THE MAXIMA OF THE MEAN LARGEST VALUE AND OF THE RANGE

By E. J. Gumbel

Coltimbia University

Summary and Introduction. R. L. Plackett derived the maximum of the ratio of mean range to the standard deviation as function of the sample size, and gave the initial (symmetrical) distribution for which this maximum is actually reached. On the other hand, Moriguti derived the maximum for the mean largest value under the assumption that the distribution from which the maximum is taken is symmetrical. His mean value turned out to be one half of the value given by Plackett.

In the following, these results will be generalized for an arbitrary (not necessarily symmetrical) continuous variate. The mean and the standard deviation of the largest value and the mean range will be given for two distributions: one where the mean largest value is a maximum, and another one where the mean range is a maximum.

Obviously, a mean largest value can exist if and only if the initial mean exists. In addition we postulate in both cases the existence of the second moment.

1. Initial distribution which maximizes the mean largest value. The initial population mean $\bar{x}$, mean square $\overline{x^2}$, variance σ^2, the mean $\overline{x_n}$, mean square $\overline{x_n^2}$, variance σ_n^2 and the generating function $G_n(t)$ of the largest among n independent observations, finally the mean range $\overline{w_n}$ for n observations are written as

$$\bar{x} = \int_0^1 x\,dF, \qquad \overline{x^2} = \int_0^1 x^2\,dF, \qquad \sigma^2 = \overline{x^2} - \bar{x}^2,$$

$$\bar{x}_n = n\int_0^1 xF^{n-1}\,dF,\quad \overline{x_n^2} = n\int_0^1 x^2F^{n-1}\,dF,\quad \sigma_n^2 = n\int_0^1 (x-\bar{x}_n)^2F^{n-1}\,dF, \tag{1.1}$$

$$G_n(t) = n\int_0^1 e^{xt}F^{n-1}\,dF, \qquad \overline{w}_n = n\int_0^1 x\{F^{n-1} - (1-F)^{n-1}\}\,dF,$$

where $x = x(F)$. Corresponding designations will be used for transformed variates.

In order to derive the initial distribution which maximizes the mean largest value $\bar{x}_n$ for given values of the initial mean and standard deviation, we look for the corresponding variate $x(F)$, and put the first variation of

$$\int_0^1 [nxF^{n-1} - \lambda_1x^2 - \lambda_2x]\,dF$$

with respect to x equal to zero. Here λ_1 and λ_2 are constant factors which will take on the role of parameters. The operation leads to

$$nF^{n-1} - 2\lambda_1x - \lambda_2 = 0:$$

whence

$$x = (nF^{n-1} - \lambda_2)/2\lambda_1\,. \tag{1.2}$$

Received 7/30/52, revised 7/23/53.

To obtain the maximum of $\bar{x}_n$ we have to eliminate the parameters λ_1 and λ_2. The mean largest value becomes from (1.1) after integration of (1.2)

$$\bar{x}_n = \frac{1}{2\lambda_1}\frac{n^2}{2n-1} - \frac{\lambda_2}{2\lambda_1}. \tag{1.3}$$

The initial mean obtained for $n = 1$ is

$$\bar{x} = (1 - \lambda_2)/2\lambda_1. \tag{1.4}$$

The initial standard deviation becomes from (1.2) and (1.4)

$$\sigma = \frac{1}{2\lambda_1}\frac{n-1}{\sqrt{2n-1}}. \tag{1.5}$$

Combination of the three preceding equations leads to the mean largest value

$$\bar{x}_n = \bar{x} + \sigma\frac{n-1}{\sqrt{2n-1}}. \tag{1.6}$$

Therefore, the mean largest value for any continuous distribution possessing the first two moments increases more slowly than $\sqrt{n/2}$ times the initial standard deviation.

