Characterizations based on regression assumptions of order statistics

Wen-Jang Huang* and Nan-Cheng Su†

*Department of Applied Mathematics, National University of Kaohsiung, Kaohsiung 81148, Taiwan;
†Department of Statistics, National Taipei University, New Taipei City 23741, Taiwan

Abstract

Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ be the order statistics from independent and identically distributed random variables $\{X_i,\ 1 \leq i \leq n\}$ with a common absolutely continuous distribution function. In this work, first a new characterization of distributions based on order statistics is presented. Next we review some conditional expectation properties of order statistics, which can be used to establish some equivalent forms for conditional expectations for sum of random variables based on order statistics. Using these equivalent forms, some known results can be extended immediately.

Keywords: Characterization; exponential distribution; gamma distribution; inverse gamma distribution; normal distribution; order statistics; regression function; student t distribution; uniform distribution.
AMS 2000 Subject Classifications: Primary: 62E10; Secondary: 62G30

1 Introduction

For a fixed $n \geq 1$, let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ be the order statistics based on a sequence of independent and identically distributed (i.i.d.) random variables (r.v.'s) $\{X_i,\ 1 \leq i \leq n\}$, which has a common absolutely continuous distribution function F and probability density function f. Also assume that F has support (a, b), where $-\infty \leq a < b \leq \infty$. It is known that order statistics play an important role in statistics, reliability theory, and many applied areas. There

*Corresponding author. E-mail: huangwj@nuk.edu.tw; Tel: 886-7-5919169; Fax: 886-7-5919360. Support for this research was provided in part by the National Science Council of the Republic of China, Grant No. NSC 100-2118-M-390-002-MY2.

†E-mail: sunanchen@gmail.com; Tel: 886-2-86741111 ext 66778; Fax: 886-2-86715907. Support for this research was provided in part by the National Science Council of the Republic of China, Grant No. NSC 101-2118-M-305-003

are a lot of studies on characterizing F based on conditional moments of order statistics. For example, each of the following assumptions:

$$E(X_{(i)}|X_{(i+1)} = x) = \eta x + \theta,$$

$$E(X_1|X_{(n)} = x) = \eta x + \theta,$$

and

$$E(X_1|X_{(1)} = x, X_{(n)} = y) = \frac{1}{2}(x + y), \tag{1}$$

where $1 \leq i \leq n-1$, and η, θ are constants, can be used to characterize F, see Ferguson (1967), Beg and Kirmani (1974), and Das Gupta *et al.* (1993), respectively. On the other hand, Huang and Su (1999) showed that

$$E\left(\frac{1}{k_2 - k_1 - 1}\sum_{i=k_1+1}^{k_2-1} X_{(i)} \middle| X_{(k_1)} = u, X_{(k_2)} = x\right) = P(u, x),\ u < x, \tag{2}$$

where $1 \leq k_1 < k_2 \leq n$ and $k_2 - k_1 \geq 2$, can determine F. Note that (2) with $k_1 = k$, $k_2 = k+2$, and $P(u, x) = (u + x)/2$, is reduced to

$$E\left(X_{(k+2)} - X_{(k+1)}|X_{(k)} = u, X_{(k+2)} = x\right) = \frac{1}{2}(x - u),\ u < x, \tag{3}$$

a condition which can be used to characterize the uniform distribution. Inspired by this, in this work we will characterize F by a condition which is similar to (3) yet with quadratic functions. We state the result as the following theorem, and the proof will be given in Section 2.

Theorem 1. *Let $E(X_1^2) < \infty$ and f be differentiable. Assume that*

$$E\left(\left(X_{(l)} - X_{(l-1)}\right)^2 |X_{(k)} = u, X_{(l)} = x\right) = \eta(x - u)^2,\ a < u < x < b, \tag{4}$$

for some fixed integers $1 \leq k < l \leq n$, where $l - k \geq 2$.

