

STICHTING  
MATHEMATISCH CENTRUM  
2e BOERHAAVESTRAAT 49  
AMSTERDAM

S 405

SP 110

Some remarks concerning the quotient of sample  
median and sample range for a sample of size  $2n+1$   
from a normal distribution

L. de Haan and J.Th. Runnenburg

Reprinted from  
Statistica Neerlandica, 23(1969), p 227-234



1969

# Some remarks concerning the quotient of sample median and sample range for a sample of size $2n+1$ from a normal distribution

by L. DE HAAN\* and J. TH. RUNNENBURG\*\*

## 0. Introduction and summary

Consider an ordered sample  $x_{(1)}, x_{(2)}, \dots, x_{(2n+1)}$  of size  $2n+1$  from the normal distribution with parameters  $\mu$  and  $\sigma$ . We then have with probability one

$$x_{(1)} < x_{(2)} < \dots < x_{(2n+1)}.$$

The random variable

$$h_n = \frac{x_{(n+1)}}{x_{(2n+1)} - x_{(1)}}, \quad (1)$$

that can be described as the quotient of the sample median and the sample range, provides us with an estimate for  $\mu/\sigma$ , that is easy to calculate. To calculate the distribution of  $h_n$  is quite a different matter\*\*\*. The distribution function of  $h_1$  and the density of  $h_2$  are given in section 1. Our results seem hardly promising for general  $h_n$ . In section 2 it is shown that  $h_n$  is asymptotically normal.

In the sequel we suppose  $\mu = 0$  and  $\sigma = 1$ , i.e. we consider only the „central” distribution. Note that  $h_n$  can be used as a test statistic replacing Student's  $t$ . In that case the central  $h_n$  is all that is needed.

This research was suggested by Prof. Dr. J. HEMELRIJK.

## 1. Exact distribution

### 1.1 Existence of moments

We define

$$t_n = \frac{\bar{x}}{\sqrt{\frac{1}{2n+1} \sum_{i=1}^{2n+1} (x_i - \bar{x})^2}}$$

with

$$\bar{x} = \frac{1}{2n+1} \sum_{i=1}^{2n+1} x_i.$$

---

Report S 405 (SP 110), Statistical Department, Mathematical Centre, Amsterdam.

\* Mathematical Centre, Amsterdam.

\*\* University of Amsterdam.

\*\*\* A different attack from ours is given in [1].

We first show that for each  $\alpha > 0$

$$E|t_n|^\alpha < \infty \iff E|h_n|^\alpha < \infty. \quad (2)$$

This is an immediate consequence of the inequalities

$$|h_n| - 1 < |t_n| < 2\sqrt{2n+1}(|h_n| + 1), \quad (3)$$

which we shall now prove: From

$$|x_k - \bar{x}| < x_{(2n+1)} - x_{(1)} \quad \text{for } k = 1, 2, \dots, 2n+1 \quad (4)$$

it follows that

$$|\bar{x}| - |x_{(n+1)}| < x_{(2n+1)} - x_{(1)}$$

and

$$|x_{(n+1)}| - |\bar{x}| < x_{(2n+1)} - x_{(1)},$$

hence

$$|x_{(n+1)}| \left\{ 1 - \frac{1}{h_n} \right\} < |\bar{x}| < |x_{(n+1)}| \left\{ 1 + \frac{1}{h_n} \right\}. \quad (5)$$

Obviously

$$\frac{1}{2n+1} \sum_{i=1}^{2n+1} (x_i - \bar{x})^2 < x_{(2n+1)} - x_{(1)}. \quad (6)$$

For  $k = 1, 2, \dots, 2n+1$  we have

$$|x_k - \bar{x}| < \sqrt{\sum_{i=1}^{2n+1} (x_i - \bar{x})^2},$$

hence

$$x_{(2n+1)} - x_{(1)} < |x_{(2n+1)} - \bar{x}| + |x_{(1)} - \bar{x}| < 2\sqrt{\sum_{i=1}^{2n+1} (x_i - \bar{x})^2}. \quad (7)$$

Combining (6) and (7) we obtain

$$\frac{1}{x_{(2n+1)} - x_{(1)}} < \frac{1}{\sqrt{\frac{1}{2n+1} \sum_{i=1}^{2n+1} (x_i - \bar{x})^2}} < \frac{2\sqrt{2n+1}}{x_{(2n+1)} - x_{(1)}}. \quad (8)$$

Now (3) is an easy consequence of (5) and (8).

