Asymptotic Behavior of Characteristic Function of Simple Serial Rank Statistics

Chikhi El Mokhtar
Centre d’orientation et de planification de l’éducation
Avenue Allal El Fassi, Hay Riad
BP: 6222 - Rabat Instituts, Rabat, Morocco

Hammou El Hachmi
High School of Technology
P.O. Box 473, 60000 Oujda, Morocco
LaMSD, Faculty of Sciences
University Mohamed Ier, Oujda, Morocco

Rifi Khalid
High School of Technology, Km 7, Road El Jadida
P.O. Box 8012 Oasis Casablanca, Morocco

Copyright © 2017 Chikhi El Mokhtar, Hammou El Hachmi and Rifi Khalid. This article is distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract
Under suitable assumptions, verified for a wide class of score generating functions including van der Waerden, Wilcoxon and Spearman scores, we establishes an asymptotic bound on the characteristic function of serial linear rank statistics. It generalizes the result of van Zwet [13] and constitutes an essential step to the elaboration an Edgeworth expansion for distribution function of these statistics.

Mathematics Subject Classification: 62G10

Keywords: Edgeworth expansion, characteristic function, serial linear rank statistics
1 Introduction

Denotes by $H_0$ the hypothesis under which $X = (X_1, \ldots, X_n)$ is an $n$-tuple of independent and identically distributed random variables (i.i.d.), and $R = (R_1, \ldots, R_n)$ a vector of ranks.

It is well-known [5] that locally asymptotically optimal tests in linear models can be based on nonserial linear rank statistics of type

$$T_n = \sum_{i=1}^{n} c_i a_n(R_i),$$

where $a_n(1), \ldots, a_n(n)$, and $c_1, \ldots, c_n$, respectively denote a collection of scores and a triangular array of regression constants. This statistic has been studied extensively. In particular, Edgeworth expansion has been obtained for the distribution function of (1). Particularly, the problem for two-sample has been treated in Albers [1], Bickel and van Zwet [3], the same result has been obtained in Robinson [11] under null hypothesis. Robinson [12] has established an asymptotic expansion for the rank tests of several samples.

For a wide class of score generating functions including, van der Waerden, Wilcoxon, Spearman score, Does [4] has established an Edgeworth expansion for nonserial rank statistic $T_n$ with remainder $o(n^{-1})$. The proof of this result itself relies on earlier work in Albers, Bickel and van Zwet [2] and a bound on the characteristic function of (1) which is due to van Zwet [13].

However, in the statistical analysis of times series and other stochastic processes, the observations are no longer independent and more general rank-based statistic. Taking into account the serial dependence structure of the data, are need, the serial rank statistic.

Hallin and al. [6] have introduced a serial version of $T_n$ of the form

$$T_n = \frac{1}{\sqrt{n-k}} \sum_{t=k+1}^{n} a_n(R_t, \ldots, R_{t-k}),$$

where $a_n(\ldots)$ is a sequence of scores.

Moreover, they are established [7, 8] that the locally asymptotically optimal tests in the general context of linear models with ARMA error can be based on simple serial linear rank statistics of the form

$$T_n = \frac{1}{\sqrt{n-k}} \sum_{t=k+1}^{n} a_n^{(1)}(R_t) a_n^{(2)}(R_{t-k}),$$

where $a_n^{(1)}(\cdot)$ and $a_n^{(2)}(\cdot)$ are two sequence of scores.

The asymptotic normality of serial rank statistics of the form (2) or (3) has been established in Hallin and al. [6]. Hallin and Rifi [10] derived a Berry-Esséen bound for these statistics whose proof is based essentially on the works
Asymptotic behavior of characteristic of simple serial rank statistics

of Hallin and Rifi [9] and van Zwet [13].