(i) *If (4) holds for a fixed u, and every $u < x < b$, then $0 < \eta < 1$, $b \in R$, and*

$$F(x) = F(u) + \frac{(1 - F(u))(x - u)^r}{(b - u)^r},\ u < x < b, \tag{5}$$

where $r = (\sqrt{1 + 8/\eta} - 3)/(2(l - k - 1)) > 0$.

(ii) *If (4) holds for $u = u_i$, and every $u_i < x < b$, $i = 1, 2$, where $a < u_1 < u_2 < b$, then $\eta = 2/((l - k + 1)(l - k))$, $b \in R$, and*

$$F(x) = F(u_1) + \frac{(1 - F(u_1))(x - u_1)}{(b - u_1)},\ u_1 < x < b. \tag{6}$$

Remark 1. In (ii) if $u_1 = a$, then F is the distribution function of the uniform distribution over (a, b).

Remark 2. It can be seen easily that there is a parallel characterization by using

$$E\left(\left(X_{(k+1)} - X_{(k)}\right)^2 \middle| X_{(k)} = x, X_{(l)} = u\right) = \eta(u-x)^2,\ a < x < u < b,$$

a condition similar to (4). We omit the statement of this result.

In the past decade, characterizations of the Student's t_v distribution with $v > 0$ degrees of freedom based on some simple properties of certain regression functions associated with the order statistics have been investigated by many authors. Among them, for $n = 3$, Nevzorov *et al.* (2003) showed that the regression relation

$$E(X_{(2)} - X_{(1)} | X_{(2)} = x) = E(X_{(3)} - X_{(2)} | X_{(2)} = x), \tag{7}$$

characterizes the t_2 distribution, and Akhundov and Nevzorov (2010) used

$$E((X_{(2)} - X_{(1)})^2 | X_{(2)} = x) = E((X_{(3)} - X_{(2)})^2 | X_{(2)} = x), \tag{8}$$

to determine the t_3 distribution. Later, Huang and Su (2010) pointed out that (7) and (8) can be extended to the general case with any sample size n, and used a weaker than (8) assumption

$$E\left(\left(\frac{\eta}{2} X_{(2)} - X_{(1)} + \frac{\theta}{2}\right)^2 \middle| X_{(2)} = x\right) = E\left(\left(X_{(n)} - \frac{\eta}{2} X_{(n-1)} - \frac{\theta}{2}\right)^2 \middle| X_{(n-1)} = x\right), \tag{9}$$

to characterize a large class of distributions including not only the t_v distribution but also the normal, gamma, exponential, inverse gamma and uniform distributions. The special case of $a = -\infty$, $b = \infty$, $n = 3$, $\theta = 0$, and certain types of η, were studied by Wang (2011). Recently, for $a = -\infty$, $b = \infty$, $n \geq 3$, $2 \leq k \leq n-1$ and $v \geq 3$, Yanev and Ahsanullah (2012) characterized the t_2 and t_v distributions by the regression relations

$$E\left(\frac{1}{k-1} \sum_{i=1}^{k-1} \left(X_{(k)} - X_{(i)}\right) \middle| X_{(k)} = x\right) = E\left(\frac{1}{n-k} \sum_{i=k+1}^{n} \left(X_{(i)} - X_{(k)}\right) \middle| X_{(k)} = x\right) \tag{10}$$

and

$$\begin{aligned} & E\left(\frac{1}{k-1} \sum_{i=1}^{k-1} \left(\frac{v-1}{2} X_{(k)} - (v-2) X_{(i)}\right)^2 \middle| X_{(k)} = x\right) \\ = \ & E\left(\frac{1}{n-k} \sum_{i=k+1}^{n} \left((v-2) X_{(i)} - \frac{v-1}{2} X_{(k)}\right)^2 \middle| X_{(k)} = x\right), \end{aligned} \tag{11}$$

which are reduced to (7) and (8) as $n = 3, k = 2, v = 2$, and $n = 3, k = 2, v = 3$, respectively. Other related characterizations of distributions based on order statistics can be found in Balakrishnan and Akhundov (2003), Akhundov *et al.* (2004), Nevzorova *et al.* (2007), Huang *et al.* (2007) and Akhundov and Nevzorov (2010), etc.