As  $\sqrt{2n} \cdot t_n$  has Student's distribution with  $2n$  degrees of freedom, we obtain from (2)

$$\left. \begin{aligned} E|h_n|^{2n} &= \infty \\ E|h_n|^\alpha &< \infty \quad \text{for } 0 < \alpha < 2n. \end{aligned} \right\} \quad (9)$$

## 1.2 Distribution of $\underline{h}_1$ and $\underline{h}_2$

We shall compute the distribution function of  $\underline{h}_1$  and the density of  $\underline{h}_2$ . As a first step we write the integral representing the distribution function  $F_n(t)$  of  $\underline{h}_n$  in a convenient form. The details are given for  $n = 1$  only, for other values of  $n$  an analogous result applies.

$$\begin{aligned} F_1(t) &= \frac{\int \int \int_{\substack{x < y < z \\ y < t(z-x)}} \exp\{-\tfrac{1}{2}(x^2 + y^2 + z^2)\} dx dy dz}{\int \int \int_{\substack{x < y < z \\ 0 < r < \infty}} \exp\{-\tfrac{1}{2}(x^2 + y^2 + z^2)\} dx dy dz} = \\ &= \frac{\int \int \int_{\substack{R_1 \cap R_2 \\ 0 < r < \infty}} \exp(-\tfrac{1}{2}r^2) r^2 \sin \theta dr d\phi d\theta}{\int \int \int_{\substack{R_2 \\ 0 < r < \infty}} \exp(-\tfrac{1}{2}r^2) r^2 \sin \theta dr d\phi d\theta} = \frac{\int \int_{R_1 \cap R_2} \sin \theta d\phi d\theta}{\int \int_{R_2} \sin \theta d\phi d\theta}, \end{aligned}$$

where we have applied the transformation

$$\begin{cases} x = r \cos \phi \sin \theta \\ y = r \sin \phi \sin \theta \\ z = r \cos \theta \end{cases}$$

and then integrated over  $r$ . The region  $R_1(\phi, \theta)$  is determined by

$$\sin \phi \sin \theta < t(\cos \theta - \cos \phi \sin \theta)$$

and the region  $R_2(\phi, \theta)$  by

$$\begin{cases} \cos \phi < \sin \phi \\ \sin \phi \sin \theta < \cos \theta \\ 0 < \phi \leq 2\pi \\ 0 < \theta \leq \pi \end{cases}$$

This last quotient for  $F_1(t)$  equals

$$\frac{\int \int \int_{\substack{R_1 \cap R_2 \\ 0 < r \leq 1}} r^2 \sin \theta d\phi d\theta dr}{\int \int \int_{\substack{R_2 \\ 0 < r \leq 1}} r^2 \sin \theta d\phi d\theta dr} = \frac{\int \int \int_{\substack{x < y < z \\ y < t(z-x) \\ x^2 + y^2 + z^2 \leq 1}} dx dy dz}{I_3}, \quad (10)$$

where ([2] part II p. 304)

$$I_k = \int \int \int \dots \int_{\substack{x_1 < x_2 < \dots < x_k \\ x_1^2 + x_2^2 + \dots + x_k^2 \leq 1}} dx_1 dx_2 \dots dx_k = \frac{\pi^{k/2}}{k! \Gamma\left(\frac{k+2}{2}\right)}; \quad \left(I_3 = \frac{4\pi}{3 \cdot 3!}, I_5 = \frac{8\pi^2}{15 \cdot 5!}\right).$$

To determine  $F_n(t)$  for arbitrary  $n$  we have to compute the content of a part of the unit hypersphere in  $R^{2n+1}$  limited by  $2n+1$  hyperplanes. This can also be seen without any calculation. Equivalently we can compute the area of a part of the surface of the unit hypersphere in  $R^{2n+1}$  limited by  $2n+1$  hyperplanes. We choose the latter formulation for  $n=1$ : we need only compute the surface of a spherical triangle. The result is wellknown (see [7] p. 216). In this particular case we have

$$F_1(t) = \frac{\frac{2\pi}{3} - \frac{1}{3} \left\{ \arccos(-\frac{1}{2}) + \arccos\left(\frac{t-1}{\sqrt{4t^2+2}}\right) + \arccos\left(\frac{t+1}{\sqrt{4t^2+2}}\right) \right\}}{\frac{4\pi}{3 \cdot 3!}},$$

i.e.