In this paper, we consider the simple serial rank statistic $T_n$ in the case where the score $a_n^{(1)}(.)$ and $a_n^{(2)}(.)$ are derived from score generating functions $J_1$ and $J_2$ defined on $]0,1[$ by $a_n^{(i)}(j) = J_i(\frac{j}{n+1})$, $j = 1,...,n$ (approximate scores). This statistic take the following form

$$T_n = \frac{1}{\sqrt{n-k}} \sum_{t=k+1}^{n} J_1 \left( \frac{R_t}{n+1} \right) J_2 \left( \frac{R_{t-k}}{n+1} \right).$$

(4)

2 Assumptions and main theorem

Throughout this paper, we make the following assumptions, on score generating functions $J_1$ and $J_2$.

Assumption (A) The score generating functions $J_i$, $i = 1, 2$ are non-constant and three times differentiable on $]0,1[$ such that, for $i = 1, 2$,

$$\int_0^1 J_i(t) \, dt = 0 \quad \text{and} \quad \int_0^1 J_i^2(t) \, dt = 1.$$ 

There exist positive numbers $\Gamma > 0$ and $3 < \alpha < 3 + \frac{1}{14}$, such that, $\forall t \in ]0,1[,$

$$| J_i^{(3)}(t) | \leq \Gamma (t(1-t))^{-(3+\frac{1}{14})+\delta},$$

(5)

where $0 < \delta < \frac{1}{14}$.

Assumption (B) The score generating functions $J_1$ and $J_2$ are concordant, i.e., $\forall u, v \in ]0,1[, \quad J_1(u) \leq J_1(v) \Leftrightarrow J_2(u) \leq J_2(v).$  

(6)

Our main theorem is as follows.

Theorem 2.1 Under hypothesis $H_0$ and the assumptions (A) and (B), there are positive numbers $\gamma, B$ and $\beta$ such that, for $\log n \leq |u| \leq \gamma n^{3/2}$

$$| \phi_n(u) | \leq Bn^{-\beta \log n},$$

(7)

where $\phi_n$ denote the characteristic function of centered statistic (4).

3 Preliminary lemma

Lemma 3.1 (Does [4]) If the score generating function $J_i$, $i = 1, 2$ satisfies the assumption (A), then

$$\sum_{j=1}^{n} \left( J_i \left( \frac{j}{n+1} \right) - \bar{J}_i \right)^2 = n + O(n^{4-2\delta}),$$

where $\bar{J}_i = \frac{1}{n} \sum_{j=1}^{n} J_i \left( \frac{j}{n+1} \right)$ and $\delta$ is given in assumption (A).
Lemma 3.2 If the score generating functions $J_i$, $i = 1, 2$, satisfies the assumption (A), then there are positive numbers $a$ and $A$ such that
\[
\sum_{j=1}^{n} \left| J_i \left( \frac{j}{n+1} \right) - \bar{J}_i \right| > an,
\]
and
\[
\sum_{j=1}^{n} \left( J_i \left( \frac{j}{n+1} \right) - \bar{J}_i \right)^2 \leq An,
\]
where $\bar{J}_i = \frac{1}{n} \sum_{j=1}^{n} J_i \left( \frac{j}{n+1} \right)$.

Proof. According to Lemma 3.1, we have
\[
\sum_{j=1}^{n} \left( J_i \left( \frac{j}{n+1} \right) - \bar{J}_i \right)^2 = n + O \left( n^{\frac{3}{2} - 2\delta} \right). \tag{8}
\]
Then from positive number $c_i$, we have
\[
\sum_{j=1}^{n} \left( J_i \left( \frac{j}{n+1} \right) - \bar{J}_i \right)^2 \geq c_in. \tag{9}
\]
Otherwise, the generating functions $J_i$, $i = 1, 2$, are integrable to order 4, which tends the existence of positive number $C_i$, such that
\[
\sum_{j=1}^{n} \left( J_i \left( \frac{j}{n+1} \right) - \bar{J}_i \right)^4 \leq C_in. \tag{10}
\]
The last two relations and by using the Cauchy-Schwartz inequality, we obtain the proof of last lemma.

