ORDER STATISTICS IN OUTLIER MODELS

KEREM TÜRKYILMAZ

JUNE 2009

ORDER STATISTICS IN OUTLIER MODELS

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF

NATURAL AND APPLIED SCIENCES OF

IZMIR UNIVERSITY OF ECONOMICS

BY

KEREM TÜRKYILMAZ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

IN THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

JUNE 2009

M.S. THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"ORDER STATISTICS IN OUTLIER MODELS"** completed by **KEREM TÜRKYILMAZ** under supervision of **Prof. Dr. İsmihan Bayramoğlu** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Prof. Dr. İsmihan Bayramoğlu
Supervisor

Thesis Committee Member

Thesis Committee Member

Director

ABSTRACT

ORDER STATISTICS IN OUTLIER MODELS

KEREM TÜRKYILMAZ
M.S. in Applied Statistics
Graduate School of Natural and Applied Sciences
Supervisor: Prof. Dr. İsmihan Bayramoğlu
June 2009

In this study, order statistics from single and multiple outlier models are considered. The marginal and joint distributions of the corresponding order statistics are derived. Robust estimations for normal distribution in single outlier model are investigated, numerical results and Bias and MSE tables of these estimators are obtained. Moreover, probability of rth order statistic being outlier is derived whenever there is one or two outlier in the sample. A robust estimator based on this probability is provided and MSE, Bias results of this estimator of mean for normal distribution are presented. Conditional probablity of maximum and minimum order statistics given that rth order statistic is outlier is derived. Also, the empirical distribution function for single outlier model is provided.

Keywords: order statistics, outliers, single-outlier model, multiple-outlier model, robust estimators, location outlier, scale outlier.

ÖZ

SAPAN DEĞER MODELLERİNDE SIRA İSTATİSTİKLERİ

KEREM TÜRKYILMAZ
Uygulamalı İstatistik, Yüksek Lisans
Fen Bilimleri Enstitüsü
Tez Yöneticisi: Prof. Dr. İsmihan Bayramoğlu
Haziran 2009

Bu çalışmada, tekli ve çoklu sapan değer modellerinde sıra istatistiklerinin üzerinde durulmuştur. Bu sıra istatistiklerinin dağılım ve ortak dağılım fonksiyonları elde edilmiştir. Tekli sapan değer modelinde, normal dağılım için sağlam tahmin ediciler araştırılmıştır. Bu tahmin ediciler için sayısal sonuçlar, Bias ve ortalama hata kareleri tabloları elde edilmiştir. Ayrıca, bir veya iki sapan değerli modelde r. sıra istatistiğinin sapan değer olma olasılığı hesaplanmıştır. Bu olasılığa dayanarak normal dağılımın ortalaması için, bir sağlam tahmin edici önerilip Bias, ortalama hata kareleri değerleri bulunmuştur. r. sıra istatistiğinin sapan değer olma olasılığı koşulu altında minimum ve maksimum sıra istatistiklerinin dağılımları elde edilmiştir. Buna ek olarak, tekli sapan değer modeli için ampirik dağılım fonksiyonu hesaplanmıştır.

Anahtar Kelimeler: sıra istatistikleri, sapan değerler, tekli sapan değer modeli, çoklu sapan değer modeli, sağlam tahmin ediciler, yer sapan değeri, ölçek sapan değeri.

ACKNOWLEDGEMENT

I would like to thank a number of people who helped and supported me. I express my deepest gratitude to my supervisor and mentor Prof. Dr. İsmihan Bayramoğlu for continued and expanded with excellent guidance and encouragements.

I would like to extend my appreciation to my committee members: Assoc. Prof. Dr. Serkan Eryılmaz, Asst. Prof. Dr. Özlem Ege Oruç.

My sincere appreciation is also to all my friends and colleague, especially Cihangir Kan, Hande Tunçel, Ceki Franko and Göknur Yapakçı. Also my family who always with me whenever I needed.

Contents

Introduction

In classical statistics, an outlier is an observation that lies numerically distant from the rest of data in a random sample from a population. Since the earliest attempts to interpret data, there has been a concern for outlying observations in data sets. These outliers are generally considered as reducer of information about data. Therefore, it is reasonable to attempt to interpret means and to seek methods for handling outliers. Sometimes rejecting outliers may be improve fitness of the data, or applying methods of decreasing their effect in statistical analysis.

Peirce has stated the concept of outlier and outlier problem in 1852 by his following words: "In almost every true series of observations, some are found, which differ so much from the others as to indicate some abnormal source of error not contemplated in the theoretical discussions, and the introduction of which into the investigations can only serve to perplex and mislead the inquirer." The earliest method for dealing with outliers was introduced by Chauvenet in 1863.

Outlier definition can be defined in terms of distributions rather than numerical distance between observations. Assume that an experimenter wants to obtain n observations from population with distribution function F. It may happen that one or more observations among this sample is obtained from population with distribution function G. These observations are called outliers. In this case, in ordered sample, outliers may not be extremes. More precisely, outliers are observations only having different distributions. For example, in a population with continuous distribution with p.d.f. having two modes, the outliers may fall into interval, where the p.d.f. has minimum value between two modes. Clearly, none of these outliers will be extreme value of the sample. Therefore, distribution of order statistics from independent non-identical random variables are closely related with the outlier models.

Since the early 20th century, important studies on order statistics and their properties have been presented. The first fundamental book describing this theory is David (1981). Arnold et al. (1992) and David and Nagaraja (2003) include new developments on order statistics from independent and identically distributed (i.i.d.) and independent but not necessarily identically distributed (i.n.i.d.) random variables. The distribution theory of order statistics from i.n.i.d. random variables were first described in Vaughan and Venables (1972) by involving permanent, a concept defined similar to the determinant except that it does not have alternating sign, i.e. taking all terms in the summation of the definition of determinant to be positive. For a recent review describing the theory of order statistics from i.n.i.d. case and also including interesting results on outliers and robustness, we refer Balakrishnan (2007). Permanent expressions for the distribution function of i.n.i.d. order statistics allow to obtain some recurrence relations, using the expansion of the permanent by some of the rows. However, in some cases, where the applications of order statistics from the i.n.i.d. random variables are considered, the usage of the permanent expressions for the distributions of i.n.i.d. order statistics causes some difficulties connected with the complexity of operations. Despite researches are generally focused on the order statistics from i.i.d. variables, after 1970's order statistics and outlier models are considered together under robust estimation subject. Early studies were on single outlier model by H. A. David, V. S. Shu, V. Barnett and T. Lewis but in the last two decades, by the help of researches on order statistics from independent non identical random variables, important contributions on multiple outlier models have been made by N. Balakrishnan, A. Childs, H. A. David.

Chapter 1

Order statistics from single outlier model

The distribution theory of order statistics from independent identically distributed random variables has been well studied in the literature. However, in the case of non-identically distributed random variables the situation becomes complex, and the distribution theory of order statistics, in this case, still has problems to be solved. A single-outlier model can be considered as follows. Assume that a collection of n independent random variables $X_1, ..., X_n$ is considered. Furthermore, $n-1$ of these random variables, say, $X_1, ..., X_{n-1}$ have cumulative distribution function $F(x)$ and one of them, say, X_n has different distribution function $G(x)$. Let $X_{1:n} \leq ... \leq X_{n:n}$ be the order statistics constructed from sample $X_1, ..., X_n$ containing one outlier. In this chapter, we describe the distribution theory of order statistics from single outlier model.

1.1 Distributions of order statistics

By considering combinatorial arguments and the outlier X_n may fall in the intervals $(-\infty, x]$, $(x, x+\Delta x]$ and $(x+\Delta x, \infty]$. The density function of $X_{r:n}$ $(1 \leq$

$r \leq n$) can be obtained as

$$\begin{aligned} f_{r:n}(x) = & \frac{(n-1)!}{(r-2)!(n-r)!}\{F(x)\}^{r-2}G(x)f(x)\{1-F(x)\}^{n-r} \\ & + \frac{(n-1)!}{(r-1)!(n-r)!}\{F(x)\}^{r-1}g(x)\{1-F(x)\}^{n-r} \\ & + \frac{(n-1)!}{(r-1)!(n-r-1)!}\{F(x)\}^{r-1}f(x) \\ & \times \{1-F(x)\}^{n-r-1}\{1-G(x)\},\ x \in \mathbb{R} \end{aligned}$$

when $r = 1$ and $r = n$, the first and last terms do not appear in the formula respectively. Similar argument can be given for finding the joint density function of $X_{r:n}$ and $X_{s:n}$ $(1 \leq r < s \leq n)$ as

$$\begin{aligned} f_{r,s:n}(x,y) = & \frac{(n-1)!}{(r-2)!(s-r-1)(n-s)!}\{F(x)\}^{r-2}G(x)f(x) \\ & \times \{F(y)-F(x)\}^{s-r-1}f(y)\{1-F(y)\}^{n-s} \\ & + \frac{(n-1)!}{(r-1)!(s-r-1)(n-s)!}\{F(x)\}^{r-1}g(x) \\ & \times \{F(y)-F(x)\}^{s-r-1}f(y)\{1-F(y)\}^{n-s} \\ & + \frac{(n-1)!}{(r-1)!(s-r-2)(n-s)!}\{F(x)\}^{r-1}f(x) \\ & \times \{F(y)-F(x)\}^{s-r-2}\{G(y)-G(x)\}f(y)\{1-F(y)\}^{n-s} \\ & + \frac{(n-1)!}{(r-1)!(s-r-1)(n-s)!}\{F(x)\}^{r-1}f(x) \\ & \times \{F(y)-F(x)\}^{s-r-1}g(y)\{1-F(y)\}^{n-s} \\ & + \frac{(n-1)!}{(r-1)!(s-r-1)(n-s-1)!}\{F(x)\}^{r-1}f(x) \\ & \times \{F(y)-F(x)\}^{s-r-1}f(y)\{1-F(y)\}^{n-s-1}\{1-G(y)\} \\ & -\infty < x < y < \infty \end{aligned}$$

where the first, middle and last terms do not appear when $r = 1,\ s = r+1$ and $s = n$, respectively.

1.2 Robust estimation

Statistical methods heavily depend on a number of assumptions. These assumptions generally aim at formalizing statistical model, at the same time, aim at making result of the statistical model manageable from the computational and theoretical points of view. Usually, it is thought that the formalized models are simple forms of reality, and that they are best approximations. The generally used model formalization is the assumption of the observed data obtained from the population which has normal distribution. This assumption constitutes the basis of the classical statistical methods. The classical statistics are quite easy to compute with the modern computational methods. Unfortunately, computational and theoretical easiness is not always sufficient for practice of statistics and data analysis.

In practice, it is usually encountered that some observations may violate normality assumption of classical statistical models. Such data are called outliers and even one outlier can lead the classical methods to have poor results. Moreover, the power of classical tests can be quite low, and their confidence level may be unreliable for the classical confidence level.

