

# Externally studentized normal midrange distribution

## Distribuição da midrange estudentizada externamente

Ben Dêvide de Oliveira Batista<sup>1\*</sup>, Daniel Furtado Ferreira<sup>2</sup>, Lucas Monteiro Chaves<sup>2</sup>

<sup>1</sup>Universidade Federal de São João del-Rei/UFESJ, Departamento de Matemática e Estatística/DEMAT, São João del-Rei, MG, Brasil

<sup>2</sup>Universidade Federal de Lavras/UFLA, Departamento de Estatística/DES, Lavras, MG, Brasil

\*Corresponding author: ben.devide@gmail.com

Received in December 15, 2016 and approved in February 15, 2017

---

### ABSTRACT

The distribution of externally studentized midrange was created based on the original studentization procedures of Student and was inspired in the distribution of the externally studentized range. The large use of the externally studentized range in multiple comparisons was also a motivation for developing this new distribution. This work aimed to derive analytic equations to distribution of the externally studentized midrange, obtaining the cumulative distribution, probability density and quantile functions and generating random values. This is a new distribution that the authors could not find any report in the literature. A second objective was to build an R package for obtaining numerically the probability density, cumulative distribution and quantile functions and make it available to the scientific community. The algorithms were proposed and implemented using Gauss-Legendre quadrature and the Newton-Raphson method in R software, resulting in the SMR package, available for download in the CRAN site. The implemented routines showed high accuracy proved by using Monte Carlo simulations and by comparing results with different number of quadrature points. Regarding to the precision to obtain the quantiles for cases where the degrees of freedom are close to 1 and the percentiles are close to 100%, it is recommended to use more than 64 quadrature points.

**Index terms:** Distribution function; density function; Gauss-Legendre quadrature; Newton-Raphson method; R.

---

### RESUMO

A distribuição da midrange estudentizada externamente foi criada com base nos procedimentos de estudentização de *Student* e foi inspirada na distribuição da amplitude estudentizada externamente. O amplo uso da amplitude estudentizada externamente em comparações múltiplas também foi uma das motivações para desenvolver esta nova distribuição. Neste trabalho objetivou-se derivar expressões analíticas da distribuição da midrange estudentizada externamente, obtendo a função de distribuição, função densidade de probabilidade, função quantil e geradores de números aleatórios. Essa é uma nova distribuição que os não há relatos na literatura especializada. Um segundo objetivo foi construir um pacote R para obter numericamente as funções mencionadas e torná-las disponíveis para a comunidade científica. Os algoritmos foram propostos e implementados usando os métodos de quadratura Gauss-Legendre e Newton-Raphson no software R, resultando no pacote *SMR*, disponível para baixar na página do CRAN. As rotinas implementadas apresentaram alta acurácia, sendo verificadas usando simulação Monte Carlo e pela comparação com diferentes pontos de quadratura. Quanto a precisão para se obter os quantis da distribuição da midrange estudentizada externamente para 1 grau de liberdade ou percentis próximo de 100%, é sugerido utilizar mais do que 64 pontos de quadratura.

**Termos para indexação:** Função de distribuição; função densidade; quadratura Gauss-Legendre; método Newton-Raphson; algoritmo; R.

---

### INTRODUCTION

Many problems on statistical investigations are based on studies of sample order statistics. Important cases of the order statistics are the minimum and maximum value of a sample. Among other functions of order statistics, the range and midrange are of special interest here, which correspond to the difference between the maximum and the minimum and to the average from these two extreme values, respectively. Several authors studied and applied this subject (Tippet, 1925; Pearson and Hartley, 1942;

Gumbel, 1944, 1946; Wilks, 1948; Yalcin et. al., 2014; Wan et. al., 2014; Barakat et. al., 2015; Bland, 2015; Mansouri, 2015; Li and Mansouri, 2016) and some of these studies are discussed in the sequence.

Tippet (1925) studied the first four moments of the range. Pearson and Hartley (1942) obtained tabulated values of the cumulative probabilities for several range values in small samples ( $n = 2$  to 20) drawn from normal populations. Gumbel (1944, 1946) established the independence of extreme values for large samples from several continuous distribution, as well as the distribution

of the range and midrange. Wilks (1948) reviewed several articles relating to the order statistics and suggested several examples of their applications to statistical inference.

Studies related with studentization were initially proposed by Student (1927) and the particular case of the studentized range distribution have been largely used in multiple comparison procedures in different areas of scientific research since the pioneering works in this area (Duncan, 1952, 1955; Tukey, 1949, 1953). The studentized range refers to the random variable defined simply as the range divided by the sample standard deviation, considering that both terms of this ratio are random variables independently distributed and computed in samples drawn from the normal distribution.

Let  $Y_1, Y_2, \dots, Y_n$  be the order statistics in a sample of size  $n$ , that are defined by sorting the original sample variables  $X_1, X_2, \dots, X_n$  in increasing order. The sample  $X_1, X_2, \dots, X_n$  are drawn from a population with distribution function  $F(x)$ . The range is defined by  $W = Y_n - Y_1$ . The cumulative distribution and the probability density functions (cdf and pdf) of the range are the Equations 1 and 2,

$$F_w(w) = n \int_{-\infty}^{\infty} f(z)[F(w+z) - F(z)]^{n-1} dz, \tag{1}$$

and

$$f_w(w) = \int_{-\infty}^{\infty} n(n-1)f(z)f(w+z) \times [F(w+z) - F(z)]^{n-2} dz, \tag{2}$$

respectively, as showed in David and Nagaraja (2003) and in Gumbel (1947).

Considering now samples of size  $n$  from the normal distribution with standard deviation  $\sigma$  and mean  $\mu \in \mathbb{R}$ , the externally studentized range is defined by the ratio  $Q = \frac{W'}{X} = \frac{W}{S}$  where  $W' = W/\sigma$  is the sample standard range and  $S^2$  is an independent and unbiased estimator of  $\sigma^2$ , associated with  $v$  degrees of freedom. The cumulative distribution and the probability density functions, according to David and Nagaraja (2003), are given by Equations 3 and 4,

$$F(q; n, v) = \int_0^{\infty} \int_{-\infty}^{\infty} n\phi(y)[\Phi(xq+y) - \Phi(y)]^{n-1} \times f(x; v) dy dx, \tag{3}$$

and

$$f(q; n, v) = \int_0^{\infty} \int_{-\infty}^{\infty} n(n-1)x\phi(y)\phi(xq+y) \times [\Phi(xq+y) - \Phi(y)]^{n-2} f(x; v) dy dx, \tag{4}$$

where  $\phi(y)$  and  $\Phi(y)$  are the probability density and cumulative distribution functions from a standard normal population evaluated at  $y$ , with  $y \in ]-\infty, \infty[$  and  $f(x; v)$  is the probability density function of  $X = S/\sigma$ . Considering that  $X$  is obtained in a sample of size  $v+1$  from the normal distribution, then it is well known the fact that  $\frac{vS^2}{\sigma^2} \sim \chi_v^2$ , i.e., it has a chi-square distribution with  $v$  degrees of freedom (Mood et al., 1974). Hence,  $\frac{vS^2}{v\sigma^2} = \frac{S^2}{\sigma^2} \sim \chi_v^2/v$ . Therefore, it can be concluded that  $X = \frac{S}{\sigma} \sim \sqrt{\chi_v^2/v}$ .