$$F_1(t) = 2 - \frac{3}{2\pi} \left\{ \arccos \frac{t-1}{\sqrt{3t^2+2}} + \arccos \frac{t+1}{\sqrt{4t^2+2}} \right\}. \quad (11)$$

For  $n=2$  we compute the content of the part of the unit hypersphere in  $R^5$  determined by the linear inequalities

$$\begin{cases} x_i < x_{i+1} & \text{for } i = 1, 2, 3, 4, \\ x_3 < tx_5 - tx_1. \end{cases}$$

To this end we use some results from VAN DER VAART [7]. We only quote those parts of the definitions and two theorems in his paper that we need.

Consider a given set of  $r$  independent vectors  $b_1, b_2, \dots, b_r \in R^r$  with  $b_i' b_i = 1$  for  $i = 1, 2, \dots, r$  (a prime is used to denote the transpose of a vector or a matrix). We consider the hyperspherical simplex

$$\{x | x \in R^r, b_1' x \geq 0, b_2' x \geq 0, \dots, b_r' x \geq 0, x' x \leq 1\}. \quad (12)$$

The content of this simplex depends only on the matrix

$$C = BB',$$

where  $B$  is the matrix with  $b_i$  as  $i$ -th row ( $i = 1, 2, \dots, r$ ). The content of the simplex (12) is denoted by

$$V_r(C; 1, 2, \dots, r),$$

whereas for  $1 \leq k < l \leq r$

$$\begin{aligned} V_{r-2}(C; \overline{k, l}; 1, 2, \dots, k-1, k+1, \dots, l-1, l+1, \dots, r) = \\ = \text{content} \{x | x \in R^r, b'_1 x \geq 0, \dots, b'_{k-1} x \geq 0, b'_k x = 0, b'_{k+1} x \geq 0, \dots, \\ b'_{l-1} x \geq 0, b'_l x = 0, b'_{l+1} x \geq 0, \dots, b'_r x \geq 0, x'x \leq 1\}. \end{aligned} \quad (13)$$

The index  $r-2$  is used to indicate that the simplex can be imbedded in an  $(r-2)$ -dimensional space.

**Theorem 1** (SCHLÄFLI)

$$\frac{\partial}{\partial c_{kl}} V_r(C; 1, 2, \dots, r) = \frac{1}{r} \cdot \frac{1}{\sqrt{1-c_{kl}^2}} V_{r-2}(C; \overline{k, l}; 1, 2, \dots, k-1, k+1, \dots, l-1, l+1, \dots, r),$$

where  $c_{kl}$  is the  $(k, l)$ -element of  $C$  (with  $1 \leq k < l \leq r$ ).

The set (13) can be imbedded in  $R^{r-2}$ ; a formula giving the matrix  $'C$  of an  $(r-2)$ -dimensional hyperspherical simplex with the same content as (13) is given in the next theorem.

**Theorem 2** (VAN DER VAART)

$$V_{r-2}(C; \overline{k, l}; 1, 2, \dots, k-1, k+1, \dots, l-1, l+1, \dots, r) = V_{r-2}('C; 1, 2, \dots, r-2)$$

with for  $p, q = 1, 2, \dots, r-2$

$$'c_{pq} = \frac{\det(c_{k,l,u}^{k,l,u})}{\sqrt{\det(c_{k,l,u}^{k,l,u}) \det(c_{k,l,v}^{k,l,v})}},$$

where the matrix  $c_{k,l,v}^{k,l,u}$  is the  $3 \times 3$ -submatrix of  $C$  where only the rows  $k, l$  and  $u$  and the columns  $k, l$  and  $v$  are maintained; here  $u$  is the  $p^{\text{th}}$  element and  $v$  the  $q^{\text{th}}$  element of the sequence  $1, 2, \dots, k-1, k+1, \dots, l-1, l+1, \dots, r$ .