Robust statistics provide an alternative approach to the classical statistical methods. The aim of this approach is to find methods that produce reliable parameter estimations and corresponding tests, confidence intervals, even if classical approach assumptions are violated. If there is no outlier in the sample, robust method and classical method give approximately same results.

1.3 Robust estimation in the presence of outliers

Let us have n independent continuous random variables X_j $(j = 1, ..., n-1)$ and Y, such that

$$X_j \text{ has cdf } F(x) \text{ and pdf } f(x)$$
$$Y \text{ has cdf } G(x) \text{ and pdf } g(x),$$

where Y represents an outlier. Let $Z_{r:n}$, $r = 1, ..., n$, denote rth order statistic of the combined sample. Then the pdf of $Z_{r:n}$ is given by

$$\begin{aligned} h_{r:n}(x) &= f_{r-1:n-1}(x)G(x) + \binom{n-1}{r-1} F^{r-1}(x)[1-F(x)]^{n-r}g(x) \\ &\quad + f_{r:n-1}(x)[1-G(x)] \end{aligned}$$

where, $f_{r:n-1}(x)$ is the pdf of $X_{r:n-1}$.
We consider the location shift case, $G(x) = F(x-\lambda)$. Then we can write $Y = X_n + \lambda$, where X_n has cdf $F(x)$ and independent of $X_1, ..., X_{n-1}$ then we can write the dependence on λ as

$$\begin{aligned} h_{r:n}(x;\infty) &= f_{r:n-1}(x) \qquad r = 1, ..., n-1 \\ h_{r:n}(x;-\infty) &= f_{r-:n-1}(x) \qquad r = 2, ..., n. \end{aligned}$$

To see how $Z_{r:n}(\lambda)$ behaves as a function of λ. Lowercase x, y, z will as usual denote realizations of X, Y, Z. Adding $y = x_n + \lambda$ into the ordered sample of size $n-1$. Then for any fixed values of $x_1, ..., x_n$ we have

$$z_{1:n}(\lambda) = \begin{cases} x_n + \lambda & if \;\; x_n + \lambda \le x_{1:n-1} \\ x_{1:n-1} & if \;\; x_n + \lambda > x_{1:n-1} \end{cases}$$

and for $r = 2, ..., n-1$

$$z_{r:n}(\lambda) = \begin{cases} x_{r-1:n-1} & if \;\; x_n + \lambda \le x_{r-1:n-1} \\ x_n + \lambda & if \;\; x_{r-1:n-1} < x_n + \lambda \le x_{r:n-1} \\ x_{r:n-1} & if \;\; x_n + \lambda > x_{r:n-1} \end{cases}$$

and

$$z_{n:n}(\lambda) = \begin{cases} x_{n-1:n-1} & if \quad x_n + \lambda \leq x_{n-1:n-1} \\ x_n + \lambda & if \quad x_n + \lambda > x_{n-1:n-1} \end{cases}$$

Hence $z_{r:n}(\lambda)$ is a nondecreasing function of λ with $z_{n:n}(\infty) = \infty$, $z_{1:n}(-\infty) = -\infty$ and otherwise $z_{r:n}(\infty) = x_{r:n-1}$, $z_{r:n-1}(-\infty) = x_{r-1:n-1}$.
For the finite λ if $E(X)$ exists so does $\mu_{r:n}(\lambda) = E[Z_{r:n}(\lambda)]$, $r = 1, ..., n$. We write $\mu_{r:n}(0) = \mu_{r:n}$, etc. Using the monotone convergence theorem it follows that, for $r = 1, ..., n-1$,

$$\begin{aligned} \lim_{\lambda \to \infty} E[Z_{r:n}(\lambda)] &= E[\lim_{\lambda \to \infty} Z_{r:n}(\lambda)], \\ \mu_{r:n}(\infty) &= E[X_{r:n-1}] \equiv \mu_{r:n-1} \end{aligned}$$

Similarly, for $r = 2, ..., n$

$$\begin{aligned} \lim_{\lambda \to -\infty} E[Z_{r:n}(\lambda)] &= E[\lim_{\lambda \to -\infty} Z_{r:n}(\lambda)], \\ \mu_{r:n}(-\infty) &= E[X_{r-1:n-1}] \equiv \mu_{r-1:n-1} \end{aligned}$$

and

$$\mu_{1:n}(-\infty) = -\infty, \quad \mu_{n:n}(\infty) = \infty.$$

1.4 Sensitivity curves

It is reasonable to look at the difference $t_n(x_1, ..., x_{n-1}, x) - t_{n-1}$ for evaluate how sensitive an estimate $t_{n-1} = t_{n-1}(x_1, ..., x_{n-1})$ is to the values of an additional observation x.

Obviously, for an estimator to be robust, this difference should remain within reasonable bounds as x ranges through its possible values.
The graph of $n[t_n(x) - t_{n-1}]$ against x is called as a *sensitivity curve*. By replacing $x_1, ..., x_{n-1}$ by the expected values of the order statistics in samples of $n - 1$, *stylized sensitivity curves* can be obtained.

1.5 Robust estimation for normal distribution

In the case of the normal distribution, location and scale outlier model can be considered as:

i. Location-outlier model:

$$X_1, ..., X_{n-1} \stackrel{d}{=} N(0,1) \;\; and \;\; X_n \stackrel{d}{=} N(\lambda, 1)$$

ii. Scale-outlier model:

$$X_1, ..., X_{n-1} \stackrel{d}{=} N(0,1) \;\; and \;\; X_n \stackrel{d}{=} N(0, \sigma^2)$$

For the sample size up to 20, the values of means, variances and covariances of order statistics for different selection of λ and σ were tabulated by H. A. David (1977). By the help of these tables, several linear estimators of the normal mean established by Arnold and Balakrishnan (1989), such as

i. Sample mean:

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_{i:n}$$

ii. Trimmed means:

$$T_n(r) = \frac{1}{n-2r}\sum_{i=r+1}^{n-r} X_{i:n}$$

iii. Winsorized means:

$$W_n(r) = \frac{1}{n}\left[\sum_{i=r+2}^{n-r-1} X_{i:n} + (r+1)[X_{r+1:n} + X_{n-r:n}]\right]$$

iv. Modified maximum likelihood estimators:

$$M_n(r) = \frac{1}{m}\left[\sum_{i=r+2}^{n-r-1} X_{i:n} + (1+r\beta)[X_{r+1:n} + X_{n-r:n}]\right]$$

where $m = n - 2r + 2r\beta$, $\beta = (g(h_2) - g(h_1))/(h_2 - h_1)$, $h_1 = F^{-1}(1 - q - \sqrt{q(1-q)/n})$, $h_2 = F^{-1}(1-q+\sqrt{q(1-q)/n})$, $q = r/n$, $F(h) = \int_{-\infty}^{h} f(z)dz$, $f(z) = \frac{1}{\sqrt{2\pi}}e^{-z^2}/2$, and $g(h) = f(h)/(1 - F(h))$.

v. Linearly weighted means:

$$L_n(r) = \frac{1}{2(\frac{n}{2} - r)^2} \left[\sum_{i=1}^{\frac{n}{2}-r} (2i-1)[X_{r+i:n} + X_{n-r-i+1:n}] \right]$$

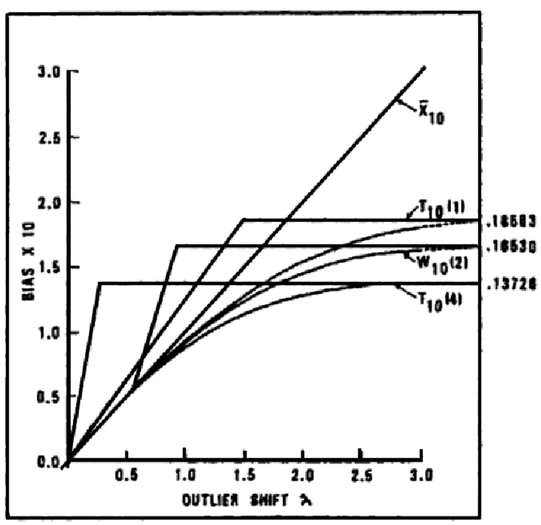
for even values of n;

vi. Gastwirth mean:

$$G_n = 0.3(X_{[\frac{n}{3}]+1:n} + X_{n-[\frac{n}{3}]:n}) + 0.2(X_{\frac{n}{2}:n} + X_{\frac{n}{2}+1:n})$$

for even values of n, where $[\frac{n}{3}]$ denotes the integer part of $\frac{n}{3}$.

The plot of bias versus λ obviously has some similarity with the sensitivity curve, and for $n = 10$ is compared with the corresponding stylized sensitivity curve in figure below for four well known estimators ($\bar{X}_{10}$, $T_{10}(1)$, $W_{10}(2)$, $T_{10}(4)$)



The median $T_{10}(4)$ has, uniformly minimum bias in the class of L *estimators*. It is easy to see that the bias is monotonically increasing in λ. But the median has uniformly larger MSE than the less severely trimmed means.

By using the tables of means, variances and covariances of order statistics from a single location outlier normal model by David(1977), in the tables below, bias and MSE of all these estimators are presented (Balakrishnan(2007)).

				λ				
Estimator	0.0	0.5	1.0	1.5	2.0	3.0	4.0	∞
$\bar{X}_{10}$	0.10000	0.10250	0.11000	0.12250	0.14000	0.19000	0.26000	∞
$T_{10}(1)$	0.10534	0.10791	0.11471	0.12387	0.13285	0.14475	0.14865	0.14942
$T_{10}(2)$	0.11331	0.11603	0.12297	0.13132	0.13848	0.14580	0.14730	0.14745
Med_{10}	0.13833	0.14161	0.14964	0.15852	0.16524	0.17072	0.17146	0.17150
$W_{10}(1)$	0.10437	0.10693	0.11403	0.12405	0.13469	0.15039	0.15627	0.15755
$W_{10}(2)$	0.11133	0.11402	0.12106	0.12995	0.13805	0.14713	0.14926	0.14950
$M_{10}(1)$	0.10432	0.10688	0.11396	0.12385	0.13430	0.14950	0.15513	0.15581
$M_{10}(2)$	0.11125	0.11395	0.12097	0.12974	0.13770	0.14649	0.14853	0.14876
$L_{10}(1)$	0.11371	0.11644	0.12337	0.13169	0.13882	0.14626	0.14797	0.14820
$L_{10}(2)$	0.12097	0.12386	0.13105	0.13933	0.14598	0.15206	0.15310	0.15318
G_{10}	0.12256	0.12549	0.13276	0.14111	0.14777	0.15376	0.15472	0.15479

Table 1. MSE of various estimators of μ for $n = 10$ when a single outlier is from $N(\mu + \lambda, 1)$ and the others from $N(\mu, 1)$

				λ				
Estimator	0.0	0.5	1.0	1.5	2.0	3.0	4.0	∞
$\bar{X}_{10}$	0.0	0.05000	0.10000	0.15000	0.20000	0.30000	0.40000	∞
$T_{10}(1)$	0.0	0.04912	0.09325	0.12870	0.15400	0.17871	0.18470	0.18563
$T_{10}(2)$	0.0	0.04869	0.09023	0.12041	0.13904	0.15311	0.15521	0.15538
Med_{10}	0.0	0.04832	0.08768	0.11381	0.12795	0.13642	0.13723	0.13726
$W_{10}(1)$	0.0	0.04938	0.09506	0.13368	0.16298	0.19407	0.20239	0.20377
$W_{10}(2)$	0.0	0.04889	0.09156	0.12389	0.14497	0.16217	0.16504	0.16530
$M_{10}(1)$	0.0	0.04934	0.09484	0.13311	0.16194	0.19229	0.20037	0.20169
$M_{10}(2)$	0.0	0.04886	0.09137	0.12342	0.14418	0.16091	0.16369	0.16394
$L_{10}(1)$	0.0	0.04869	0.09024	0.12056	0.13954	0.15459	0.15727	0.15758
$L_{10}(2)$	0.0	0.04850	0.08892	0.11700	0.13328	0.14436	0.14576	0.14585
G_{10}	0.0	0.04847	0.08873	0.11649	0.13237	0.14285	0.14407	0.14414

Table 2. Bias of various estimators of μ for $n = 10$ when a single outlier is from $N(\mu + \lambda, 1)$ and the others from $N(\mu, 1)$

It can be seen that from the tables above, even if median provides the best prediction in single outlier model in terms of bias, it causes a higher MSE than other robust estimators. The trimmed mean, modified maximum likelihood and linearly weighted mean estimators seem to be more robust and efficient.

