the k^{th} functional relationship between the response and the explanatory variables, where $k = 1, 2, \ldots, m$ is the number of profiles, $i = 1, 2, \ldots, n$ is the number of observations in the k^{th} profile, and $j = 1, 2, \ldots, l$ is the number of explanatory variables in the k^{th} profile. It is assumed that the form of functional relationship f is linear and e_{ki}'s is iid random variable with mean zero and variance σ^2 from t distribution. We have a linear model with unknown parameters $\theta = (\beta_j, \sigma^2)'$ given by

$$y_{ki} = \sum_{j=0}^{l} \beta_{kj} x_{kij} + e_{ki} \text{ ; where } x_{ki0} = 1$$

This can be re-written in a matrix form $Y_k = X_k \beta + \varepsilon_k$, $Y_k = (y_{k1}, y_{k2}, \ldots, y_{kn})'$ is iid having mean $X_k \beta$ and variance σ_k^2, $\varepsilon_k = (e_{k1}, e_{k2}, \ldots, e_{kn})'$ where $\beta = (\beta_{k0}, \beta_{k1}, \ldots, \beta_{kl})'$ $X_k = n_k \times l_k$ matrix of explanatory variables.

We consider a simple linear relationship between the response variable Y_k and the explanatory variable X_k given by $Y_k = X_k \beta_k + \varepsilon_k$, where $X_k = (1_{ki0}, x_{ki1})$ and $\beta_k = (\beta_0, \beta_1)'$. y_{ki}'s follows univariate $t-$distribution with mean $X_k \beta_k$, variance σ_k^2, and degree of freedom v which is given by

$$f(y; \beta_k, \sigma_k^2, v) = \frac{\Gamma(\frac{v+1}{2})}{(v\pi)^{1/2} \Gamma(\frac{v}{2}) (\sigma_k^2)^{1/2}} \left(1 + \frac{(y - X_k \beta_k)^2}{v \sigma_k^2}\right)^{-(\frac{v+1}{2})}$$

$$-\infty < y < \infty$$

$$\sigma^2 > 0, v > 0$$

A. ESTIMATION METHOD

The maximum likelihood (ML) estimators of the k^{th} profile parameters of the univariate $t-$distribution are given by

$$\hat{\beta}_k = (X_k' W_k X_k)^{-1} X_k' W_k Y_k, \qquad (1)$$

$$\hat{\sigma}_k^2 = \frac{(Y_k - X_k \hat{\beta}_k)' W_k (Y_k - X_k \hat{\beta}_k)}{n}$$

where

$$W_k = \begin{bmatrix} w_{k1} & 0 & & . & . & . & 0 \\ 0 & & w_{k2} & 0 & 0 & 0 & 0 \\ 0 & 0 & & . & 0 & 0 & . \\ . & . & & 0 & . & 0 & . \\ . & . & & . & 0 & . & . \\ 0 & 0 & & . & . & 0 & w_{kn} \end{bmatrix}, \text{ and}$$

$$w_{ki} = \frac{v+1}{v + \frac{(Y_k - X_k \beta_k)^2}{\sigma_k^2}}$$

is the weight assigned to each i which down-weights outlying observations. [15] noted that the degree of down-weighting of outliers increases as the degree of freedom v decreases and the estimation is form of $M-$estimation of [11], yielding robust estimates.

The ML estimators above correspond to system of nonlinear equations. The solution to this is achieved through iterative methods. The expectation maximization (EM) algorithm is used to determine the ML estimates. According to [16], the EM algorithm provides a general approach of computing the maximum likelihood of iteratively reweighted least squares among many others. The algorithm involves two steps; the expectation step followed by the maximization step.