Theorem 1. If  $X=S/\sigma$  is computed in a sample of size  $v+1$  from a normal distribution with mean  $\mu$  and variance  $\sigma^2$ , then its probability density function is given by Equation 5,

$$f(x; v) = \frac{v^{v/2}}{\Gamma(v/2)2^{v/2-1}} x^{v-1} e^{-vx^2/2}, x \geq 0. \tag{5}$$

Proof. If  $U \sim \chi_v^2$ , as discussed above, then the distribution of  $X = S/\sigma$  can be obtained from the transformation  $U = v X^2$ . The Jacobian of this transformation is given by

$$J = \frac{du}{dx} = 2vx,$$

for  $x \geq 0$ .

The density function of  $X$ , from samples of a normal population, is obtained from the chi-square distribution by

$$\begin{aligned} f(x; v) &= f_U(u; v) |J| = \frac{1}{2^{v/2} \Gamma(v/2)} u^{v/2-1} e^{-u/2} |J| \\ &= \frac{1}{2^{v/2} \Gamma(v/2)} (vx^2)^{v/2-1} e^{-vx^2/2} 2vx \\ &= \frac{2v^{v/2-1}}{2^{v/2} \Gamma(v/2)} x(vx^2)^{v/2-1} e^{-vx^2/2} \end{aligned}$$

resulting in Equation 5.

Another very interesting order statistic is the midrange and introduced, among others, by Gumbel (1947) and Rider (1957).

Definition 1. The midrange is defined as the mean between the minimum and the maximum order statistics by Equation 6,

$$\bar{R} = (Y_1 + Y_n) / 2. \quad (6)$$

The externally studentized midrange is defined considering the original studentization procedures of Student (1927) and was inspired in the externally studentized range definition.

Definition 2. The externally studentized midrange are defined by

$$\bar{Q} = \bar{R} / S,$$

where  $\bar{R}$  is expressed in Equation 6 and  $S$  is an estimator of the population standard deviation  $\sigma$  with  $v$  degrees of freedom obtained independently from  $\bar{R}$ . It should be noticed that  $\bar{Q}$  is a random variable.

However, few studies address the midrange distribution and none was found on the externally studentized midrange, considering normal or non-normal populations. The importance of studies about the distribution of the externally studentized midrange could be enormous in the analysis of experiments. Rider (1957), among others, proved that the midrange estimator is more efficient than the sample mean in platykurtic distributions. Another important aspect that could be useful is the proposition of multiple comparison procedures based on externally studentized midrange, that could potentially show better results than the traditional tests based on the studentized range.

This work aimed to obtain the externally studentized midrange distribution. It intended to develop analytic equations to distribution of  $\bar{Q}$ , obtaining the cumulative distribution, probability density and quantile functions and generating random values. A second objective was to build an R package (R Development Core Team, 2017) for obtaining numerically the probability density, cumulative distribution and quantile functions using Gaussian quadratures and Newton-Raphson method and make the R package available to the scientific community.

## MATERIAL AND METHODS

### The externally studentized normal midrange distribution

Let  $Y_1, Y_2, \dots, Y_n$  be the order statistics in a sample of size  $n$ , that are defined by sorting the original sample variables  $X_1, X_2, \dots, X_n$ , in increasing order. The sample  $X_1,$

$X_2, \dots, X_n$  are drawn from a population with distribution function  $F(x)$ . Therefore, the distribution of the midrange is given by the following theorem.

Theorem 2. The probability density function and the cumulative distribution function of  $\bar{R}$ , Definition 1, from a random sample  $X_1, X_2, \dots, X_n$ , of size  $n$ , where  $X_j$  has distribution function  $F(x)$  and probability density function  $f(x)$ ,  $j = 1, 2, \dots, n$  given by Equations 7 and 8,

$$f_{\bar{R}}(\bar{r}) = \int_{-\infty}^{\bar{r}} 2n(n-1)f(z)f(2\bar{r}-z) \times [F(2\bar{r}-z) - F(z)]^{n-2} dz, \quad (7)$$

and

$$F_{\bar{R}}(\bar{r}) = n \int_{-\infty}^{\bar{r}} f(z)[F(2\bar{r}-z) - F(z)]^{n-1} dz, \quad (8)$$

respectively.

Proof. Let the joint distribution of  $Y_1$  and  $Y_n$  (David and Nagaraja, 2003) given by Equation 9,

$$f_{y_1, y_n}(u, x) = n(n-1)f(u)f(x) \times [F(x) - F(u)]^{n-2}, \quad (9)$$

for  $u < x$ , then to obtain the distribution of  $\bar{R}$ , the transformations  $Y_1 = g_1^{-1}(\bar{R}, Z) = Z$  and  $Y_n = g_2^{-1}(\bar{R}, Z) = 2\bar{R} - Z$  should be considered. The Jacobian of these transformations is

$$J = \begin{vmatrix} \frac{\partial y_1}{\partial z} & \frac{\partial y_1}{\partial \bar{r}} \\ \frac{\partial y_n}{\partial z} & \frac{\partial y_n}{\partial \bar{r}} \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix} = 2.$$

Therefore, the joint density of  $\bar{R}$  and  $Z$ , using 9, is given by

$$f_{\bar{R}, Z}(\bar{r}, z) = |J| f_{y_1, y_n}(g_1^{-1}(\bar{r}, z), g_2^{-1}(\bar{r}, z)) \\ = 2n(n-1)f(z)f(2\bar{r}-z) \times [F(2\bar{r}-z) - F(z)]^{n-2}.$$

The required density  $f_{\bar{R}}(\bar{r})$  can be obtained by integrating the above joint density in relation to  $z$ , resulting in

$$f_{\bar{R}}(\bar{r}) = \int_{-\infty}^{\bar{r}} 2n(n-1)f(z)f(2\bar{r}-z) \times [F(2\bar{r}-z) - F(z)]^{n-2} dz.$$

The distribution function  $\bar{R}$  can be obtained by Equation 10,

$$F_{\bar{R}}(\bar{r}) = \int_{-\infty}^{\bar{r}} \int_{-\infty}^t 2n(n-1)f(z)f(2t-z) \times [F(2t-z) - F(z)]^{n-2} dz dt. \tag{10}$$