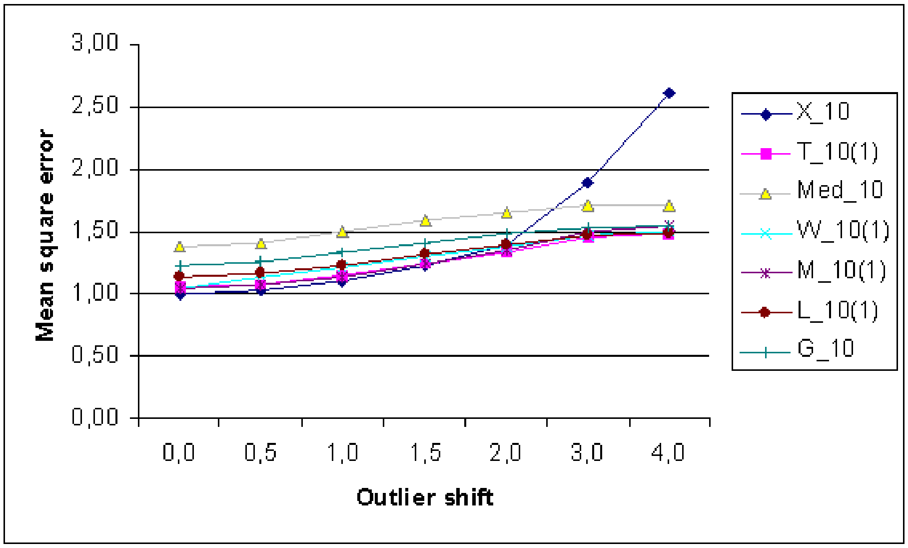


Figure 2. MSE of various estimators of μ for $n = 10$ when a single outlier is from $N(\mu + \lambda, 1)$ and the others from $N(\mu, 1)$

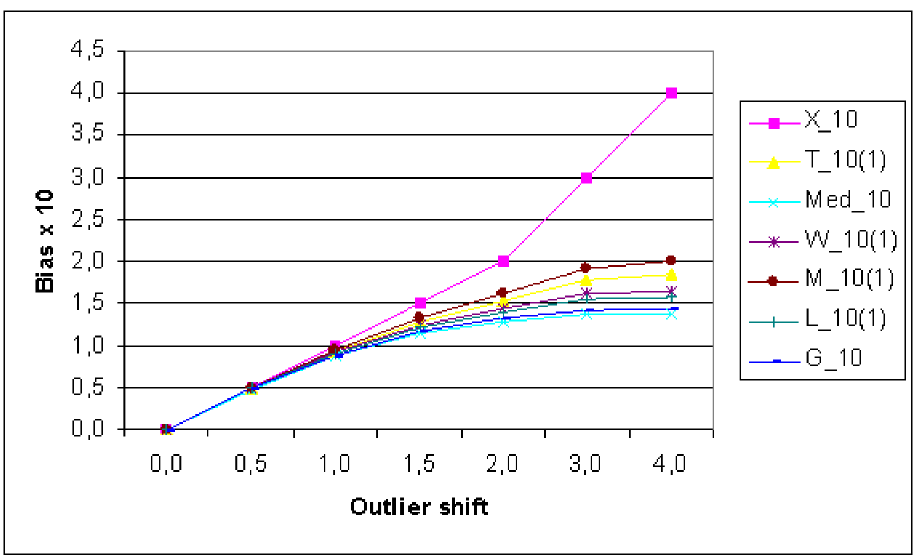


Figure 3. Bias of various estimators of μ for $n = 10$ when a single outlier is from $N(\mu + \lambda, 1)$ and the others from $N(\mu, 1)$

Similarly, estimators of the location parameter μ can be considered in a single scale outlier normal model and results for several estimators have been obtained in the following table. In this situation, because of the estimators are unbiased, it is sufficient to evaluate variances of them to compare mean square errors. The trimmed mean, modified maximum likelihood and linearly weighted mean estimators again seem to be quite robust according to this table.

			τ			
Estimator	0.5	1.0	2.0	3.0	4.0	∞
$\bar{X}_{10}$	0.09250	0.10000	0.13000	0.18000	0.25000	∞
$T_{10}(1)$	0.09491	0.10534	0.12133	0.12955	0.13417	0.14942
$T_{10}(2)$	0.09953	0.11331	0.12773	0.13389	0.13717	0.14745
Med_{10}	0.11728	0.13833	0.15375	0.15953	0.16249	0.17150
$W_{10}(1)$	0.09571	0.10437	0.12215	0.13221	0.13801	0.15754
$W_{10}(2)$	0.09972	0.11133	0.12664	0.13365	0.13745	0.14950
$M_{10}(1)$	0.09548	0.10432	0.12187	0.13171	0.13735	0.15581
$M_{10}(2)$	0.09940	0.11125	0.12638	0.13328	0.13699	0.14876
$L_{10}(1)$	0.09934	0.11371	0.12815	0.13436	0.13769	0.14820
$L_{10}(2)$	0.10432	0.12097	0.13531	0.14101	0.14398	0.15318
G_{10}	0.10573	0.12256	0.13703	0.14270	0.14565	0.15479

Table 3. Variance of various estimators of μ for $n = 10$ when a single outlier is from $N(\mu, \tau^2)$ and the others from $N(\mu, 1)$

Chapter 2

Order statistics from multiple outlier model

In single outlier model, density function of $X_{r:n}$ and joint density function of $(X_{r:n},\, X_{s:n})$ can be evaluated by direct approach. It can be observed that in the expressions of density functions of order statistics, they have three and five terms respectively. However, if we consider two outliers in the sample, the marginal density of $X_{r:n}$ has five terms and joint density of $(X_{r:n},\, X_{s:n})$ have thirteen terms. For this reason, the theory of order statistics in the presence of two or more outliers remains many unsolved problems. Hence, in multiple outlier models, we need different special methods. Permanents, described in the following section are useful tool to deal with these models.

2.1 Permanents

The permanent of an $n \times n$ matrix $A = (a_{i,j})$ is defined as

$$Per(A) = \sum_{P} \prod_{j=1}^{n} a_{j,i_j},$$

where $\sum_P$ represents the sum of all $n!$ permutations $(i_1, i_2, ..., i_n)$ from $(1, 2, ..., n)$. The definition of the permanent of a matrix A differs from determinant of A in that the signatures of the permutations are not taken into account. Some properties of permanents can be given as follows;

i. If columns or rows of A are permuted, $Per(A)$ does not change.

ii. Let $A(i, j)$ show the sub-matrix of A that obtained by deleting ith row and jth column, then

$$\begin{aligned} Per(A) &= \sum_{i=1}^{n} a_{i,j} Per(A(i, j)), \quad j = 1, 2, ..., n \\ &= \sum_{j=1}^{n} a_{i,j} Per(A(i, j)), \quad i = 1, 2, ..., n \end{aligned}$$

iii. If we change ith row of matrix A by $c \times a_{i,j}, \;\; j = 1, 2, ..., n$ and new matrix A^* have the property

$$Per(A^*) = cPer(A)$$

2.2 Distribution of order statistics in terms of the symmetric functions

Let $X_1, X_2, ..., X_n$ be independent but not necessarily identically distributed random variables with cumulative distribution functions (cdf) $F_1(x), F_2(x), ..., F_n(x)$ and $X_{1:n}, X_{2:n}, ..., X_{n:n}$ be corresponding order statistics. If $F_1, F_2,...,F_n$ are absolutely continuous with corresponding probability density functions (pmf) $f_1, f_2, ..., f_n$, then the joint pmf of $X_{1:n}, X_{2:n}, ..., X_{n:n}$ is

$$f_{1,2,...,n}(x_1, x_2, ..., x_n) = \sum_{\wp} \prod_{j=1}^{n} f_{i_j}(x_j),$$

where the summation $\wp$ extends over all permutations $(i_1, i_2, ..., i_n)$ of $1, 2, ..., n$. For any borel set $B \in \Re$, where $\Re$ is the Borel $\sigma-$algebra of subsets of the set of real numbers $\mathbb{R}$ consider indicators $I_{X_i}(B) = \begin{cases} 1, & X_i \in B \\ 0, & X_i \notin B \end{cases}$, $i = 1, 2, ..., n$ and let $\nu^*(B) = \sum_{i=1}^{n} I_{X_i}(B)$. Define the empirical distribution of the I.N.I.D. sample $X_1, X_2, ..., X_n$ as $P_n^*(B) = \frac{\nu^*(B)}{n}$. It is clear that $EI_{X_i}(B) = P_i\{X_i \in B\} = \int_B dF_i(x) \equiv P_i(B)$ and $var(I_{X_i}(B)) = P_i(B)(1 - P_i(B))$ and $EP_n^*(B) = \sum_{i=1}^{n} P_i(B)$ and $var(P_n^*(B)) = \sum_{i=1}^{n} P_i(B)(1-P_i(B))$. The empirical distribution function of the I.N.I.D. sample then is defined as $F_n^*(x) = P_n^*((-\infty, x])$. Since $\frac{1}{n^2}\sum_{i=1}^{n} I_{X_i}(B) \to 0$ as $n \to \infty$, then the sequence of independent random variables obeys the strong low of large numbers, *i.e.* for any $\varepsilon > 0$ and $\eta > 0$ there exists n_0 such that for arbitrary s and for all n, satisfying $n_0 \leq n \leq n_0 + s$, the probability of the inequality

$$\max_{n_0 \leq n \leq n_0+s} \left| P_n^*(B) - \frac{1}{n}\sum_{i=1}^{n} P_i(B) \right| < \varepsilon \quad \text{and} \quad \max_{n_0 \leq n \leq n_0+s} \left| F_n^*(x) - \frac{1}{n}\sum_{i=1}^{n} F_i(x) \right| < \varepsilon$$

is greater than $1 - \eta$, for any $B \in \Re$ and $x \in \mathbb{R}$.