B. CONTROL CHART

Consider the construction of control limits from historical data set (HDS) to determine whether the profiles are statistically in-control. The k^{th} profile $y_{ki} = \beta_{k0} + \beta_{k1}x_{ki1} + e_{ki}$ parameters which are used in construction of the control limits are usually unknown but are estimated from the HDS. A particular profile will be statistically in-control only when the estimated parameters of the profile are within the control limit. If a particular profile is found to be outside the limits, the profile is removed and the control limits is re-constructed. The estimators of the parameters for the control chart are given as

$$\hat{\beta} = \frac{\sum_{k=1}^{m} \hat{\beta}_k}{m}, \text{ and } \hat{\sigma}^2 = \frac{\sum_{k=1}^{m} \hat{\sigma}_k^2}{m}$$ where $\hat{\beta}_k$ and $\hat{\sigma}_k^2$ are estimated parameters from the k^{th} profile obtained from equation (1) and (2).

The covariance matrix S between the intercepts β_{ko} and the slope β_{k1} is given as

$$S = \sum_{k=1}^{m} S_k \quad \text{where} \quad S_k = \frac{(v+3)}{(v+1)}(X'X)^{-1}\sigma_k^2$$ is given by (Lange et al. 1989).

III. RESULT AND DISCUSSION

The partial regression adjusted axial response and axial forces data set of [17] were used to test this new approach. The first 10 profiles of the data set are used and each profile is of the first 63 observations. This is to ensure a balanced data set. Q-Q plot is used to test the normality of the observations of the profiles. The graph of the residuals Q-Q plot (Appendix A) of each profile indicates that the observations of each profile do not follow normal distribution and there is presence of outlier(s). The degree of freedom of the $t-$distribution is considered to be known and it is fixed at $v = 3$. Lange et al (1989) noted that fixing v priori at some reasonable value serves as robustness tuning parameter. However, as $v \to \infty$ the $t-$distribution tends to normal distribution as shown in Table 1'

Table1: Shows the estimates of the intercept and slope of each profile assuming normality of the profile observations, Huber Psi function and t-distribution.

Profile	Normal		Huber Psi function		t- distribution (df = 3)		t- distribution (df = 10000)	
	Intercept (s.d. error)	Slope (s.d. error)	Intercept (s.d. error)	Slope (s.d. error)	Intercept (s.d. error)	Slope (s.d. error)	Intercept (s.d. error)	Slope (s.d. error)
1	9.19067 (1.53698)	21.01043 (0.09406)	9.914273 (1.147741)	21.075233 (0.071669)	9.59333 (1.07012)	21.04818 (0.06652)	9.19139 (1.48589)	21.01047 (0.09398)
2	12.51793 (1.32720)	21.01441 (0.07912)	12.25573 (1.220862)	21.03470 (0.071643)	12.34472 (0.98961)	21.03412 (0.05726)	12.51778 (1.29584)	21.01443 (0.07907)
3	15.43731 (1.21173)	21.03126 (0.07135)	15.40820 (1.100840)	21.04764 (0.063526)	15.3572 (0.8953)	21.0470 (0.0512)	15.43728 (1.18769)	21.03128 (0.07133)
4	13.37306 (1.23833)	21.02059 (0.07615)	13.48609 (1.115475)	21.05278 (0.065750)	13.48371 (0.91658)	21.04278 (0.05406)	13.37335 (1.26556)	21.02062 (0.07611)
5	15.66052 (1.23819)	21.01520 (0.07393)	15.62588 (1.030532)	21.04116 (0.061113)	15.68596 (0.89595)	21.03573 (0.05215)	15.66055 (1.21206)	21.01522 (0.07392)
6	13.30175 (1.26554)	21.14188 (0.07525)	13.43622 (1.091846)	21.17619 (0.062791)	13.40757 (0.92749)	21.16643 (0.05335)	13.30201 (1.23816)	21.14190 (0.07522)
7	10.47084 (1.26585)	21.14060 (0.07196)	10.01981 (0.950704)	21.16393 (0.055044)	10.15082 (0.86593)	21.16528 (0.04985)	10.47017 (1.18726)	21.14063 (0.07192)
8	12.05117 (1.29595)	21.04610 (0.07433)	11.58730 (1.081016)	21.06525 (0.061868)	11.63786 (0.96328)	21.06775 (0.05489)	12.0507 (1.2951)	21.0461 (0.0743)
9	7.59535 (1.21100)	21.04789 (0.06759)	7.002119 (1.031929)	21.057667 (0.055534)	7.14891 (0.90495)	21.05841 (0.04826)	7.59473 (1.21152)	21.04790 (0.06757)
10	9.30926 (1.21112)	21.04804 (0.06804)	8.94910 (1.079191)	21.06202 (0.058145)	9.04297 (0.90508)	21.06478 (0.04853)	9.30893 (1.21112)	21.04805 (0.06802)