The simplification of the equation 10 can be performed by inversion of the order of the integrals. There is a dependency between  $\bar{R}$  and  $Z$ , then by fixing  $z$ ,  $t$  will vary in the interval  $[z, \bar{r}]$ . Note that the upper limit of the interval,  $\bar{r}$ , refers to the definition of the cumulative distribution function. Therefore, the results of the change of integration order is

$$F_{\bar{R}}(\bar{r}) = \int_{-\infty}^{\bar{r}} \left[ \int_z^{\bar{r}} 2n(n-1)f(z)f(2t-z) \times [F(2t-z) - F(z)]^{n-2} dt \right] dz.$$

Note that

$$\frac{d}{dt} [F(2t-z) - F(z)]^{n-1} = 2(n-1)f(2t-z) \times [F(2t-z) - F(z)]^{n-2}.$$

So,

$$F_{\bar{R}}(\bar{r}) = \int_{-\infty}^{\bar{r}} n f(z) \left\{ [F(2t-z) - F(z)]^{n-1} \right\}_{t=z}^{t=\bar{r}} dz.$$

As,

$$\begin{aligned} \left\{ [F(2t-z) - F(z)]^{n-1} \right\}_{t=z}^{t=\bar{r}} &= [F(2\bar{r}-z) - F(z)]^{n-1} - \\ &\quad - [F(2z-z) - F(z)]^{n-1} \\ &= [F(2\bar{r}-z) - F(z)]^{n-1}, \end{aligned}$$

the result of the cumulative distribution function  $f_{\bar{R}}(\bar{r})$  is

$$F_{\bar{R}}(\bar{r}) = n \int_{-\infty}^{\bar{r}} f(z) [F(2\bar{r}-z) - F(z)]^{n-1} dz,$$

as showed in Gumbel (1958).

### RESULTS AND DISCUSSION

In the particular case of standard normal population, the Equations 7 and 8 can be rewritten by Equations 11 and 12,

$$f_{\bar{R}}(\bar{r}) = \int_{-\infty}^{\bar{r}} 2n(n-1)\phi(z)\phi(2\bar{r}-z) \times [\Phi(2\bar{r}-z) - \Phi(z)]^{n-2} dz, \tag{11}$$

and

$$F_{\bar{R}}(\bar{r}) = n \int_{-\infty}^{\bar{r}} \phi(z) [\Phi(2\bar{r}-z) - \Phi(z)]^{n-1} dz, \tag{12}$$

respectively.

If samples from a normal distribution with mean 0 and variance  $\sigma^2$  are considered, then the distribution of  $\bar{R}$  will depend on  $\sigma$ . The cumulative distribution function of  $\bar{R}$  in this case is obtaining directly from 12 by Equation 13,

$$F_{\bar{R}}(\bar{r}) = n \int_{-\infty}^{\bar{r}} \phi_{0,\sigma^2}(z) [\Phi_{0,\sigma^2}(2\bar{r}-z) - \Phi_{0,\sigma^2}(z)]^{n-1} dz, \tag{13}$$

where  $\phi_{0,\sigma^2}(z)$  and  $\Phi_{0,\sigma^2}(z)$  are the probability density and cumulative distribution functions, respectively, of the normal distribution with mean 0 and variance  $\sigma^2$ . The probability density function  $\Phi_{0,\sigma^2}(z)$  is related to the probability density function of the standard normal distribution by  $\phi_{0,\sigma^2}(z) = \phi(z/\sigma)/\sigma$ . In the same way, the relationship between the cumulative probability functions is  $\Phi_{0,\sigma^2}(z) = \Phi(z/\sigma)$ . Hence, if  $Z/\sigma$  is denote by  $Y$ , the cumulative distribution function, Equation 13, can be rewritten by Equation 14,

$$F_{\bar{R}}(\bar{r} / \sigma) = n \int_{-\infty}^{\bar{r}/\sigma} \frac{1}{\sigma} \phi(y) [\Phi(2\bar{r} / \sigma - y) - \Phi(y)]^{n-1} dy. \tag{14}$$

Similarly, the same procedure can be realized for Equation 11. Definition 3. The standardized midrange is defined by Equation 15,

$$\bar{W} = \frac{\bar{R}}{\sigma}, \tag{15}$$

where  $\bar{R}$  is the midrange from Definition 1 and  $\sigma$  is the population standard deviation.

Theorem 3. The probability density function and cumulative distribution function of  $\bar{W}$ , from Equation 15 are given by Equations 16 and 17,

$$f_{\bar{W}}(\bar{w}) = \int_{-\infty}^{\bar{w}} 2n(n-1)\phi(y)\phi(2\bar{w}-y) \times [\Phi(2\bar{w}-y) - \Phi(y)]^{n-2} dy, \tag{16}$$

and

$$F_{\bar{w}}(\bar{w}) = n \int_{-\infty}^{\bar{w}} \phi(y) [\Phi(2\bar{w} - y) - \Phi(y)]^{n-1} dy, \quad (17)$$

respectively.

Proof. The cumulative distribution function of  $\bar{W}$  is obtained from the cumulative distribution function given in 14. Hence, making the transformations of variables  $\bar{W} = \bar{R}/\sigma$  and  $Y$ , with the Jacobian  $J = \sigma$ , the cumulative distribution function of  $\bar{W}$  is given by

$$F_{\bar{w}}(\bar{w}) = n \int_{-\infty}^{\bar{w}} \phi(y) [\Phi(2\bar{w} - y) - \Phi(y)]^{n-1} dy$$

and the probability density function of  $\bar{W}$  is obtained by  $dF_{\bar{w}}(\bar{w}) / d\bar{w} = f_{\bar{w}}(\bar{w})$ , that is,

$$f_{\bar{w}}(\bar{w}) = n \int_{-\infty}^{\bar{w}} \phi(y) \frac{d}{d\bar{w}} [\Phi(2\bar{w} - y) - \Phi(y)]^{n-1} dy.$$

Solving

$$\frac{d}{d\bar{w}} [\Phi(2\bar{w} - y) - \Phi(y)]^{n-1} = 2(n-1)\phi(2\bar{w} - y) \times [\Phi(2\bar{w} - y) - \Phi(y)]^{n-2},$$

the probability density function of  $\bar{W}$  is

$$f_{\bar{w}}(\bar{w}) = \int_{-\infty}^{\bar{w}} 2n(n-1)\phi(y)\phi(2\bar{w} - y) \times [\Phi(2\bar{w} - y) - \Phi(y)]^{n-2} dy,$$

as expected.

It is worth noting that the probability density functions 11 and 16 and the cumulative probability functions 12 and 17 are equals. Therefore, the standardized midrange distribution obtained in normal populations with mean zero and variance  $\sigma^2$  is the same of the midrange from standard normal populations.