Motivated by the above results, in Section 3 we give a lemma for order statistics, which will enable us to establish some equivalent forms for conditional expectations for sum of r.v.'s based on order statistics. Using this lemma, many of the characterization results in the literature for the order statistics immediately have the corresponding characterizations by regression of sum of conditional moments of order statistics. As an example, we use the result of Huang and Su (2010) to obtain characterization of a large class of distributions by using the corresponding regression assumption, which is also more general than the characterization result of Yanev and Ahsanullah (2012) given in (11).

Finally, some conclusions are discussed in Section 4.

2 The proof of Theorem 1

We need a lemma which can be found in Boyce and DiPrima (1997).

Lemma 1. *Consider the Euler equation*:

$$t^2 g''(t) + \alpha t g'(t) + \beta g(t) = 0, \tag{12}$$

where t belongs to an interval not containing the origin, and α and β are some fixed real numbers. Then

$$g(t) = \begin{cases} c_1|t|^{(1-\alpha+\sqrt{(1-\alpha)^2-4\beta})/2} + c_2|t|^{(1-\alpha-\sqrt{(1-\alpha)^2-4\beta})/2}, & \text{if } (1-\alpha)^2 > 4\beta, \\ (c_3 + c_4 \log|t|)|t|^{(1-\alpha)/2}, & \text{if } (1-\alpha)^2 = 4\beta, \end{cases}$$

where c_1, c_2, c_3 and c_4 are arbitrary constants.

Proof of Theorem 1. (i) First $\eta > 0$ is obvious. From (4), we have

$$\begin{aligned} &(l-k-1)\int_u^x (x-w)^2 f(w)(F(w)-F(u))^{l-k-2}dw \\ =\ & \eta(x-u)^2(F(x)-F(u))^{l-k-1},\ u < x < b. \end{aligned} \tag{13}$$

Taking the second derivatives of both sides of (13) with respect to x and after some manipulations, we arrive at

$$\begin{aligned} &\eta(x-u)^2\frac{d^2}{dx^2}(F(x)-F(u))^{l-k-1} + 4\eta(x-u)\frac{d}{dx}(F(x)-F(u))^{l-k-1} \\ &+2(\eta-1)(F(x)-F(u))^{l-k-1} = 0,\ u < x < b, \end{aligned} \tag{14}$$

which can be rewritten as

$$t^2 G''(t) + 4tG'(t) + 2(1-1/\eta)G(t) = 0,\ 0 < t < b-u, \tag{15}$$

where $t = x - u$, and $G(t) = (F(t+u) - F(u))^{l-k-1} = (F(x) - F(u))^{l-k-1}$, $t > 0$.

If $\eta = 1$, by Lemma 1, the solution of (15) is

$$G(t) = c_1 + \frac{c_2}{t^3},\ 0 < t < b - u, \tag{16}$$

where c_1 and c_2 are constants. This contradicts to the fact that $\lim_{t\downarrow 0} G(t) = 0$. Hence $\eta \neq 1$.

Now consider $\eta \neq 1$. Again by Lemma 1, the solution of (15) is

$$G(t) = c_3 t^{(-3+\sqrt{1+8/\eta})/2} + c_4 t^{(-3-\sqrt{1+8/\eta})/2},\ 0 < t < b - u,$$

where c_3 and c_4 are constants. As $\lim_{t\downarrow 0} G(t) = 0$ and $G(t) > 0, 0 < t < b - u$, we obtain $c_3 > 0$, $c_4 = 0$, and $0 < \eta < 1$. Consequently,

$$F(x) = F(u) + c_3^{1/(l-k-1)}(x-u)^r,\ u < x < b, \tag{17}$$

where $r = (\sqrt{1+8/\eta} - 3)/(2(l-k-1)) > 0$. By letting $x \to b$ in (17), we obtain $b < \infty$ and (5) follows immediately. On the other hand, it can be shown easily that the F given in (5) satisfies (13). The proof of (i) is completed.