The result of the intercepts and slopes of the ten profiles as indicated in Table 1 shows that the estimates of $t-$distribution with $v=3$ are more efficient than the Huber psi function, least square approach and $t-$distribution of $v=10000$. This is evident as the standard error of the estimates of the $t-$distribution with $v=3$ is smallest than that of the Huber Psi function estimates, least square and $t-$distribution of $v=10000$.

VI. Conclusion

This paper has considered the use of $t-$distribution to model simple linear profile as robust approach when the profile data are not normally distributed usually caused by outliers. The data set of the partial adjusted axial response and axial force of[17] was used and the estimates of the $Y-$intercept and slope of the simple linear profile were evaluated using the $t-$distribution, Huber psi function and least square approach. The results indicate that the standard error of the estimates of $t-$distribution with 3-degree of freedom is the smallest when compared with the Huber psi function, the least square method. This shows that at 3-degree of freedom, using $t-$distribution to model linear profile when the data are not normally distributed produces better estimates. However, at 10000-degree of freedom $t-$distribution estimates tend to estimates obtained using the least square approach.

References

[1] Kang L. and Albin S. L.(2000). On-line monitoring when the process yields a linear profile. *Journal of Quality Technology, 32.418-426.*

[2] Kim K., Mahmoud A. M., and Woodall W. H. (2003). On the monitoring of linear profiles. *Journal of Quality Technology, 35: 317-328.*

[3] Mahmoud M.A. (2008). Phase I analysis of multiple linear regression profiles. *Communications in Statistics, Simulation and computation, 37: 2106-2130.*

[4] Gupta S., Montgomery D.C., and Woodall W. H. (2006). Performance evaluation of two methods for online monitoring of linear calibration profiles. *International journal of production research, 44: 1927-1942*

[5] Kazemzadeh R.B., Noorossana R., and Amiri A. (2008). Phase I monitoring of polynomial profiles. *Communications in Statistics-Theory and Methods, 37:1671-1686.*

[6] Jureckova J. and Piecek J. (2005). *Robust Statistics with R.* Chapman and Hall/CRC

[7] Jensen, W.A., Birch, J.B., and Woodall, W.H. (2008). Profile Monitoring via Linear Mixed Models. Journal of Quality Technology

[8] Williams D.J., Woodall W.H., and Birch B.J. (2007). Statistical monitoring of Nonlinear product and process quality profiles. *Quality and Reliability Engineering international, 23:925-941*

[9] Noorossana R., Vaghefi A., and Dorri M. (2010). Effect of non-normality on the monitoring of simple linear profile. *Quality and reliability engineering international, 27: 425-436.*

[10] Noorossana R., Saghaei A., and Dorri M. (2010). Linear profile monitoring in the presence of non-normality and autocorrelation. *International journal of Industrial Engineering and Production research, 21.4.*

[11] Huber P. J. (1986). *Robust Statistics*. John wiley &sons Inc.

[12] Hampel F. R., Ronchetti E. M., Rousseeuw P. J., and Stahel W. A. (1986). *Robust Statistics: The Approach based on influence function*. John wiley & sons Inc

[13] Wei Y., Zhao Z., and Lin D. K. J. (2012) Profile control charts based on nonparametric $L-1$ regression methods. *The annals of Applied statistics, 6.1:409-427*

[14] Ebadi M. and Shahriari H. (2014). Robust estimation of parameters in simple linear profiles using M-estimators. *Journal of communications in statistics-theory and Methods, 43:4308-4323*