However, the objective is to find the externally studentized normal midrange ( $\bar{Q}$ , Definition 2) distribution, where  $\bar{Q}$  can also be defined by Equation 18,

$$\bar{Q} = \frac{\bar{R} / \sigma}{S / \sigma} = \frac{\bar{R}}{S} = \frac{\bar{W}}{X}, \quad (18)$$

where  $\bar{W}$  and  $S$ , with  $\nu$  degrees of freedom, are independently distributed. This occurs, for example, when  $\bar{W}$  is obtained from a random sample of size  $n$  from a normal population,

and  $S$ , the standard deviation, is from another random sample of size  $\nu + 1$ . The independence also occurs when  $\bar{W}$  is a function of factor means in an experimental design, with  $n$  levels and  $S = \sqrt{MSE / r}$ , where  $MSE$  is the mean square error with  $\nu$  degrees of freedom and  $r$  is the number of replications associated with each treatment mean, as can be found, e.g., in Searle (1987).

Theorem 4. The probability density function and cumulative distribution function of  $\bar{Q}$  from Definition 2 are given by Equations 19 and 20,

$$f(\bar{q}; n, \nu) = \int_0^{\infty} \int_{-\infty}^{x\bar{q}} 2n(n-1)x\phi(y)\phi(2x\bar{q} - y) \times [\Phi(2x\bar{q} - y) - \Phi(y)]^{n-2} f(x; \nu) dy dx, \quad (19)$$

and

$$F(\bar{q}; n, \nu) = \int_0^{\infty} \int_{-\infty}^{x\bar{q}} n\phi(y) [\Phi(2x\bar{q} - y) - \Phi(y)]^{n-1} \times f(x; \nu) dy dx, \quad (20)$$

respectively.

Proof. The distribution of  $\bar{Q}$  given in (18), is obtained from the joint distribution of  $\bar{W}$  given in 16 and  $X = S/\sigma$  given in 5. As the variables  $\bar{W}$  and  $X$  are independently distributed, the joint density is the product of their marginal densities, given by Equation 21,

$$f(\bar{w}, x; n, \nu) = \int_{-\infty}^{\bar{w}} 2n(n-1)\phi(y)\phi(2\bar{w} - y) \times [\Phi(2\bar{w} - y) - \Phi(y)]^{n-2} f(x; \nu) dy, \quad (21)$$

where  $f(x; \nu)$  is given in 5.

Considering the transformations  $\bar{Q} = \bar{W}/X$  and  $X$ , the Jacobian is

$$J = \begin{vmatrix} \frac{\partial \bar{w}}{\partial \bar{q}} & \frac{\partial \bar{w}}{\partial x} \\ \frac{\partial x}{\partial \bar{q}} & \frac{\partial x}{\partial x} \end{vmatrix} = \begin{vmatrix} x & \bar{q} \\ 0 & 1 \end{vmatrix} = x, \quad x > 0.$$

Therefore, from the above transformations  $\bar{W} = X\bar{Q}$  and using the joint distribution of  $\bar{W}$  and  $X$ , given in 21, the joint distribution of  $X$  and  $\bar{Q}$  is given by Equation 22,

$$f(\bar{q}, x; n, \nu) = f(\bar{w}, x; n, \nu) |J|, \\ = \int_{-\infty}^{x\bar{q}} 2n(n-1)\phi(y)\phi(2x\bar{q} - y) \times [\Phi(2x\bar{q} - y) - \Phi(y)]^{n-2} f(x; \nu) |x| dy. \quad (22)$$

Integrating the Equation 22 with respect to  $x$ , for  $x \in [0, \infty]$ , the probability density function of  $\bar{Q}$  is obtained. Thus,

$$f(\bar{q}; n, \nu) = \int_0^{\infty} \int_{-\infty}^{\bar{q}} 2n(n-1)x\phi(y)\phi(2x\bar{q} - y) \times [\Phi(2x\bar{q} - y) - \Phi(y)]^{n-2} f(x; \nu) dy dx.$$

The cumulative distribution function  $\bar{Q}$  can be obtained by performing the integration of 19 in  $\bar{q}$  over  $[-\infty, \bar{q}]$ . Changing the variable to  $z$ , the cumulative distribution function of  $\bar{Q}$  is given by

$$F(\bar{q}; n, \nu) = \int_{-\infty}^{\bar{q}} \int_0^{\infty} \int_{-\infty}^{xz} 2n(n-1)x\phi(y)\phi(2xz - y) \times [\Phi(2xz - y) - \Phi(y)]^{n-2} f(x; \nu) dy dx dz.$$

Changing the order of the integrals, and fixing the smallest studentized order statistic, the lower limit with respect to  $\bar{Q}$  is now  $y/x$ . So,

$$F(\bar{q}; n, \nu) = \int_0^{\infty} \int_{-\infty}^{x\bar{q}} n\phi(y) \left\{ \int_{y/x}^{\bar{q}} 2x(n-1)\phi(2xz - y) \times [\Phi(2xz - y) - \Phi(y)]^{n-2} dz \right\} f(x; \nu) dy dx.$$

Since

$$\frac{d}{dz} [\Phi(2xz - y) - \Phi(y)]^{n-1} = 2x(n-1)\phi(2xz - y) \times [\Phi(2xz - y) - \Phi(y)]^{n-2},$$

then,

$$F(\bar{q}; n, \nu) = \int_0^{\infty} \int_{-\infty}^{x\bar{q}} n\phi(y) \left[ \Phi(2xz - y) - \Phi(y) \right]^{n-1} \Big|_{z=y/x}^{z=\bar{q}} \times f(x; \nu) dy dx.$$

Solving,

$$\begin{aligned} \left[ \Phi(2xz - y) - \Phi(y) \right]^{n-1} \Big|_{z=y/x}^{z=\bar{q}} &= \{\Phi(2x\bar{q} - y) - \Phi(y)\}^{n-1} - \\ &\quad - \left\{ \Phi\left(2x \cdot \frac{y}{x} - y\right) - \Phi(y) \right\}^{n-1} \\ &= \{\Phi(2x\bar{q} - y) - \Phi(y)\}^{n-1}, \end{aligned}$$

the cumulative distribution function of  $\bar{Q}$  is given by

$$F(\bar{q}; n, \nu) = \int_0^{\infty} \int_{-\infty}^{x\bar{q}} n\phi(y) [\Phi(2x\bar{q} - y) - \Phi(y)]^{n-1} \times f(x; \nu) dy dx,$$

as expected.

The Equations 19 and 20 are the probability density and cumulative distribution functions, respectively, of  $\bar{Q}$ , which is the externally studentized normal midrange. This is a novel distribution that the authors could not find any report in the literature.

In the next section, numerical methods for obtaining the probability density and cumulative distribution functions are described. Also, the quantile functions are also taken into account. The same approaches will be considered for the distribution of  $\bar{W}$ , Equations 16 and 17, which is the standard normal midrange.