Lemma 1. For any $B \in \Re$ and $x \in \mathbb{R}$

$$P\{nP_n^*(B) = k\} = \sum_{S_k} \prod_{i=1}^{k} P_{j_i}(B) \prod_{i=k+1}^{n} (1 - P_{j_i}(B))$$

and

$$P\{nF_n^*(x) = k\} = \sum_{S_k} \prod_{i=1}^{k} F_{j_i}(x) \prod_{i=k+1}^{n} (1 - F_{j_i}(x)),$$

where the summation S_k extends over all permutations $j_1, j_2, ..., j_n$ of $1, 2, ..., n$ for which $j_1 < j_2 < ... < j_k$ and $j_{k+1} < j_{k+2} < ... < j_n$.

Denote now

$$B(n, k; x) = \binom{n}{k} x^k (1 - x)^{n-k}$$

and the symmetric function

$$B(n, k; x_1, x_2, ..., x_n) = \sum_{S_k} \prod_{i=1}^{k} x_{j_i} \prod_{i=k+1}^{n} (1 - x_{j_i}), 1 \leq k \leq n.$$

It is clear that $B(n, k; x_{j_1}, x_{j_2}, ..., x_{j_n}) = B(n, k; x_1, x_2, ..., x_n)$ for all $n!$ permutations $(j_1, j_2, ..., j_n)$ of $(1, 2, ..., n)$.

$$P\{nF_n^*(x) = k\} = B(n, k; F_1(x), F_2(x), ..., F_n(x)).$$

It is clear that if $F_1 = F_2 = \cdots = F_n = F$ then

$$P\{nF_n^*(x) = k\} = B(n, k; F(x)).$$

The following recurrence relation will be useful.

Lemma 2.

$$\begin{aligned} B(n, k; x_1, x_2, ..., x_n) &= B(n-1, k; x_1, x_2, ..., x_{n-1})\bar{x}_n \\ &\quad + B(n-1, k-1; x_1, x_2, ..., x_{n-1})x_n, \end{aligned}$$

where $\bar{x} = 1 - x$.

The cdf of $r-$th order statistic $X_{r:n}$ is

$$\begin{aligned} F_r(x) &= P\{X_{r:n} \le x\} \\ &= \sum_{i=r}^{n} \sum_{S_k} \prod_{i=1}^{k} F_{j_i}(x) \prod_{i=k+1}^{n} (1 - F_{j_i}(x)) \end{aligned} \tag{2.1}$$

(see David and Nagaraja (2003)) and in terms of $B(n, k; x_1, x_2, ..., x_n)$ it can be written as

$$F_r(x) = \sum_{i=r}^{n} B(n, i, F_1(x), F_2(x), ..., F_n(x)). \tag{2.2}$$

Using Lemma 2 we can write

$$\begin{aligned} F_r(x) &= \sum_{i=r}^{n} B(n, i, F_1(x), F_2(x), ..., F_n(x)) \\ &= \bar{F}_n(x) \sum_{i=r}^{n} B(n-1, i, F_1(x), F_2(x), ..., F_{n-1}(x)) \\ &\quad + F_n(x) \sum_{i=r}^{n} B(n-1, i-1, F_1(x), F_2(x), ..., F_{n-1}(x)) \end{aligned}$$

$$= F_{r:n-1}(x)\bar{F}_n(x) + F_n(x)F_{r-1:n-1}(x), \tag{2.3}$$

where $F_{i:n-1}$ denotes the cdf of the $i-$th order statistic from I.N.I.D. random variables $X_1, X_2, ..., X_{n-1}$ with corresponding cdf's $F_1, F_2, ..., F_{n-1}$. Note that (2.3) and related recurrence equalities can be found in David and Nagaraja (2003, p. 105)). Since,

$$\begin{aligned} P\{nF_n^*(x) &= i\} = B(n, i, F_1(x), F_2(x), ..., F_n(x)), \\ i &= 0, 1, 2, ..., n \end{aligned}$$

then

$$\sum_{i=0}^{n} B(n, i, F_1(x), F_2(x), ..., F_n(x)) = 1. \tag{2.4}$$

We have also,

$$P\{X_{r:n} \leq x\} = \sum_{i=r}^{n} P\{nF_n^*(x) = i\}.$$

2.3 Log-concavity

Log-concavity of the distribution functions of the order statistics can be showed by Alexandroff inequality which is important result in permanents theory of non-negative matrices. A sequence of non-negative numbers $\alpha_1, \alpha_2, ..., \alpha_n$ is log-concave if $\alpha_1^2 \geq \alpha_{i-1}\alpha_{i+1}$ $(i = 2, 3, ...n - 1)$. Some properties of log-concavity can be given as follows; Let $\alpha_1, \alpha_2, ..., \alpha_n$ and $\beta_1, \beta_2, ..., \beta_n$ be two log-concave sequences. Then the statements below hold.

i. 1. If $\alpha_i > 0$ for $i = 1, 2, ..., n$, then

$$\frac{\alpha_i}{\alpha_{i-1}} \geqslant \frac{\alpha_{i+1}}{\alpha_i}, \quad i = 2, ..., n-1$$

which means, $\frac{\alpha_i}{\alpha_{i-1}}$ is non-increasing in i.

ii. if $\alpha_i > 0$ for $i = 1, 2, ..., n$, then $\alpha_1, \alpha_2, ..., \alpha_n$ is unimodal, i.e.

$$\alpha_1 \leq \alpha_2 \leq ... \leq \alpha_k \geq a_{k+1} \geq ... \geq a_n$$

for some k $(1 \leq k \leq n)$

iii. The sequence $\alpha_1\beta_1, \alpha_2\beta_2, ..., \beta_n\alpha_n$ is log-concave.

iv. The sequence $\gamma_1, \gamma_2, ..., \gamma_n$ is log-concave, where

$$\gamma_k = \sum_{i=1}^{k} \alpha_i \beta_{k+1-i} \quad k = 1, 2, ..., n.$$

v. The sequences $\alpha_1, \alpha_1 + \alpha_2, ..., \sum_{i=1}^{n} a_i$ and $\alpha_n, \alpha_{n-1} + \alpha_n, ..., \sum_{i=1}^{n} a_i$ are both log-concave.

vi. The sequence of combinatorial coefficients $\begin{pmatrix} n \\ i \end{pmatrix}$, $i = 0, 1, ..., n$ is log-concave.

2.4 Alexandroff's inequality

$$A = \begin{pmatrix} a_1 \\ ... \\ a_n \end{pmatrix}$$

be a non-negative square matrix of order n. Then,

$$(PerA)^2 \geq Per \begin{pmatrix} a_1 \\ ... \\ a_{n-2} \\ a_{n-1} \end{pmatrix} \}2 \quad Per \begin{pmatrix} a_1 \\ ... \\ a_{n-2} \\ a_n \end{pmatrix} \}2$$

2.5 Order statistics from independent non-identical variables in terms of permanents

2.5.1 Distributions and joint distributions

Let $X_1, X_2, ..., X_n$ be independent random variables from the population where each X_i has cdf $F_i(x)$ and pdf $f_i(x)$, $i = 1, 2, ..., n$. $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ be the order statistics from these n variables. Then, to obtain probability density function of $X_{r:n}$;

$$\begin{aligned} P(x &< X_{r:n} \leq x + \Delta x) \\ &= \frac{1}{(r-1)!(n-r)!} \sum_P F_{i_1}(x)...F_{i_{r-1}}(x)\{F_{i_r}(x+\Delta x) - F_{i_r}(x)\} \\ &\quad \times\{1 - F_{i_{r+1}}(x+\Delta x)\}...\{1 - F_{i_n}(x+\Delta x)\} + O((\Delta x)^2) \end{aligned}$$

where $\sum\limits_P$ represents the sum of all $n!$ permutations $(i_1, i_2, ..., i_n)$ from $(1, 2, ..., n)$. Dividing both side of equality by Δx and letting Δx go to zero, density function of $X_{r:n}$ is obtained $(1 \leq r \leq n)$ as

$$\begin{aligned} f_{r:n}(x) = \frac{1}{(r-1)!(n-r)!} \sum_P F_{i_1}(x)...F_{i_{r-1}}(x) f_{i_r}(x) \\ \times\{1 - F_{i_{r+1}}(x)\}...\{1 - F_{i_n}(x)\}, \quad x \in \mathbb{R} \end{aligned}$$

Permanent representation of $f_{r:n}(x)$ can be rewritten as

$$f_{r:n}(x) = \frac{1}{(r-1)!(n-r)!} Per A_1, \quad x \in \mathbb{R}$$

where

$$A_1 = \begin{pmatrix} F_1(x) & F_2(x) & ... & F_n(x) \\ f_1(x) & f_2(x) & ... & f_n(x) \\ 1 - F_1(x) & 1 - F_2(x) & ... & 1 - F_n(x) \end{pmatrix} \begin{array}{l} \}r-1 \\ \}1 \\ \}n-r \end{array}$$

For finding the joint density function $X_{r:n}$ and $X_{s:n}$ $(1 \leq r < s \leq n)$;

$$\begin{aligned} P(x &< X_{r:n} \leq x + \Delta x, \ y < X_{s:n} \leq y + \Delta y) \\ &= \frac{1}{(r-1)!(s-r-1)!(n-s)!} \sum_P F_{i_1}(x)...F_{i_{r-1}}(x)\{F_{i_r}(x+\Delta x) - F_{i_r}(x)\} \\ &\quad \times\{F_{i_{r+1}}(y) - F_{i_{r+1}}(x+\Delta x)\}...\{F_{i_{s-1}}(y) - F_{i_{s-1}}(x+\Delta x)\} \\ &\quad \times\{F_{i_s}(y+\Delta y) - F_{i_s}(y)\}\{1 - F_{i_{n+1}}(y+\Delta y)\}...\{1 - F_{i_n}(y+\Delta y)\} \\ &\quad + O((\Delta x)^2 \Delta y) + O((\Delta y)^2 \Delta x) \end{aligned}$$

$O((\Delta x)^2 \Delta y)$ represents terms which of X_i's falling exactly one in $(y, y + \Delta y]$ and more than one falling in $(x, x + \Delta x]$, and $O((\Delta y)^2 \Delta x)$ represents terms which of X_i's falling exactly one in $(x, x + \Delta x]$ and more than one falling in $(y, y + \Delta y]$, dividing both side of equation by $\Delta x \Delta y$ and both goes to zero, density function of $X_{r:n}$ and $X_{s:n}$ $(1 \leq r < s \leq n)$ is obtained as

$$\begin{aligned} & f_{r,s:n}(x,y) \\ = & \frac{1}{(r-1)!(s-r-1)!(n-s)!}\sum_P F_{i_1}(x)...F_{i_{r-1}}(x)f_{i_r}(x) \\ & \times\{F_{i_{r+1}}(y)-F_{i_{r+1}}(x)\}...\{F_{i_{s-1}}(y)-F_{i_{s-1}}(x)\} \\ & \times f_{i_s}(y)\{1-F_{i_{n+1}}(y)\}...\{1-F_{i_n}(y)\}, \quad -\infty < x < y < \infty. \end{aligned}$$