According to (10) we have to compute

$$\begin{aligned} V_5(C; 1, 2, \dots, 5) = \\ \text{content} \left\{ (x_1, x_2, \dots, x_5) | x_1 < x_2 < \dots < x_5; x_2 \leq t(x_5 - x_1); \sum_{i=1}^5 x_i^2 \leq 1 \right\} \end{aligned} \quad (14)$$

The matrix  $C$  is here

$$\begin{pmatrix} 1 & -\frac{1}{2} & 0 & 0 & \frac{t}{\sqrt{4t^2+2}} \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & \frac{-1}{\sqrt{4t^2+2}} \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & \frac{1}{\sqrt{4t^2+2}} \\ 0 & 0 & -\frac{1}{2} & 1 & \frac{t}{\sqrt{4t^2+2}} \\ \frac{t}{\sqrt{4t^2+2}} & \frac{-1}{\sqrt{4t^2+2}} & \frac{1}{\sqrt{4t^2+2}} & \frac{t}{\sqrt{4t^2+2}} & 1 \end{pmatrix}.$$

It is wellknown ([5] p. 171) that the content of a hyperspherical simplex in  $R^r$  with  $r > 3$  cannot be expressed in terms of elementary functions of the elements of  $C$ . In our case however it is possible to compute the derivative of  $V_5$  as a function of  $t$  with the aid of the two preceding theorems. We have

$$\begin{aligned} \frac{dV_5(C(t); 1, \dots, 5)}{dt} &= \sum_{i=1}^4 \frac{dc_{i5}(t)}{dt} \cdot \frac{\partial V_5(C(t); 1, \dots, 5)}{\partial c_{i5}} = \\ &= \sum_{i=1}^4 \frac{dc_{i5}(t)}{dt} \cdot \frac{1}{\sqrt{1-c_{i5}^2(t)}} \cdot V_3(C; i, 5; 1, \dots, i-1, i+1, \dots, 4) = \\ &= \sum_{i=1}^4 \frac{dc_{i5}(t)}{dt} \cdot \frac{1}{\sqrt{1-c_{i5}^2(t)}} \cdot V_3(C^{(i)}; 1, 2, 3). \end{aligned}$$

The last quantity only involves a three-dimensional spherical simplex, the content of which can be computed. This computation gives us for the density of  $h_2$

$$\begin{aligned} \frac{dF_2(t)}{dt} &= \frac{15}{\pi^2} \frac{1}{2t^2+1} \cdot \\ &\cdot \left[ \frac{1}{\sqrt{3t^2+2}} \left\{ 2\pi - \arccos\left(\frac{-\frac{1}{2}(3t^2+t)}{\sqrt{(2t^2+t+\frac{1}{2})(3t^2+1)}}\right) - \arccos\left(\frac{-\frac{1}{2}(t^2-2t)}{\sqrt{(2t^2+t+\frac{1}{2})(2t^2+2)}}\right) + \right. \right. \\ &- \arccos\left(\frac{-\frac{1}{2}(3t^2+2t+2)}{\sqrt{(3t^2+1)(2t^2+2)}}\right) \left. \right\} + \frac{2t}{\sqrt{4t^2+1}} \left\{ 2\pi - \arccos\left(\frac{-\frac{1}{2}(2t^2+t)}{\sqrt{(2t^2+t+\frac{1}{2})(3t^2+\frac{1}{2})}}\right) + \right. \\ &- \arccos\left(\frac{-\frac{1}{2}(2t^2-t)}{\sqrt{(2t^2+t+\frac{1}{2})(3t^2+1)}}\right) - \arccos\left(\frac{-\frac{1}{2}(4t^2+t+1)}{\sqrt{(3t^2+\frac{1}{2})(3t^2+1)}}\right) \left. \right\} + \end{aligned}$$