### Gauss-Legendre quadrature

The basic idea of Gauss-Legendre quadrature of a function  $f(x)$  is to write equal the Equation 23,

$$\int_{-1}^1 f(x) dx = \int_{-1}^1 w(x)g(x) dx \approx \sum_{k=1}^s w_k g(x_k), \tag{23}$$

where  $f(x) \equiv w(x)g(x)$  and  $w(x)$  is the weight function in the Gaussian quadrature,  $w_k$  and  $x_k$  are the nodes and weights, respectively, in an  $s$ -point Gaussian quadrature rule, for  $k = 1, 2, \dots, s$ . The weight function is  $w(x)=1$  in the Gauss-Legendre quadrature, thus  $f(x)=g(x)$ . The set  $\{x_k, w_k\}$  should be determined such that equation 23 yields an exact result for polynomials of degree  $2s-1$  or less. For non-polynomial function the Gauss-Legendre quadrature error is defined by Equation 24,

$$\mathcal{E}_s \approx \left| \sum_{k=1}^i w_k g(x_k) - \sum_{k=1}^s w_k g(x_k) \right|, \quad i > s. \tag{24}$$

The Gauss-Legendre quadrature was used to compute the functions 16, 17, 19 and 20. However, these functions depend on integrals over infinite intervals. The integral over an infinite range must be changed into an integral over  $[-1, 1]$  by using the Equations 25, 26 and 27,

$$\int_a^{\infty} f(x) dx = \int_{-1}^1 f\left(a + \frac{1+t}{1-t}\right) \frac{2}{(1-t)^2} dt, \tag{25}$$

$$\int_{-\infty}^b f(x) dx = \int_{-1}^1 f\left(b + \frac{1+t}{t-1}\right) \frac{2}{(t-1)^2} dt, \tag{26}$$

$$\int_a^b f(x) dx = \frac{b-a}{2} \int_{-1}^1 f\left(\frac{(b-a)}{2}t + \frac{(b+a)}{2}\right) dt. \tag{27}$$

Therefore, the integrals 25, 26 and 27 were computed by applying the Gauss-Legendre quadrature rule in these transformed variables by Equations 28, 29 and 30.

$$\int_a^{\infty} f(x) dx \approx \sum_{i=1}^s w_i \frac{2}{(1-x_i)^2} g\left(a + \frac{1+x_i}{1-x_i}\right), \quad (28)$$

$$\int_{-\infty}^b f(x) dx \approx \sum_{i=1}^s w_i \frac{2}{(x_i-1)^2} g\left(b + \frac{1+x_i}{x_i-1}\right), \quad (29)$$

$$\int_a^b f(x) dx \approx \frac{b-a}{2} \sum_{i=1}^s w_i g\left(\frac{b-a}{2} x_i + \frac{a+b}{2}\right). \quad (30)$$

Details of how to compute the nodes and weights to apply the Gauss-Legendre quadrature can be found in Hildebrand (1974).

An R package (R Development Core Team, 2017) denoted by SMR (Batista; Ferreira, 2012) was developed to apply the Gauss-Legendre quadratures to compute the cumulative distribution functions 17 and 20 and the probability density functions 16 and 19. The numerical transformations given by Equations 28, 29 and 30 and the methods of computation of the nodes and weights of Gauss-Legendre quadrature, as described in Hildebrand (1974) and Gil, Segura and Temme (2007) were used in the R codes of the implemented package. The quantiles were computed using the Newton-Raphson method solving equations formed by equating the functions 17 or 20 to  $p$ , where  $0 < p < 1$  is the cumulative probability (Gil, Segura, Temme, 2007), which is known. These computations make use of numerical quadratures of the respectively probability density functions 16 or 19, which are the first derivatives of 17 or 20.

Besides computing the cumulative distribution, probability density and quantile functions, the SMR package generates random samples of size  $N$  by Monte Carlo simulation. For this, random samples of size  $n$ ,  $X_1, X_2, \dots, X_n$ , are generated, where the  $X_i$ 's are independent and identically distributed standard normal variables,  $N(0, 1)$ , for  $i = 1, 2, \dots, n$ . A random variable  $U$ , distributed as a chi-square variable,  $\chi^2_\nu$ , is simulated, which is independently distributed of the  $X_i$ 's, where  $\nu > 0$  is the degrees of freedom. Finally, the following transformation is performed

$$\bar{Q} = \frac{[\max(X_i) + \min(X_i)] / 2}{\sqrt{\frac{U}{\nu}}}.$$

With infinity degrees of freedom  $\nu = \infty$ , it is suffice to compute

$$\bar{W} = \frac{\max(X_i) + \min(X_i)}{2}.$$

The process was repeated  $N$  times and the required sample values of  $\bar{Q}$  or of  $\bar{W}$  were obtained. The package SMR provides the following functions, where  $np$  is the number of nodes and weights of the Gauss-Legendre quadrature:

- dSMR( $x, n, nu, np=32$ ): computes values of the probability density function, given in (16) or (19);
- pSMR( $x, n, nu, np=32$ ): computes values of the cumulative distribution function, given in (17) or (20);
- qSMR( $p, n, nu, np=32$ ): computes quantiles of the externally studentized normal midrange;
- rSMR( $N, n, nu=Inf$ ): drawn random samples of the externally studentized normal midrange.

The user can choose the argument  $nu$  as finity or infinity value. If  $nu=Inf$ , values of the probability density, cumulative distribution and quantile functions of the normal midrange (standard normal midrange) are computed. If the argument  $nu$  is not specified in the rSMR function, the default value  $Inf$  is used and random samples from the normal midrange distribution are drawn.

## Performance

Evaluations of the accuracy of quadratures for computing the cumulative distribution functions 17 and 20 and for computing quantiles were performed. The SMR package was used for this purpose. There are no cumulative probabilities or quantiles of the externally studentized normal midrange to be compared, since reports of this distribution were not found in the literature. Two strategies were proposed to verify the accuracy. First, Monte Carlo simulations were used to obtain quantiles and cumulative probabilities and to compare them to those obtained by using the Gauss-Legendre quadratures from the SMR package. Second, two different number of quadrature points were used to compute these quantities. Thus, the quadrature errors were computed by comparing these two values, as showed in 24.

Validations of the algorithm by Monte Carlo simulation were done using the R software with the SMR package (Batista; Ferreira, 2012). A random sample of externally studentized normal midrange,  $\bar{Q}$ , of size  $N = 1,000,001$  was simulated using the rSMR function of the SMR package, following the procedure described above. If the degrees of freedom were  $\infty$ , then this function generates random sample from the standard normal midrange,  $\bar{W}$ .