(ii) First from (i) we have

$$F(x) = F(u_i) + \frac{(1-F(u_i))(x-u_i)^r}{(b-u_i)^r},\ u_i < x < b,\ i = 1, 2. \tag{18}$$

By comparing $F(x)$ for $i = 1, 2$ in (18), it yields

$$\frac{(1-F(u_1))(x-u_1)^r}{(b-u_1)^r} = \frac{(1-F(u_1))(u_2-u_1)^r}{(b-u_1)^r} + \frac{(1-F(u_2))(x-u_2)^r}{(b-u_2)^r}, u_2 < x < b.$$

This in turn implies

$$\frac{(x-u_1)^r - (u_2-u_1)^r}{(x-u_2)^r} = \frac{(1-F(u_2))(b-u_1)^r}{(1-F(u_1))(b-u_2)^r},\ u_2 < x < b. \tag{19}$$

That the right hand side of (19) is a constant implies $r = 1$, or equivalently $\eta = 2/((l-k+1)(l-k))$. Again it can be shown that the F given in (6) satisfies (4). The proof of (ii) is finished.

3 One lemma and its implications

First we give a useful lemma.

Lemma 2. *Let g and h be Borel measurable functions of three variables and two variables, respectively.* (i) *For every $1 \le k_1 < k_2 \le n$, where $k_2 - k_1 \ge 2$, $1 \le k \le n-2$, and $a \le x < y \le b$, we have*

$$\begin{aligned} & E\left(\frac{1}{k_2-k_1-1}\sum_{i=k_1+1}^{k_2-1} g(X_{(i)}, x, y)\middle| X_{(k_1)} = x, X_{(k_2)} = y\right) \\ = & \frac{1}{F(y)-F(x)}\int_x^y g(t, x, y)\,dF(t) \\ = & E\left(g(X_{(k+1)}, x, y)\middle| X_{(k)} = x, X_{(k+2)} = y\right). \end{aligned} \tag{20}$$

(ii) *For $2 \le k \le n$, and $a \le x \le b$, we have*

$$E\left(\frac{1}{k-1} \sum_{i=1}^{k-1} h(X_{(i)}, x) \middle| X_{(k)} = x \right) = \frac{1}{F(x)} \int_a^x h(t,x) dF(t) \tag{21}$$
$$= \quad E\left(h(X_{(1)}, x) \middle| X_{(2)} = x \right).$$

(iii) *For $1 \le k \le n-1$, and $a \le x \le b$, we have*

$$E\left(\frac{1}{n-k} \sum_{i=k+1}^{n} h(X_{(i)}, x) \middle| X_{(k)} = x \right) = \frac{1}{1-F(x)} \int_x^b h(t,x) dF(t) \tag{22}$$
$$= \quad E\left(h(X_{(n)}, x) \middle| X_{(n-1)} = x \right).$$

Although it was mentioned by Das Gupta *et al.* (1993), that the proof of the above lemma is elementary, as it may not be so well-known, we briefly sketch the proof of (20) in the following. First it is known that given $X_{(k_1)} = x$, and $X_{(k_2)} = y$, $(X_{(k_1+1)}, \cdots, X_{(k_2-1)})$ are distributed as the order statistics of $k_2 - k_1 - 1$ i.i.d. r.v.'s, say $Z_1, \cdots, Z_{k_2-k_1-1}$, with the common distribution function

$$\frac{F(t) - F(x)}{F(y) - F(x)}, \ t \in [x, y],$$

see, e.g., Theorem 2.5 of David and Nagaraja (2003). Consequently, given $X_{(k_1)} = x$, and $X_{(k_2)} = y$,

$$\sum_{i=k_1+1}^{k_2-1} g(X_{(i)}, x, y) \overset{d}{=} \sum_{i=1}^{k_2-k_1-1} g(Z_i),$$

where $\overset{d}{=}$ denotes equality in distribution. Hence

$$E\left(\frac{1}{k_2 - k_1 - 1} \sum_{i=k_1+1}^{k_2-1} g(X_{(i)}, x, y) \middle| X_{(k_1)} = x, X_{(k_2)} = y \right) = E(g(Z_1)).$$

The rest of the proof follows immediately.