Lemma 3.3 If the score generating functions $J_i$, $i = 1, 2$, satisfies the assumption (A), then there are positive numbers $\delta$ and $\zeta \geq n^{-3/2} \log(n)$ such that
\[
\sum_{j=1}^{n} \left| J_i \left( \frac{j}{n+1} \right) - \bar{J}_i \right| > an,
\]
and
\[
\gamma(J_i(\frac{1}{n+1}), \ldots, J_i(\frac{n}{n+1}), \zeta) \geq \delta \zeta n,
\]
where $\gamma$ is the Lebesgue measure of $\zeta$- neighborhood of the set $\{J_i(\frac{1}{n+1}), J_i(\frac{2}{n+1}), \ldots, J_i(\frac{n}{n+1})\}$, defined by
\[
\gamma(J_i(\frac{1}{n+1}), \ldots, J_i(\frac{n}{n+1}), \zeta) = \lambda \left( \bigcup_{j=1}^{n} [J_i(\frac{j}{n+1}) - \zeta, J_i(\frac{j}{n+1}) + \zeta] \right)
\]
Proof. From the assumption (A), the function $J_i$ isn’t constant and three times differentiable on $(0,1)$, then it exist a real number $\zeta$ in $(0,1)$ such that $J'_i(\zeta) \neq 0$, without loss of generality, we can assume that $J'_i(\zeta) > 0$. Then we can find a $\zeta$-neighborhood denoted by $]\zeta_1, \zeta_2[$ such that for all $t \in ]\zeta_1, \zeta_2[$, we have

$$J'_i(\zeta) \geq \eta > 0.$$  \hspace{1cm} (11)

Put $\zeta_0 = (\zeta_2 - \zeta_1)/3$, and let

$$n_0 = \min \left\{ n \text{ such that } n^{-1} \leq \zeta_0 \text{ and } \frac{\sqrt{n} \eta}{2 \log(n)} \geq 1 \right\}.$$  

If $n \geq n_0$, we have

$$\frac{j}{n+1} \in ]\zeta_1 + \zeta_0, \zeta_2 - \zeta_0[ \text{ for at least } [(n+1)\zeta_0] \text{ index } j,$$  \hspace{1cm} (12)

where $[.]$ denotes integer part. By using the Taylor’s expansion, we have

$$J_i \left( \frac{j+1}{n+1} \right) - J_i \left( \frac{j}{n+1} \right) = \frac{1}{n+1} J'_i(\theta_j),$$  \hspace{1cm} (13)

where $\theta_j = \alpha \frac{j}{n+1} + (1 - \alpha) \frac{j+1}{n+1}$, with $\alpha \in (0,1)$.

From the relations (11), (12) and (13), we deduce, for $\zeta \geq n^{-\frac{3}{2}} \log(n)$, the proof of Lemma (3.3).

4 Proof of main theorem

Since $X_1, X_2, \cdots, X_n$ are i.i.d. then the assumption (A4) is trivial. Furthermore, the two assumptions (A2) and (B) are identical.

Lemmas (3.2) and (3.3) ensure that the assumption (A) leads the two assumptions (A1) and (A3) given in Hallin and Rifi [9].

References


[3] P.J. Bickel and W.R. van Zwet, Asymptotic expansions for the power of

https://doi.org/10.1214/aos/1176346166

1967.


[7] M. Hallin and M. L. Puri, Optimal rank-based procedures for time-
series analysis: testing an ARMA model against other ARMA models, *The
https://doi.org/10.1214/aos/1176350712

[8] M. Hallin and M. L. Puri, Rank tests for time series analysis: a survey,
In *New Directions In Time Series Analysis*, D. Brillinger, E. Parzen and

[9] M. Hallin and Kh. Rifi, The asymptotic behavior of the characteristic func-
(1996), 199 - 213.

[10] M. Hallin and Kh. Rifi, A Berry-Esséen theorem for serial rank statistics,
https://doi.org/10.1023/a:1003286814679

[11] J. Robinson, An asymptotic expansion for samples from a finite popula-
https://doi.org/10.1214/aos/1176344306

[12] J. Robinson, An asymptotic expansion for permutation tests with several
https://doi.org/10.1214/aos/1176345078

statistic, In: *Nonparametric Statistical Inference*, Vol. II, B. V. Gnedenko,
M. L. Puri and I. Vincze, Eds., North-Holland, Amsterdam, 1980, 889 -
909.

Received: April 11, 2017; Published: May 12, 2017