It can also be rewritten as permanent form

$$\begin{aligned} & f_{r,s:n}(x,y) \\ = & \frac{1}{(r-1)!(s-r-1)!(n-s)!}Per A_2, \quad -\infty < x < y < \infty. \end{aligned}$$

where

$$A_2 = \begin{array}{cc} \begin{pmatrix} F_1(x) & F_2(x) & ... & F_n(x) \\ f_1(x) & f_2(x) & ... & f_n(x) \\ F_1(y)-F_1(x) & F_2(y)-F_2(x) & ... & F_n(y)-F_n(x) \\ f_1(y) & f_2(y) & ... & f_n(y) \\ 1-F_1(y) & 1-F_2(y) & ... & 1-F_n(y) \end{pmatrix} & \begin{matrix} \}r-1 \\ \}1 \\ \}s-r-1 \\ \}1 \\ \}n-s \end{matrix} \end{array}$$

By similar consideration, joint density function of $X_{r_1:n}, X_{r_2:n}, ..., X_{r_k:n}$ ..$(1 \leq r_1 < r_2 < ... < r_k \leq n.)$ can be obtain as

$$\begin{aligned} & f_{r_1,r_2,...,r_k:n}(x_1,x_2,...,x_k) \\ = & \frac{1}{(r_1-1)!(r_2-r_1-1)!...(r_k-r_{k-1}-1)!(n-r_k)!}Per A_k, \\ -\infty < & x_1 < ... < x_k < \infty. \end{aligned}$$

where

$$A_k = \begin{array}{cc} \begin{pmatrix} F_1(x_1) & ... & F_n(x_1) \\ f_1(x_1) & ... & f_n(x_1) \\ F_1(x_2)-F_1(x_1) & ... & F_n(x_2)-F_n(x_1) \\ f_1(x_2) & ... & f_n(x_2) \\ ... & ... & ... \\ F_1(x_k)-F_1(x_{k-1}) & ... & F_n(x_k)-F_n(x_{k-1}) \\ f_1(x_k) & ... & f_n(x_k) \\ 1-F_1(x_k) & ... & 1-F_n(x_k) \end{pmatrix} & \begin{matrix} \}r_1-1 \\ \}1 \\ \}r_2-r_1-1 \\ \}1 \\ \}... \\ \}r_k-r_{k-1}-1 \\ \}1 \\ \}n-r_k \end{matrix} \end{array}$$

Moreover, cumulative distribution function $F_{r:n}(x)$ can be computed by

$$\begin{aligned}
F_{r:n}(x) &= P(X_{r:n} \leq x) \\
&= \sum_{i=r}^{n} P(\text{exactly } i \text{ of } X\text{'s are } \leq x) \\
&= \sum_{i=r}^{n} \frac{1}{i!(n-i)!} \sum_{P} F_{j_1}(x)...F_{j_i}(x)\{1 - F_{j_{i+1}}(x)\}...\{1 - F_{j_n}(x)\}
\end{aligned}$$

where $\sum\limits_{P}$ represents the sum of all $n!$ permutations $(j_1, j_2, ..., j_n)$ from $(1, 2, ..., n)$. Also it can be written by using permanent representation as;

$$F_{r:n}(x) = \sum_{i=r}^{n} \frac{1}{i!(n-i)!} Per B_1, \quad x \in \mathbb{R}$$

where

$$B_1 = \begin{array}{c} \begin{pmatrix} F_1(x) & F_2(x) & ... & F_n(x) \\ 1 - F_1(x) & 1 - F_2(x) & ... & 1 - F_n(x) \end{pmatrix} \end{array} \begin{array}{l} \}i \\ \}n - i \end{array}$$

Considering similarly, the joint cumulative distribution function of $X_{r_1:n}, X_{r_2:n}, ..., X_{r_k:n}$ $(1 \leq r_1 < r_2 < ... < r_k \leq n)$ may be obtained as

$$\begin{aligned}
F_{r_1, r_2, ..., r_k:n}(x_1, x_2, ..., x_n) &= P(X_{r_1:n} \leq x_1, ..., X_{r_k:n} \leq x_k) \\
&= \sum \frac{1}{j_1! j_2! ... j_{k+1}!} Per B_k \quad -\infty < x_1 < ... < x_k < \infty,
\end{aligned}$$

where

$$B_k = \begin{pmatrix} F_1(x_1) & ... & F_n(x_1) \\ F_1(x_2) - F_1(x_1) & ... & F_n(x_2) - F_n(x_1) \\ ... & ... & ... \\ F_1(x_k) - F_1(x_{k-1}) & ... & F_n(x_k) - F_n(x_{k-1}) \\ 1 - F_1(x_k) & ... & 1 - F_n(x_k) \end{pmatrix} \begin{array}{l} \}j_1 \\ \}j_2 \\ ... \\ \}j_k \\ \}j_{k+1} \end{array}$$

and $\sum$ is over $j_1, j_2, ..., j_{k+1}$ with $j_1 \geq r_1$, $j_1 + j_2 \geq r_2$, ..., $j_1 + ... + j_k \geq r_k$ and $j_1 + ... + j_k + j_{k+1} = n$. It can be clearly seen that by using independent non-identical distribution of order statistics, distribution of order statistics from multiple outlier model may be obtained. For example; Three outliers model wherein $X_1, X_2, ..., X_n$ are independent random variables with $X_1, X_2, ...X_{n-3}$ being from a population with cumulative distribution function $F(x)$ and probability density function $f(x)$

and X_{n-2}, X_{n-1}, X_n being outliers from a different population with cumulative distribution $G(x)$ and probability density function of $g(x)$.

$$f_{r:n}(x) = \frac{1}{(r-1)!(n-r)!} Per E_1, \quad x \in \mathbb{R}$$

where

$$E_1 = \begin{array}{c} \left(\begin{array}{cccccc} F(x) & ... & F(x) & G(x) & G(x) & G(x) \\ f(x) & ... & f(x) & g(x) & g(x) & g(x) \\ \underbrace{1-F(x) \quad ... \quad 1-F(x)}_{n-3} & & & 1-G(x) & 1-G(x) & 1-G(x) \end{array} \right) \end{array} \begin{array}{l} \}r-1 \\ \}1 \\ \}n-r \end{array}$$

2.5.2 Log-concavity

Alexandroff's inequality leads the log-concavity of distribution functions of order statistics. The following theorem presents this interesting result.

Theorem 2.1 *Let $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ show the order statistics from n independent non-identical variables with cumulative distribution functions $F_1(x), F_2(x), ..., F_n(x)$. Then for fixed x, the sequences $\{F_{r:n}(x)\}_{r=1}^n$ and $\{1 - F_{r:n}(x)\}_{r=1}^n$ are both log-concave. Moreover, if the underlying variables are all continuous with respective densities $f_1(x), f_2(x), ..., f_n(x)$ then the sequence $\{f_{r:n}(x)\}_{r=1}^n$ is also log-concave. (see Balakrishnan (2007)).*

Proof. For $i = 1, 2, ..., n$,

$$\alpha_i = Per \begin{pmatrix} F_1(x) & F_2(x) & ... & F_n(x) \\ 1-F_1(x) & 1-F_2(x) & ... & 1-F_n(x) \end{pmatrix} \begin{array}{l} \}i \\ \}n-i \end{array}$$

Since the matrix above is non-negative, an application of Alexandroff's inequality leads;

$$\alpha_i^2 \geq \alpha_{i-1}\alpha_{i+1}, \quad i = 2, 3, ..., n-1$$

Therefore, the sequence $\{\alpha_i\}_{i=1}^n$ is log-concave. The coefficients $\left\{\frac{1}{i!(n-i)!}\right\}_{i=1}^n$ form a log-concave sequence, this leads the sequence $\left\{\frac{\alpha_i}{i!(n-i)!}\right\}_{i=1}^n$ is also log-concave.

From the permanent representation of the cumulative distribution function of $X_{r:n}$ and property **(v),** we have the log-concavity of the sequence $\{F_{r:n}(x)\}_{i=1}^{n}$. Also according to property **(v)** partial sums of $\left\{\frac{\alpha_i}{i!(n-i)!}\right\}_{i=1}^{n}$ from the left form a log-concave sequence that leads the log-concavity of $\{1-F_{r:n}(x)\}_{i=1}^{n}$. By similar consideration it is easy to see that $\{f_{r:n}(x)\}_{i=1}^{n}$ is also log-concave. $\square$

Chapter 3

Main results

In the robust estimation for the normal distribution, the constructions of estimators for the parameter μ are based on the idea of removing r maximum and minimum terms from ordered sample. The reason for this procedure is to eliminate outliers from the sample, and provide more robust estimators for the parameters by neutralizing the influence of outliers. However, outliers in the sample are not always the extremes. The outliers in the sample are those observations that have different distributions. Therefore, they can fall in the middle part of ordered sample. For example, let $X_1, X_2, ...X_{n-1}$ be a sample from a population with cumulative distribution function $F(x)$ and X_n be a sample value from a population with cumulative distribution function $G(x)$. In ordered sample, $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ outlier X_n may take place of kth order statistics $X_{k:n}, 1 < k < n$. Estimators of μ for the normal distribution; sample mean($\bar{X}$), trimmed mean($T_n(r)$) , winsorised mean($W_n(r)$), modified maximum likelihood($M_n(r)$), linearly weighted mean($L_n(r)$) are unable to eliminate this type of outliers.

3.1 Probability of "$X_{r:n}$ is outlier"

Let $X_1, ..., X_{n-1}$ be from a population with cumulative distribution function $F(x)$ and X_n from a population with cumulative distribution function $G(x)$. To find an estimation of the parameter μ for the normal distribution by considering rth order statistic is outlier, probability that $X_{r:n}$ is outlier required to be obtained.

$$
\begin{aligned}
P\{X_{r:n} \text{ is outlier}\} &= \binom{n-1}{r-1} P\{X_1 < X_r^{'}, X_2 < X_r^{'}, ..., \\
X_{r-1} &< X_r^{'}, X_{r+1} > X_r^{'}, ..., X_n > X_r^{'}\} = \\
&= \binom{n-1}{r-1} \int_{-\infty}^{\infty} P\{X_1 < X_r^{'}, X_2 < X_r^{'}, ..., \\
X_{r-1} &< X_r^{'}, X_{r+1} > X_r^{'}, ..., X_n > X_r^{'} | X_r^{'} = t\} P\{X_r^{'} = t\} dt \\
&= \binom{n-1}{r-1} \int_{-\infty}^{\infty} P\{X_1 < t, X_2 < t, ..., \\
X_{r-1} &< t, X_{r+1} > t, ..., X_n > t\} P\{X_r^{'} = t\} dt \\
&= \binom{n-1}{r-1} \int_{-\infty}^{\infty} F^{r-1}(t)(1-F(t))^{n-r} dG(t)
\end{aligned}
$$

where $X_r^{'}$ denotes the random variable which have distribution function $G(x)$.