By Case (i) of Lemma 2, the following condition

$$E\left(\frac{1}{k_2 - k_1 - 1} \sum_{i=k_1+1}^{k_2-1} \left(X_{(k_2)} - X_{(i)} \right)^2 \middle| X_{(k_1)} = u, X_{(k_2)} = x \right) = \eta(x-u)^2, a < u < x < b, \tag{23}$$

where $1 \le k_1 < k_2 \le n$, $k_2 - k_1 \ge 2$, is equivalent to

$$E\left(\left(X_{(k+2)} - X_{(k+1)} \right)^2 \middle| X_{(k)} = u, X_{(k+2)} = x \right) = \eta(x-u)^2, \ a < u < x < b,$$

for every $1 \le k \le n-2$, a special form of (4) with $l = k+2$. Hence, according to Theorem 1, there is a corresponding characterization of F based on (23).

Let $E(X_1) = \mu_1 < \infty$, $E(X_1^2) = \mu_2 < \infty$. As X_1 is nondegenerate, $\mu_2 - \mu_1^2 = \mathrm{Var}(X_1) > 0$. Assume for some fixed $n \ge 2$, $2 \le k \le n$, $1 \le l \le n-1$, and constants η and θ,

$$\begin{aligned} & E\left(\frac{1}{k-1} \sum_{i=1}^{k-1} \left(\frac{\eta}{2} X_{(k)} - X_{(i)} + \frac{\theta}{2} \right)^2 \middle| X_{(k)} = x \right) \\ = \; & E\left(\frac{1}{n-l} \sum_{i=l+1}^{n} \left(X_{(i)} - \frac{\eta}{2} X_{(l)} - \frac{\theta}{2} \right)^2 \middle| X_{(l)} = x \right), \; a < x < b. \end{aligned} \tag{24}$$

By Cases (ii) and (iii) of Lemma 2, (24) implies

$$\begin{aligned} & E\left(\left(\frac{\eta}{2} X_{(2)} - X_{(1)} + \frac{\theta}{2} \right)^2 \middle| X_{(2)} = x \right) \\ = \; & E\left(\left(X_{(n)} - \frac{\eta}{2} X_{(n-1)} - \frac{\theta}{2} \right)^2 \middle| X_{(n-1)} = x \right), \; a < x < b. \end{aligned} \tag{25}$$

which is exactly (16) of Huang and Su (2010), hence the solution of (24) follows immediately. Table 1 taken from Huang and Su (2010), lists some widely used distributions that can be characterized by using (24).

Note that in the case $(\eta, \theta, a, b) = (\eta, (2-\eta)\mu_1, -\infty, \infty)$, if $\mu_1 = 0$, then $\theta = 0$, and (24) becomes

$$\begin{aligned} & E\left(\frac{1}{k-1} \sum_{i=1}^{k-1} \left(\frac{\eta}{2} X_{(k)} - X_{(i)} \right)^2 \middle| X_{(k)} = x \right) \\ = \; & E\left(\frac{1}{n-l} \sum_{i=l+1}^{n} \left(X_{(i)} - \frac{\eta}{2} X_{(l)} \right)^2 \middle| X_{(l)} = x \right), \; -\infty < x < \infty. \end{aligned} \tag{26}$$

Hence the characterization of t_v distribution by Yanev and Ahsanullah (2012), is a special case of the present characterization with $\eta = (v-1)/(v-2)$ and $k = l$.