$$\begin{aligned}
& -\frac{2t}{\sqrt{4t^2+1}} \left\{ 2\pi - \arccos\left(\frac{-\frac{1}{2}(2t^2-t)}{\sqrt{(3t^2+\frac{1}{2})(2t^2-t+\frac{1}{2})}}\right) - \arccos\left(\frac{-\frac{1}{2}(4t^2-t+1)}{\sqrt{(3t^2+1)(3t^2+\frac{1}{2})}}\right) + \right. \\
& \left. - \arccos\left(\frac{-\frac{1}{2}(2t+t)}{\sqrt{(3t^2+1)(2t^2-t+\frac{1}{2})}}\right) \right\} + \frac{1}{\sqrt{3t^2+2}} \left\{ 2\pi - \arccos\left(\frac{-\frac{1}{2}(3t_2-2t+2)}{\sqrt{(2t^2+2)(3t^2+1)}}\right) + \right. \\
& \left. - \arccos\left(\frac{-\frac{1}{2}(t^2+2t)}{\sqrt{(2t^2+2)(2t^2-t+\frac{1}{2})}}\right) - \arccos\left(\frac{-\frac{1}{2}(3t^2-t)}{\sqrt{(3t^2+1)(2t^2-t+\frac{1}{2})}}\right) \right\} \Big]. \quad (15)
\end{aligned}$$

Some percentage points of the distribution of  $\underline{h}_2$  obtained by numerical integration are

$$\begin{cases} P\{\underline{h}_2 > 0.560\} = 0.025 \\ P\{\underline{h}_2 > 0.735\} = 0.01 \\ P\{\underline{h}_2 > 0.888\} = 0.005. \end{cases}$$

Theoretically it is possible to do a similar computation for  $n = 3$ , but the result seems hardly worth the effort.

This method of computation can also be applied to other quotients of linear combinations of order statistics, e.g. the so-called extremal quotient  $-\underline{x}_{(2n+1)} \cdot \underline{x}_{(1)}^{-1}$ .

## 2. Asymptotic distribution

GNEDENKO [4] proved that if  $a_n$  satisfies

$$\Phi(a_n) = 1 - \frac{1}{n},$$

where  $\Phi$  is the standard normal distribution function, then

$$\lim_{n \rightarrow \infty} P\{a_{2n+1}^{-1} \underline{x}_{(2n+1)} \leq t\} = \begin{cases} 0 & \text{when } t < 1 \\ 1 & \text{when } t > 1. \end{cases}$$

The same relation holds for  $-\underline{x}_{(1)}$ . Hence

$$\lim_{n \rightarrow \infty} P\{2^{-1} a_{2n+1}^{-1} (\underline{x}_{(2n+1)} - \underline{x}_{(1)}) \leq t\} = \begin{cases} 0 & \text{when } t < 1 \\ 1 & \text{when } t > 1. \end{cases}$$

CRAMÉR [3] proved that

$$\lim_{n \rightarrow \infty} P\{2\sqrt{2n+1} \cdot \underline{x}_{(n+1)} \leq t\} = \Phi(t) \quad \text{for all } t.$$

As  $a_n \sim \sqrt{2 \log n}$  for  $n \rightarrow \infty$ , these results imply

$$\lim_{n \rightarrow \infty} P\{8\sqrt{n \log n} \cdot \underline{h}_n \leq t\} = \lim_{n \rightarrow \infty} P\{4\sqrt{2n+1} \cdot \sqrt{2 \log(2n+1)} \cdot \underline{h}_n \leq t\} = \Phi(t)$$

for all  $t$ . But then  $\underline{h}_n$  is asymptotically normal with parameters  $\mu_n = 0$  and  $\sigma_n = (8\sqrt{n \log n})^{-1}$ .

## References

- [1] N. BOUMA en A. VEHMEYER, Percentiles and approximations of the sample median over the sample range for samples of size 3, 5, 7 and 9 from a standard normal distribution. *Statistica Neerlandica* **23** (1969) 235–239.
- [2] R. COURANT, *Integral and differential calculus*, I and II. New York 1936.
- [3] H. CRAMÉR, *The elements of probability theory and some of its applications*. New York 1955.
- [4] B. GNEDENKO, Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Math.* **44** (1943) 423–453.
- [5] A. SARHAN and B. GREENBERG, *Contributions to order statistics*. New York 1962.
- [6] L. SCHLÄFLI, *Theorie der Vielfachen Kontinuität*. *Gesammelte Mathematische Abhandlungen*, Band I. Basel 1950.
- [7] H. VAN DER VAART, The content of some classes of non-euclidian polyhedra for any number of dimensions, with several applications, I and II. *Indag. Math.* **17** (1955) 199–221.