Example 3.1. Let $X_1, X_2, ..., X_{n-1}$ be from a population with cumulative distribution function $F(x)$ which is $Uniform(0, 1)$ and outlier X_n has cumulative distribution function $G(x)$ where

$$
G(x) = \begin{cases} 0 & x < 0 \\ x^{\theta} & 0 \le x < 1 \\ 1 & x > 1 \end{cases} \qquad F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x < 1 \\ 1 & x > 1 \end{cases}
$$

$$
\begin{aligned}
P\{X_{r:n} \text{ is outlier}\} &= \binom{n-1}{r-1} \int_{-\infty}^{\infty} F^{r-1}(t)(1-F(t))^{n-r} dG(t) \\
&= \binom{n-1}{r-1} \int_{0}^{1} x^{r-1}(1-x)^{n-r} \theta x^{\theta-1} dx \\
&= \binom{n-1}{r-1} \theta \int_{0}^{1} x^{r+\theta-2}(1-x)^{n-r} dx \\
&= \binom{n-1}{r-1} \theta\beta(r+\theta-1, n-r+1)
\end{aligned}
$$

Example 3.2. Let $X_1, X_2, ..., X_{n-1}$ be a sample from a population with cumulative distribution function $F(x)$ and outlier X_n has cumulative distribution function $G(x)$ where

$$
F(x) = \begin{cases} 1 - e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad G(x) = \begin{cases} 1 - e^{-\dfrac{x}{\delta}} & x \geq 0 \\ 0 & x < 0 \end{cases}
$$

The probability that $X_{r:n}$ is the outlier is given by

$$
P\{X_{r:n} \text{ is outlier}\} = \frac{\Gamma(n)\Gamma(n-i+(1/\delta))}{\delta\Gamma(n+(1/\delta))\Gamma(n-i+1)}
$$

(Kale and Sinha, 1971) (Numerical verification of this result with our formula can be found in appendix section 2.)

Similarly, probability that rth order statistic is outlier, whenever two outlier in the sample can be obtained. Let $X_1, ..., X_{r-2}$ be from a population with cumulative distribution function $F(x)$ and X_{n-1}, X_n from a population with cumulative distribution function $G(x)$.

$$P\{X_{r:n} \text{ is outlier}\} =$$
$$= \binom{n-2}{r-2} P\{X_1 < X_r', ..., X_{r-2} < X_r', X_{r-1} < X_r', X_{r+1} > X_r', ..., X_n > X_r'\} +$$
$$+ \binom{n-2}{r-1} P\{X_1 < X_r', ..., X_{r-2} < X_r', X_{r+1} < X_r', X_{r-1} > X_r', X_{r+2} > X_r', ..., X_n > X_r'\} +$$
$$+ \binom{n-2}{r-2} P\{X_1 < X_{r-1}', ..., X_{r-2} < X_{r-1}', X_r < X_{r-1}', X_{r+1} > X_{r-1}', ..., X_n > X_{r-1}'\} +$$
$$+ \binom{n-2}{r-1} P\{X_1 < X_{r-1}', ..., X_{r-2} < X_{r-1}', X_{r+1} < X_{r-1}', X_r > X_{r-1}',$$
$$X_{r+2} > X_{r-1}', ..., X_n > X_{r-1}'\}$$
$$= 2\binom{n-2}{r-2} \int_{-\infty}^{\infty} F^{r-2}(x) G(x) (1-F(x))^{n-r} dG(x) +$$
$$+2\binom{n-2}{r-1} \int_{-\infty}^{\infty} F^{r-1}(x) (1-G(x)) (1-F(x))^{n-r-1} dG(x)$$

where X_{r-1}', X_r' denote the random variables which have distribution function $G(x)$.

Example 3.3. Let $X_1, X_2, ..., X_{n-2}$ be from a population with cumulative distribution function $F(x)$ and outlier X_{n-1}, X_n has cumulative distribution function $G(x)$ where

$$F(x) = \begin{cases} 1 - e^{-x} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad G(x) = \begin{cases} 1 - e^{-\dfrac{x}{\delta}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

The probability that $X_{r:n}$ is the outlier is given by

$$
\begin{aligned}
&P\{X_{r:n} \text{ is outlier}\} = \\
&= 2\binom{n-2}{r-2}\int_{-\infty}^{\infty} F^{r-2}(x)G(x)(1-F(x))^{n-r}dG(x)+ \\
&+2\binom{n-2}{r-1}\int_{-\infty}^{\infty} F^{r-1}(x)(1-G(x))(1-F(x))^{n-r-1}dG(x) \\
&= 2\binom{n-2}{r-2}\int_{0}^{\infty}(1-e^{-x})^{r-2}(1-e^{-\frac{x}{\delta}})(e^{-x})^{n-r}d(1-e^{-\frac{x}{\delta}})+ \\
&+2\binom{n-2}{r-1}\int_{0}^{\infty}(1-e^{-x})^{r-1}(e^{-\frac{x}{\delta}})(e^{-x})^{n-r-1}d(1-e^{-\frac{x}{\delta}})
\end{aligned}
$$

3.2 An estimator for the mean of normal distribution

Multiplying order statistics with the respected probability that is not outlier and dividing the number of non-outlier observations gives an estimator of parameter μ for the normal distribution.

$$X^* = \frac{1}{n-1}\sum_{i=1}^{n}(1-\alpha_i)X_{i:n}$$

where α_i denotes the $P\{X_{i:n}$ is outlier$\}$. It can be easily seen that if there is no outlier in the sample,

$$
\begin{aligned}
\alpha_i &= \binom{n-1}{r-1}\int_{-\infty}^{\infty} F^{r-1}(t)(1-F(t))^{n-r}dF(t) \\
&= \binom{n-1}{r-1}\int_{0}^{1} u^{r-1}(1-u)^{n-r}du \\
&= \frac{(n-1)!}{(r-1)!(n-r)!}\frac{(r-1)!(n-r)!}{n!} = \frac{1}{n}
\end{aligned}
$$

Hence, the estimator X^* turns into sample mean $(\bar{X})$

$$\begin{aligned} X^* &= \frac{1}{n-1}\sum_{i=1}^{n}(1-\frac{1}{n})X_{i:n} \\ &= \frac{1}{n}\sum_{i=1}^{n}X_{i:n} = \bar{X} \end{aligned}$$

The tables below show the bias, mean squared error and variance of X^* that enable us to have an idea about this estimator and others. Mathematical software Mathcad is used to computation of bias, mean squared error and variance of X^*. Mathcad codes are included in appendix section 1.

					λ			
Estimator	0.0	0.5	1.0	1.5	2.0	3.0	4.0	∞
X^*_{10}	0.0	0.0061	0.013	0.018	0.021	0.014	0.0041	0.0
$\bar{X}_{10}$	0.0	0.05000	0.10000	0.15000	0.20000	0.30000	0.40000	∞
$T_{10}(1)$	0.0	0.04912	0.09325	0.12870	0.15400	0.17871	0.18470	0.18563
$T_{10}(2)$	0.0	0.04869	0.09023	0.12041	0.13904	0.15311	0.15521	0.15538
Med_{10}	0.0	0.04832	0.08768	0.11381	0.12795	0.13642	0.13723	0.13726
$W_{10}(1)$	0.0	0.04938	0.09506	0.13368	0.16298	0.19407	0.20239	0.20377
$W_{10}(2)$	0.0	0.04889	0.09156	0.12389	0.14497	0.16217	0.16504	0.16530
$M_{10}(1)$	0.0	0.04934	0.09484	0.13311	0.16194	0.19229	0.20037	0.20169
$M_{10}(2)$	0.0	0.04886	0.09137	0.12342	0.14418	0.16091	0.16369	0.16394
$L_{10}(1)$	0.0	0.04869	0.09024	0.12056	0.13954	0.15459	0.15727	0.15758
$L_{10}(2)$	0.0	0.04850	0.08892	0.11700	0.13328	0.14436	0.14576	0.14585
G_{10}	0.0	0.04847	0.08873	0.11649	0.13237	0.14285	0.14407	0.14414

Table 4. Bias of various estimators of μ for $n = 10$ when a single outlier is from $N(\mu+\lambda, 1)$ and the others from $N(\mu, 1)$

As it is seen from the table above, bias of the X^* is significantly smaller than other estimators whenever location outlier observed in the sample. If the distance between outlier's location and other sample observations increases, bias of the X^* is decreasing and X^* is becoming unbiased for large shiftings of outlier location.

	λ							
Estimator	0.0	0.5	1.0	1.5	2.0	3.0	4.0	∞
X^*_{10}	0.10000	0.10000	0.10100	0.10200	0.10500	0.10900	0.11100	0.11100
$\bar{X}_{10}$	0.10000	0.10250	0.11000	0.12250	0.14000	0.19000	0.26000	∞
$T_{10}(1)$	0.10534	0.10791	0.11471	0.12387	0.13285	0.14475	0.14865	0.14942
$T_{10}(2)$	0.11331	0.11603	0.12297	0.13132	0.13848	0.14580	0.14730	0.14745
Med_{10}	0.13833	0.14161	0.14964	0.15852	0.16524	0.17072	0.17146	0.17150
$W_{10}(1)$	0.10437	0.10693	0.11403	0.12405	0.13469	0.15039	0.15627	0.15755
$W_{10}(2)$	0.11133	0.11402	0.12106	0.12995	0.13805	0.14713	0.14926	0.14950
$M_{10}(1)$	0.10432	0.10688	0.11396	0.12385	0.13430	0.14950	0.15513	0.15581
$M_{10}(2)$	0.11125	0.11395	0.12097	0.12974	0.13770	0.14649	0.14853	0.14876
$L_{10}(1)$	0.11371	0.11644	0.12337	0.13169	0.13882	0.14626	0.14797	0.14820
$L_{10}(2)$	0.12097	0.12386	0.13105	0.13933	0.14598	0.15206	0.15310	0.15318
G_{10}	0.12256	0.12549	0.13276	0.14111	0.14777	0.15376	0.15472	0.15479

Table 5. MSE of various estimators of μ for $n = 10$ when a single outlier is from $N(\mu + \lambda, 1)$ and the others from $N(\mu, 1)$

Mean squared errors of X^* the for all outlier shifting values of parameter μ are smaller than other estimators which indicate X^* predicts μ with better accuracy than others in location outlier case of the normal distribution.