Huang and Su (2010) also characterized distributions by either of the two assumptions

$$E(X_{(1)}^2 - (\eta X_{(2)} + \theta) X_{(1)} | X_{(2)} = x) = \gamma x + \delta, \; a < x < b, \tag{27}$$

$$E(X_{(n)}^2 - (\eta X_{(n-1)} + \theta) X_{(n)} | X_{(n-1)} = x) = \gamma x + \delta, \; a < x < b. \tag{28}$$

As (27) and (28) are equivalent to

$$E\left(\frac{1}{k-1} \sum_{i=1}^{k-1} \left(X_{(i)}^2 - \left(\eta X_{(k)} + \theta \right) X_{(i)} \right) \middle| X_{(k)} = x \right) = \gamma x + \delta, \; a < x < b, \tag{29}$$

where $2 \le k \le n$, and

$$E\left(\frac{1}{n-k} \sum_{i=k+1}^{n} \left(X_{(i)}^2 - \left(\eta X_{(k)} + \theta \right) X_{(i)} \right) \middle| X_{(k)} = x \right) = \gamma x + \delta, \; a < x < b, \tag{30}$$

Table 1
Characterization of distributions by using assumption (24) for certain (η, θ, a, b).

(η, θ, a, b)	Distribution	$f(x)$
$(1, \mu_1, -\infty, \infty)$	Normal $\mathcal{N}(\mu_1, \sigma^2)$	$\frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu_1)^2}{2\sigma^2}\right\}$, where $\sigma^2 = \mu_2 - \mu_1^2$.
$(1, \frac{\mu_2}{\mu_1}, 0, \infty)$	Gamma $\mathcal{G}amma(\alpha, \beta)$	$\frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp\left\{-\frac{x}{\beta}\right\}$, where $\alpha = \frac{\mu_1^2}{\mu_2-\mu_1^2}$ and $\beta = \frac{\mu_2-\mu_1^2}{\mu_1}$.
$(1, 2\mu_1, 0, \infty)$	Exponential $\mathcal{E}(\lambda)$	$\lambda e^{-\lambda x}$, where $\lambda = \frac{1}{\mu_1}$.
$(\eta, (2-\eta)\mu_1, -\infty, \infty)$, where $\eta > 1$.	Student t $t_v(\mu_1, \rho)$	$\frac{\Gamma((v+1)/2)}{\rho\sqrt{\pi v}\Gamma(v/2)} \left(\left(\frac{x-\mu_1}{\rho}\right)^2 \frac{1}{v} + 1\right)^{-(v+1)/2}$, where $v = \frac{2\eta-1}{\eta-1}$ and $\rho = \sqrt{\frac{\mu_2-\mu_1^2}{2\eta-1}}$.
$(\frac{2}{3}, \frac{4}{3}\mu_1, a, b)$	Uniform $\mathcal{U}(a, b)$	$\frac{1}{b-a}$
$(\frac{\mu_2}{\mu_1^2}, \frac{\mu_2}{\mu_1}, 0, \infty)$	Inverse Gamma $\mathcal{IG}amma(\alpha, \beta)$	$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp\left\{-\frac{\beta}{x}\right\}$, where $\alpha = \frac{2\mu_2-\mu_1^2}{\mu_2-\mu_1^2}$ and $\beta = \frac{\mu_1\mu_2}{\mu_2-\mu_1^2}$.

where $1 \leq k \leq n-1$, respectively, it can be expected that there are corresponding characterizations of distributions by using either one of the assumptions (29) and (30). We omit the statements of the results.

In Theorem 1 of Yanev and Ahsanullah (2012), the condition

$$\lambda E\left(\frac{1}{k-1}\sum_{i=1}^{k-1} X_{(i)} \middle| X_{(k)} = x\right) + (1-\lambda)E\left(\frac{1}{n-k}\sum_{i=k+1}^{n} X_{(i)} \middle| X_{(k)} = x\right) = x, \qquad (31)$$

was used to characterize F belonging to the so-called Q-family. It should be mentioned here, again by Cases (ii) and (iii) of Lemma 2, the above condition is equivalent to

$$\lambda E(X_{(1)}|X_{(2)} = x) + (1-\lambda)E(X_{(n)}|X_{(n-1)} = x) = x,$$

and

$$\lambda E(X_{(1)}|X_{(2)} = x) + (1-\lambda)E(X_{(3)}|X_{(2)} = x) = x,$$

a relation which has already been used by Akhundov *et al.* (2004) to obtain the same characterization.