The initial probability function $F(x, n)$ for which the bound (1.6) is actually reached is from (1.2)

$$F(x, n) = \left(\frac{2\lambda_1 x + \lambda_2}{n}\right)^{1/(n-1)}.$$

The parameters λ_1 and λ_2 are from (1.4) and (1.5)

$$2\lambda_1 = \frac{n-1}{\sigma\sqrt{2n-1}}, \qquad \lambda_2 = 1 - \frac{\bar{x}(n-1)}{\sigma\sqrt{2n-1}};$$

whence

$$F(x, n) = \left(\frac{(n-1)(x-\bar{x})/\sigma}{n\sqrt{2n-1}} + \frac{1}{n}\right)^{1/(n-1)}. \tag{1.7}$$

If we introduce the standardized variate z with zero mean and unit variance,

$$z = (x - \bar{x})/\sigma \tag{1.8}$$

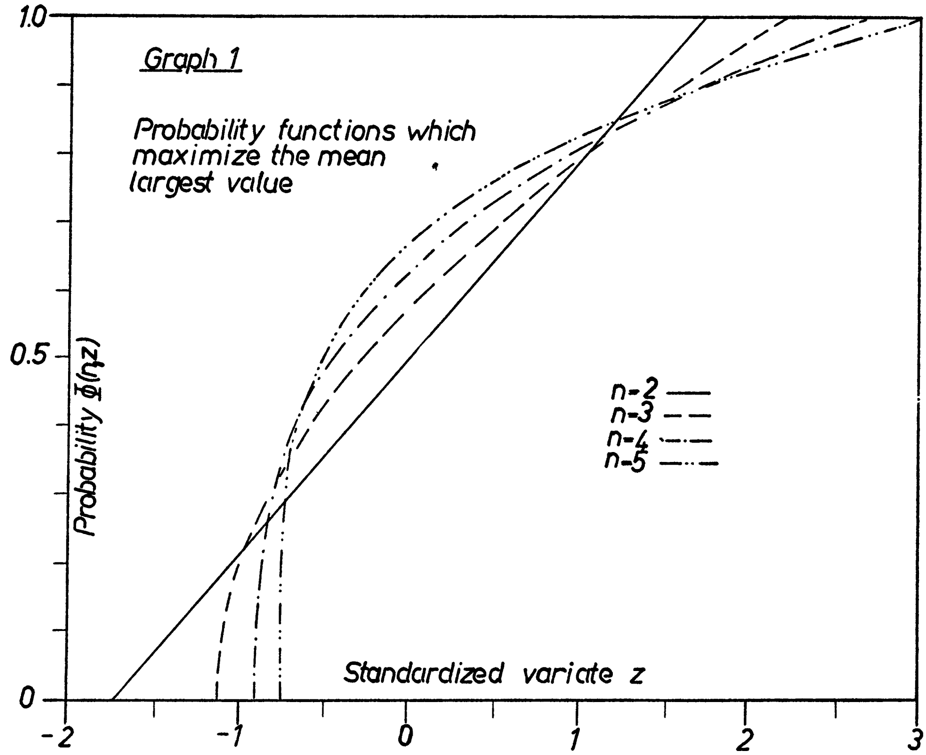
its bounds z_0 and z_ω are for $F(x) = 0$ and $F(x) = 1$

$$z_0 = -\frac{\sqrt{2n-1}}{n-1} \leqq z \leqq \sqrt{2n-1} = z_\omega. \tag{1.9}$$

Therefore the domain of variation spreads with n increasing and the lower bound z_0 approaches zero in the negative domain.

The probability function $\Phi(z, n)$ of the reduced variate z is from (1.7)

$$\Phi(z, n) = \left(\frac{z}{\sqrt{2n-1}}\frac{n-1}{n} + \frac{1}{n}\right)^{1/(n-1)}, \qquad z_0 \leqq z \leqq z_\omega. \tag{1.10}$$



For $n = 2$ the probability function is linear within the domain $-\sqrt{3} < z < \sqrt{3}$. For n increasing the initial median

$$\check{z} = \frac{\sqrt{2n-1}}{n-1}(n2^{1-n} - 1)$$

starts with zero for $n = 2$, is negative for $n \geqq 3$ and diminishes with increasing n to the lower bound z_0. Therefore, for $z < \check{z}$ the probability function approaches a straight line, perpendicular to the z axis, as shown in Graph 1.