Therefore, given a quantile  $\bar{q}$  or  $\bar{w}$ , the cumulative probability is computed, respectively, by

$$P(\bar{Q} \leq \bar{q}) = \frac{\sum_{i=1}^N I(\bar{Q}_i \leq \bar{q})}{N}, \nu < \infty \text{ or}$$

$$P(\bar{W} \leq \bar{w}) = \frac{\sum_{i=1}^N I(\bar{W}_i \leq \bar{w})}{N}, \nu = \infty.$$

Table 1 shows the values of the distribution function computed by 64 Gauss-Legendre quadrature points and by Monte Carlo simulations (with  $N=1,000,001$  observations in the Monte Carlo sample). Several combinations of  $n$  and  $\nu$  were used for some particular choices of the quantiles  $\bar{q}$  or  $\bar{w}$ . The cumulative probabilities computed by Gauss-Legendre quadrature were close to those obtained by Monte Carlo simulations (MC). The MC error is proportional to  $1/\sqrt{N}$  (Ciftja and Wexler, 2003), which in this case is 0.001. The two methods showed differences in the third and fourth decimal places (Table 1), as expected. These results show, in principle, that the Gauss-Legendre quadrature has at least three significant digits of accuracy.

The quadrature error  $\varepsilon$  can also be computed increasing the number of points and calculating the difference between the two results. Cumulative probabilities computed by using  $s=64$  and  $i=250$  quadrature points

and the errors obtained by calculating the differences between these values are shown in Table 2. The maximum observed error is of the order of  $10^{-10}$ , showing that with 64 quadrature points a high precision was achieved, for computing 17 or 20. Several other combinations of  $n$ ,  $\nu$  and  $\bar{q}$  or  $\bar{w}$  were used to compute the Monte Carlo errors and their maximum value is still the same.

Table 3 shows quantiles for a settled value of the cumulative probabilities of 0.95, using the qSMR function of SMR package (Batista; Ferreira, 2012, 2014). In this case, almost all the quantiles were computed using 64 quadrature points, except, for the case of  $\nu = 1$ , where this number of points was insufficient. In such circumstances, there are two alternatives: a) refine the quadrature, dividing the integral interval into small subintervals and approximate the integral by a sum of computation on each subinterval; or b) increasing the number of quadrature points. In this case, we opted for the latter, using 250 quadrature points. The results were shown using 3 decimal places, for  $\nu$  degrees of freedom, of 1(1)20(5)30, 50(50)200, 300, 1,000 and  $\infty$ , and for sample sizes  $n$  of 2(1)10(5)50(25)100. When  $\nu \rightarrow \infty$ , thus  $s^2 \rightarrow \sigma^2$  and the studentized midrange cumulative distribution function from standard normal populations  $F(\bar{q}; n, \nu)$ , given in 20, tends to  $F_{\bar{w}}(\bar{w}; n)$ , the standardized midrange cumulative distribution function, given in 17. If the desired value of  $\bar{q}$  can not be found in Table 3, the SMR package can be used or linear interpolations can be applied to compute this quantile.

**Table 1:** Computed values of  $P(\bar{Q} \leq \bar{q})$ , using the Gauss-Legendre Quadrature and by simulation using Monte Carlo method, with sample size  $n$ , and  $\nu$  degrees of freedom, considering  $N=1,000,001$  simulations.

Values			Method	
$n$	$\nu$	$\bar{q}$ or $\bar{w}$	Quadrature	MC
15	4	4.000	0.999712051088144	0.999719000280999
30	7	2.000	0.999521701960583	0.999520000479999
45	10	4.000	0.999999689708233	0.999999999999999
60	25	1.000	0.996635357142795	0.996665003334996
20	2	1.000	0.941476242577670	0.941947058052942
30	5	0.300	0.786876942543113	0.786846213153786
90	40	0.200	0.748082418017427	0.748314251685748
30	10	0.000	0.500000000000000	0.500382499617500
40	5	-1.000	0.016709246604515	0.016733983266017
20	20	-0.400	0.147628604257637	0.147586852413147
15	$\infty$	0.300	0.778662390742543	0.778745221254779
60	$\infty$	0.800	0.991576069342297	0.991477008522991



**Table 2:** Computed values of  $P(\bar{Q} \leq \bar{q})$ , using the Gauss-Legendre Quadrature for  $s=64$  and  $i=250$  points of quadrature, with sample size  $n$ , and  $\nu$  degrees of freedom.

Values			Quadrature		
$n$	$\nu$	$\bar{q}$ or $\bar{w}$	$s = 64$ points	$i = 250$ points	$\epsilon$
15	4	4.000	0.999712051088145	0.99971205062441360	$4.637313 \times 10^{-10}$
30	7	2.000	0.999521701960583	0.9995217019605828	$1.110223 \times 10^{-10}$
45	10	4.000	0.999999689708230	0.99999968970816311	$6.694644 \times 10^{-10}$
60	25	1.000	0.996635357142795	0.99663535714279172	$3.330669 \times 10^{-10}$
20	2	1.000	0.941476242577670	0.94147624257766860	$1.443290 \times 10^{-10}$
30	5	0.300	0.786876942543113	0.78687694254311258	$4.440892 \times 10^{-10}$
90	40	0.200	0.748082418017427	0.74808241801730846	$1.185718 \times 10^{-10}$
30	10	0.000	0.5000000000000002	0.500000000000000200	$1.000000 \times 10^{-10}$
40	5	-1.000	0.016709246604515	0.01670924660451461	$3.920475 \times 10^{-10}$
20	20	-0.400	0.147628604257637	0.14762860425763724	$2.498002 \times 10^{-10}$
15	$\infty$	0.300	0.778662390742543	0.77866239074254306	$3.330669 \times 10^{-10}$
60	$\infty$	0.800	0.991576069342296	0.99157606934229103	$5.662137 \times 10^{-10}$

**Table 3:** Upper quantile of the distribution of externally studentized normal midrange ( $\bar{q}$  or  $\bar{w}$ ) for different degrees of freedom  $\nu$  and sample sizes ( $n$ ), according the following probability event:  $P(\bar{Q} \leq \bar{q}) = 0.95$  or  $P(\bar{W} \leq \bar{w}) = 0.95$ .