4 Conclusion

We are interested in knowing which distributions can be characterized by using the following more general than (4) form

$$\begin{aligned} & E\left(X_{(l)}^2 + p_1 X_{(l-1)}^2 + p_2 X_{(l-1)} X_{(l)} + p_3 X_{(l)} + p_4 X_{(l-1)} \middle| X_{(k)} = u, X_{(l)} = x\right) \\ = \quad & q_0 x^2 + q_1 u^2 + q_2 ux + q_3 x + q_4 u, \ a < u < x < b, \end{aligned} \tag{32}$$

where $p_1, \cdots, p_4, q_0, q_1, \cdots, q_4$ are constants. That (4) corresponds to $p_1 = 1$, $p_2 = -2$, $p_3 = p_4 = 0$, $q_0 = q_1 = \eta$, $q_2 = -2\eta$, and $q_3 = q_4 = 0$. Suppose u is fixed. By (20), (32) implies

$$\begin{aligned} & (l-k-1) \int_u^x (x^2 + p_1 w^2 + p_2 wx + p_3 x + p_4 w) f(w) (F(w) - F(u))^{l-k-2} dw \\ = \quad & (q_0 x^2 + q_1 u^2 + q_2 ux + q_3 x + q_4 u)(F(x) - F(u))^{l-k-1}, u < x < b. \end{aligned} \tag{33}$$

Taking the second derivatives of both sides of (33) with respect to x and after some manipulations, it yields

$$P(t)G''(t) + Q(t)G'(t) + 2(q_0 - 1)G(t) = 0, 0 < t < b - u, \tag{34}$$

where $t = x - u$, $G(t) = (F(t+u) - F(u))^{l-k-1}, t > 0$, $P(t) = (q_0 - p_1 - p_2 - 1)t^2 + (q_0 + q_1 + q_2 - p_1 - p_2 - 1)u^2 + (q_3 - p_3 - p_4)t + (q_3 + q_4 - p_3 - p_4)u + (2q_0 + q_2 - 2p_1 - 2p_2 - 2)tu$, and $Q(t) = (4q_0 - 2p_1 - 3p_2 - 4)t + (4q_0 + 2q_2 - 2p_1 - 3p_2 - 4)u$. Although for differential equations of second order, there may have solutions for forms different from (12), for (34), it can be shown easily that only if $p_2 = -2p_1$, $q_1 = p_1 + q_0 - 1 \neq 0$, $q_2 = -2q_1$, $q_3 = p_3 + p_4$, and $q_4 = 0$, will reduce to

$$t^2 G''(t) + 4tG'(t) + 2\left(q_0/q_1 - 1/q_1\right) G(t) = 0, 0 < t < b - u, \tag{35}$$

a form which can be solved by Lemma 1. In this case the condition (32) can be simplified to

$$\begin{aligned} & E\left(X_{(l)}^2 + p_1 X_{(l-1)}^2 - 2p_1 X_{(l-1)} X_{(l)} + p_3 X_{(l)} + p_4 X_{(l-1)} \middle| X_{(k)} = u, X_{(l)} = x\right) \\ & = q_0 x^2 + q_1 u^2 - 2q_1 ux + (p_3 + p_4)\, x, a < u < x < b, \end{aligned} \tag{36}$$

which nevertheless is more general than (4).

It is also interesting to characterize the common distribution of $\{X_i, i \geq 1\}$ by using the more general relation

$$E\left(g_1\left(X_{(i)}, u, x\right) | X_{(k)} = u, X_{(l)} = x\right) = g_2(u, x), \ a < u < x < b, \tag{37}$$

where $k < i < l$, $g_1 : R^3 \to R$ and $g_2 : R^2 \to R$. In this work, we have offered $i = l - 1$, $k + 1$ and some functions g_1, g_2 for which there are solutions. Other possible i's, g_1's and g_2's which may lead to (37) having solutions will be studied in the future.