[15] Lange K. L., Little R. J.A., and Taylor J. M. G. (1989). Robust Statistical Modeling using t distribution. *Journal of American statistical Association, 84: 881-896.*

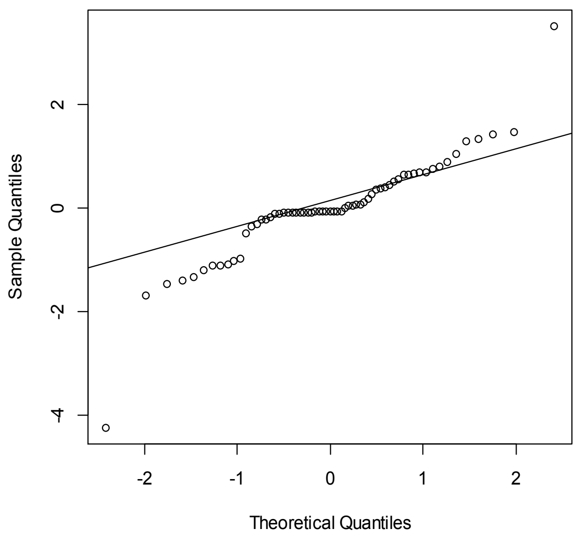
[16] Dempster A. P., Laird N. M., and Rubin D. B. (1977) Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the royal statistical society,39:1-38.*

[17] Noorossana R., Saghaei A., Amiri A. (2011). *Statistical analysof linear profile*. John wiley &sons, New Jersey.

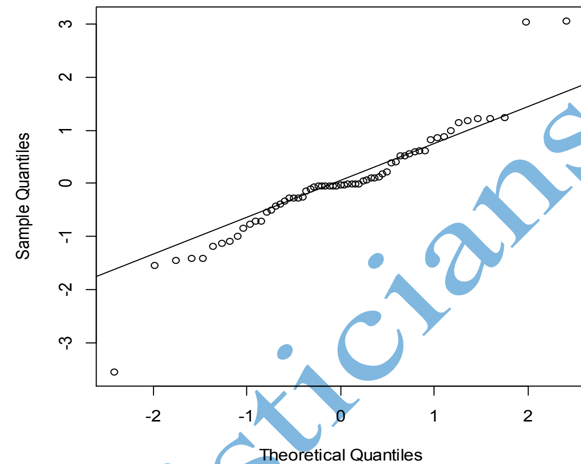
[18] Zou C., Tsung F., and Wang Z. (2008). Monitoring profile on nonparametric regression methods. *Technometrics, 50:512-526*

APPENDIX A

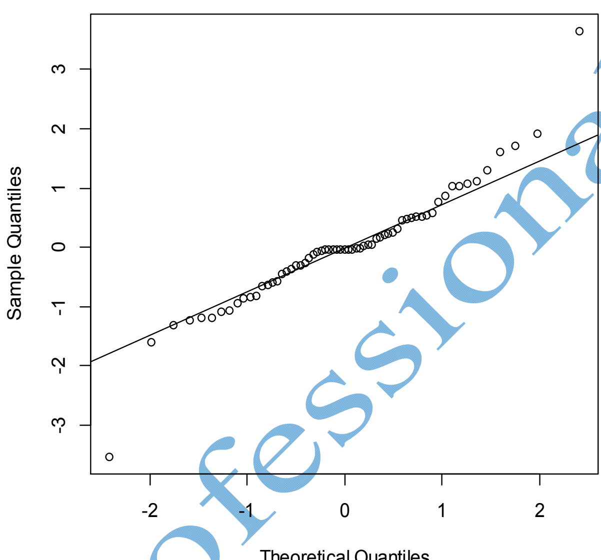
Standardized Residuals Q-Q plot of profile 1