$\nu$	Sample size ( $n$ )								
	2	3	4	5	6	7	8	9	10
1	4.464	3.799	3.445	3.219	3.060	2.941	2.846	2.770	2.705
2	2.065	1.757	1.594	1.490	1.417	1.362	1.319	1.284	1.254
3	1.664	1.416	1.285	1.201	1.143	1.098	1.064	1.035	1.012
4	1.507	1.283	1.164	1.088	1.035	0.995	0.964	0.938	0.917
5	1.425	1.213	1.100	1.029	0.978	0.941	0.911	0.887	0.867
6	1.374	1.169	1.061	0.992	0.944	0.907	0.879	0.855	0.836
7	1.340	1.140	1.034	0.967	0.920	0.885	0.857	0.934	0.815
8	1.315	1.119	1.015	0.949	0.903	0.868	0.841	0.819	0.800
9	1.296	1.103	1.001	0.936	0.890	0.856	0.829	0.807	0.789
10	1.282	1.091	0.990	0.925	0.880	0.846	0.820	0.798	0.780
15	1.240	1.055	0.957	0.895	0.851	0.819	0.793	0.772	0.754
16	1.235	1.051	0.953	0.891	0.848	0.815	0.789	0.769	0.751
17	1.230	1.047	0.950	0.888	0.845	0.812	0.787	0.766	0.748
18	1.226	1.044	0.947	0.885	0.842	0.810	0.784	0.763	0.746
19	1.223	1.041	0.944	0.883	0.840	0.807	0.782	0.761	0.744
20	1.220	1.038	0.942	0.881	0.838	0.805	0.780	0.759	0.742
25	1.208	1.028	0.933	0.872	0.830	0.798	0.772	0.752	0.735
30	1.200	1.021	0.927	0.867	0.824	0.793	0.768	0.747	0.730
50	1.185	1.009	0.915	0.856	0.814	0.783	0.758	0.738	0.721

Continue...

**Table 3:** Continuation...

$\nu$	Sample size ( $n$ )								
	2	3	4	5	6	7	8	9	10
100	1.174	0.999	0.906	0.848	0.806	0.775	0.751	0.731	0.714
150	1.170	0.996	0.904	0.845	0.804	0.773	0.748	0.729	0.712
200	1.168	0.994	0.902	0.844	0.803	0.772	0.747	0.727	0.711
300	1.167	0.993	0.901	0.842	0.801	0.770	0.746	0.726	0.710
1000	1.164	0.991	0.899	0.841	0.800	0.769	0.744	0.725	0.708
$\infty$	1.163	0.990	0.898	0.840	0.799	0.768	0.744	0.724	0.708

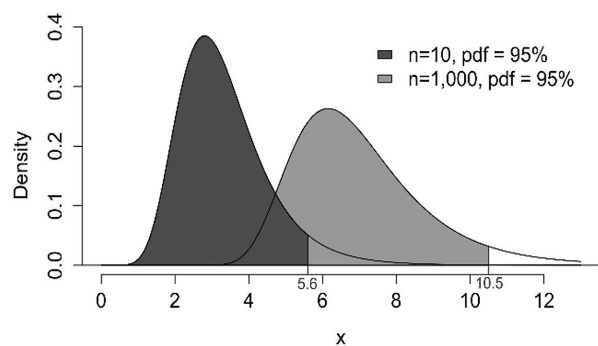
$\nu$	Sample size ( $n$ )									
	15	20	25	30	35	40	45	50	75	100
1	2.492	2.366	2.280	2.215	2.165	2.124	2.090	2.060	1.957	1.893
2	1.156	1.099	1.059	1.030	1.007	0.988	0.972	0.959	0.912	0.882
3	0.933	0.887	0.855	0.832	0.813	0.798	0.785	0.774	0.736	0.713
4	0.846	0.804	0.775	0.754	0.737	0.723	0.712	0.702	0.668	0.646
5	0.800	0.760	0.733	0.713	0.697	0.684	0.673	0.664	0.632	0.611
6	0.771	0.733	0.707	0.687	0.672	0.660	0.649	0.640	0.609	0.590
7	0.752	0.715	0.689	0.670	0.656	0.643	0.633	0.625	0.594	0.575
8	0.738	0.702	0.677	0.658	0.643	0.632	0.622	0.613	0.583	0.564
9	0.728	0.692	0.667	0.649	0.634	0.623	0.613	0.604	0.575	0.556
10	0.719	0.684	0.660	0.641	0.627	0.616	0.606	0.598	0.568	0.550
15	0.696	0.661	0.638	0.620	0.607	0.596	0.586	0.578	0.550	0.532
16	0.693	0.659	0.635	0.618	0.604	0.593	0.584	0.576	0.548	0.530
17	0.691	0.656	0.633	0.616	0.602	0.591	0.582	0.574	0.546	0.528
18	0.688	0.645	0.631	0.614	0.600	0.589	0.580	0.572	0.544	0.527
19	0.686	0.652	0.629	0.612	0.598	0.587	0.578	0.570	0.542	0.525
20	0.685	0.651	0.628	0.610	0.597	0.586	0.577	0.569	0.541	0.524
25	0.678	0.645	0.622	0.605	0.591	0.580	0.571	0.563	0.536	0.519
30	0.674	0.640	0.618	0.601	0.587	0.577	0.568	0.560	0.532	0.515
50	0.665	0.632	0.610	0.593	0.580	0.569	0.560	0.553	0.526	0.509
100	0.659	0.626	0.604	0.588	0.575	0.564	0.555	0.548	0.521	0.504
150	0.657	0.625	0.602	0.586	0.573	0.562	0.553	0.546	0.519	0.503
200	0.656	0.624	0.601	0.585	0.572	0.561	0.553	0.545	0.518	0.502
300	0.655	0.623	0.601	0.584	0.571	0.561	0.552	0.544	0.518	0.501
1000	0.654	0.621	0.599	0.583	0.570	0.559	0.551	0.543	0.517	0.500
$\infty$	0.653	0.621	0.599	0.582	0.569	0.559	0.550	0.543	0.516	0.500

In Table 3, if the number of degrees of freedom and the cumulative probability is settled, the quantiles decrease as the sample size increases. Comparing these results of the studentized normal midrange with quantiles of the studentized normal range in the same

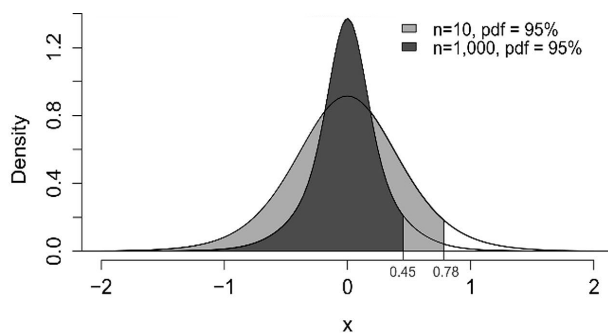
circumstances (Newman, 1939), this is not observed. Fixing the number of degrees of freedom and the probability, the studentized normal range quantiles increase as the sample size increases. This difference can be observed in Figure 1, where the studentized

range and midrange probability density functions were plotted, for  $\nu = 10$  with  $n = 10$  and  $n = 1,000$ . Another interesting observation is that fixing the sample size and the cumulative probability, the studentized normal midrange quantiles decrease as the number of degrees of freedom increases (Table 3). The same can be observed for the studentized normal range quantiles. In Figure 1(a), for  $n=10$  and  $n=1,000$ , with  $\nu = 10$ , the quantiles for  $P(Q \leq q) = F(q; n, \nu) = 0.95$  are  $q_{n=10} = 5.6$  and  $q_{n=1,000} = 10.5$ , respectively. Increasing the sample size there is a translation of the range probability density function to the right. This occurs because the range increases as sample size increases.