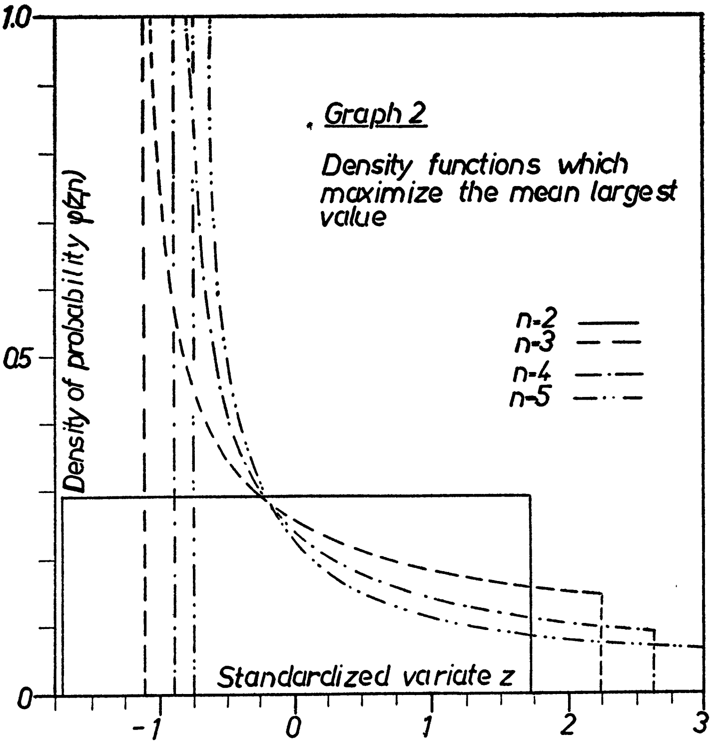
Since, from (1.10) the variate z is

$$z = (n\Phi^{n-1} - 1)\frac{\sqrt{2n-1}}{n-1}, \tag{1.11}$$

the density function $\varphi(z, n)$ obtained from $\varphi(z) = 1/(dz)/(d\Phi)$ is

$$\varphi(z, n) = 1/(n\sqrt{2n-1}\Phi^{n-2}). \tag{1.12}$$

This is uniform for $n = 2$. For $n \geqq 3$, the density is infinite at the lower bound z_0 and decreases uniformly with increasing values of z. At the upper bound the value of the density function $\varphi(z_\omega, n) = 1/n\sqrt{2n-1}$ does not vanish, but is very small. For increasing values of n, the density functions approach more and more parallels to the vertical axis located at z_0, and spread at the same time along the horizontal axis. The probability function $\Phi(z, n)$ and the density func-



tion $\varphi(z, n)$ which maximize the mean largest value are traced in Graphs 1 and 2. It is not to be wondered that a distribution obtained from a queer condition should show strange properties.

2. Extremes. Consider now the extreme values of the initial distribution (1.10). The mean largest reduced value itself is, of course,

$$\bar{z}_n = \frac{n-1}{\sqrt{2n-1}} \tag{2.1}$$

in accordance with (1.6).

The characteristic largest value u_n, which, in previous publications [3], was called the expected largest value defined by $\Phi(u_n) = 1 - 1/n$ becomes from (1.10), even for moderate values of n

$$u_n \sim \frac{n/e - 1}{n - 1} \sqrt{2n - 1}. \tag{2.2}$$

It is smaller than the mean largest value and increases asymptotically as $\sqrt{2n}/e$ that is, more slowly than the mean largest value. The mean square of the reduced largest value is obtained from (1.1) and (1.11) as

$$\overline{z_n^2} = \frac{n(2n-1)}{(n-1)^2} \int_0^1 (n^2\Phi^{3n-3} - 2n\Phi^{2n-2} + \Phi^{n-1})\, d\Phi = \frac{2n^2}{3n-2} - 1.$$

Formula (2.1) leads to the standard deviation σ_n of the largest value

$$\sigma_n = n\sqrt{n/(2n-1)(3n-2)} \tag{2.3}$$

which converges to $\sqrt{n/6}$. Consequently, the coefficient of variation converges to $1/\sqrt{3}$.

In addition to the standard deviation of the largest value, we calculate the mean range for the variate, which shows a maximum of the mean largest value. The mean range $\overline{w}_n$ for the variate z in standard units becomes from (1.1) and (1.11)

$$\overline{w}_n = \frac{n\sqrt{2n-1}}{n-1}\int_0^1 [n\Phi^{2n-2} - \Phi^{n-1} - n\Phi^{n-1}(1-\Phi)^{n-1} + (1-\Phi)^{n-1}]\, d\Phi.$$

Therefore, the mean range is obtained after trivial calculations as

$$\overline{w}_n = \frac{n^2}{(n-1)\sqrt{2n-1}}\left(1 - \frac{(n-1)!^2}{(2n-2)!}\right). \tag{2.4}$$

These values are traced in graph (3) and converge to

$$\overline{w}_n = \sqrt{n/2}. \tag{2.4'}$$

Thus $\overline{w_n}$ converges from above toward $\overline{z_n}$. This is not to be wondered at, since z_0 converges toward zero, and, therefore, $\overline{z_n}$ constitutes a larger and larger part of the range. It follows that the quotients $\sigma_n/\overline{w_n}$ converge toward the same value $1/\sqrt{3}$ as the coefficient of variation. Finally, the quotient $\overline{w_n}/\overline{z_n}$ converges towards unity.