	τ					
Estimator	0.5	1.0	2.0	3.0	4.0	∞
X^*_{10}	0.09334	0.09978	0.12821	0.16632	0.20968	∞
$\bar{X}_{10}$	0.09250	0.10000	0.13000	0.18000	0.25000	∞
$T_{10}(1)$	0.09491	0.10534	0.12133	0.12955	0.13417	0.14942
$T_{10}(2)$	0.09953	0.11331	0.12773	0.13389	0.13717	0.14745
Med_{10}	0.11728	0.13833	0.15375	0.15953	0.16249	0.17150
$W_{10}(1)$	0.09571	0.10437	0.12215	0.13221	0.13801	0.15754
$W_{10}(2)$	0.09972	0.11133	0.12664	0.13365	0.13745	0.14950
$M_{10}(1)$	0.09548	0.10432	0.12187	0.13171	0.13735	0.15581
$M_{10}(2)$	0.09940	0.11125	0.12638	0.13328	0.13699	0.14876
$L_{10}(1)$	0.09934	0.11371	0.12815	0.13436	0.13769	0.14820
$L_{10}(2)$	0.10432	0.12097	0.13531	0.14101	0.14398	0.15318
G_{10}	0.10573	0.12256	0.13703	0.14270	0.14565	0.15479

Table 6. Variance of various estimators of μ for $n = 10$ when a single outlier is from $N(\mu, \tau^2)$ and the others from $N(\mu, 1)$

Since estimators of the parameter μ for the scale outlier model is unbiased, it is sufficient to compare variances of estimators. From the table above, we observe that the estimator X^* is not accurate for the scale outlier case as in the location outlier. But it gives better estimate than $\bar{X}$ for the large values of scale shift.

3.3 Conditional distributions of maximum and minimum order statistics

Distributions of order statistics can be recomputed whenever we know that the rth order statistic is outlier. Conditional distribution of maximum order statistic may be found as;

$$P\{X_{n:n} \le t|\ X_{r:n}\ is\ outlier\} = \frac{P\{X_{n:n} \le t,\ X_{r:n}\ is\ outlier\}}{P\{X_{r:n}\ is\ outlier\}}$$

$$\begin{aligned} P\{X_{n:n} &\le t,\ X_{r:n}\ is\ outlier\} = C_{n-1}^{r-1} P\{X_1 \le t, X_2 \le t, ..., X_n \le t \\ , X_1 &\le X_n, ..., X_{r-1} \le X_n, X_r > X_n, X_{r+1} > X_n, ..., X_{n-1} > X_n\} = \\ &= C_{n-1}^{r-1} P\{X_1 \le X_n, ..., X_{r-1} \le X_n, X_n \le X_r \le t, ..., X_n \le X_{n-1} \le t\} \\ &= C_{n-1}^{r-1} \int_{-\infty}^{t} F^{r-1}(x) \left(F(t) - F(x)\right)^{n-r} dG(x) \end{aligned}$$

$$P\{X_{n:n} \le t|\ X_{r:n}\ is\ outlier\} = \frac{\int_{-\infty}^{t} F^{r-1}(x) \left(F(t) - F(x)\right)^{n-r} dG(x)}{\int_{-\infty}^{\infty} F^{r-1}(x) \left(1 - F(x)\right)^{n-r} dG(x)}$$

On the other hand, distribution of minimum order statistic when $X_{r:n}\ is\ outlier$ is given can be evaluated as;

$$P\{X_{1:n} \le t|\ X_{r:n}\ is\ outlier\} = \frac{P\{X_{1:n} \le t,\ X_{r:n}\ is\ outlier\}}{P\{X_{r:n}\ is\ outlier\}}$$

$$P\{X_{1:n} \le t,\ X_{r:n}\ is\ outlier\} = P\{X_{r:n}\ is\ outlier\} - P\{X_{1:n} > t,\ X_{r:n}\ is\ outlier\}$$

$$\begin{aligned} P\{X_{1:n} &> t,\ X_{r:n}\ is\ outlier\} = C_{n-1}^{r-1} P\{X_1 > t, X_2 > t, ..., X_n > t \\ , X_1 &\le X_n, ..., X_{r-1} \le X_n, X_r > X_n, X_{r+1} > X_n, ..., X_{n-1} > X_n\} = \\ &= C_{n-1}^{r-1} P\{t \le X_1 \le X_n, ..., t \le X_{r-1} \le X_n, X_r > t, ..., X_{n-1} > t\} \\ &= C_{n-1}^{r-1} \int_{t}^{\infty} \left(F(x) - F(t)\right)^{r-1} \left(1 - F(x)\right)^{n-r} dG(x) \end{aligned}$$

$$P\{X_{1:n} \leq t,\ X_{r:n}\ is\ outlier\} =$$

$$= \frac{C_{n-1}^{r-1} \int\limits_{-\infty}^{\infty} F^{r-1}(x)\,(1-F(x))^{n-r}\,dG(x) - C_{n-1}^{r-1} \int\limits_{t}^{\infty} (F(x)-F(t))^{r-1}\,(1-F(x))^{n-r} dG(x)}{C_{n-1}^{r-1} \int\limits_{-\infty}^{\infty} F^{r-1}(x)\,(1-F(x))^{n-r}\,dG(x)}$$

3.4 Empirical distribution function

Let $X_1, X_2, ..., X_n$ be random variables with realizations $x_i = X_1(\omega) \in \mathbb{R}, i = 1, 2, ..., n$ where $X_1, X_2, ..., X_{n-1}$ from the population with cumulative distribution function $F(x)$ and X_n has cumulative distribution function $G(x)$.Empirical distribution function $F_n^*(x,\omega)$ based on $x_1, ..., x_n$

$$F_n^*(x,\omega) = \begin{cases} 0 & , x < x_{1:n} \\ \frac{1}{n-1} \sum\limits_{k=1}^{i} (1-S_k) & , x_{i:n} \leq x < x_{i+1:n} \quad , i = 1, 2, ..., n-1 \\ 1 & , x > x_{n:n} \end{cases}$$

where $S_k = P\{X_{k:n}$ is outlier$\}$ and $x_{i:n}$ denotes the realization of the random variable $X_{i:n}$ with outcome ω.
If there is no outlier in the population which means $F(x) = G(x)$, we have $S_1 = S_2 = ... = S_n = \frac{1}{n}$. Hence, empirical distribution based on iid sample is obtained as;

$$F_n(x,\omega) = \begin{cases} 0 & , x < x_{1:n} \\ \frac{i}{n} & , x_{i:n} \leq x < x_{i+1:n} \quad , i = 1, 2, ..., n-1 \\ 1 & , x > x_{n:n} \end{cases}$$

Let us denote the jump points at the point $X_{i:n}$ by P_i.

$$\begin{aligned} P_i &= F_n^*(x_{i:n},\omega) - F_n^*(x_{i-1:n},\omega) = \\ &= \frac{1}{n-1} \sum_{k=1}^{i} (1-S_k) - \frac{1}{n-1} \sum_{k=1}^{i-1} (1-S_k) \\ &= \frac{1-S_i}{n-1} \end{aligned}$$

If $S_i = \frac{1}{n}$ then $P_i = \frac{1-\frac{1}{n}}{n-1} = \frac{1}{n}, i = 1, 2, .., n$ as in the independent identical case of distributions.
More precisely;

$$\begin{aligned} F_n^*(x_{i:n} + 0) - F_n^*(x_{i:n} - 0) &\equiv P_i = \frac{1 - S_i}{n - 1} \\ &\cong P(X_F \in (X_{i:n} - 0, X_{i:n} + 0)) \\ &\cong P(X_F = X_{i:n}) = P(X_F \leq X_{i:n}) - P(X_F < X_{i:n}) \end{aligned}$$

where X_F is observation with distribution function F. If $X_{i:n}$ is and outlier with a large probability, then we take its effect to be small.
The summation of probabilities at the jump points gives

$$\begin{aligned} P_1 + ... + P_n &= \frac{1}{n-1}((1 - S_1) + ... + (1 - S_n)) = \\ &= \frac{1}{n-1}((1 + ... + 1) - (S_1 + ... + S_n)) \\ &= 1 \end{aligned}$$

Consider $\alpha_F = E_{F_n}(X) = \int x dF_n(x)$, which is the natural estimation by definition of the Stieltjes integral. Then

$$\begin{aligned} E_{F_n^*}(X) &= \int x dF_n^*(x) = \sum_{i=1}^{n} X_{i:n} P_i \\ &= \sum_{i=1}^{n} X_{i:n} \frac{1 - S_i}{n - 1} = \frac{1}{n-1} \sum_{i=1}^{n} (1 - S_i) X_{i:n} = \alpha_{F^*} \end{aligned}$$

which leads to our estimator X^*.
Similarly, $\sigma_F^2 = \int (x - \alpha_F)^2 dF_n(x)$ can be redefined by our consideration as

$$\hat{\sigma}_F^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_{i:n} - \alpha_{F^*})^2 (1 - S_i)$$

Chapter 4

Appendix

1. Bias, MSE and Variance of estimator X^* for Normal distribution.

$n := 10 \qquad rs := n \qquad r1 := 1 \qquad r := 1..n \qquad f1 := 0 \qquad f2 := 1$

$s := 1..n \qquad g1 := 4 \qquad g2 := 1$

$F(x) := \mathrm{pnorm}(x, f1, f2) \qquad f(x) := \mathrm{dnorm}(x, f1, f2)$

$G(x) := \mathrm{pnorm}(x, g1, g2) \qquad g(x) := \mathrm{dnorm}(x, g1, g2)$

$$S_r := \frac{1 - \left[\mathrm{combin}(n-1, r-1) \cdot \int_{-\infty}^{\infty} (F(x))^{r-1} \cdot (1 - F(x))^{n-r} \cdot \mathrm{dnorm}(x, g1, g2)\, dx\right]}{n-1}$$

$$d := \sum_{r=2}^{n-1} \left[S_r \cdot \left[\int_{-\infty}^{\infty} \left[\frac{(n-1)!}{(r-2)!\cdot(n-r)!} \cdot x \cdot (F(x))^{r-2} \cdot G(x) \cdot f(x) \cdot (1-F(x))^{n-r} \right] dx \right.\right.$$

$$+ \int_{-\infty}^{\infty} \left[\frac{(n-1)!}{(r-1)!\cdot(n-r)!} \cdot x \cdot (F(x))^{r-1} \cdot g(x) \cdot (1-F(x))^{n-r} \right] dx$$

$$\left.\left. + \int_{-\infty}^{\infty} \left[\frac{(n-1)!}{(r-1)!\cdot(n-r-1)!} \cdot x \cdot (F(x))^{r-1} \cdot f(x) \cdot (1-F(x))^{n-r-1} \cdot (1-G(x)) \right] dx \right]\right]$$

$$a := S_{rs} \cdot \left[\int_{-\infty}^{\infty} \left[\frac{(n-1)!}{(rs-2)!\cdot(n-rs)!} \cdot x \cdot (F(x))^{rs-2} \cdot G(x) \cdot f(x) \cdot (1-F(x))^{n-rs} \right] dx \right.$$

$$\left. + \int_{-\infty}^{\infty} \left[\frac{(n-1)!}{(rs-1)!\cdot(n-rs)!} \cdot x \cdot (F(x))^{rs-1} \cdot g(x) \cdot (1-F(x))^{n-rs} \right] dx \right]$$