In Figure 1(b), for  $n=10$  and  $n=1,000$ , with  $\nu = 10$ , the quantiles for  $P(\bar{Q} \leq \bar{q}) = F(\bar{q}; n, \nu) = 0.95$  are  $q_{n=10} = 0.78$  and  $q_{n=1,000} = 0.45$ , respectively. When the sample size increases there is an increased concentration of data around its center without change its central position. This concentration of data decreases the value of the quantile in larger sample, considering the same cumulative probability.



(a) studentized range density function.



(b) Studentized midrange density function.

**Figure 1:** Probability density function (pdf) of the externally studentized normal range (a) and midrange (b), considering  $\nu = 10$  with  $n = 10$  and  $n = 1,000$ .

## Applications

The development of externally studentized normal midrange distribution allows that several applications can be performed. One of the already results of this distribution is the midrangeMCP package that performs four multiple comparison tests based on this distribution (Batista; Ferreira, 2014). The tests developed and implemented in this package are versions similar to the Tukey, SNK and Scott-Knott tests. The first two, in their original version, are based on the distribution of externally studentized normal range. The Scott-Knott test is based on the likelihood ratio as a criterion for separating groups of means. However, in the midrangeMCP package, these tests were based on the distribution of externally studentized normal midrange, which present high performance when evaluated for the type I error rates and power, although results are in the *publishing* process. Another interesting application of this distribution is in the building of control charts in the statistical area of quality control, replacing the normal standardized range or the studentized range.

## CONCLUSIONS

The analytical equations of the distribution of the externally studentized normal midrange ( $\bar{Q}$ ), as well as of the distribution of normal midrange ( $\bar{W}$ ), were achieved. Probability density, cumulative distribution and quantiles functions were obtained computationally for  $\bar{Q}$  and  $\bar{W}$ . The algorithms were proposed and implemented using Gauss-Legendre quadrature and the Newton-Raphson method in R software, resulting in the SMR package, available for download in the CRAN site. The implemented routines showed high accuracy proved by using Monte Carlo simulations and by comparing results with different number of quadrature points. Regarding to the precision to obtain the quantiles for cases where the degrees of freedom are close to 1 and the percentiles are close to 100%, it is recommended to use more than 64 quadrature points.

## ACKNOWLEDGMENTS

We would like to thank CNPq and CAPES for their financial support.

## REFERENCES

BATISTA, B. D. O.; FERREIRA, D. F. SMR: An R package for computing the externally studentized normal midrange distribution. *The R Journal*, 6(2):123-136, 2014.

- BATISTA, B. D. O.; FERREIRA, D. F. **SMR: Externally Studentized Midrange Distribution**, 2012. R package version 1.0.0.
- BARAKAT, H. M.; NIGM, E. M.; ELSAWAH, A. M. Asymptotic distributions of the generalized range, midrange, extremal quotient, and extremal product, with a comparison study. **Communications in Statistics - Theory and Methods**, 44(5):900-913, 2015.
- BLAND, M. Estimating mean and standard deviation from the sample size, three quartiles, minimum, and maximum. **International Journal of Statistics in Medical Research**, 4:57-64, 2015.
- CIFTJA, O.; WEXLER, C. Monte Carlo simulation method for laughlin-like states in a diskgeometry. **Physical Review B**, 67(7):075304, 2003.
- DAVID, H. A.; NAGARAJA, H. N. **Order Statistics**. John Wiley & Sons, Canada, 2003. p.458.
- DUNCAN, D. B. Multiple range and multiple F tests. **Biometrics**, 11:1-42, 1955.
- DUNCAN, D. B. On the properties of the multiple comparisons test. **Virginia Journal of Science**, 3:50-67, 1952.
- GIL, A.; SEGURA, J.; TEMME, N. M. **Numerical methods for special functions**. SIAM, 2007. 417p.
- GUMBEL, E. J. On the independence of the extremes in a sample. **The Annals of Mathematical Statistics**, 17(1):78-80, 1946.
- GUMBEL, E. J. **Statistics of Extremes**. Columbia University Press, New York, 1958. p.375
- GUMBEL, E. J. Range and midranges. **The Annals of Mathematical Statistics**, 15(4):414-422, 1944.
- GUMBEL, E. J. The distribution of the range. **The Annals of Mathematical Statistics**, 18(3):384-412, 1947.
- HILDEBRAND, F. B. **Introduction to Numerical Analysis**. McGraw-Hill, New York, 2 edition, 1974. p.669.
- YALCIN, G. C.; ROBLEDO, A.; GELL-MANN, M. Incidence of q statistics in rank distributions. **Proceedings of the National Academy of Sciences**, 111(39):14082-14087, 2014.
- LI, B.; MANSOURI, H. Simultaneous rank tests for detecting differentially expressed genes. **Journal of Statistical Computation and Simulation**, 86(5):959-972, 2016.
- MANSOURI, H. Simultaneous inference based on rank statistics in linear models. **Journal of Statistical Computation and Simulation**, 85(4):660-674, 2015.
- MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the Theory of Statistics**. McGraw-Hill, New York, 1974. p.564.
- NEWMAN, D. The distribution of range in samples from a normal population, expressed in terms of an independent estimate of standard deviation. **Biometrika**, 31(1-2):20-30, 1939.
- PEARSON, E. S.; HARTLEY, H. O. The probability integral of the range in samples of n observations from a normal population. **Biometrika**, 32:301-310, 1942.
- RDEVELOPMENT CORE TEAM. **R: A Language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2017.
- RIDER, P. R. The midrange of a sample as an estimator of the population midrange. **Journal of the American Statistical Association**, 52(280):537-542, Dec. 1957.
- SEARLE, S. R. **Linear models for unbalanced data**. J. New York: Wiley, 1987. 560p.
- STUDENT. Errors in routine analysis. **Biometrika**, 19:151-164, 1927.
- TIPPET, L. H. C. On the extreme individuals and the range of samples taken from a normal population. **Biometrika**, 17:365-387, 1925.
- TUKEY, J. W. Comparing individual means in the analysis of variance. **Biometrics**, 5:99-114, 1949.
- TUKEY, J. W. **The problem of multiple comparisons**. Unpublished Dittoed Notes, Princeton University, 1-300, 1953.
- WAN, X. et al. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. **BMC Medical Research Methodology**, 14:135, 2014.
- WILKS, S. S. Order statistics. Bulletin of the **American Mathematical Society**, 54:6-50, 1948.