The probability function $\Phi_n(z)$ of the largest value defined by

$$\Phi_n(z) = \Phi^n(z,n) \tag{2.5}$$

becomes from (1.10) and (2.1)

$$\Phi_n(z) = [(1 + z\cdot z_n)/n]^{n/(n-1)}. \tag{2.6}$$

The probability $\Phi_n(\bar{z}_n)$ at the mean largest value converges toward $\frac{1}{2}$. The density function $\varphi_n(z)$ of the largest value obtained from (2.6) and (1.12)

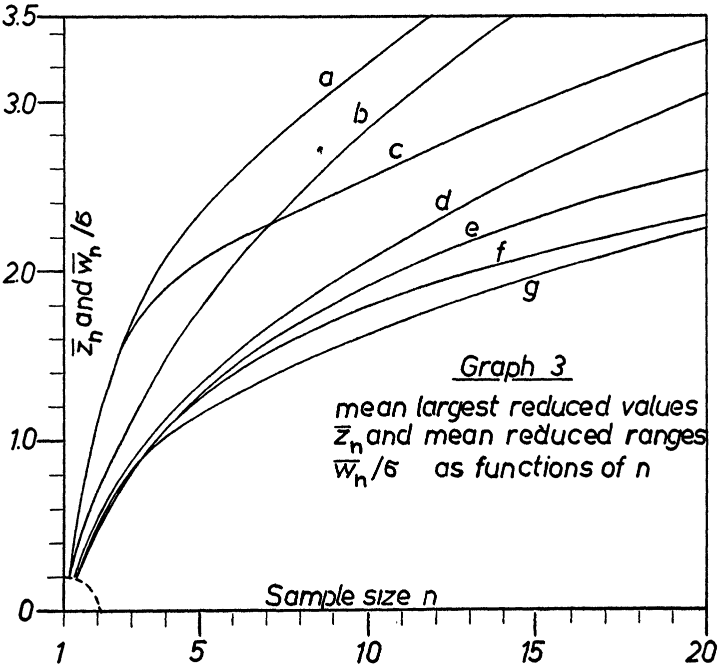
$$\varphi_n(z) = \Phi(z,n)\bar{z}_n/(n-1) \tag{2.7}$$

is equal to the initial probability function reduced by a factor. It increases with z up to z_ω. Hence no mode exists. The median largest value $\check{z}_n$ obtained from (2.6) and (2.1)

$$\check{z}_n = (n2^{1/n-1} - 1)\frac{\sqrt{2n-1}}{n-1} \tag{2.8}$$

is smaller than the mean largest value, converges toward it for n increasing and increases asymptotically as $\sqrt{n/2}$. The probability of the largest value becomes at the characteristic largest value u_n from (2.2) and (2.6)

$$\Phi_n(u_n) \sim e^{-1-1/(n-1)}. \tag{2.9}$$



a) Upper bound for $\bar{w}_n$ (4.7)
b) Upper bound for $(x_n - \bar{x}_o)/s$ (3.6)
c) $\bar{w}_n$ for distribution (1.11) which maximizes $\bar{z}_n$ (2.4)
d) Upper bound for $\bar{z}_n$ (2.1)
e) $\bar{z}_n$ for exponential distribution (3.2)
f) $\bar{z}_n$ for double exponential distribution (3.4)
g) Upper bound $\bar{z}_n$ for symmetrical distribution (4.7)