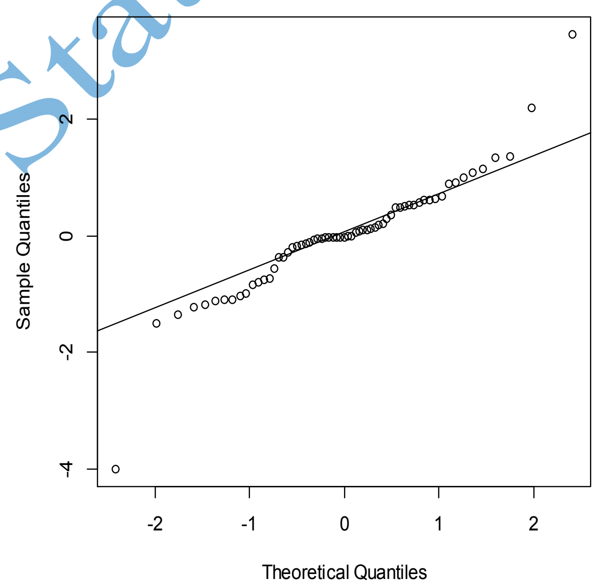
Standardized Residuals Q-Q plot of profile 3



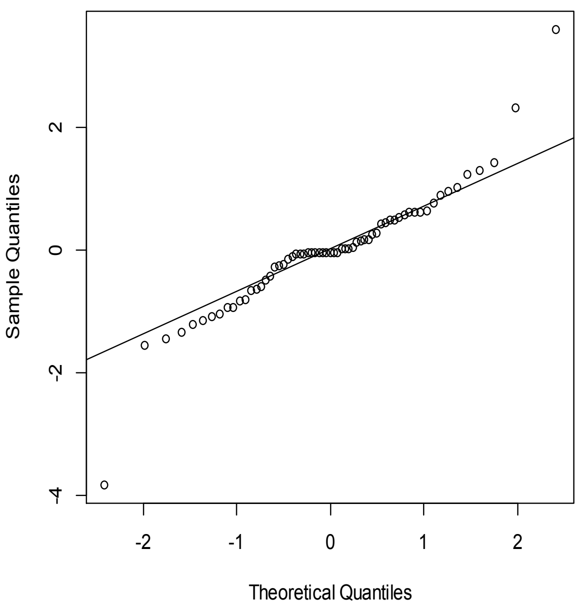
Standardized Residuals Q-Q plot of profile 2



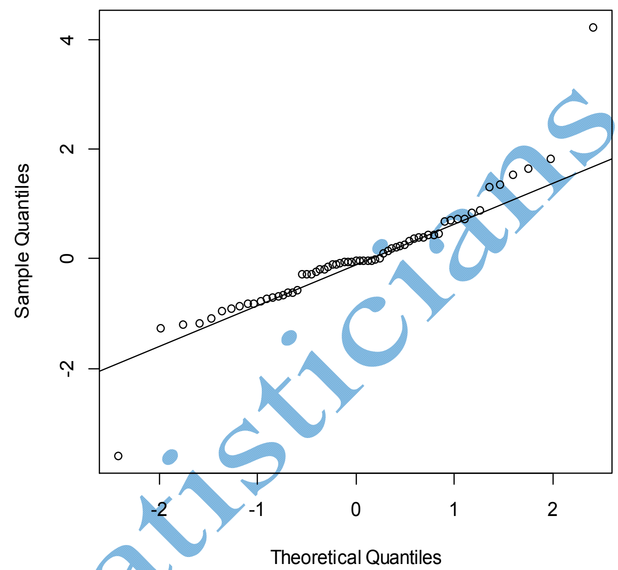
Standardized Residuals Q-Q plot of profile 4



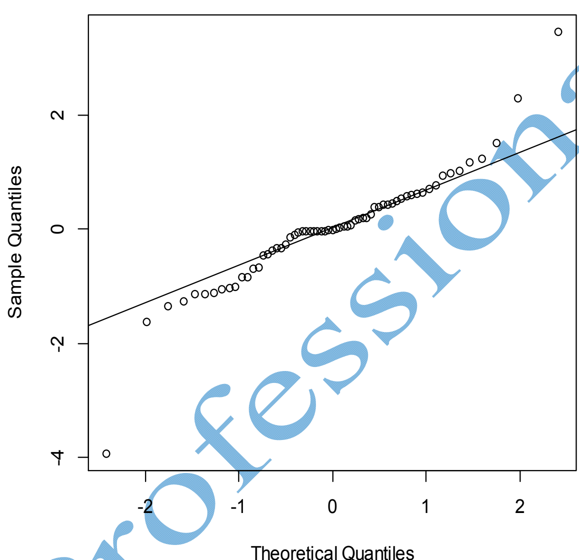
Standardized Residuals Q-Q plot of profile 5