$$b := S_{r1} \cdot \left[\int_{-\infty}^{\infty} \left[\frac{(n-1)!}{(r1-1)!\cdot(n-r1)!} \cdot x \cdot (F(x))^{r1-1} \cdot g(x) \cdot (1-F(x))^{n-r1} \right] dx \right.$$

$$\left. + \int_{-\infty}^{\infty} \left[\frac{(n-1)!}{(r1-1)!\cdot(n-r1-1)!} \cdot x \cdot (F(x))^{r1-1} \cdot f(x) \cdot (1-F(x))^{n-r1-1} \cdot (1-G(x)) \right] dx \right]$$

$$\begin{aligned} E(X^*) &= E(\sum_{r=1}^{n} S_r X_{r:n}) = a + b + d = 4.119 \times 10^{-3} \\ k1 &= (E(X^*))^2 = (a+b+d)^2 = 1.697 \times 10^{-5} \end{aligned}$$

$$x := \sum_{r=2}^{n-1}\left[(S_r)^2\cdot\left[\int_{-\infty}^{\infty}\left[\frac{(n-1)!}{(r-2)!\cdot(n-r)!}\cdot x^2\cdot(F(x))^{r-2}\cdot G(x)\cdot f(x)\cdot(1-F(x))^{n-r}\right]dx\right.\right.$$

$$+\int_{-\infty}^{\infty}\left[\frac{(n-1)!}{(r-1)!\cdot(n-r)!}\cdot x^2\cdot(F(x))^{r-1}\cdot g(x)\cdot(1-F(x))^{n-r}\right]dx$$

$$\left.\left.+\int_{-\infty}^{\infty}\left[\frac{(n-1)!}{(r-1)!\cdot(n-r-1)!}\cdot x^2\cdot(F(x))^{r-1}\cdot f(x)\cdot(1-F(x))^{n-r-1}\cdot(1-G(x))\right]dx\right]\right]$$

$$y := (S_{rs})^2\cdot\left[\int_{-\infty}^{\infty}\left[\frac{(n-1)!}{(rs-2)!\cdot(n-rs)!}\cdot x^2\cdot(F(x))^{rs-2}\cdot G(x)\cdot f(x)\cdot(1-F(x))^{n-rs}\right]dx\right.$$

$$\left.+\int_{-\infty}^{\infty}\left[\frac{(n-1)!}{(rs-1)!\cdot(n-rs)!}\cdot x^2\cdot(F(x))^{rs-1}\cdot g(x)\cdot(1-F(x))^{n-rs}\right]dx\right]$$

$$z := (S_{r1})^2\cdot\left[\int_{-\infty}^{\infty}\left[\frac{(n-1)!}{(r1-1)!\cdot(n-r1)!}\cdot x^2\cdot(F(x))^{r1-1}\cdot g(x)\cdot(1-F(x))^{n-r1}\right]dx\right.$$

$$\left.+\int_{-\infty}^{\infty}\left[\frac{(n-1)!}{(r1-1)!\cdot(n-r1-1)!}\cdot x^2\cdot(F(x))^{r1-1}\cdot f(x)\cdot(1-F(x))^{n-r1-1}\cdot(1-G(x))\right]dx\right]$$

$$k2 = \sum_{r=1}^{n} S_r^2 E(X_{r:n})^2 = x + y + z = 0.11$$

$$a_{r,s} := \begin{cases} 0 & \text{if } r = 1 \vee r \geq s \\ Ca\cdot\int_{-\infty}^{\infty}\int_{-\infty}^{y} x\cdot y\cdot(F(x))^{r-2}\cdot G(x)\cdot f(x)\cdot(F(y)-F(x))^{s-r-1}\cdot f(y)\,(1-F(y))^{n-s}\,dx\,dy & \text{otherwise}\end{cases}$$

$$\text{where } Ca = \frac{(n-1)!}{(r-2)!(s-r-1)!(n-s)!}$$

$$b_{r,s} := \begin{cases} 0 & \text{if } r \ge s \\ Cb \cdot \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{y} x \cdot y \cdot (F(x))^{r-1} \cdot g(x) \cdot (F(y) - F(x))^{s-r-1} \cdot f(y)\,(1 - F(y))^{n-s}\,dx\,dy & \text{otherwise} \end{cases}$$

$$\text{where } Cb = \frac{(n-1)!}{(r-1)!(s-r-1)!(n-s)!}$$

$$c_{r,s} := \begin{cases} 0 & \text{if } s = r+1 \vee r \ge s \\ Cc \cdot \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{y} x \cdot y \cdot (F(x))^{r-1} \cdot f(x) \cdot (F(y) - F(x))^{s-r-2} \cdot (G(y) - G(x)) \cdot f(y)\,(1 - F(y))^{n-s}\,dx\,dy & \text{otherwise} \end{cases}$$

$$\text{where } Cc = \frac{(n-1)!}{(r-1)!(s-r-2)!(n-s)!}$$

$$d_{r,s} := \begin{cases} 0 & \text{if } r \ge s \\ Cd \cdot \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{y} x \cdot y \cdot (F(x))^{r-1} \cdot f(x) \cdot (F(y) - F(x))^{s-r-1} \cdot g(y)\,(1 - F(y))^{n-s}\,dx\,dy & \text{otherwise} \end{cases}$$

$$\text{where } Cd = \frac{(n-1)!}{(r-1)!(s-r-1)!(n-s)!}$$

$$f_{r,s} := \begin{cases} 0 & \text{if } s = n \vee r \ge s \\ Cf \cdot \displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{y} x \cdot y \cdot (F(x))^{r-1} \cdot f(x) \cdot (F(y) - F(x))^{s-r-1} \cdot f(y)\,(1 - F(y))^{n-s-1} \cdot (1 - G(y))\,dx\,dy & \text{otherwise} \end{cases}$$

$$\text{where } Cf = \frac{(n-1)!}{(r-1)!(s-r-1)!(n-s-1)!}$$

$$k3 := 2 \cdot \sum_{r=1}^{n} \sum_{s=1}^{n} \left[S_r \cdot S_s \cdot \left(a_{r,s} + b_{r,s} + c_{r,s} + d_{r,s} + f_{r,s} \right) \right] = 9.465 \times 10^{-4}$$

$$\text{where } k3 = 2 \sum_{r<s}\sum S_r S_s E(X_{r:n} X_{s:n})$$

$$\begin{aligned}
Var(X^*) &= Var(\sum_{r=1}^{n} S_r X_{r:n}) = E((\sum_{r=1}^{n} S_r X_{r:n})^2) - (E(\sum_{r=1}^{n} S_r X_{r:n}))^2 \\
&= \sum_{r=1}^{n} S_r^2 E(X_{r:n})^2 + 2\sum\sum_{r<s} S_r S_s E(X_{r:n} X_{s:n}) - (E(\sum_{r=1}^{n} S_r X_{r:n}))^2 \\
&= k2 + k3 - k1 = 0.11078
\end{aligned}$$

$$MSE(X^*) = Var(X^*) + Bias^2(X^*)$$

Since, we considered $\mu = 0$ the $Bias(X^*)$ is equal to $E(X^*)$ so,

$$MSE(X^*) = k2 + k3 = 0.1108$$

2. $P\{X_{r:n}$ is outlier$\}$ for exponential distribution

$n := 15 \qquad r := 2 \qquad F(x) := 1 - e^{-x} \qquad G(x) := 1 - e^{-3x}$

$f(x) := e^{-x} \qquad g(x) := 3 \cdot e^{-3x}$

$$\mathrm{combin}(n-1, r-1) \cdot \int_0^{\infty} (F(x))^{r-1} \cdot (1 - F(x))^{n-r} \cdot (g(x))\, dx = 0.154$$

$$\frac{\Gamma(n) \cdot \Gamma[n - r + (3)]}{\frac{1}{3}\, \Gamma(n+3) \cdot \Gamma(n - r + 1)} = 0.154$$

Bibliography

[1] **Andrews, D. F., Bickel, P. J., Hampel, F. R. Huber, P. J. Rogers, W. H. and Tukey, J. W.** (1972). *Robust estimates of location: Survey and advances*, Princeton University Press, Princeton, N.J.

[2] **Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N.** (1992). *A first course in order statistics*, John Wiley & Sons Inc., New York.

[3] **Balakrishnan, N. and Rao, C. R.(eds.)** (1998). *Order statistics: Theory & methods, Handbook of Statistics*, vol. 16, North-Holland, Amsterdam.

[4] **Balakrishnan, N.** (2007). *"Permanents, order statistics, outliers, and robustness"*, Revista Matematica Complutense, vol. 20, pp. 7-107.

[5] **Balasubramanian, K. and Balakrishnan, N.** (1993). *"A log-concavity property of probability of occurrence of exactly r arbitrary events"*, Statist. Probab. Lett. 16 no. 3, 249–251.

[6] **Bapat, R. B. and Beg, M. I.** (1989). *"Order statistics for nonidentically distributed variables and permanents"*, Sankhya Ser. A 51, no. 1, 79–93.

[7] **Barnett, V. and Lewis, T.** (1994). *Outliers in statistical data*, 3rd ed., John Wiley & Sons Ltd., Chichester.

[8] **David, H. A. and Nagaraja, H. N.** (2003). *Order statistics*, 3rd ed., Wiley-Interscience [John Wiley & Sons], Hoboken, NJ.

[9] **David, H. A., Kennedy, W. J. and Knight, R. D.** (1977). *"Means, variances, and covariances of normal order statistics in the presence of an outlier"*, Selected Tables in Mathematical Statistics, vol. 5, pp. 75–204.

[10] **David, H. A. and Shu, V. S.** (1978). *"Robustness of location estimators in the presence of an outlier"*, Contributions to survey sampling and applied statistics: Papers in honour of H. O. Hartley (H. A. David, ed.), Academic Press, New York, pp. 235–250.

[11] **Kale, B. K. and Sinha, S. K.** (1971). *"Estimation of expected life in the presence of an outlier observation"*, Technometrics 13, 755–759.

[12] **Maronna, R. A., Martin, R. D. and Yohai, V. J.** (2006). *Robust Statistics*, John Wiley & Sons Ltd, Chichester.

[13] **Vaughan, R. J. and Venables, W. N.** (1972). *"Permanent expressions for order statistic densities"*, J. Roy. Statist. Soc. Ser. B 34, 308–310.

VITA

Kerem Türkyılmaz was born in Bornova, İzmir, Turkey, on May 07, 1984, the son of Mahmut and Saadet Türkyılmaz. He began his B.S. degree in 2002, İzmir University of Economics. Since 2004, he has continued his academic studies with Prof. Dr. İsmihan Bayramoğlu in the same department. After receiving his B.S. degree in 2006 from the Department of Mathematics, he has begun his M.S. degree in Applied Statistics in İzmir University of Economics. He is still working as research assistant in the Department of Mathematics in İzmir University of Economics.