This probability converges toward $1/e$ which is (cf. [3]), the value of the first asymptotic probability of extremes, at the characteristic largest value. This is an analogy between the probability function of the largest value (2.6) and the asymptotic probability functions of largest values obtained by Fréchet [1], Fisher and Tippett [2]. If we introduce a reduced largest value

$$y = (1 + z\cdot\bar{z}_n)/n = [1 + (x - \bar{x})(\bar{x}_n - \bar{x})/\sigma^2]/n, \tag{2.10}$$

the probability (2.6) of the largest value converges to the simple expression

$$\Phi(y) = y; \qquad 0 \leqq y \leqq 1. \tag{2.11}$$

The rectangular distribution is thus the asymptotic distribution of the extreme value of the variate z, given by (2.6). This case cannot be derived from the stability postulate used by Fréchet, Fisher and Tippett to construct the three asymptotic distributions of extreme values. Inversely, the asymptotic probability function (2.11) is not stable in the sense used by these authors.

3. Comparisons. It is interesting to compare the maxima (2.1) of the mean largest value to the mean largest values obtained from actual asymmetrical distributions. To this end, consider first the exponential distribution

$$f(x) = 1 - F(x) = e^{-x}. \tag{3.1}$$

The generating function (1.1) of the largest value is

$$G_n(t) = \int_0^1 n(1 - F)^{-t} F^{n-1}\, dF; \qquad t < 1.$$

The right side becomes

$$\Gamma(n + 1)\Gamma(1 - t)/\Gamma(n + 1 - t) = 1/\prod_{\nu=1}^{n} (1 - t/\nu).$$

The usual procedure leads to the mean largest value $\bar{x}_n = \sum_1^n 1/\nu$. The mean largest reduced value

$$\bar{z}_n = \sum_2^n 1/\nu \tag{3.2}$$

converges for large n to

$$\bar{z}_n \sim \gamma + \lg n$$

where γ is Euler's number. Some numerical values of (3.2) taken from Karl Pearson's table [5] are traced in graph 3.

As a further example, take the first asymptotic distribution of the largest value [3].

$$F(x) = \exp [-e^{-x}], \tag{3.3}$$

which is closely related to the exponential function. The initial standard deviation σ is $\pi/\sqrt{6}$ and the mean largest value is

$$\bar{x}_n = \bar{x} + \lg n.$$

The standardized mean largest value

$$\bar{z}_n = (\lg n)\sqrt{6}/\pi \tag{3.4}$$

is also traced in graph 3. For small samples, the mean largest values taken from the exponential and double exponential distributions (3.1) and (3.3) are only slightly short of the maxima for any asymmetrical distribution.

The factors $\bar{x}$ and σ in the Theorem (1.6) and the mean largest value $\bar{x}_n$ are population values. We now establish a similar theorem[1] which deals with sample values. Let x_n stand for the observed largest value, $\bar{x}_0$ for the sample mean and s for the sample standard deviation calculated with the usual correction. Then $(x_n - \bar{x}_0)^2 \leqq \sum_{\nu=1}^n (x_\nu - \bar{x}_0)^2$ where the equality holds if, and only if, all observations x_ν are the same. It follows immediately that

$$\frac{(x_n - \bar{x}_0)^2}{n - 1} \leqq s^2 \tag{3.5}$$

or

$$x_n \leqq \bar{x}_0 + s\sqrt{n - 1}. \tag{3.6}$$

Formula 3.5 may be used to control the calculation of the sample standard deviation.

[1] This simplification of a different proof was suggested by the referee of this paper. The author is also obliged to Mr. W. Hoeffding for pointing out a slight inaccuracy in the original draft.

4. The maximum of the mean range. Plackett [6] has calculated the maximum of the mean range as function of the sample size and constructed the initial distribution where the maximum is actually reached. In the following it will be shown that this maximum holds for any continuous variate possessing the first two moments.

The procedure of Section 1 is now used for the mean range $\bar{w}_n$ given in (1.1). To find the unknown probability function F_2 which maximizes $\bar{w}_n$ for constant values of the initial mean and standard deviation the first variation of

$$\int_0^1 [nx\{F_2^{n-1} - (1 - F_2)^{n-1}\} - k_1x^2 - k_2x]\, dF_2$$

with respect to x is to put equal to zero. The operation leads to

$$n\{F_2^{n-1} - (1 - F_2)^{n-1}\} - 2k_1x - k_2 = 0 \tag{4.1}$$

whence

$$x = \frac{n\{F_2^{n-1} - (1 - F_2)^{n-1}\} - k_2}{2k_1}. \tag{4.2}$$

The mean is from (1.1)