For an integer $n \geq 1$, in the following let $X_{1,n} \leq X_{2,n} \leq \cdots \leq X_{n,n}$ be the order statistics based on $\{X_1, X_2, \cdots, X_n\}$. Now for integers $2 \leq k \leq n$ and $2 \leq l \leq m$, it can be seen easily, that (21) implies

$$E\left(\frac{1}{k-1} \sum_{i=1}^{k-1} h(X_{i,n}, x) \middle| X_{k,n} = x \right) = E\left(\frac{1}{l-1} \sum_{i=1}^{l-1} h(X_{i,m}, x) \middle| X_{l,m} = x \right),$$

and there is a similar equation based on (22). Consequently, some results in this work can be generalized to the situation with different sample sizes of order statistics.

Finally, when viewed as point process, the sequence of order statistics $\{X_{i,n}, 1 \leq i \leq n\}$ forms a sample process, and has the so-called order statistics property, see Puri (1982). Along the lines of Huang and Su (1999) and Huang *et al.* (2007), it is expected that there is a point process version for Theorem 1. This and some other related works will also be studied in the future.

Acknowledgements

The authors are very grateful to the anonymous associate editor and two referees for their careful review of this paper and many valuable comments and suggestions, which have greatly improved the presentation of this paper.

References

[1] Akhundov, I. and Nevzorov, V.B. (2010). A simple characterization of Student's t_3 distribution. *Statist. Prob. Lett.* 80, 293-295.

[2] Akhundov, I.S., Balakrishnan, N. and Nevzorov, V.B. (2004). New characterizations by properties of midrange and related statistics. *Commun. Statist. Theory Meth.* 33, 3133-3143.

[3] Balakrishnan, N. and Akhundov, I.S. (2003). A characterization by linearity of the regression function based on order statistics. *Statist. Prob. Lett.* 63, 435-440.

[4] Beg, M.I. and Kirmani, S.N.U.A. (1974). On a characterization of exponential and related distributions. *Austral. J. Statist.* 16, 163-166.

[5] Boyce, W.E. and DiPrima, R.C. (1997). *Elementary Differential Equations and Boundary Value Problems.* Wiley, New York.

[6] Das Gupta, S., Goswami, A. and Rao, B.V. (1993). On a characterization of uniform distributions. *J. Multivariate Anal.* 44, 102-114.

[7] David, H.A. and Nagaraja, H.N. (2003). *Order Statistics, 3rd ed.* Wiley, New York.

[8] Ferguson, T.S. (1967). On characterizing distributions by properties of order statistics. *Sankhyā A* 29, 265-278.

[9] Huang, W.J. and Su, J.C. (1999). On certain problems involving order statistics - a unified approach through order statistics property of point processes. *Sankhyā A* 61, 36-49.

[10] Huang, W.J. and Su, N.C. (2010). Characterizations based on certain regression assumptions of adjacent order statistics. *Statist. Prob. Lett.* 80, 1700-1704.

[11] Huang, W.J., Su, N.C. and Su, J.C. (2007). Characterizations of the order statistics point process by the relations between its conditional moments. *J. Chinese Statist. Assoc.* 45, 38-54.

[12] Nevzorov, V.B., Balakrishnan, N., and Ahsanullah, M. (2003). Simple characterizations of Student's t_2 distribution. *Statistician* 52, 395-400.

[13] Nevzorova, L., Nevzorov, V.B. and Akhundov, I. (2007). A simple characterization of Student's t_2 distribution. *Metron-International J. Statist.* LXV (1), 53-57.

[14] Puri, P.S. (1982). On the characterization of point processes with the order statistic property without the moment condition. *J. Appl. Prob.* 19, 39-51.

[15] Wang, J. (2011). A simple characterization of Student's distributions and normal distributions. *Statist. Prob. Lett.* 81, 903-906.

[16] Yanev, G.P. and Ahsanullah, M. (2012). Characterizations of Student's t-distribution via regressions of order statistics. *Statistics* 46, 429-435.