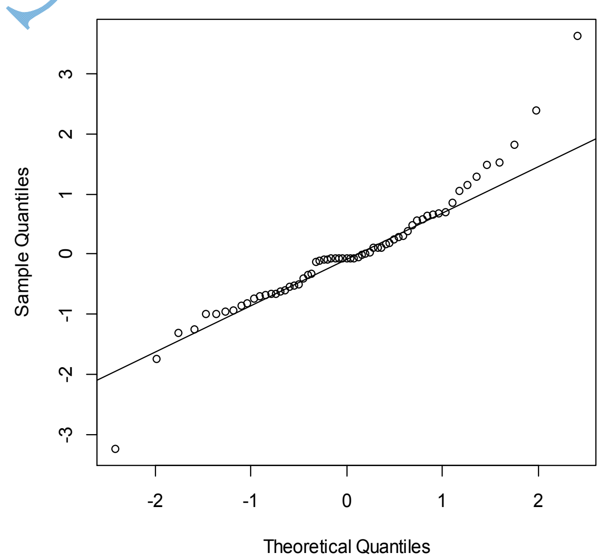
Standardized Residuals Q-Q plot of profile 7



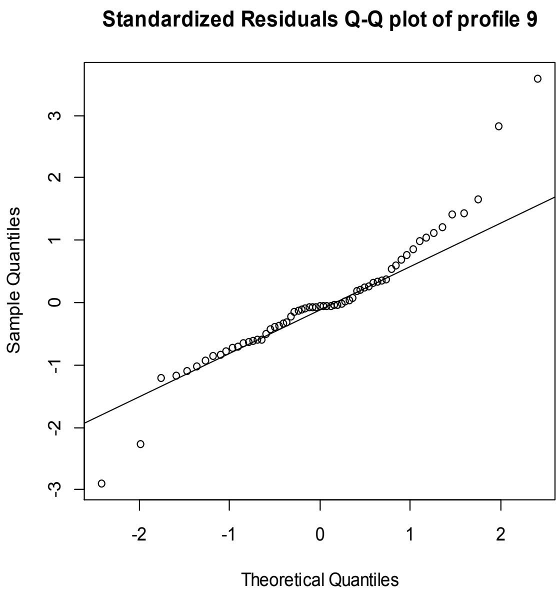
Standardized Residuals Q-Q plot of profile 6

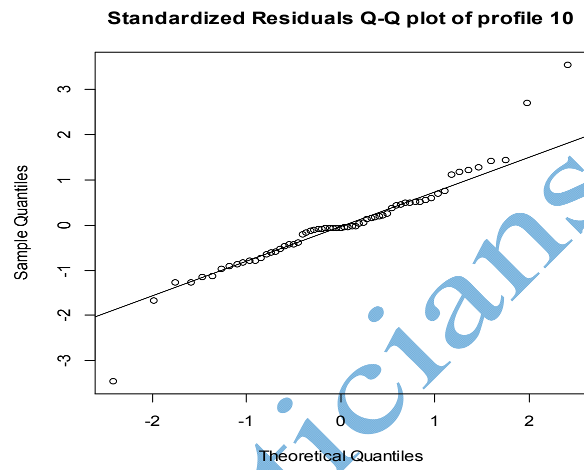


Standardized Residuals Q-Q plot of profile 8



,


Standardized Residuals Q-Q plot of profile 9
Sample Quantiles
Theoretical Quantiles


Standardized Residuals Q-Q plot of profile 10
Sample Quantiles
Theoretical Quantiles