$$\bar{x} = -k_2/2k_1 . \tag{4.3}$$

The variance σ^2 is obtained from (1.1) as

$$\sigma^2 = \frac{2n^2(1 - \epsilon_n)}{4k_1^2(2n - 1)} \tag{4.4}$$

where the factor

$$\epsilon_n = (n - 1)!^2/(2n - 2)! \tag{4.5}$$

vanishes as 2^{-2n}. The standardized variate z defined in (1.8) becomes from (4.2) and (4.3)

$$z = \sqrt{\frac{2n - 1}{2(1 - \epsilon_n)}}\,\{F_2^{n-1} - (1 - F_2)^{n-1}\}. \tag{4.6}$$

It follows from (1.1) that the mean range is

$$\bar{w}_n = n\sqrt{\frac{2(1 - \epsilon_n)}{2n - 1}} \tag{4.7}$$

which is Plackett's formula for the maximum of the mean range. From (4.6) it follows that the variate z for which this maximum is reached has a symmetrical distribution. Therefore, the mean of the largest value which is one-half of the value given by (4.7) is the maximum which this mean can reach for any symmetrical distribution. This result was derived by Moriguti [4] from a different approach.

It remains to calculate the standard deviation of the largest value for the distribution (4.6). From (1.1) it follows that the mean square of the largest value is

$$\overline{z_n^2} = \frac{n(2n - 1)(1 - \delta_n)}{2(3n - 2)(1 - \epsilon_n)}$$

where

(4.8) $$\delta_n = (2n - 2)!(n - 1)!/(3n - 3)!$$

converges to zero as $2^{2n}3^{-3n}$. Since from (4.7) $\bar{z}_n^2 = n^2(1 - \epsilon_n)/2(2n - 1)$, the standard deviation of the largest value for z approaches

(4.9) $$\sigma_n \sim (n - 1)\sqrt{n/[2(2n - 1)(3n - 2)]}$$

and for large values of n

(4.9′) $$\sigma_n \sim \tfrac{1}{2}\sqrt{\frac{n}{3}}.$$

The coefficient of variation $\sigma_n/\bar{z}_n$ of the largest value approaches

(4.10) $$\sigma_n/\bar{z}_n \sim 1/\sqrt{3}.$$

Of course, the quotient $\sigma_n/\bar{w}_n$ approaches one-half this amount.

5. Conclusions. It is interesting to compare the asymptotic properties of the reduced values for the two distributions (1.10) and (4.6) which lead to the maximum of the mean largest value and to the maximum of the mean range respectively. The maximum of the mean largest value for an asymmetrical distribution increases as $\sqrt{n/2}$, while for symmetrical distributions it increases as $\sqrt{n}/2$. The maximum of the mean largest value for an asymmetrical distribution is 41 per cent larger than for a symmetrical one. Inversely, the asymmetrical distribution for which the mean largest value is a maximum yields a mean range which is only 71 per cent of the maximum of the mean range. However, the standard deviation of the largest value for the distribution with maximum mean range is only 71 per cent of the corresponding value for the distribution with maximum mean largest value. The coefficient of variation of the extreme value for the two distributions is the same. The quotients $\bar{w}_n/\sigma_n$ and $\bar{w}_n/\bar{z}_n$ for the distribution with maximum mean largest value are one-half of the values for the other distribution.

REFERENCES

[1] M. Fréchet, "Sur la loi de probabilité de l'écart maximum," *Ann. Soc. Polon. Math.*, Vol. 6 (1927), pp. 93–122.

[2] R. A. Fisher and L. H. C. Tippett, "Limiting forms of the frequency distribution of the smallest and the largest member of a sample," *Proc. Cambridge Philos. Soc.*, Vol. 24 (1928), pp. 180–190.

[3] E. J. Gumbel, "The return period of flood flows," *Ann. Math. Stat.*, Vol. 12 (1941), pp. 163–190.

[4] Sigeiti Moriguti, "Extremal properties of extreme value distributions," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 523–536.

[5] K. Pearson and M. V. Pearson, "On the mean character and variance of a ranked individual, and on the mean and variance of the intervals between ranked individuals, II," *Biometrika*, Vol. 24 (1932), pp. 203–279.

[6] R. L. Plackett, "Limits of the ratio of mean range to standard deviation," *Biometrika*, Vol. 34 (1947), pp. 120–122.